NOTES 1

SCOT ADAMS

Contents

Date: Printout date: May 24, 2018.

1. Class 1 on 5 September 2017, Tu of Week 1

Some abbreviations:

&	:=	and					
$\Leftrightarrow$	:=	iff	:=	if and only if			
$\Rightarrow$	:=	implies	$\neq$	therefore	=:	$\therefore$	

COURSE GOAL: Present math as absolute truth.
We present **Truth Tables**.
We defined NOT X via a truth table.
We defined $X\&Y$, X or Y, $X \Rightarrow Y$, $X \Leftrightarrow Y$ via truth tables.

THEOREM 1.1. *Let A be a proposition. Then: A or (NOT A).*

The last column of the truth table below shows that, in call cases, the proposition A or (NOT A) is true. When a proposition is always true, we call it a theorem.

Proof.

A	NOT A	A or (NOT A)
T	F	T
F	T	T

$\square$

THEOREM 1.2. *Let A and B be propositions. Then:*

$$(\; A \; \Rightarrow B \;) \qquad \Leftrightarrow \qquad (\; (\mathit{NOT}\, A) \; \mathit{or} \; B \;).$$

Proof. Let $C := (A \Rightarrow B)$ and $D := (\text{NOT}\, A)$. We wish to show:

$$C \qquad \Leftrightarrow \qquad (D \text{ or } B).$$

Let $E := (D \text{ or } B)$. We wish to show: $C \Leftrightarrow E$.

A	B	C	D	E	$C \Leftrightarrow E$
T	T	T	F	T	T
T	F	F	F	F	T
F	T	T	T	T	T
F	F	F	T	T	T

$\square$

Assigned HW#1.

THEOREM 1.3. *Let A, B and C be propositions. Then:*

$$(\; A \; \Rightarrow (\; B \,\&\, C \;) \;) \qquad \Leftrightarrow \qquad (\; (\; A \; \Rightarrow B \;) \;\&\; (\; A \Rightarrow C \;) \;).$$

Proof. Let $D := (B\&C)$, $E := (A \Rightarrow B)$ and $F := (A \Rightarrow C)$. We wish to show: $[A \Rightarrow D] \Leftrightarrow [E\&F]$. Let $G := (A \Rightarrow D)$ and $H := (E\&F)$. We wish to show: $G \Leftrightarrow H$.

A	B	C	D	E	F	G	H	$G \Leftrightarrow H$
T	T	T	T	T	T	T	T	T
T	T	F	F	T	F	F	F	T
T	F	T	F	F	T	F	F	T
T	F	F	F	F	F	F	F	T
F	T	T	T	T	T	T	T	T
F	T	F	F	T	T	T	T	T
F	F	T	F	T	T	T	T	T
F	F	F	F	T	T	T	T	T

$\square$

Assigned HW#2 and HW#3.

Homework 3 asks for a proof that, for any propositions P and Q, we have: $[P\&(P \Rightarrow Q)] \Rightarrow Q$. I sometimes call this the "**bad form**" theorem; in practice, it's bad form to use it, because: If, in a proof, we write $P \Rightarrow Q$, it conveys the thought, "If we could somehow just show that P is true, then we would know Q is true", and it's a little odd to say this if we already know that P is true. We'll see how this kind of situation can come up, and, when it does, we'll talk about how to rewrite the proof.

Some abbreviations:

$\forall$	:=	for all/any/every	"universal quantifier"
$\exists$	:=	there exists	"existential quantifier"
$\exists!$	:=	$\exists!$:=	there exists a unique
s.t.	:=	such that	
$\in$	:=	is an element of	
$\subseteq$	:=	is a subset of	

We will not develop all of the rules, conventions and axioms of mathematics here; this course is not an *ab ovo* development of mathematics. Nevertheless, certain axioms deserve to be highlighted. For example, we'll call the following the **Axiom of Equality**:

Axiom: $\forall x,\ x = x$.

The above axiom tells us that, for every object x in our mathematical universe, we have $x = x$. This begs the quesion: What are all the objects that we will consider in this course? If this were a course in foundations of mathematics, we would answer this in a precise fashion, but here we will just give examples:

- every complex number is an object,
- ∞ and $-\infty$ are objects,
- every set is an object and
- ☹ is an object.

NOTE TO SELF: Next year, let's also have "classes" as objects. The "common objects" will be complex numbers, $\pm\infty$ and sets. The "exotic objects" will be ☹ and classes. Every element of a class must be a common object. Every set is a class, but some classes are not sets. A class that is not a set is a "big class". A "small class" is the same as a set. A set that is contained in a class will be called a "small subclass" of the class. Classes will be denoted by : $\{\cdots\}$. For any well-formed condition $\phi(x)$ on x, $\{x \,|\, \phi(x)\}$ is the class of all common objects x such that $\phi(x)$. Can use any unbound variable in place of x. For any well-formed condition $\phi(x)$ on x, for any class C,

$$\{x \in C \,|\, \phi(x)\} \;:=\; \{x \,|\, (x \in C)\&(\phi(x))\}$$

denotes the class of all objects $x \in C$ such that $\phi(x)$. For any well-formed condition $\phi(x)$ on x, for any set S, $\{x \in S \,|\, \phi(x)\}$ is a set. Define $\{\text{comob}\} := \{x \,|\, x = x\}$; then $\{\text{comob}\}$ is the class of all common objects. Also, $\{\text{sets}\} := \{A \,|\, A \text{ is a set}\}$ is the class of all sets. Also, $\{\text{singleton sets}\}$ is the class of all singleton sets, *i.e.*, sets with exactly one element. Also, $\{\text{nonempty sets}\} := \{A \,|\, A \text{ is a nonempty set}\}$ is the class of all nonempty sets.

By Russell's Paradox, $\{A \,|\, (A \text{ is a set}) \text{ and } (A \notin A)\}$ is a not a set. So it's a class that's not a small class. That is, it's a big class. Thus $\{\text{sets}\}$ and $\{\text{comob}\}$ are also big classes.

A "class-function" is a class of ordered pairs satisfying the vertical line test. For example, $\bigcup$ is a class-function and

$$\bigcup \;:\; \{\text{sets of sets}\} \;\to\; \{\text{sets}\}.$$

By "function", we mean a class-function whose domain is a set. By axiom, the image of a function is also a set.

Let f be a function. The notation $f : A \to B$ is okay if A is a set and B is class, but, since domain of a function must be a set, we see that A cannot be a big class. The notation $f : A \dashrightarrow B$ is okay if A and B are both big classes, but keep in mind that $\mathrm{dom}[f]$ and $\mathrm{im}[f]$ are both sets, so $\mathrm{dom}[f]$ will be a small subclass of A and $\mathrm{im}[f]$ will be a small subclass of B.

The Axiom of Choice asserts: $\forall$set D, $\forall S : D \to \{\text{nonempty sets}\}$, $\exists$function $c : D \to \{\text{lobjs}\}$ s.t., $\forall j \in D$, $c_j \in S_j$. That is, "every set valued function admits a choice function".

Or, even easier: There is a class function

$$\text{CHOOSE} \;:\; \{\text{nonempty sets}\} \;\to\; \{\text{common objects}\}$$

such that, for any set S, $\text{CHOOSE}(S) \in S$. So, here, ELT gives us the unique element from any singleton set, and CHOOSE picks some element from any nonempty set. We could even define ELT to be the restriction of CHOOSE to the class {singleton sets}. Maybe we might use Ξ_S or $\Xi(S)$ instead of $\text{CHOOSE}(S)$, just to save writing. Then, in the last paragraph, we can simply define $c_\bullet$ by $c_j = \Xi(S_j)$.
END OF NOTE TO SELF

The last object in the list above, ☹, is a frownie face. It will play a special role in our course and is nontraditional; most courses would omit it. For us, whenever we have a computation that does not lead to an answer, like $0/0$, we will define the answer to be ☹. Moreover, we assert that ☹ is NOT a set and is not even allowed to be an element of a set. That is, in our course, we have the nontraditional

Axiom: $\forall$set S, [(☹ $\neq S$) & (☹ $\notin S$)].

To some logic purists, every quantifier should be followed by a variable and the quantified clause ends there. To them "$\forall$ set S" is not allowed, and they would rewrite this as:

Axiom: $\forall S$, $\Big($ [S is a set] $\Rightarrow$ [(☹ $\neq S$) & (☹ $\notin S$)] $\Big)$.

There are reasons to adopt such a view of quantification, but the level of increased notational complexity makes it infeasible for us; we will not be purists.

This now begs the question of what objects we DO allow to live in a set, if ☹ is *verboten.* Again, without all the work of setting up foundational mathematics, it's difficult to give a precise answer to this question, but, in essence, any mathematical object can be put into a set, *except* ☹. For example, each of the following is a set, and is therefore an object:

- the set $\mathbb{N} := \{1, 2, 3, \ldots\}$ of positive integers,
- the set $\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\}$ of semipositive (a.k.a. nonnegative) integers,
- the set $\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ of integers,
- the set $\mathbb{Q} := \{m/n \,|\, m \in \mathbb{Z}, n \in \mathbb{N}\}$ of rational numbers,
- the set $\mathbb{R}$ of real numbers (in 1-1 correspondence with the points of any coordinatized line),
- the set $\mathbb{R}^* := \mathbb{R} \cup \{\infty, -\infty\}$ and
- the set $\mathbb{C}$ of complex numbers (in 1-1 correspondence with the points of any complex coordinatized line).

We have the following inclusions

$$\mathbb{N} \quad \subseteq \quad \mathbb{N}_0 \quad \subseteq \quad \mathbb{Z} \quad \subseteq \quad \mathbb{Q} \quad \subseteq \quad \mathbb{R}.$$

We also have $\mathbb{R} \subseteq \mathbb{R}^*$ and $\mathbb{R} \subseteq \mathbb{C}$. Note, however, that $\mathbb{R}^*$ and $\mathbb{C}$ are inclusion-incomparable, meaning that neither is a subset of the other:

$$(\ \mathrm{NOT}(\mathbb{R}^* \subseteq \mathbb{C})\ ) \quad \& \qquad (\ \mathrm{NOT}(\mathbb{C} \subseteq \mathbb{R}^*)\ ).$$

For any sets A and B, by $A \nsubseteq B$, we will mean: $\mathrm{NOT}(A \subseteq B)$. Following this, we write: $\mathbb{R}^* \nsubseteq \mathbb{C}$ and $\mathbb{C} \nsubseteq \mathbb{R}^*$.

For any two objects a and b, by $a \neq b$ we mean: $\mathrm{NOT}(a = b)$. For any object a, for any set S, by $a \notin S$, we mean: $\mathrm{NOT}(a \in S)$.

We will not develop the four basic operations of addition, subtraction, multiplication, division (denoted, repectively, $+$, $-$, $\cdot$ and $/$) on $\mathbb{C}$. Again, this is not a foundational, or *ab ovo*, course, and we don't have time to set up all the necessary definitions and basic theorems of such a course. Instead, we will simply assume that you understand how these four operations work; we rely on your intuition and earlier education. One caveat: In most courses, certain computations, like $0/0$ are simply said to be undefined; for us they will be set to ☹. We also extend addition, subtraction, multiplication, division to the set $\mathbb{R}^*$ of extended reals. Here are some results that we will not prove, and will simply assume as basic knowledge:

- $2 + 2 = 4$,
- $1/0 =$ ☹,

- $0/0 = ☹$,
- $\infty + \infty = \infty$,
- $-\infty - \infty = -\infty$,
- $\infty - \infty = ☹$,
- $-\infty + \infty = ☹$,
- $(0.00001) \cdot \infty = \infty$,
- $(-0.00001) \cdot \infty = -\infty$,
- $0 \cdot \infty = ☹$,
- $0/\infty = 0$,
- $1000/\infty = 0$,
- $\infty/\infty = ☹$,
- $\infty/(-\infty) = ☹$.

If you have any questions about any of these, or about any other arithmetic computations in $\mathbb{C}$ or in $\mathbb{R}^*$, please feel free to contact me.

We will also assume that you know the basic properties of addition, subtraction, multiplication and division on $\mathbb{C}$ and on $\mathbb{R}^*$. Here are some more results that we will not prove, and will simply assume as basic knowledge:

(1) $\forall x, y \in \mathbb{R},\ x + y = y + x$.
(2) $\forall x \in \mathbb{R},\ x^2 \neq -1$.
(3) $\forall x \in \mathbb{Q},\ x^2 \neq 2$.
(4) $\forall x \in \mathbb{R}^*,\ x/0 = ☹$.
(5) $\forall x \in \mathbb{R},\ x/\infty = x/(-\infty) = 0$.
(6) $\forall x, y \in \{\infty, -\infty\},\ x/y = ☹$.

The logic purist would rewrite these as:

(1) $\forall x, \forall y,\ [(x, y \in \mathbb{R}) \Rightarrow (x + y = y + x)]$.
(2) $\forall x,\ [(x \in \mathbb{R}) \Rightarrow (x^2 \neq -1)]$.
(3) $\forall x,\ [(x \in \mathbb{Q}) \Rightarrow (x^2 \neq 2)]$.
(4) $\forall x,\ [(x \in \mathbb{R}^*) \Rightarrow (x/0 = ☹)]$.
(5) $\forall x,\ [(x \in \mathbb{R}) \Rightarrow (x/\infty = x/(-\infty) = 0)]$.
(6) $\forall x, \forall y,\ [(x, y \in \{\infty, -\infty\}) \Rightarrow (x/y = ☹)]$.

Note that we can extend (1) above to $\mathbb{C}$ and to $\mathbb{R}^*$. That is:

$$(\forall x, y \in \mathbb{C},\ x + y = y + x) \quad \& \quad (\forall x, y \in \mathbb{R}^*,\ x + y = y + x).$$

Note that (6) above is equivalent to

$$\infty/\infty \ = \ (-\infty)/\infty \ = \ \infty/(-\infty) \ = \ (-\infty)/(-\infty) \ = \ ☹.$$

We will also assume you know basic properties of $<$, $\leqslant$, $>$, $\geqslant$ on $\mathbb{R}^*$. Here are some more results that we will not prove, and will simply assume as basic knowledge:

(1) $-2 < 1, \quad 5 > 3, \quad 5 \geqslant 3, \quad 5 \geqslant 5, \quad -\infty < \infty, \quad \infty \leqslant \infty.$
(2) $\forall x \in \mathbb{R}^*,\ x^2 \geqslant 0.$
(3) $\forall x \in \mathbb{R}^*,\ -\infty \leqslant x \leqslant \infty.$
(4) $\forall x \in \mathbb{R},\ -\infty < x < \infty.$
(5) $\forall a, b, y, z \in \mathbb{R},\ [\ (\ (\,a < b\,)\ \&\ (\,y < z\,)\ )\ \Rightarrow\ (\,a + y < b + z\,)\ ].$

For some aspects of set theory, we will rely on your prior knowledge. In this course, we will not prove, for example, that there exists a unique set whose only elements are 1 and 2. That is, we will not prove

$$\exists!\text{ set } S \text{ s.t. } [\ \forall x,\ (\ [\ x \in S\ ]\ \Leftrightarrow\ [\ (x = 1) \text{ or } (x = 2)\ ]\ )\ ].$$

Instead, we rely on the reader to know that such a set exists, is unique, and is denoted $\{1, 2\}$. Similarly, there's a set $\{3, 4\}$ and we will simply expect you to know that it exists and has only two elements, namely 3 and 4. Every set is an object, so $\{1, 2\}$ and $\{3, 4\}$, being sets, are objects. We can now put those two objects together to form a new set $\{\{1, 2\}, \{3, 4\}\}$. Again, we will not prove that this set exists, but it has exactly two elements, each of which is a set of numbers. Thus, $\{\{1, 2\}, \{3, 4\}\}$ is a "set of sets of numbers". In this course, if a variable is a number, it will typically be denoted by a small roman letter, like a or s or x. If a variable is a set, we would usually use a capital roman letter, like A or S or X. For a set of sets of numbers, I will usually use a script capital roman letter, like $\mathcal{A}$ or $\mathcal{S}$ or $\mathcal{X}$. Thus, if I want to say that there's a unique set whose elements are exactly $\{1, 2\}$ and $\{3, 4\}$, I would write

$$\exists!\text{ set } \mathcal{S} \text{ s.t. } [\ \forall X,\ (\ [\ X \in \mathcal{S}\ ]\ \Leftrightarrow\ [\ (X = \{1, 2\}) \text{ or } (X = \{3, 4\})\ ]\ )\ ].$$

Formally, we can change the variables and write

$$\exists!\text{ set } S \text{ s.t. } [\ \forall x,\ (\ [\ x \in S\ ]\ \Leftrightarrow\ [\ (x = \{1, 2\}) \text{ or } (x = \{3, 4\})\ ]\ )\ ],$$

without affecting the meaning, but it often helps the reader if we make a few conventions about which alphabets we use for which kinds of objects. In this class, integers are typically denoted by small roman letters between i and n, although we will sometimes need more than six integer variables, so we'll have to make exceptions.

We will give a few axioms of set theory, but will not try to be comprehensive. Most mathematicians follow a kind of axiomatic set theory called "**ZFC**", or "**Zermelo-Fraenkel with Choice**". In this class, we'll follow a slight variant of ZFC, which could possibly be called "ZFC with a frownie and happy sets". By this, I mean that, while we allow ☹ as an object, it is *NOT* a set, and, in fact, it's not allowed even to be an element of any set. That is, while we allow sadness in our formulas, we don't allow any sadness to infect our sets. More precisely, one of the axioms that we'll be using says:

$$\forall \text{set } S, \; [\; (☹ \neq S) \text{ and } (☹ \notin S) \;].$$

A logic purist would rewrite this as:

$$\forall S, \; (\; [\; S \text{ is a set }] \; \Rightarrow \; [\; (☹ \neq S) \text{ and } (☹ \notin S) \;] \;).$$

So ☹ is an object, but not a set. On the other hand keep in mind that every set is an object.

Even though ☹ is never an element of a set, we do allow ourselves to enclose ☹ in braces, *e.g.*, $\{1, 2, 3, ☹\}$; however anytime we see ☹ in between braces, the result is equal to ☹. So, in the notation that we develop in this course, we have: $\{1, 2, 3, ☹\} = ☹$. Thus, contrary to appearance, $\{1, 2, 3, ☹\}$ is *NOT* a set. Similarly, $\{\{1, 2\}, \{3, 4, ☹\}\} = ☹$, and so the equation

$$\{\,\{1, 2\}\,,\,\{3, 4, ☹\}\,\} \quad = \quad \{\,1,\; 2,\; 3,\; ☹\},$$

is true, even if it looks strange.

We will come back to talk more about set theory later, but we first turn our attention to general issues about proofs. Since Math is Truth, it follows that Math is error-free. However, most of us have little experience with error-free thought in our lives. An exception occurs when we play games, where error can lead to defeat. Consequently, playing games like chess or checkers or even tic-tac-toe leads naturally to a mode of rigorous thought. Playing off of that, I will, in this course, often discuss theorems as games.

For example, here's a theorem that is very much in the spirit of a real analysis course:

THEOREM 1.4. *$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\; 0 < x < 2\delta \;] \quad \Rightarrow \quad [\; x + x^2 < \varepsilon \;].$$

The "$\varepsilon > 0$" indicates both that $\varepsilon \in \mathbb{R}$ and that $\varepsilon > 0$. Similarly, "$\delta > 0$" indicates both that $\delta \in \mathbb{R}$ and that $\delta > 0$. The compound inequality "$[0 < x < 2\delta]$" is shorthand for $[(0 < x) \text{ and } (x < 2\delta)]$.

Theorem 1.4 is an example of a "triply-quantified" result, because it has three quantifiers ($\forall$, then $\exists$, then another $\forall$) appearing in it. This kind of result would become unreadable if we were to write like logic purists. It's even difficult to understand (much less prove!) as written.

In class, we tried to clarify Theorem 1.4 by considering a related game. In this game, you make the first move by choosing and telling me a positive number which we call ε. I then choose a $\delta > 0$. You then choose $x \in \mathbb{R}$. We then check whether the impliaction

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ]$$

is TRUE. If it is, I win. Otherwise, you win.

In class, someone chose $\varepsilon = 0.1$. I chose

$$\delta \quad = \quad 0.000000000000000001 \quad = \quad 10^{-18}.$$

I made the point that, if you choose, say $x = 5$, then you lose immediately becuase, with that choice of x, the compound inequality $0 < x < 2\delta$ would be FALSE, and so the implication

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ]$$

would be TRUE. In order to have a chance, you have to choose an $x \in \mathbb{R}$ such that $0 < x < 0.000000000000000002$. Someone chose

$$x \quad = \quad 0.0000000000000000019.$$

We then verified that the inequality $x + x^2 < \varepsilon$ is TRUE. That is,

$$0.0000000000000000019 \ + \ 0.0000000000000000019^2 \quad < \quad 0.1.$$

Since the inequality $x + x^2 < \varepsilon$ is TRUE, it follows that the implication

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ]$$

is TRUE, and I win.

I offered to play again, but no one took me up on it. We agreed that THE GAME IS RIGGED! That is, we agreed that Theorem 1.4 is correct. If you disagree, please get in touch with me, and we'll play.

Taking the perspective that theorems are games, to be successful in this course, you'll need three skills:

(1) Finding plausible strategies.

(2) Proving to yourself that they win.
(3) Communicating the strategy and proof.

Each of these is quite difficult. Also, I can attest that students are often frustrated when they master (1) and (2), but not (3), especially if the standards of communication are not made clear. I want to mention that, in my own work, many has been the time that I finished (1) and (2) to my satisfaction, but then, in attempting (3), found an error in my thinking. So (3) is crucial to the mathematical process.

For today, we focus on (1) and (2); we will talk about (3) in the next class. Everything from here to the end of this class should be thought of as "scratch work". We will not be following carefully any hard and fast rules of exposition for today.

The game of Theorem 1.4 seems too hard. Let's simplify to

THEOREM 1.5. *$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x < \varepsilon\ ].$$

Then only difference between Theorem 1.4 and Theorem 1.5 is that "$x + x^2$" is changed to "x". For this new game, someone suggested the strategy $\delta := \varepsilon/2$. To prove to myself that this works, I noted that $2\delta = \varepsilon$, and then wrote down

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ 0 < x < \varepsilon\ ],$$

and then erased "$0 <$" on the RIGHT side of $\Rightarrow$, and felt persuaded that the strategy will work. That is, if you and I play the game, if I use the strategy $\delta := \varepsilon/2$, and if you choose an x satisfying $0 < x < 2\delta$, then your x will automatically satisfy $x < \varepsilon$. So either ($[0 < x < 2\delta]$ will be FALSE) or ($[x < \varepsilon]$ will be TRUE). In either case, ($[0 < x < 2\delta] \Rightarrow [x < \varepsilon]$) will be TRUE, and I win.

Next, we looked at

THEOREM 1.6. *$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\ 0 < x < \delta\ ] \quad \Rightarrow \quad [\ x^2 < \varepsilon\ ].$$

Here, we replaced "2δ" by "δ", and, also, we're using "x^2" instead of "$x + x^2$" or "x". For this new game, someone suggested the strategy $\delta := \sqrt{\varepsilon}$. To prove to myself that this works, I wrote down

$$[\ 0 < x < \delta\ ] \quad \Rightarrow \quad [\ 0^2 < x^2 < \delta^2\ ].$$

I also made the point that, if you square an inequality, you need to be sure that every item you're squaring is semipositive (a.k.a. nonnegative). Since 0 is semipositive and since $0 < x < \delta$, it follows that 0 and x and δ are all three semipositive, and so it's all right to square the inequality. Remember that, to take an EVEN power of an inequality, each quantity must be semipositive. By contrast, you can take ODD powers, even if some or all of the quantities are negative. I then noted that $\delta^2 = \varepsilon$. I then changed "0^2" to "0", and changed "δ^2" to "ε", obtaining the implication

$$[\ 0 < x < \delta\ ] \quad \Rightarrow \quad [\ 0 < x^2 < \varepsilon\ ].$$

I then erased "$0 <$" on the right side of $\Rightarrow$, and felt persuaded that the strategy will work. it's fine to put a "proof" like this down on scratch paper, but you shouldn't turn in work like this. (We will explain why later, but we're only working on (1) and (2) now.)

Next, we looked at

THEOREM 1.7. *$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x^2 < \varepsilon\ ].$$

This is the same as Theorem 1.6, except that we changed "δ" back to "2δ". Someone suggested the strategy $\delta := \sqrt{\varepsilon}/2$. To prove to myself that this works, I noted that $(2\delta)^2 = \varepsilon$, and then wrote down

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ 0 < x^2 < \varepsilon\ ],$$

and then erased "$0 <$" on the right side of $\Rightarrow$, and felt persuaded that the strategy will work.

The goal is to get Theorem 1.4, which has the inequality "$x + x^2 < \varepsilon$". Note that

$$\Big[\ (x < \varepsilon/2) \ \text{ and } \ \big(x^2 < \varepsilon/2\big)\ \Big] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

So, if I can force you to choose an x that is so small that

$$(x < \varepsilon/2) \ \text{ and } \ \big(x^2 < \varepsilon/2\big)\,,$$

then I'll win. I therefore focus on replacing ε by $\varepsilon/2$ in Theorem 1.5 and Theorem 1.7, obtaining:

THEOREM 1.8. *$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x < \varepsilon/2\ ].$$

THEOREM 1.9. *$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x^2 < \varepsilon/2\ ].$$

I need strategies for Theorem 1.8 and for Theorem 1.9. I know by experience that, if I take the strategies for Theorem 1.5 and for Theorem 1.7, and replace ε by $\varepsilon/2$ in each of them, then I'll obtain the desired strategies.

The strategy for Theorem 1.5 was $\delta := \dfrac{\varepsilon}{2}$. So a good strategy for Theorem 1.8 should be $\delta := \dfrac{\varepsilon/2}{2}$. That is, we try $\delta := \dfrac{\varepsilon}{4}$. Noting that $2\delta = \varepsilon/2$, I wrote

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ 0 < x < \varepsilon/2\ ],$$

and then erased "0 <" on the RIGHT side of $\Rightarrow$, and felt persuaded that the strategy will work.

The strategy for Theorem 1.7 was $\delta := \dfrac{\sqrt{\varepsilon}}{2}$. So a good strategy for Theorem 1.9 should be $\delta := \dfrac{\sqrt{\varepsilon/2}}{2}$. Noting that $(2\delta)^2 = \varepsilon/2$, I wrote

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ 0 < x^2 < \varepsilon/2\ ],$$

and then erased "0 <" on the RIGHT side of $\Rightarrow$, and felt persuaded that the strategy will work.

Finally, we returned to the original problem, of Theorem 1.4. The implication in that result reads

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

In playing Theorem 1.4, I try to force a win by forcing you to choose x so small that both $x < \varepsilon/2$ and $x^2 < \varepsilon/2$. To do this, I look at the last two strategies (for Theorem 1.8 and for Theorem 1.9), which involved the expressions $\dfrac{\varepsilon}{4}$ and $\dfrac{\sqrt{\varepsilon/2}}{2}$. With experience as my guide, I settled on a strategy: $\delta := \min\left\{\dfrac{\varepsilon}{4}, \dfrac{\sqrt{\varepsilon/2}}{2}\right\}$. That is, my strategy is: compute both $\dfrac{\varepsilon}{4}$ and $\dfrac{\sqrt{\varepsilon/2}}{2}$, and then let δ be the smaller of those two numbers.

We played the game with me using that strategy. Someone suggested $\varepsilon = 42$. I computed $\delta := \min\left\{\dfrac{42}{4}, \dfrac{\sqrt{42/2}}{2}\right\} = \dfrac{\sqrt{21}}{2}$. I noted that

$\delta \leqslant \dfrac{\varepsilon}{4}$ and that $\delta \leqslant \dfrac{\sqrt{\varepsilon/2}}{2}$, and so

$$2\delta \leqslant \varepsilon/2 \qquad \text{and} \qquad 2\delta \leqslant \sqrt{\varepsilon/2}.$$

I don't remember what x you chose, but it satsified $0 < x < 2\delta$, and so

$$0 < x < \varepsilon/2 \qquad \text{and} \qquad 0 < x < \sqrt{\varepsilon/2}.$$

Leaving the first part as is, and squaring the second part, we get

$$0 < x < \varepsilon/2 \qquad \text{and} \qquad 0 < x^2 < \varepsilon/2.$$

So, with the x that you gave me, we calculated:

$$x \ + \ x^2 \quad < \quad (\varepsilon/2) \ + \ (\varepsilon/2) \quad = \quad \varepsilon.$$

And so it came to pass that . . . I won.

Enough with scratch work. In the next class, we focus on communicating this strategy and proof for Theorem 1.4. We will follow the many strict rules from our exposition handout.

2. Class 2 on 7 September 2017, Th of Week 1

We reformulate theorems as games and, seek to:

(1) Find plausible strategies.
(2) Prove to ourselves that they win.
(3) Communicate the strategy and proof to others.

In the last class, we made a good start at discussing (1) and (2), in the context of

THEOREM 1.4. $\forall \varepsilon > 0$, $\exists \delta > 0$ *s.t.*, $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

We now discuss (3), the communication part of this course. The most common way students lose credit in their first proofs is by failing to follow the **Cardinal Binding Rule**:

You must **bind** a variable before you use it.

In any proof, *ANY* time you use a variable, you *MUST* be able to tell me where that variable is **bound**, and the binding must happen *before* the variable is used. If you can't point to where the binding happens, then you lose credit. Some bindings are temporary, and only last until the end of the sentence in which they appear. In that case, you can't use the variable after the sentence, unless it gets rebound somehow.

The past participle of "to bind" is "**bound**"; it is NOT "bounded". After you bind a variable, it becomes bound, NOT bounded. Confusion arises because "to bound" is another verb used frequently in mathematics, and is quite different from "to bind". The past participle of "to bound" is "bounded". After you *bound* a variable, it becomes bounded. In this lecture, we will bound no variables; we only bind them. So, in this lecture, no variables become bounded; they become bound.

One of my pet peeves is the common confusion between "bound" and "bounded". Mostly, in this course, you should say "bound", but, if, in some future lecture, I want to stress that some variable is trapped between two real numbers (*e.g.*, if I have a line in a proof that reads "$-1000 < x < 1000$"), then it's reasonable to say that it's bounded (*e.g.*, I might say "x is bounded").

Free is the opposite of bound. To say that a variable is **free** is to say that it is unbound.

In Theorem 1.4, we discussed the binding of ε, δ and x. We noted that by (5) of the exposition handout, all three of these variables become bound before use. However, this binding is only temporary, and lasts only until the end of the sentence. So, when we begin the proof, all three are *free* variables. Therefore, you cannot begin the proof with a statement like: "We know that $\varepsilon > 0$". If you do that, you lose some credit because the ε in that statement is free, and you can NEVER use a variable that's free.

Instead, we examine the statement we are trying to prove. It reads:

$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

Note that it starts with "$\forall$". We therefore go to (10) of the exposition handout. This is the **Want: $\forall$ template**. The instructions in (10) say that we should begin our proof as follows:

Proof:
Given $\varepsilon > 0$.
Want: $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

By (2) of the exposition handout, "Given" is a permanent binding word, so ε is permanently bound. We are therefore able to use it all the way to the end of the proof-section that we are in. We'll talk later about how some proofs are broken into sections. In our present proof, there is only one section, so ε is bound until the end of the proof.

Next examine the statement we are trying to prove. It reads

$\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

Note that it starts with "$\exists$". We therefore go to (11) of the exposition handout. This is the **Want: $\exists$ template**. The instructions in (11) indicate that we should leave a blank space, and that, eventually we will fill in that blank space with mathematical statements that permanently bind δ, *and* prove $\delta > 0$. For now, though, there's just a blank space, followed by:

Want: $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

Next, examine the statement we are trying to prove. It reads

$\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

Note that it starts with "$\forall$". We therefore go to (10) of the exposition handout. This is the "Want: $\forall$" template. The instructions in (10) say that we should continue our proof as follows:

Given $x \in \mathbb{R}$.
Want: $[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ]$.

By (2) of the exposition handout, x is now permanently bound.

Next, examine the statement we are trying to prove. It reads:

$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ]$.

Note that it has a $\Rightarrow$ in the middle. We therefore to to (12) of the exposition handout. This is the **Want: $\cdots \Rightarrow \cdots$ template**. The insturctions in (12) say that we should continue our proof as follows:

Assume $0 < x < 2\delta$.
Want: $x + x^2 < \varepsilon$.

Next examine the statement we are trying to prove. It reads:

$x + x^2 < \varepsilon$.

Note the lack of $\forall$, of $\exists$ and of $\Rightarrow$ in that statement. We'll say that a statement is an **atomic statement** if it lacks all three of these. When you get to proving an atomic statement, you can't rely on templates anymore. Instead, simply leave a big blank space followed by a small box. (That box indicates the end of the proof.) Eventually, we'll fill in that blank space by a with mathematical statements that end with "$x + x^2 < \varepsilon$".

At this point, we have **structured the proof**. This means we've used the templates (10), (11) and (12) as much as we can to set up the proof. The finish is to fill in the two blank spaces. Even if you can't finish a proof, if you can structure it properly, then I will give you substantial credit.

Now go back to the first blank space, in which we must permanently bind δ and show that $\delta > 0$. We have already discussed our strategy for finding δ from ε, with some notes that would typically be on scratch paper – paper that we don't turn in. Recall that, with experience as my guide, I settled on a strategy: $\delta := \min\left\{\frac{\varepsilon}{4}, \frac{\sqrt{\varepsilon/2}}{2}\right\}$. Into the first blank space, we write:

Let $\delta := \min\left\{\frac{\varepsilon}{4}, \frac{\sqrt{\varepsilon/2}}{2}\right\}$.
Then $\delta > 0$.

Remember that we were trying to permanently bind δ and prove $\delta > 0$. By (1) of the exposition handout, "Let" is a permanent binding word, so δ is permanently bound. We also asserted that $\delta > 0$, but, in fact, I might even leave out the statement "Then $\delta > 0$", treating it as obvious, because: First, I consider it to be obvious, from the rules of arithmetic that, since $\varepsilon > 0$, we know that both $\frac{\varepsilon}{4}$ and $\frac{\sqrt{\varepsilon/2}}{2}$ are positive. Second, I also consider it obvious that a minimum of two positive numbers, being equal to one of them, is necessarily positive.

We now have to prove that our strategy works. This proof goes in the second of the blank spaces. There are a number of possibilities for exactly how to handle this, but one approach reads:

$\left[\ \delta \leqslant \frac{\varepsilon}{4}\ \right]$ and $\left[\ \delta \leqslant \frac{\sqrt{\varepsilon/2}}{2}\ \right]$.
$\left[\ 2\delta \leqslant \frac{\varepsilon}{2}\ \right]$ and $\left[\ 2\delta \leqslant \sqrt{\frac{\varepsilon}{2}}\ \right]$.
$0 < x < 2\delta$.
$\left[\ 0 < x < 2\delta \leqslant \frac{\varepsilon}{2}\ \right]$ and $\left[\ 0 < x < 2\delta \leqslant \sqrt{\frac{\varepsilon}{2}}\ \right]$.
$\left[\ 0 < x < \frac{\varepsilon}{2}\ \right]$ and $\left[\ 0 < x < \sqrt{\frac{\varepsilon}{2}}\ \right]$.
$\left[\ 0 < x < \frac{\varepsilon}{2}\ \right]$ and $\left[\ 0^2 < x^2 < \frac{\varepsilon}{2}\ \right]$.
$\left[\ x < \frac{\varepsilon}{2}\ \right]$ and $\left[\ x^2 < \frac{\varepsilon}{2}\ \right]$.
$x + x^2 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.
$x + x^2 < \varepsilon$.

Remember: in this sequence of statements, we were trying to prove $x + x^2 < \varepsilon$. Once we get "$x + x^2 < \varepsilon$" as a *known* statement, we STOP. The proof is complete.

3. Class 3 on 12 September 2017, Tu of Week 2

Assigned HW#6 and HW#7.

Recall that we proved:

THEOREM 1.4. $\forall \varepsilon > 0$, $\exists \delta > 0$ *s.t.*, $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon\ ].$$

The strategy involved setting $\delta := \min\left\{\dfrac{\varepsilon}{4}, \dfrac{\sqrt{\varepsilon/2}}{2}\right\}$.

We focus now on using Theorem 1.4 to prove:

THEOREM 3.1. *$\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\ 0 < x < 6\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

Of course, one way to prove Theorem 3.1 is to simply mimic, *mutatis mutandis*, the proof of Theorem 1.4. That may, in fact, be the quickest way to get a proof of Theorem 3.1. However, our goal is broader than just proving one result: We want to demonstrate, by example, how to use one theorem to prove another. Specifically, we will show how to use Theorem 1.4 to prove Theorem 3.1.

First, we use **Replacement Rule (20)** from the exposition handout, applied to Theorem 1.4. We replace ε by μ, and δ by λ, and obtain:

THEOREM 3.2. *$\forall \mu > 0$, $\exists \lambda > 0$ s.t., $\forall x \in \mathbb{R}$,*

$$[\ 0 < x < 2\lambda\ ] \quad \Rightarrow \quad [\ x + x^2 < \mu\ ].$$

The strategy for this game involves setting $\lambda := \min\left\{\dfrac{\mu}{4}, \dfrac{\sqrt{\mu/2}}{2}\right\}$.

Now imagine that we wish to play the game associated to Theorem 3.1. You give me $\varepsilon > 0$. Here's what I do:

Let $\mu := \dfrac{\varepsilon}{5}$. I then use the strategy from Theorem 3.2, and set $\lambda := \min\left\{\dfrac{\mu}{4}, \dfrac{\sqrt{\mu/2}}{2}\right\}$. Because I know this strategy works for Theorem 3.2, I know, for every $x \in \mathbb{R}$, that

$$[\ 0 < x < 2\lambda\ ] \quad \Rightarrow \quad [\ x + x^2 < \mu\ ].$$

Next, I set $\delta := \dfrac{\lambda}{3}$, and I report this to you.

Because $2\lambda = 6\delta$ and because $\mu = \dfrac{\varepsilon}{5}$, I know, for every $x \in \mathbb{R}$, that

$$[\ 0 < x < 6\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \frac{\varepsilon}{5}\ ].$$

That means that I know that, whatever $x \in \mathbb{R}$ you choose, I will win!

I now have a plausible strategy, and I'm convinced that it will work. I must now communicate this strategy and proof, using the rules of exposition in the exposition handout.

I begin by structuring the proof of Theorem 3.1. For this particular triply-quantified theorem, I would write:

Proof of Theorem Theorem 3.1: Given $\varepsilon > 0$.
Want: $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 6\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

BLANK SPACE A: PERMANENTLY BIND δ, AND PROVE $\delta > 0$.

Want: $\forall x \in \mathbb{R}$, ($[0 < x < 6\delta] \Rightarrow [x + x^2 < \varepsilon/5]$).
Given $x \in \mathbb{R}$.
Want: $[0 < x < 6\delta] \Rightarrow [x + x^2 < \varepsilon/5]$.
Assume $0 < x < 6\delta$. Want: $x + x^2 < \varepsilon/5$.

BLANK SPACE B: PROVE $x + x^2 < \varepsilon/5$.

We end the proof with a small box: □

The proof is now structured, and, were this an exam problem, you could get substantial credit if you can just write down this much. The remaining credit would be for filling in the two BLANK SPACES.

For BLANK SPACE A, on scratch paper, we apply **Replacement Rule (22)** to Theorem 1.4, and replace ε by the bound expression $\varepsilon/5$. Remember that, according to **Catch (22)**, this strips off the "$\forall \varepsilon > 0$", and we get: $\exists \delta > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\delta\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

To this we apply Replacement Rule (20), and replace δ by λ, yielding: $\exists \lambda > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\lambda\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

To this apply **Replacement Rule (21)**; replace $\exists$ by "choose", yielding: choose $\lambda > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\lambda\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

Next, remember that our strategy for Theorem 3.1 was to set $\delta := \lambda/3$.

So, inside BLANK SPACE A, we might write:

By Theorem 1.4 (with ε replaced by $\varepsilon/5$, and δ by λ),
choose $\lambda > 0$ s.t., $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\lambda\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

Let $\delta := \lambda/3$.

This finishes BLANK SPACE A.

In BLANK SPACE B, we might write:

Since $0 < x < 6\delta$, and since $6\delta = 6 \cdot (\lambda/3) = 2\lambda$,
it follows that $0 < x < 2\lambda$.
Then, by our choice of λ,
we see that $x + x^2 < \varepsilon/5$, as desired.

This finishes BLANK SPACE B.

Remember that when Know=Want, we stop. This is **Stopping Rules (23,24)** in the exposition handout. In our proof, "$x+x^2 < \varepsilon/5$" was both known and wanted, so we stopped there.

We left out a couple of steps in BLANK SPACE B. Specifically, if you look back into BLANK SPACE A at how we chose λ, what it actually tells us is: $\forall x \in \mathbb{R}$,

$$[\ 0 < x < 2\lambda\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

Unfortunately, we can't say this in BLANK SPACE B, because x is already bound, and "$\forall x \in \mathbb{R}$" would cause a **double-binding**, which is NEVER allowed. Without explictly saying so, we used Replacement Rule (22) from the exposition handout, and replaced x by the bound expression x. (It's not uncommon to replace a variable by itself.) Remember that, according to Catch (22), this strips off the "$\forall x \in \mathbb{R}$", and we get:

$$[\ 0 < x < 2\lambda\ ] \quad \Rightarrow \quad [\ x + x^2 < \varepsilon/5\ ].$$

It would actually be bad form to write this at the end of BLANK SPACE B, because: We already observed that $0 < x < 2\lambda$, and we typically don't use $P \Rightarrow Q$, when P is known. Instead, we can either write "P, therefore Q", or, if it seems clear enough, simply write "Q". In this case, we simply wrote "$x + x^2 < \varepsilon/5$".

We will stop proving triply-quantified statements for the moment, and move on to discussing SET THEORY BASICS. We'll list some, but not all of the axioms of set theory. If you want to look them up, you should look for Zermelo-Fraenkel, with Choice, sometimes called ZFC set theory. We modify that slightly by adding a frownie that is not allowed to be equal to a set and is not allowed to be an element of a set. Our sets are "happy":

Axiom: $\forall S, \quad \Bigg(\quad [S \text{ is a set}] \quad \Rightarrow \quad [\ (\ \frown \neq S\ )\ \&\ (\ \frown \notin S\ )\ ] \quad \Bigg).$

That is, two sets are equal iff they have the same elements. So, in a proof, if we have a statement that reads "Want: $S = T$", then we can, if we wish, proceed as follows:

Want: $\forall x$, [$(x \in S) \Leftrightarrow (x \in T)$].
Given x.
Want: $(x \in S) \Leftrightarrow (x \in T)$.

At that point, still have work to do, but at least equality of sets has been "unwound" a little and we now have an object x to work with.

We define **subset** as follows:

DEFINITION 3.3. *For any sets S, T, $S \subseteq T$ means:*

$$\forall x, \quad [\ (x \in S) \ \Rightarrow\ (x \in T)\ ].$$

We recalled HW#1: For any propositions A and B, we have

$$(\ A \ \Leftrightarrow\ B\ ) \qquad \Leftrightarrow \qquad (\ (\ A \ \Rightarrow\ B\ ) \quad \& \quad (\ B \ \Rightarrow\ A\ )\ ).$$

THEOREM 3.4. *For any sets S and T, we have*

$$[\ S = T\ ] \quad \Leftrightarrow \quad [\ (S \subseteq T)\ \&\ (T \subseteq S)\ ].$$

In what follows, any text between "((" and "))" doesn't belong in the proof. It is there only to explain some ideas of exposition.

Proof. Given sets S and T.
Want: $[S = T] \Leftrightarrow [(S \subseteq T) \& (T \subseteq S)]$.
Let $A := [S = T]$, $B := [S \subseteq T]$, $C := [T \subseteq S]$.

Want: $A \Leftrightarrow [B\&C]$.
By HW#1 (replacing A by A, and B by $B\&C$), want:
(1) $A \Rightarrow [B\&C]$ and
(2) $[B\&C] \Rightarrow A$.
((By HW#1 with the given replacements, we know:

$$[\ A \Leftrightarrow [B\&C]\ ] \qquad \Leftrightarrow \qquad [\ (1)\,\&\,(2)\ ].$$

So, by **Rule (31)** of the exposition handout we can make the transition from "Want: $A \Leftrightarrow [B\&C]$" to "Want: (1)&(2)".

Next, we apply the **Want: . . . & . . . template**, which is Rule (14) from the exposition handout, and proceed as follows.))
Proof of (1): BLANK SPACE A *End of proof of (1).*
Proof of (2): BLANK SPACE B *End of proof of (2).* □

It remains to fill in the blank spaces. Each one is a **section of a proof**. This has the effect that varibles that are bound permanently inside that section become free at the end of that section. So "permanent" is relative.
In BLANK SPACE A, we might write:
Assume A.
Want: $B\&C$.
((We just followed Rule (12).))
Know: A
((Once you assume something, you know it. Typically, we wouldn't bother to say this, but there's nothing incorrect about it.))
Know: $S = T$.
((Since A is, by definition, equivalent to $S = T$, we may transition from "Know: A" to "Know: $S = T$". Each known statement should follow from earlier known statements. By contrast, according to Rule (31), each wanted statement (except those coming from templates) should imply the preceding wanted statement.

Next, we use Rule (31), keeping in mind that $B\&C$ is equivalent to $(S \subseteq T)\&(T \subseteq S)$.))
Want: $(S \subseteq T)\&(T \subseteq S)$.
((Next, we use the "Want: . . . & . . . " template.))
Proof of $S \subseteq T$: BLANK SPACE A1 *End of proof of $S \subseteq T$.*
Proof of $T \subseteq S$: BLANK SPACE A2 *End of proof of $T \subseteq S$.*
End of BLANK SPACE A

Each of A1 and A2 is a **subsection of a proof**. This has the effect that varibles that are bound permanently inside that subsection become free at the end of that subsection. So "permanent" is relative.

In BLANK SPACE A1, we might write:
((We look up the definition of subset.))
Want: $\forall x,\ [\,(x \in S) \Rightarrow (x \in T)\,]$.
((Next, the "Want: $\forall$" template.))
Given x.
Want $(x \in S) \Rightarrow (x \in T)$.
((Next, the "Want: $\cdots \implies \cdots$" template.))
Assume $xinS$.
Want: $x \in T$.
((Recalling that $S = T$, we apply the Axiom of Extensionality, but replacing x by y, since x is bound.))
Know: $\forall y,\ [\,(y \in S) \Leftrightarrow (y \in T)\,]$ ((Next replace y by x, using Replacement Rule (22), remembering Catch (22). This would give $(x \in S) \Leftrightarrow (x \in T)$. However, we don't write this because we *KNOW* that $x \in S$, so "$x \in S$" should not appear as the left side of an implication. Instead we wite the following.))
$(x \in S) \therefore (x \in T)$.
((Since we now know that $x \in T$, we STOP, by Stopping Rules (23,24) from the exposition handout.))
End of BLANK SPACE A1.

In BLANK SPACE A2, we might write:
((We look up the definition of subset.))
Want: $\forall x,\ [\,(x \in T) \Rightarrow (x \in S)\,]$.
((Next, the "Want: $\forall$" template.))
Given x.
Want $(x \in T) \Rightarrow (x \in S)$.
((Next, the "Want: $\cdots \implies \cdots$" template.))
Assume $xinT$.
Want: $x \in S$.
((Recalling that $S = T$, hence $T = S$, we apply the Axiom of Extensionality, but replacing x by y, since x is bound.))
Know: $\forall y,\ [\,(y \in T) \Leftrightarrow (y \in S)\,]$ ((Next replace y by x, using Replacement Rule (22), remembering Catch (22). This would give $(x \in T) \Leftrightarrow (x \in S)$. However, we don't write this because we *KNOW*

that $x \in T$, so "$x \in T$" should not appear as the left side of an implication. Instead we wite the following.))
$(x \in T) \therefore (x \in S)$.
((Since we now know that $x \in S$, we STOP, by Stopping Rules (23,24) from the exposition handout.))
End of BLANK SPACE A2.

In BLANK SPACE B, we might write:
Assume $B\&C$.
Want: A.
Know: $[S \subseteq T]\&[T \subseteq S]$.
Want: $S = T$.
((Use the Axiom of Extensionality.))
Want: $\forall x,\ ([x \in S] \Leftrightarrow [x \in T])$.
((Use the "Want: $\forall$" template.))
Given x.
Want $[x \in S] \Leftrightarrow [x \in T]$.
((We next use the definition of $\subseteq$, combined with Catch (22).))
Since $S \subseteq T$, we have: $(x \in S) \Rightarrow (x \in T)$.
((We again use the definition of $\subseteq$, combined with Catch (22).))
Since $T \subseteq S$, we have: $(x \in T) \Rightarrow (x \in S)$.
((We next use:
$\forall$propositions P, Q, $[(P \Rightarrow Q)\&(Q \Rightarrow P)] \Leftrightarrow [P \Leftrightarrow Q]$.))
Since $[\,(x \in S) \Rightarrow (x \in T)\,]$ and $[\,(x \in T) \Rightarrow (x \in S)\,]$,
we get: $(x \in S) \Leftrightarrow (x \in T)$. ((By Stopping Rules (23,24), we STOP.))

Here's our next axiom, called the **Axiom of Specification**:

Axiom: $\forall$set S, $\forall$well-formed condition P on S, there exists a unique set R such that: $\qquad \forall x, \qquad (\,[\,x \in \mathbb{R}\,] \Leftrightarrow [\,(x \in S)\,\&(P(x))\,]\,)$.

In the preceding Axiom of Specification, the set R is typically denoted either by $\{x \in S \,|\, P(x)\}$ or by $\{x \in S \text{ s.t. } P(x)\}$. That is:

DEFINITION 3.5. *For any set S, for any well-formed condition P on S, by either $\{x \in S \,|\, P(x)\}$ or $\{x \in S \text{ s.t. } P(x)\}$, we mean the unique set R such that:* $\qquad \forall x, \qquad (\,[\,x \in \mathbb{R}\,] \Leftrightarrow [\,(x \in S)\,\&(P(x))\,]\,)$.

For example, $\{x \in \mathbb{N} \,|\, x/2 \in \mathbb{N}\} = \{2, 4, 6, \ldots\}$.
For example, $\{x \in \mathbb{Z} \,|\, -2 \leqslant x < 3\} = \{-2, -1, 0, 1, 2\}$.

For a non-example, $\{x \in \mathbb{N} \,|\, x \text{ is cool, man}\}$ is not a set because "x is cool, man" is not a well-formed condition. (To get a more precise idea of what well-formed means, you'll need to take a course on logic.)

Another non-example: Define a set B by

$$B \quad := \quad \{ \ A \mid (A \text{ is a set}) \ \& \ (A \notin A) \ \}.$$

Then $(B \in B) \Rightarrow (B \notin B)$, and, also $(B \notin B) \Rightarrow (B \in B)$. Therefore both $B \in B$ and $B \notin B$ are impossible, but one of them must be true. This is called **Russel's Paradox**, and the point is that you can't use the Axiom of Specification to define a set of the form $\{A \,|\, (A \text{ is a set}) \,\&\, (A \notin A)\}$. If you have a set $\mathcal{S}$ of sets, then you could let $B := \{A \in \mathcal{S} \,|\, (A \text{ is a set}) \,\&\, (A \notin A)\}$. However, if you did this then you couldn't prove that $(B \notin B) \Rightarrow (B \in B)$. In fact, you could only prove that $((B \notin B)\&(B \in \mathcal{S})) \Rightarrow (B \in B)$.

We assigned HW#8.

Next, we define the **empty set**:

DEFINITION 3.6. $\varnothing := \{x \in \mathbb{N} \,|\, x \neq x\}$.

FACT 3.7. $\forall x,\ x \notin \phi$.

Proof. Given x.
Want: $x \notin \varnothing$.
((We next use the **contradiction template**,
which is Rule (13) from the exposition handout.))
Assume $x \in \varnothing$.
Want: $\rightarrow\leftarrow$.
By definition of $\varnothing$, since $x \in \varnothing$,
we know both that $x \in \mathbb{N}$ and that $x \neq x$.
By the Axiom of Equality (with x replaced by x),
we know that $x = x$.
Then $x = x$ and $x \neq x$.
$\rightarrow\leftarrow$. □

FACT 3.8. $\forall x \in \varnothing,\ x = 5$.

Proof. Given $x \in \varnothing$.
Want: $x = 5$.
Assume $x \neq 5$.
Want: $\rightarrow\leftarrow$

By definition of $\varnothing$, since $x \in \varnothing$,
we know both that $x \in \mathbb{N}$ and that $x \neq x$.
By the Axiom of Equality (with x replaced by x),
we know that $x = x$.
Then $x = x$ and $x \neq x$.
$\rightarrow\leftarrow$. $\square$

FACT 3.9. $\forall x \in \varnothing$, $x \neq 5$.

Proof. Given $x \in \varnothing$.
Want: $x \neq 5$.
Assume $x = 5$.
Want: $\rightarrow\leftarrow$
By definition of $\varnothing$, since $x \in \varnothing$,
we know both that $x \in \mathbb{N}$ and that $x \neq x$.
By the Axiom of Equality (with x replaced by x),
we know that $x = x$.
Then $x = x$ and $x \neq x$.
$\rightarrow\leftarrow$. $\square$

The next fact asserts that $\varnothing$ is a set of subsets of $\mathbb{R}$.

FACT 3.10. $\forall A \in \varnothing$, $A \subseteq \mathbb{R}$.

Proof. Given $A \in \varnothing$.
Want: $A \subseteq \mathbb{R}$.
Assume $A \nsubseteq \mathbb{R}$.
Want: $\rightarrow\leftarrow$
By definition of $\varnothing$, since $A \in \varnothing$,
we know both that $A \in \mathbb{N}$ and that $A \neq A$.
By the Axiom of Equality (with A replaced by A),
we know that $A = A$.
Then $A = A$ and $A \neq A$.
$\rightarrow\leftarrow$. $\square$

FACT 3.11. *$\forall$set A, we have: $\varnothing \subseteq A$.*

Now that we have introduced most of the points in the exposition handout, we will start typing our proofs in paragraph format, to save paper. We encourage students to stick with our prior writing conventions, when they turn in written work.

Proof. Let a set A be given. We wish to prove: $\varnothing \subseteq A$. By definition of $\subseteq$, we wish to prove: $\forall x \in \varnothing$, $x \in A$. Let $x \in \varnothing$ be given. We wish to prove: $x \in A$. Assume that $x \notin A$. We seek a contradiction.

Since $x \in \varnothing$, we conclude that $x \in \mathbb{N}$ and that $x \neq x$. By the Axiom of Equality, we conclude that $x = x$. Contradiction. $\square$

4. Class 4 on 14 September 2017, Th of Week 2

Recall:

- $\forall x \in \varnothing$, $x = 5$.
- $\forall x \in \varnothing$, $x \neq 5$.
- $\forall A \in \varnothing$, $A \subseteq \mathbb{R}$.
- $\forall$set S, $\varnothing \subseteq S$.

Here's our next axiom, called the **Axiom of Union**:

Axiom: $\forall$set $\mathcal{S}$ of sets, there exists a unique set V such that:

$$\forall x, \qquad (\, [\, x \in V \,] \;\Leftrightarrow\; [\, \exists A \in \mathcal{S} \text{ s.t. } x \in A \,] \,).$$

That is, the elements of V are the objects that are elements of at least one set in $\mathcal{S}$. In the preceding Axiom of Union, the set V is typically denoted by $\bigcup \mathcal{S}$. That is:

DEFINITION 4.1. *For any set $\mathcal{S}$ of sets, by $\bigcup \mathcal{S}$ we mean the unique set V such that:* $\qquad \forall x, \qquad (\, [\, x \in V \,] \;\Leftrightarrow\; [\, \exists A \in \mathcal{S} \textit{ s.t. } x \in A \,] \,)$.

For example, $\bigcup \{\{1, 2, 3\}, \{3, 4, 5\}\} = \{1, 2, 3, 4, 5\}$. Using $\bigcup \bigcup$, we can flatten a set of sets of sets of numbers, down to a set of numbers:

$$\bigcup \bigcup \{\, \{\, \{1, 2\}\, , \{3, 4\} \,\}\, ,\; \{\, \{5, 6\}\, , \{7, 8\} \,\} \,\} \;=\; \{1, 2, 3, 4, 5, 6, 7, 8\}.$$

Note that, in Definition 4.1, the variable V is temporarily bound, and became free at the end of the sentence. This example of temporary binding isn't to be found in the exposition handout, because I don't expect that students will be using it. However, occasionally, I'll need it to make certain definitions.

Note that $\bigcup \varnothing = \varnothing$.

We now begin on intersections. These exist, not by axiom, but by a theorem, whose proof we omit:

THEOREM 4.2. *$\forall$set $\mathcal{S}$ of sets, there exists a unique set Y such that:*

$$\forall x, \qquad (\, [\, x \in Y \,] \;\Leftrightarrow\; [\, \forall A \in \mathcal{S},\, x \in A \,] \,).$$

Proof. Omitted. $\square$

That is, the elements of Y are the objects that are elements of *every* set in $\mathcal{S}$. In the preceding theorem, the set Y is typically denoted by $\bigcup \mathcal{S}$. That is:

DEFINITION 4.3. *For any set $\mathcal{S}$ of sets, by $\bigcap \mathcal{S}$ we mean the unique set Y such that:* $\quad \forall x, \qquad ([x \in Y] \Leftrightarrow [\forall A \in \mathcal{S}, x \in A])$.

For example, $\bigcap\{\{1,2,3\},\{3,4,5\}\} = \{3\}$.

Note that, in Definition 4.1, the variable Y is temporarily bound, and became free at the end of the sentence.

Note that $\bigcap \varnothing$ is undefined, so, by our conventions, we have $\bigcap \varnothing =$ ☹.

We have a simplified notation for **finite unions** and **finite intersections**.

DEFINITION 4.4. *For any $n \in \mathbb{N}$, for any sets $A_1, \ldots, A_n$, we define $A_1 \cup \cdots \cup A_n := \bigcup\{A_1, \ldots, A_n\}$ and $A_1 \cap \cdots \cap A_n := \bigcap\{A_1, \ldots, A_n\}$.*

For example,

$$\begin{array}{rrcl} & \{1,2,3\} \cup \{3,4,5\} & = & \{1,2,3,4,5\} \\ \text{and} & \{1,2,3\} \cap \{3,4,5\} & = & \{3\}. \end{array}$$

We assigned HW#9 and HW#10.

DEFINITION 4.5. *For any set S, the number of elements in S is denoted $\#S$.*

For example,

$$\begin{array}{rcl} \#\{2,4,6\} & = & 3, \\ \#\{1,2,5,9\} & = & 4, \\ \#\{\{1,2\},\{3,4\},5\} & = & 3, \\ \#\mathbb{N} & = & \infty, \\ \#\{\varnothing\} & = & 1, \\ \text{and} \quad \#\varnothing & = & 0. \end{array}$$

We compute #☹, as follows. Remember our general convention that any computation that is undefined yields ☹. In this case, ☹ is *NOT* a set, and so, when we review Definition 4.5, we see that #☹ is undefined. We therefore have: #☹ = ☹.

Next is the **Axiom of Power Set**:

Axiom: $\forall$set U, there exists a unique set $\mathcal{S}$ of sets such that:

$$\forall A, \qquad ([A \in \mathcal{S}] \Leftrightarrow [A \subseteq U]).$$

That is, the elements of $\mathcal{S}$ are the subsets of U. In the preceding Axiom of Power Set, the set $\mathcal{S}$ will be denoted by 2^U or $2\hat{\ }U$. That is:

DEFINITION 4.6. *For any set U, by 2^U or $2\hat{\ }U$, we mean the unique set $\mathcal{S}$ such that:* $\quad \forall A, \quad ([A \in \mathcal{S}] \Leftrightarrow [A \subseteq U]).$

For example, we compute

$$\begin{array}{rcll} 2^{\{2,4,6\}} & = & \{ & \varnothing, \\ & & & \{2\}, \{4\}, \{6\}, \\ & & & \{2,4\}, \{2,6\}, \{4,6\}, \\ & & & \{2,4,6\} \qquad \}. \end{array}$$

For any set U, the set 2^U is called the **power set** of U. It is not hard to show: For any set U, we have $\#(2^U) = 2^{\#U}$. So, for example, we have: $\#(2^{\{2,4,6\}}) = 2^{\#\{2,4,6\}} = 2^3 = 8$.

Next we focus on various **commutativity results**, **associativity results** and **distributivity results**. We first expose these kinds of results in propositional logic, and then expose these kinds of results in set theory.

We will need the following result from propositional logic, which asserts that & and or are commutative.

LEMMA 4.7. *Let P and Q be propositions. Then:*

(1) $[P\&Q] \Leftrightarrow [Q\&P]$ *and*
(2) $[P \textit{ or } Q] \Leftrightarrow [Q \textit{ or } P]$.

This is proved by truth tables. Since this is not a course on propositional logic, we will, in general, omit the proofs of propositional logic results. Similarly, it is an unassigned homework probem to verify by truth tables that & and or are both associative:

LEMMA 4.8. *Let P, Q and R be propositions. Then:*

(1) $[(P\&Q)\&R] \Leftrightarrow [P\&(Q\&R)]$ *and*
(2) $[(P \textit{ or } Q) \textit{ or } R] \Leftrightarrow [P \textit{ or } (Q \textit{ or } R)]$.

We have modified distributivity laws for not over both & and or:

LEMMA 4.9. *Let P and Q be propositions. Then:*

(1) $[\textit{not}(P\&Q)] \Leftrightarrow [(\textit{not } P) \textit{ or } (\textit{not } Q)]$ *and*
(2) $[\textit{not}(P \textit{ or } Q)] \Leftrightarrow [(\textit{not } P)\&(\textit{not } Q)]$.

Finally, we will need the fact that & distributes both over & and over or, as well as the fact that or distributes both over & and over or. Again we omit the proof:

LEMMA 4.10. *Let P, Q and R be propositions. Then:*

(1) $[P\&(Q\&R)] \Leftrightarrow [(P\&Q)\&(P\&R)]$,
(2) $[P\&(Q \text{ or } R)] \Leftrightarrow [(P\&Q) \text{ or } (P\&R)]$,
(3) $[P \text{ or } (Q\&R)] \Leftrightarrow [(P \text{ or } Q)\&(P \text{ or } R)]$ *and*
(4) $[P \text{ or } (Q \text{ or } R)] \Leftrightarrow [(P \text{ or } Q) \text{ or } (P \text{ or } R)]$.

We can use logic results to get set-theoretic results, like the commutativity of $\cap$ and $\cup$:

THEOREM 4.11. *Let A and B be sets. Then:*

(1) $A \cap B = B \cap A$ *and*
(2) $A \cup B = B \cup A$.

Since this is not a course on propositional logic, we will, in general, omit the proofs of set-theoretic results. However, we make exceptions to highlight certain techniques of proof.

Proof. Proof of (1): By the Axiom of Extensionality, we wish to prove: $\forall x$, $[(x \in A \cap B) \Leftrightarrow (x \in B \cap A)]$. Let x be given. We wish to prove: $(x \in A \cap B) \Leftrightarrow (x \in B \cap A)$.

Let $P := (x \in A)$ and $Q := (x \in B)$. By definition of $\cap$, we have $(x \in A \cap B) \Leftrightarrow (P\&Q)$, and we also have $(x \in B \cap A) \Leftrightarrow (Q\&P)$.

By (1) of Lemma 4.7, we have $(P\&Q) \Leftrightarrow (Q\&P)$. Then

$$\begin{aligned}(x \in A \cap B) &\Leftrightarrow (P\&Q)\\ &\Leftrightarrow (Q\&P)\\ &\Leftrightarrow (x \in B \cap A),\end{aligned}$$

as desired. *End of proof of (1).*

Proof of (2): Unassigned homework. *End of proof of (2).* □

Next, we assert associativity of $\cap$ and $\cup$:

THEOREM 4.12. *Let A, B and C be sets. Then:*

(1) $(A \cap B) \cap C = A \cap (B \cap C)$ *and*
(2) $(A \cup B) \cup C = A \cup (B \cup C)$.

Proof. Proof of (1): By the Axiom of Extensionality, we wish to prove: $\forall x$, $[(x \in (A \cap B) \cap C) \Leftrightarrow (x \in A \cap (B \cap C))]$. Let x be given. We wish to prove: $(x \in (A \cap B) \cap C) \Leftrightarrow (x \in A \cap (B \cap C))$.

Let $P := (x \in A)$, $Q := (x \in B)$ and $R := (x \in C)$. By definition of $\cap$, we have $(x \in (A \cap B) \cap C) \Leftrightarrow ((P \& Q) \& R)$, and we also have $(x \in A \cap (B \cap C)) \Leftrightarrow (P \& (Q \& R))$.

By (1) of Lemma 4.8, we have $((P \& Q) \& R) \Leftrightarrow (P \& (Q \& R))$. Then

$$\begin{aligned} (x \in (A \cap B) \cap C) &\Leftrightarrow ((P \& Q) \& R) \\ &\Leftrightarrow (P \& (Q \& R)) \\ &\Leftrightarrow (x \in A \cap (B \cap C)), \end{aligned}$$

as desired. *End of proof of (1).*

Proof of (2): Unassigned homework. *End of proof of (2).* □

We define **set-theoretic subtraction**:

DEFINITION 4.13. *For any two sets A and B, we define*

$$A \backslash B \quad := \quad \{ \, x \in A \mid x \notin B \, \}.$$

Next, we describe the modified distributivity laws for set subtraction both over & and over or:

THEOREM 4.14. *Let A, B and C be sets. Then:*

(1) $A \backslash (B \cap C) = (A \backslash B) \cup (A \backslash C)$ *and*
(2) $A \backslash (B \cup C) = (A \backslash B) \cap (A \backslash C)$.

Proof. Proof of (1): By the Axiom of Extensionality, we wish to prove: $\forall x$, $\big([x \in A \backslash (B \cap C)] \Leftrightarrow [x \in (A \backslash B) \cup (A \backslash C)]\big)$. Let x be given. We wish to prove: $[x \in A \backslash (B \cap C)] \Leftrightarrow [x \in (A \backslash B) \cup (A \backslash C)]$.

We define $P := (x \in A)$, $Q := (x \in B)$ and $R := (x \in C)$. Also, we define $Q' :=$ (not Q) and $R' :=$ (not R). By definition of $\backslash$ and of $\cap$ and of $\cup$, we have $[x \in A \backslash (B \cap C)] \Leftrightarrow [P \& (\mathrm{not}(Q \& R))]$, and we also have $[x \in (A \backslash B) \cup (A \backslash C)] \Leftrightarrow [(P \& Q')$ or $(P \& R')]$.

By (1) of Lemma 4.9, we have $(\text{not}(Q\&R)) \Leftrightarrow (Q' \text{ or } R')$. Then $[P\&(\text{not}(Q\&R))] \Leftrightarrow [P\&(Q' \text{ or } R')]$. By (2) of Lemma 4.10, we conclude that $[P\&(Q' \text{ or } R')] \Leftrightarrow [(P\&Q') \text{ or } (P\&R')]$. Then

$$\begin{aligned}[x \in A\backslash(B \cap C)] &\Leftrightarrow [P\&(\text{not}(Q\&R))]\\ &\Leftrightarrow [P\&(Q' \text{ or } R')]\\ &\Leftrightarrow [(P\&Q') \text{ or } (P\&R')]\\ &\Leftrightarrow [x \in (A\backslash B) \cup (A\backslash C)],\end{aligned}$$

as desired. *End of proof of (1).*

Proof of (2): Unassigned homework. *End of proof of (2).* □

Finally, we will assert that $\cap$ distributes both over $\cap$ and over $\cup$, as well as the fact that $\cup$ distributes both over $\cap$ and over $\cup$.

THEOREM 4.15. *Let A, B and C be sets. Then:*

(1) $[A \cap (B \cap C)] \Leftrightarrow [(A \cap B) \cap (A \cap C)]$,
(2) $[A \cap (B \cup C)] \Leftrightarrow [(A \cap B) \cup (A \cap C)]$,
(3) $[A \cup (B \cap C)] \Leftrightarrow [(A \cup B) \cap (A \cup C)]$ *and*
(4) $[A \cup (B \cup C)] \Leftrightarrow [(A \cup B) \cup (A \cup C)]$.

Proof. Unassigned homework. □

We next define **ELT**:

DEFINITION 4.16. *For any non-☹ object a, let* $\text{ELT}(\{a\}) = a$.

A set S is called a **singleton set** if $\#S = 1$, *i.e.*, if S has exactly one element. Note, for example, that $\{7\}$, $\{-\sqrt{2}\}$, $\{\{1,2\}\}$ and $\{\varnothing\}$ are all singleton sets. Remember our general convention that any computation that is undefined yields ☹. Consequently, for any object x that is not a singleton set, we have $\text{ELT}(x) = ☹$. So:

$\text{ELT}(\varnothing) = ☹$,
$\text{ELT}(\{2,3\}) = ☹$,
$\text{ELT}(\{5,8,9\}) = ☹$,
$\text{ELT}(\mathbb{Z}) = ☹$ and
$\text{ELT}(☹) = ☹$.

On the other hand, we have

$\text{ELT}(\{7\}) = 7$,
$\text{ELT}(\{-\sqrt{2}\}) = -\sqrt{2}$,
$\text{ELT}(\{\{1,2\}\}) = \{1,2\}$ and

$\mathrm{ELT}(\{\varnothing\}) = \varnothing$.

Recall that $\{1,2,3,☹\} = ☹$. Similarly, we have $\{1,☹\} = ☹$ and $\{2,☹\} = ☹$. Then $\{1,☹\} = \{2,☹\}$. Unassigned homework: using the Axiom of Extensionality, show that $\{1,2\} = \{2,1\}$. More generally, in set theory, order is unimportant. So, for example, we have the following fact.

FACT 4.17. *For any non-☹u, v, for any x, y, we have*

$$[\,\{u,v\} = \{x,y\}\,] \Leftrightarrow [\,((u = x)\&(v = y)) \text{ or } ((u = y)\&(v = x))\,].$$

Proof. Unassigned homework. □

To try to keep track of the order of a pair of objects, we make the following definition.

DEFINITION 4.18. *For all x, y, we define*

$$\langle\ x\ ,\ y\ \rangle \quad := \quad \{\ \{x\}\ ,\ \{x,y\}\ \}.$$

NOTE TO SELF: NEXT YEAR, $\langle\langle\ x\ ,\ y\ \rangle\rangle \quad := \quad \{\ \{x\}\ ,\ \{x,y\}\ \}$. END OF NOTE TO SELF

So, for example, we have

$$\begin{aligned}
\langle\ 1\ ,\ 2\ \rangle &= \{\ \{1\}\ ,\ \{1,2\}\ \},\\
\langle\ 2\ ,\ 1\ \rangle &= \{\ \{2\}\ ,\ \{2,1\}\ \},\\
\langle\ 5\ ,\ 5\ \rangle &= \{\ \{5\}\ ,\ \{5,5\}\ \}\ =\ \{\ \{5\}\ \}\\
\langle\ 1\ ,\ \{2,3\}\ \rangle &= \{\ \{1\}\ ,\ \{1,\{2,3\}\}\ \},\\
\langle\ ☹\ ,\ 5\ \rangle &= ☹ \qquad\qquad\qquad \text{and}\\
\langle\ 5\ ,\ ☹\ \rangle &= ☹.
\end{aligned}$$

Note that $\langle 1,2\rangle \neq \langle 2,1\rangle$. Generally, we have:

FACT 4.19. *For any non-☹u, v, for any x, y, we have*

$$[\ \langle u,v\rangle = \langle x,y\rangle\ ] \qquad \Leftrightarrow \qquad [\ (u = x)\ \&\ (v = y)\ ].$$

Proof. Unassigned homework. □

For any set U, for any $x, y \in U$, we have

$$\{x\}\,,\,\{x,y\} \quad \in \quad 2\hat{\ }U,$$

so $\langle x,y\rangle \subseteq 2\hat{\ }U$, so $\langle x,y\rangle \in 2\hat{\ }2\hat{\ }U$.

More ordering:

DEFINITION 4.20. *For all $n \in \mathbb{N}$, for all $x_1, \ldots, x_n$, we define*

$$(x_1 , \ldots , x_n) \quad := \quad \{ \langle 1, x_1 \rangle , \langle 2, x_2 \rangle , \ldots , \langle n, x_n \rangle \}.$$

NOTE TO SELF: NEXT YEAR,

$$\langle x_1 , \ldots , x_n \rangle \quad := \quad \{ \langle\langle 1, x_1 \rangle\rangle , \langle\langle 2, x_2 \rangle\rangle , \ldots , \langle\langle n, x_n \rangle\rangle \}.$$

This way, we have no confusion between $(a, b) = \{x \in \mathbb{R} \,|\, a < x < b\}$ and $\langle a, b \rangle = \{\langle\langle 1, a \rangle\rangle, \langle\langle 2, b \rangle\rangle\}$. END OF NOTE TO SELF

So, for example, we have

$$\begin{aligned} (1,2,3) &= \{\langle 1,1 \rangle , \langle 2,2 \rangle , \langle 3,3 \rangle \}, \\ (3,1,2) &= \{\langle 1,3 \rangle , \langle 2,1 \rangle , \langle 3,2 \rangle \} \qquad \text{and} \\ (1,2,3,\frownie) &= \frownie. \end{aligned}$$

For any $m, n \in \mathbb{N}$, for any non-$\frownie x_1, \ldots, x_m$, for any $y_1, \ldots, y_n$, we leave it as unassigned homework to show that:

$$(\qquad (x_1 , \ldots , x_m) \qquad = \qquad (y_1 , \ldots , y_n) \qquad)$$

iff $\quad (\quad [\, m = n \,] \quad \& \quad [\, (x_1 = y_1) \,\&\, \cdots \,\&\, (x_n = y_n) \,] \quad).$

For any $n \in \mathbb{N}$, for any sets $A_1, \ldots, A_n$, for any $x_1 \in A_1, \ldots, x_n \in A_n$,

$$\langle 1, x_1 \rangle, \ldots, \langle n, x_n \rangle \quad \in \quad 2\hat{\ }2\hat{\ }(\{1, \ldots, n\} \cup A_1 \cup \cdots \cup A_n),$$

so $(x_1, \ldots, x_n) \in 2\hat{\ }2\hat{\ }2\hat{\ }(\{1, \ldots, n\} \cup A_1 \cup \cdots \cup A_n)$.

For any $n \in \mathbb{N}$, for any $x_1, \ldots, x_n$, we have

$$(x_1, \ldots, x_n) \quad = \quad \{ \; \{\{1\}, \{1, x_1\}\} \; , \; \ldots \; , \; \{\{n\}, \{n, x_n\}\} \; \},$$

so

$$\bigcup (x_1, \ldots, x_n) \quad = \quad \{ \; \{1\}, \{1, x_1\} \; , \; \ldots \; , \; \{n\}, \{n, x_n\} \; \},$$

so

$$\begin{aligned} \bigcup \bigcup (x_1, \ldots, x_n) &= \{ \; 1, 1, x_1 \; , \; \ldots \; , \; n, n, x_n \; \} \\ &= \{ \; 1, \ldots, n, \; x_1, \ldots, x_n \; \}, \end{aligned}$$

so $x_1, \ldots, x_n \in \bigcup \bigcup (x_1, \ldots, x_n)$. A similar argument shows: For any $n \in \mathbb{N}$, for any $x_1, \ldots, x_n$, we have $x_1, \ldots, x_n \in \bigcup \bigcup \bigcup \{(x_1, \ldots, x_n)\}$.

DEFINITION 4.21. *For any $n \in \mathbb{N}$, for any sets $A_1, \ldots, A_n$, let*

$$A_1 \times \cdots \times A_n \quad := \quad \{ \; (x_1, \ldots, x_n) \; | \; x_1 \in A_1 \, , \; \ldots \; , \; x_n \in A_n \; \}.$$

To a logic purist,

$$\{\ (x_1,\ldots,x_n)\ |\ x_1\in A_1\ ,\ \ldots\ ,\ x_n\in A_n\ \}$$

is verboten, for two reasons. First, in the Axiom of Specification, after "{", we need a free variable, *NOT* an expression like "$(x_1,\ldots,x_n)$". Second, in the Axiom of Specification, after the free variable, we need "$\in$", *NOT* "$|$"; otherwise, we run the risk of Russel's Paradox. Mathematicans, are often sloppy, but we should always keep in mind that it's important that there be ways to rewrite our definitions and arguments to be precise. In the present situation, we should really rewrite Definition 4.21, and set $A_1\times\cdots\times A_n$ equal to

$$\{\ z\in 2\hat{\ }2\hat{\ }2\hat{\ }(\{1,\ldots,n\}\cup A_1\cup\cdots\cup A_n)\ \ |\ \ \exists x_1\in A_1,\ldots,\exists x_n\in A_n \text{ s.t. } z=(x_1,\ldots,x_n)\ \}.$$

NOTE TO SELF: IN THE FUTURE, we'll define

$$A_1\times\cdots\times A_n\quad:=\quad\{\ \langle x_1,\ldots,x_n\rangle\ |\ x_1\in A_1\ ,\ \ldots\ ,\ x_n\in A_n\ \},$$

and we'll define $\langle x_1,\ldots,x_n\rangle:=\{\langle\langle 1,x_1\rangle\rangle,\ldots\langle\langle n,x_n\rangle\rangle\}$, and we'll define $\langle\langle x,y\rangle\rangle:=\{\{x\},\{x,y\}\}$. END OF NOTE TO SELF

At this point, for any non-☹x,y, we can make an ordered pair from x and y in two ways: $\langle x,y\rangle$ or (x,y). In the future, by **ordered pair**, we always mean a $(,)$-ordered pair. That is, for all z, we say that z is an **orderd pair** if there exist non-☹x,y such that $z=(x,y)$. A set is called a **relation** if its elements are all ordered pairs. That is:

DEFINITION 4.22. *Let R be a set. By R is a* **relation***, we mean:*

$$\forall z\in R,\quad\exists x,y\quad s.t.\quad z=(x,y).$$

NOTE TO SELF: NEXT YEAR, define a **relation** as a set of $\langle\langle\rangle\rangle$ ordered pairs. Then $\bigcup\bigcup R$ is the set of underlying coordinates, *e.g.*,

$$\bigcup\bigcup\{\langle\langle 3,5\rangle\rangle,\langle\langle 7,8\rangle\rangle\}=\bigcup\{\{3\},\{3,5\},\{7\},\{7,8\}\}=\{3,5,7,8\}.$$

Also, we have $\{3\},\{3,5\}\subseteq\{3,5,7,8\}$, so $\{3\},\{3,5\}\in 2\hat{\ }\{3,5,7,8\}$, so $\{\{3\},\{3,5\}\}\subseteq 2\hat{\ }\{3,5,7,8\}$, so $\{\{3\},\{3,5\}\}\in 2\hat{\ }2\hat{\ }\{3,5,7,8\}$. That is, $\langle\langle 3,5\rangle\rangle\in 2\hat{\ }2\hat{\ }\{3,5,7,8\}$. Similarly, $\langle\langle 7,8\rangle\rangle\in 2\hat{\ }2\hat{\ }\{3,5,7,8\}$. Then

$$\{\langle\langle 3,5\rangle\rangle,\langle\langle 7,8\rangle\rangle\}\subseteq 2\hat{\ }2\hat{\ }\{3,5,7,8\},$$

so

$$\{\langle\langle 3,5\rangle\rangle,\langle\langle 7,8\rangle\rangle\}\in 2\hat{\ }2\hat{\ }2\hat{\ }\{3,5,7,8\}.$$

Thus, if S is the set of underlying coordinates of a relation R, then we have $R \subseteq 2\hat{\ }2\hat{\ }S$, and so $R \in 2\hat{\ }2\hat{\ }2\hat{\ }S$. END OF NOTE TO SELF

So, for example,

$$\begin{array}{rcl} R_0 & := & \{(x,y) \in \mathbb{R} \times \mathbb{R} \mid 4x^2 + y^2 = 1\} \qquad \text{and} \\ R_1 & := & \{2,3,4\} \times \{1,2,3\} \end{array}$$

are both relations. A logic purist would rewrite the definition of R_0:

$$R_0 := \{z \in \mathbb{R} \times \mathbb{R} \,|\, \exists x, y \in \mathbb{R} \text{ s.t. } (z = (x,y)) \,\&\, (4x^2 + y^2 = 1)\}.$$

Note that

$$R_1 = \{(2,1),(2,2),(2,3),(3,1),(3,2),(3,3),(4,1),(4,2),(4,3)\}.$$

DEFINITION 4.23. *For any relation R, we define*

$$\begin{array}{rcl} dom[R] & := & \{\ x \mid \exists y \ s.t.\ (x,y) \in R\ \}, \\ im[R] & := & \{\ y \mid \exists x \ s.t.\ (x,y) \in R\ \}, \end{array}$$

The set $\mathrm{dom}[R]$ is called the **domain** of R, and the set $\mathrm{im}[R]$ is called the **image** of R.

For any relation R, for any x, y, if $(x,y) \in R$, then $\{(x,y)\} \subseteq R$, and so $x, y \in \bigcup\bigcup\bigcup\{(x,y)\} \subseteq \bigcup\bigcup\bigcup R$. The logic purist would write:

$$\begin{array}{rcl} \mathrm{dom}[R] & := & \left\{ x \in \bigcup\bigcup\bigcup R \;\middle|\; \exists y \text{ s.t. } (x,y) \in R \right\}, \\ \mathrm{im}[R] & := & \left\{ y \in \bigcup\bigcup\bigcup R \;\middle|\; \exists x \text{ s.t. } (x,y) \in R \right\}, \end{array}$$

Going forward, we will not always rewrite every set to suit the purists. However, when we are sloppy with the Axiom of Specification, any student may, at any point, ask how the set in question might be more precisely defined.

NOTE TO SELF: IN THE FUTURE, we'll have $\bigcup\bigcup$, not $\bigcup\bigcup\bigcup$. END OF NOTE TO SELF.

We leave it as an unassigned homework to show:

$$\begin{array}{rcl} \mathrm{dom}[R_0] & = & [-1/2, 1/2] \\ \mathrm{im}[R_0] & = & [-1, 1] \\ \mathrm{dom}[R_1] & = & \{1,2,3\} \qquad \text{and} \\ \mathrm{im}[R_1] & = & \{2,3,4\} \end{array}$$

DEFINITION 4.24. *Let f be a set. By f is a* **function** *we mean*

(1) *f is a relation* *and*
(2) $\forall x \in dom[f]$, $\exists! y$ *s.t.* $(x, y) \in f$.

NOTE TO SELF: IN THE FUTURE, (2) will be replaced by

$\forall x \in \mathrm{dom}[f]$, $\exists! y$ s.t. $\langle\langle x, y\rangle\rangle \in f$.

END OF NOTE TO SELF.

In Definition 4.24, (2) is called the **vertical line test**. Note that $\{(x, y) \in \mathbb{R} \times \mathbb{R} \,|\, y = x^2\}$ and $\{(1, 4), (2, 4), (3, 4)\}$ are both functions, but that $\{(x, y) \in \mathbb{R} \times \mathbb{R} \,|\, x = y^2\}$ and $\{(1, 4), (2, 4), (3, 4), (2, 5)\}$ are both *NOT*.

DEFINITION 4.25. *For any function f, for any x, we define*

$$f(x) \quad := \quad \mathrm{ELT}\,\{y \in im[f] \,|\, (x, y) \in f\}.$$

Let $f_0 := \{(x, y) \in \mathbb{R} \times \mathbb{R} \,|\, y = x^2\}$, $f_1 := \{(1, 4), (2, 4), (3, 4)\}$. Then

$$\begin{aligned} f_0(3) &= 9, \\ f_0(-3) &= 9, \\ f_0(\{3\}) &= \frownie, \\ f_0(\frownie) &= \frownie, \\ f_1(1) &= 4, \\ f_1(2) &= 4, \\ f_1(3) &= 4, \\ f_1(\{1, 2, 3\}) &= \frownie, \\ f_1(0) &= \frownie \qquad \text{and} \\ f_1(\frownie) &= \frownie. \end{aligned}$$

5. Class 5 on 19 September 2017, Tu of Week 3

Recall $\langle x, y\rangle = \{\{x\}, \{x, y\}, \}$. Then $\langle 5, ☹\rangle = ☹ = \langle ☹, \rangle$. Also $\langle 1, 2\rangle \neq \langle 2, 1\rangle$. Recall $(x, y) = \{\langle 1, x\rangle, \langle 2, y\rangle\}$ and, more generally, $(x_1, \ldots, x_n) = \{\langle 1, x\rangle, \ldots, \langle n, x_n\rangle\}$. Warning: $(3, 7) = \{\langle 1, 3\rangle, \langle 2, 7\rangle\}$ is differemt from $(3, 7) = \{x \in \mathbb{R} \,|\, 3 < x < 7\}$. You have to figure out from context which is which.

NOTE TO SELF: IN THE FUTURE, we'll use $\langle 3, 7\rangle$ for the ordered pair, and we'll use $(3, 7)$ for the interval. IN THE FUTURE, we'll define $\langle 3, 7\rangle := \{\langle\langle 1, 3\rangle\rangle, \langle\langle 2, 7\rangle\rangle\}$. Similarly, IN THE FUTURE, we'll define $\langle 3, 7, 2, 2\rangle := \{\langle\langle 1, 3\rangle\rangle, \langle\langle 2, 7\rangle\rangle, \langle\langle 3, 2\rangle\rangle, \langle\langle 4, 2\rangle\rangle\}$. And, generally, IN THE FUTURE, we'll define $\langle x_1, \ldots, x_n\rangle := \{\langle\langle 1, x_1\rangle\rangle, \ldots, \langle\langle n, x_n\rangle\rangle\}$. NOT THIS YEAR, though. END OF NOTE TO SELF

Recall that R is a **relation** means that R is a set of () ordered pairs, *i.e.*, that $\forall z \in R$, $\exists x, y$ s.t. $z = (x, y)$.

NOTE TO SELF: IN THE FUTURE, a relation will be a set of $\langle\langle\rangle\rangle$ ordered pairs, and $\langle\langle x, y\rangle\rangle := \{\{x\}, \{x, y\}\}$. That is, IN THE FUTURE, we'll say that R is a relation means that $\forall z \in R$, $\exists x, y$ s.t. $z = \langle\langle x, y\rangle\rangle$. END OF NOTE TO SELF

Recall that, for any relation R, we defined

$$\begin{aligned} \mathrm{dom}[R] &:= \{x \,|\, \exists y \text{ s.t. } (x, y) \in R\}, \\ \mathrm{im}[R] &:= \{y \,|\, \exists x \text{ s.t. } (x, y) \in R\}. \end{aligned}$$

Recall also that, for any $n \in \mathbb{N}$, for any sets $A_1, \ldots, A_n$, we defined $A_1 \times \cdots \times A_n := \{(x_1, \ldots, x_n) \,|\, x_1 \in A_1, \ldots, x_n \in A_n\}$. Recall also that, for any $n \in \mathbb{N}$, for any set A, we defined $A^n := A \times A \times \cdots \times A$, with A repeated n times.

Recall that a **function** is a relation that satisfies the vertical line test. That is, f is a function means both that f is a relation and that $\forall x \in \mathrm{dom}[f]$, $\exists! y$ s.t. $(x, y) \in f$. Recall that, for any function f, for any x, we defined $f(x) := \mathrm{ELT}\{y \in \mathrm{im}[f] \,|\, (x, y) \in f\}$. Sometimes, we will use f_x to denote $f(x)$. For any sets A and B, for any function f, if we have $\mathrm{dom}[f] = A \times B$, then, for any $x \in A$, for any $y \in B$, we define $f(x, y) := f((x, y))$ and $xfy := f((x, y))$. The notation xfy is particularly common when f is a special symbol (not a letter). So for example, for any sets A, B, C, for any function $* : A \times B \to C$, for any $x \in A$, for any $y \in B$, we would typically use $x * y$ to denote $*(x, y)$.

NOTE TO SELF: IN THE FUTURE, we'll write $f(x, y) = f(\langle\langle x, y\rangle\rangle)$.

For example, let

$$\begin{aligned} f_0 &:= \{(x, y) \in \mathbb{R}^2 \,|\, y = x^2\} \\ f_1 &:= \{(1, 4), (2, 4), (3, 4)\}. \end{aligned}$$

We calculated the domain and image of f_0 and f_1.

DEFINITION 5.1. *Let A, B and f be sets. Then $f : A \to B$ means*

- *f is a function,*
- *$dom[f] = A$, and*
- *$im[f] \subseteq B$.*

Writing $f_0 := \{(x, y) \in \mathbb{R}^2 \,|\, y = x^2\}$ will be considered bad form in this class; even though it's technically correct, it's hard to read. Instead, we will say, "Let $f_0 : \mathbb{R} \to \mathbb{R}$ be defined by $f_0(x) = x^2$" or, equivalently, "Define $f_0 : \mathbb{R} \to \mathbb{R}$ by $f_0(x) = x^2$". We calculated

- $f_0(3) = 9 = f_0(-3)$,
- $f_0(5) = 25 = f_0(-5)$,
- $f_0(\{3, 5\}) = \frown$,
- $f_0(\infty) = \frown$ and
- $f_0(\frown) = \frown$.

DEFINITION 5.2. *Let A, B and f be sets. Then $f : A \dashrightarrow B$ means*

- *f is a function,*
- *$dom[f] \subseteq A$, and*
- *$im[f] \subseteq B$.*

By **superdomain**, we mean any superset of the domain. By **superimage**, we mean any superset of the image. Some mathematicians **co-domain** is used to mean superimage, others use **target**, and still others use **range**. Be wary of the word "range" though, since some mathematicians use **range** to mean image.

In Definition 5.1, we see that "$f : A \to B$" indicates that A is the domain of f, and that B is a superimage of f. On the other hand, in 5.2, we see that "$f : A \dashrightarrow B$" indicates that A is a superdomain of f, and that B is a superimage of f.

Writing $f_1 := \{(1,4),(2,4),(3,4)\}$ will be considered bad form in this class; even though it's technically correct, it's hard to read. Instead, we will say, "Let $f_1 : \mathbb{R} \dashrightarrow \mathbb{R}$ be defined by

$$f_1(x) = \begin{cases} 4, & \text{if } x = 1 \\ 4, & \text{if } x = 2 \\ 4, & \text{if } x = 3 \quad ", \end{cases}$$

or, equivalently, "Define $f_1 : \{1,2,3\} \to \mathbb{R}$ by $f_1(x) = 4$". We have:

- $f_1(1) = f_1(2) = f_1(3) = 4$,
- $f_1(4) = ☹$,
- $f_1(\{1,2,3\}) = ☹$,
- $f_1(\infty) = ☹$ and
- $f_1(☹) = ☹$.

For another example, let $h : \mathbb{R} \dashrightarrow \mathbb{R}$ be defined by $h(x) = 1/x$. Then, technically, we have $h = \{(x,y) \in \mathbb{R}^2 \,|\, y = 1/x\}$, but this is considered difficult to read. We have:

- $h(5) = 1/5$,
- $h(-5) = -1/5$,
- $h(0) = ☹$,
- $h(\infty) = ☹$,
- $h(\{5\}) = ☹$ and
- $h(☹) = ☹$.

For another example, let $h_1 : \mathbb{R} \dashrightarrow \mathbb{R}$ be defined by

$$h_1(x) \quad = \quad \frac{1}{x^7 + 5x^3 + 4}.$$

Note that $\mathrm{dom}[h_1] = \{x \in \mathbb{R} \,|\, x^7 + 5x^3 + 4 \neq 0\}$. Then h_1 is an example of a function whose domain is difficult to calculate precisely, and so it's nice that we have the $\dashrightarrow$ notation.

Finally, one more example. Define $q : \mathbb{R}^2 \to \mathbb{R}$ by $q(x,y) = x$. Note that, as is typical, $q(x,y)$ is an abbreviation of $q((x,y))$. Technically, we have $q = \{((x,y),z) \in \mathbb{R}^2 \times \$ \,|\, z = x\}$, but this is considered difficult to read. Question: Is $q(5,☹)$ equal to 5 or to ☹? Note that we have, as usual, $q(☹) = ☹$. Recall that $(5,☹) = ☹$. Note that, as is typical, $q(5,☹)$ is an abbreviation of $q((5,☹)$. Then

$$q(5,☹) \quad = \quad q((5,☹)) \quad = \quad q(☹) \quad = \quad ☹.$$

Recall that, for any set S, we defined $2^S := \{\text{subsets of } S\}$. Note that, for any sets A and B, for any $f : A \to B$, we have $f \subseteq A \times B$, or, equivalently, we have $f \in 2^{A \times B}$.

DEFINITION 5.3. *For any sets A and B, we define*

$$A^B \quad := \quad \{\textit{functions } B \to A\}.$$

That is, A^B denotes the set of all function with domain B and superimage A. More precisely, we should write

$$A^B \quad := \quad \{f \in 2^{A \times B} \,|\, f : B \to A\}.$$

We recalled the computation of $2^{\{3,4,5\}}$, which is a set of all the eight subsets of $\{3,4,5\}$. We computed $\{0,1\}^{\{3,4,5\}}$ and verified that these functions are in correspondence with the sets in $\{3,4,5\}$.

Recall that $A^n = A \times \cdots \times A$. We now give a second definition to A^n, as follows:

DEFINITION 5.4. *For all $n \in \mathbb{N}$, for any set A, we define $A^n := A^{\{1,\ldots,n\}}$.*

From here on out, when you see A^n, you should use Definition 5.4.

Recall that we defined $(x_, \ldots, x_n) := \{\langle 1, x_1 \rangle, \ldots, \langle n, x_n \rangle\}$. We now give a second definition to $(x_1, \ldots, x_n)$, as follows:

DEFINITION 5.5. *For all $n \in \mathbb{N}$, for all $x_1, \ldots, x_n$, we define*

$$(x_1, \ldots, x_n) \quad := \quad \{\,(1, x_1)\,,\ \ldots\ ,\ (n, x_n)\,\}.$$

NOTE TO SELF: NEXT YEAR, we won't need Definition 5.4 or Definition 5.5. We'll only have one definition of A^n and one definition of $(x_1, \ldots, x_n)$. END OF NOTE TO SELF

From here on out, when you see $(x_1, \ldots, x_n)$, you should use Definition 5.5. For example, let $v := (4,7,9)$. Then $v = \{(1,4),(2,7),(3,9)\}$. Note that $v : \{1,2,3\} \to \{4,7,9\}$ and that $v(1) = 4$, $v(2) = 7$, $v(3) = 9$. Typically, in this situation, we'll use

v_1 instead of $v(1)$,
v_2 instead of $v(2)$ and
v_3 instead of $v(3)$.

Thus, we would write $v_1 = 4$, $v_2 = 7$ and $v_3 = 9$.

Recall that we defined

$$A_1 \times \cdots \times A_n \quad := \quad \{(x_1, \ldots, x_n) \,|\, x_1 \in A_1, \ldots, x_n \in A_n\}.$$

Now that $(x_1, x_2, \ldots, x_n)$ has two definitions (new and old), we see that $A_1 \times \ldots \times A_n$ has two definitions (new and old). From here on out, the new definition is what we will mean by $A_1 \times \cdots \times A_n$.

DEFINITION 5.6. *Let A and B be sets, and let $f : A \to B$. Then*

(1) *f is* **one-to-one** *or* **injective** *means:*

$$\forall y, z \in B, \qquad (\,[f(y) = f(z)] \Rightarrow [y = z]\,).$$

(2) *f is* **onto** *B or* **surjective onto** *B means: $im[f] = B$.*
(3) *f is* **bijective** *onto B means:*

$$(f \text{ is one-to-one}) \quad \& \quad (f \text{ is onto } B).$$

REMARK 5.7. *Let A and B be sets, and let $f : A \to B$. Then:*

$$(\ f \text{ is onto } B\ ) \Leftrightarrow (\ \forall y \in B, \exists x \in A \text{ s.t. } f(x) = y\ ).$$

Proof. Proof of $\Rightarrow$: Assume that f is onto B. We wish to show:

$$\forall y \in B, \quad \exists x \in A \quad \text{s.t.} \quad f(x) = y$$

Let $y \in B$ be given. We wish to show: $\exists x \in A$ s.t. $f(x) = y$.

Since f is onto B, we have $\text{im}[f] = B$. Then

$$y \quad \in \quad B \quad = \quad \text{im}[f] \quad = \quad \{f(x) \,|\, x \in A\}.$$

Choose $x \in A$ such that $y = f(x)$. We wish to show: $f(x) = y$. By choice of x, we have $f(x) = y$, as desired. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume: $\forall y \in B$, $\exists x \in A$ s.t. $f(x) = y$. We wish to show that f is onto B. That is, we wish to show that $\text{im}[f] = B$.

Since $f : A \to B$, it follows that $\text{im}[f] \subseteq B$. We wish to show that $B \subseteq \text{im}[f]$. We wish to show, for all $y \in B$, that $y \in \text{im}[f]$. Let $y \in B$ be given. We wish to show: $y \in \text{im}[f]$.

By our assumption, choose $x \in A$ such that $f(x) = y$. Then we have $y = f(x) \in \text{im}[f]$, as desired. *End of proof of $\Leftarrow$.* □

REMARK 5.8. *Define $g : \mathbb{N}_0 \to \mathbb{N}$ by $g(x) = x + 1$. Then g is bijective onto $\mathbb{N}$.*

Proof. We wish to show:

(1) g is one-to-one and
(2) g is onto $\mathbb{N}$.

Proof of (1): We wish to show, for all $w, x \in \mathbb{N}$, that

$$(g(y) = g(z)) \quad \Rightarrow \quad (y = z).$$

Let $w, x \in \mathbb{N}$ be given. We wish to show:

$$(g(y) = g(z)) \quad \Rightarrow \quad (y = z).$$

Assume $g(y) = g(z)$. We wish to show $y = z$. We have

$$y + 1 \quad = \quad g(y) \quad = \quad g(z) \quad = \quad z + 1.$$

Then $y + 1 = z + 1$, and so $y = z$, as desired. *End of proof of (1).*

Proof of (2): By Lemma 5.7, it suffices to show:

$$\forall y \in \mathbb{N}, \quad \exists x \in \mathbb{N}_0 \quad \text{s.t.} \quad g(x) = y.$$

Let $y \in \mathbb{N}$ be given. We wish to show: $\exists x \in \mathbb{N}_0 \quad \text{s.t.} \quad g(x) = y$. Since $y \in \mathbb{N}$, it follows that $y - 1 \in \mathbb{N}_0$. Let $x = y - 1$. Then $x \in \mathbb{N}_0$, and we wish to show that $g(x) = y$.

We have $g(x) = x + 1 = (y - 1) + 1 = y$. *End of proof of (2).* $\square$

Assigned HW#11.

DEFINITION 5.9. *For all $y \geqslant 0$, we define*

$$\sqrt{y} \quad := \quad \text{ELT}\{x \geqslant 0 \,|\, y = x^2\}.$$

DEFINITION 5.10. *For any functions f and g, we define the* **composite function**, *$g \circ f$, by:* $\quad \forall x, \ (g \circ f)(x) = g(f(x))$.

For the logic purist:

$$\begin{aligned} g \circ f = \{v \in (\text{dom}[f]) \times (\text{im}[g]) \,|\, & \\ & \exists x \in \text{dom}[f],\ \exists y \in (\text{im}[f]) \cap (\text{dom}[g]),\ \exists z \in \text{im}[g] \\ & \text{s.t. } (x, y) \in f,\ (y, z) \in g,\ v = (x, z)\}. \end{aligned}$$

For example: Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^5 + 7x + 1$ and define $g : \mathbb{R} \dashrightarrow \mathbb{R}$ by $g(x) = \sqrt{x}$. Then, for all $x \in \mathbb{R}$, we have $(g \circ f)(x) = \sqrt{x^5 + 7x + 1}$. Note that $\text{dom}[g \circ f] = \{x \in \mathbb{R} \,|\, x^5 + 7x + 1 \geqslant 0\}$. This domain is a finite union of intervals, but calculating the endpoints of those intervals is difficult, and we won't attempt it.

DEFINITION 5.11. *Let A be a set. Then $\text{id}_A : A \to A$ is defined by $\text{id}_A(x) = x$.*

DEFINITION 5.12. *Let f be a function and let S be a set. Then the* **forward image** *of S under f is*

$$f_*(S) \quad := \quad \{f(x)\,|\,x \in S \cap (dom[f])\}.$$

For the logic purist:

$$f_*(S) = \{y \in \mathrm{im}[f]\,|\,\exists x \in S \cap (\mathrm{dom}[f]) \text{ s.t. } y = f(x)\}.$$

DEFINITION 5.13. *Let f be a function and let T be a set. Then the* **backward image** *or* **preimage** *of T under f is*

$$f^*(T) \quad := \quad \{x \in dom[f]\,|\,f(x) \in T\}.$$

For example, define $s : \mathbb{R} \to \mathbb{R}$ by $s(x) = x^2$. Then

$$\begin{aligned} s_*(\{1,2,3\}) &= \{1,4,9\} \qquad \text{and} \\ s^*(\{1,2,3\}) &= \{1,-1,\sqrt{2},-\sqrt{2},\sqrt{3},-\sqrt{3}\}. \end{aligned}$$

6. Class 6 on 21 September 2017, Th of Week 3

Assigned HW#12.

Recall: $\forall y \geqslant 0$, $\sqrt{y} := \mathrm{ELT}\{x \geqslant 0\,|\,y = x^2\}$.
Recall: $\forall$fns f, g, $\forall x$, $(g \circ f)(x) = g(f(x))$.
Recall: $\forall$set A, $\mathrm{id}_A : A \to A$ is defined by $\mathrm{id}_A(x) = x$.
Recall: $\forall$fn f, $\forall$set S, $f_*(S) = \{f(x)\,|\,x \in S \cap (\mathrm{dom}[f])$.
Recall: $\forall$fn f, $\forall$set T, $f^*(T) = \{x \in \mathrm{dom}[f]\,|\,f(x) \in T\}$.

DEFINITION 6.1. *For any function f, for any $S \subseteq dom[f]$, the function $f|S : S \to im[f]$ is defined by $(f|S)(x) = x$.*

Note that, in Definition 6.1, the equation "$(f|S)(x) = x$" is understood to be quantified by "$\forall x \in S$". For any $x \in (\mathrm{dom}[f])\backslash S$, we have $x \notin S = \mathrm{dom}[f|S]$, and so $(f|S)(x) = ☹$.

As an example, define $s : \mathbb{R} \to \mathbb{R}$ by $s(x) = x^2$. Then s might be called the "squaring function". Then $s(2) = 4$ and $s(-2) = 4$, so s is NOT one-to-one. However, if we restrict s to $[0,\infty)$, then we end up with a different function $r := s|[0,\infty)$, which might be called the "restricted squaring function". This *restricted* squaring function it IS one-to-one. Note, for example, that $r(2) = 4$ and $r(-2) = ☹$.

Assigned HW#13

REMARK 6.2. $\forall a, b$, $[\,(a \in \{b\}) \Leftrightarrow (a = b)\,]$.

Proof. Omitted. Follows from the definition of $\{b\}$. □

REMARK 6.3. *$\forall$fn f, $\forall y$,*

$$f^*(\{y\}) \quad = \quad \{x \in dom[f] \,|\, f(x) = y\}.$$

Proof. Let a function f be given, and let y be given. We wish to show:

$$f^*(\{y\}) \quad = \quad \{x \in \mathrm{dom}[f] \,|\, f(x) = y\}.$$

By Remark 6.2, for all x, $(f(x) \in \{y\}) \Leftrightarrow (f(x) = y)$. Then

$$\begin{aligned} f^*(\{y\}) &= \{x \in \mathrm{dom}[f] \,|\, f(x) \in \{y\}\} \\ &= \{x \in \mathrm{dom}[f] \,|\, f(x) = y\}, \end{aligned}$$

as desired. □

REMARK 6.4. *$\forall$fn f, $\forall y$,*

$$(y \in im[f]) \quad \Leftrightarrow \quad (f^*(\{y\}) \neq \varnothing).$$

Proof. Let a function f be given, and let y be given. We wish to show:

$$(y \in \mathrm{im}[f]) \quad \Leftrightarrow \quad (f^*(\{y\}) \neq \varnothing).$$

By Remark 6.3, we have:

$$f^*(\{y\}) \quad = \quad \{x \in \mathrm{dom}[f] \,|\, f(x) = y\}.$$

Then: $\quad (f^*(\{y\}) \neq \varnothing) \quad \Leftrightarrow \quad (\exists x \in \mathrm{dom}[f] \text{ s.t. } f(x) = y).$

Recall that $\mathrm{im}[f] = \{f(x) \,|\, x \in \mathrm{dom}[f]\}$. Then

$$\begin{aligned} (y \in \mathrm{im}[f]) &\Leftrightarrow (\exists x \in \mathrm{dom}[f] \text{ s.t. } f(x) = y) \\ &\Leftrightarrow (f^*(\{y\}) \neq \varnothing), \end{aligned}$$

as desired. □

Assigned HW#14.

A set is nonempty iff it has at least one element:

FACT 6.5. *$\forall$set S, we have*

$$(\#S \geqslant 1) \quad \Leftrightarrow \quad (S \neq \varnothing).$$

Proof. Omitted. □

A set has more than one element iff it has two elements that are not equal to one another:

FACT 6.6. *$\forall$set S, we have*

$$(\#S > 1) \quad \Leftrightarrow \quad (\exists w, x \in S \text{ s.t. } w \neq x).$$

Proof. Omitted. □

THEOREM 6.7. *Let f be a one-to-one function. Then, $\forall y \in im[f]$, we have: $\#[f^*(\{y\})] = 1$.*

Proof. Let $y \in \mathrm{im}[f]$ be given. We wish to prove that $\#[f^*(\{y\})] = 1$.

Since $y \in \mathrm{im}[f]$, it follows, from Remark 6.4, that $f^*(\{y\}) \neq \varnothing$. Then, by Fact 6.5, we have $\#[f^*(\{y\})] \geqslant 1$, and it remains to show: $\#[f^*(\{y\})] \leqslant 1$. Assume $\#[f^*(\{y\})] > 1$. We aim for a contradiction.

As $\#[f^*(\{y\})] > 1$, by Fact 6.6, choose $w, x \in f^*(\{y\})$ s.t. $w \neq x$. By definition of $f^*(\{y\})$, since $w, x \in f^*(\{y\})$, we get $f(w), f(x) \in \{y\}$. Then $f(w) = y$ and $f(x) = y$. Then $f(w) = f(x)$. So, since f is one-to-one, we conclude that $w = x$. However, by choice of w and x, we have: $w \neq x$. Contradiction. □

FACT 6.8. *Let f be a function. Then f is onto $im[f]$.*

Proof. Since $\mathrm{im}[f] = \mathrm{im}[f]$, it follows, from the definition of "onto $\mathrm{im}[f]$", that f is onto $\mathrm{im}[f]$. □

DEFINITION 6.9. *For any one-to-one function, the function f^{-1} is defined by*

$$\forall y, \qquad f^{-1}(y) \ = \ \mathrm{ELT}[f^*(\{y\})].$$

Note, for any one-to-one function f, that

(1) $f \ : \ \mathrm{dom}[f] \ \to \ \mathrm{im}[f]$,
(2) f is bijective onto $\mathrm{im}[f]$,
(3) $f^{-1} \ : \ \mathrm{im}[f] \ \to \ \mathrm{dom}[f]$ and
(4) f^{-1} is bijective onto $\mathrm{dom}[f]$.

All four of these are unassigned homework.

Example: Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 3x + 5$. We leave it as unassigned homework to show that f is one-to-one. We calculate

$$\begin{aligned} f^{-1}(9) &= \mathrm{ELT}[f^{-1}(\{y\})] \\ &= \mathrm{ELT}\{x \in \mathbb{R} \,|\, f(x) = 9\} \\ &= \mathrm{ELT}\{x \in \mathbb{R} \,|\, 3x + 5 = 9\} \\ &= \mathrm{ELT}\{x \in \mathbb{R} \,|\, x = 4/3\} \\ &= \mathrm{ELT}\{4/3\} \ = \ 4/3. \end{aligned}$$

Assigned HW#15.

Let $r : \mathbb{R} \dashrightarrow \mathbb{R}$ be defined by $r(x) = \sqrt{x}$. (Note that $\mathrm{dom}[r] = [0, \infty)$.) Let $s : \mathbb{R} \to \mathbb{R}$ be defined by $s(x) = x^2$. For any sets A and B, a function $f : A \to B$ is sometimes denoted $x \mapsto f(x) : A \to B$. For example, the function $x \mapsto x^2 : \mathbb{R} \to \mathbb{R}$ is equal to s. A function $f : \mathbb{R} \dashrightarrow \mathbb{R}$ is sometimes denoted $f(\bullet)$. For example, the function $(\bullet)^2$ is equal to s, and the function $\sqrt{\bullet}$ is equal to r.

Example (of an inverse function): The function $(\bullet)^2 : \mathbb{R} \to \mathbb{R}$ is not one-to-one, and so it does not have an inverse. However the restriction $(\bullet)^2|[0, \infty) : [0, \infty) \to \mathbb{R}$ *IS* one-to-one, and $[(\bullet)^2|[0, \infty)]^{-1} = \sqrt{\bullet}$.

For all $n \in \mathbb{N}$, for all $a_1, \ldots, a_n$, we say that $a_1, \ldots, a_n$ are **distinct** or **pairwise unequal** if, for all $i, j \in \{1, \ldots, n\}$, we have:

$$[\ i \neq j\ ] \quad \Rightarrow [\ a_i \neq a_j\ ].$$

For all $n \in \mathbb{N}$, for all sets $S_1, \ldots, S_n$, we say that $S_1, \ldots, S_n$ are **pairwise disjiont** if for all $i, j \in \{1, \ldots, n\}$, we have:

$$[\ i \neq j\ ] \quad \Rightarrow [\ S_i \cap S_j = \varnothing\ ].$$

For all $n \in \mathbb{N}$, for all $a_1, \ldots, a_n, b_1, \ldots, b_n$, if $a_1, \ldots, a_n$ are pairwise unequal, then we use the notation

$$\left(\begin{array}{c} a_1 \mapsto b_1 \\ \vdots \\ a_n \mapsto b_n \end{array} \right)$$

to denote the function $f : \{a_1, \ldots, a_n\} \to \{b_1, \ldots, b_n\}$ defined by

$$f(x) = \begin{cases} b_1, & \text{if } x = a_1 \\ \vdots & \qquad \vdots \\ b_n, & \text{if } x = a_n. \end{cases}$$

Example (of an inverse function): We have

$$\left(\begin{array}{c} 2 \mapsto 5 \\ 6 \mapsto 9 \end{array} \right)^{-1} = \left(\begin{array}{c} 5 \mapsto 2 \\ 9 \mapsto 6 \end{array} \right).$$

DEFINITION 6.10. *Let A, B and f be sets.*

(1) *By $f : A \hookrightarrow B$, we mean: ($f : A \to B$) and (f is one-to-one).*
(2) *By $f : A \to\!> B$, we mean: ($f : A \to B$) and (f is onto B).*
(3) *By $f : A \hookrightarrow\!> B$, we mean: ($f : A \hookrightarrow B$) and ($f : A \to\!> B$).*

DEFINITION 6.11. *Let A and B be sets.*

(1) *By $\exists A \hookrightarrow B$, we mean: $\exists f$ s.t. $f : A \hookrightarrow B$.*
(2) *By $\exists A \to> B$, we mean: $\exists f$ s.t. $f : A \to> B$.*
(3) *By $\exists A \hookrightarrow> B$, we mean: $\exists f$ s.t. $f : A \hookrightarrow> B$.*

For example,

$(\, \exists\{1,2\} \hookrightarrow \{3,4,5\}\,)$ and
$(\, \exists\{3,4,5\} \to> \{1,2\}\,)$ and
$(\, \exists\{1,2\} \hookrightarrow> \{3,4\}\,)$.

In fact, for any finite sets F and G, we have

$[\,(\,\exists F \hookrightarrow G\,) \quad \Leftrightarrow \quad (\#F \leqslant \#G)\,]$ and
$[\,(\,\exists G \to> F\,) \quad \Leftrightarrow \quad (\,(\#G \geqslant \#F)\,\&\,(F \neq \varnothing)\,)\,]$ and
$[\,(\,\exists F \hookrightarrow> G\,) \quad \Leftrightarrow \quad (\#F = \#G)\,]$ and

Reflexitivity of (1),(2),(3) of Definition 6.11: For any set A,

$$(\exists A \hookrightarrow A) \text{ and } (\exists A \to> A) \text{ and } (\exists A \hookrightarrow> A).$$

Transitivity of (1),(2),(3) of Definition 6.11: For any sets A, B, C,

$\big(\, [(\exists A \hookrightarrow B) \text{ and } (\exists B \hookrightarrow C)] \;\Rightarrow\; [\exists A \hookrightarrow C]\,\big)$ and
$\big(\, [(\exists A \to> B) \text{ and } (\exists B \to> C)] \;\Rightarrow\; [\exists A \to> C]\,\big)$ and
$\big(\, [(\exists A \hookrightarrow> B) \text{ and } (\exists B \hookrightarrow> C)] \;\Rightarrow\; [\exists A \hookrightarrow> C]\,\big)$.

Symmetry of (3) of Definition 6.11: For any sets A, B,

$$(\exists A \hookrightarrow> B) \quad \Rightarrow \quad (\exists B \hookrightarrow> A).$$

We will not, in this course, define the "cardinality" of a set, but it's worth understanding that, for any sets A and B, we have: (A has smaller cardinality than B) iff ($\exists A \hookrightarrow B$).

The following theorem is hard to prove, and we omit the proof, since it doesn't really belong in this course. However, to understand properly how sets are organized by cardinality, it's important to know all of the results in it. Even though we won't define cardinality explicitly, we nevertheless call this the **Cardinality Theorem**:

THEOREM 6.12. *Let S and T be sets. Then:*

(1) $(\exists S \hookrightarrow T)$ *or* $(\exists T \hookrightarrow S)$.
(2) $[(\exists S \hookrightarrow T)\&(\exists T \hookrightarrow S)] \Rightarrow [\exists S \hookrightarrow> T]$.
(3) $[\exists T \to> S] \Rightarrow [\exists S \hookrightarrow T]$.
(4) $[(\exists S \hookrightarrow T)\&(S \neq \varnothing)] \Rightarrow [\exists T \to> S]$.

Proof. Omitted. □

Item (2) of Theorem 6.12 is the **Schroeder-Bernstein Theorem**. The reverse implication is also true, and not hard to prove. So we can say: $\forall$sets S and T, we have: $[(\exists S \hookrightarrow T)\&(\exists T \hookrightarrow S)] \Leftrightarrow [\exists S \hookrightarrow\!\!> T]$.

It would be nice if we could omit "$\&(S \neq \emptyset)$" from (4) of Theorem 6.12, but that omission isn't possible, because the only function whose image is contained in the empty set is the empty function. So, for any set T, if $T \neq \emptyset$, then $\exists \emptyset \hookrightarrow T$, but $\nexists T \to\!\!> \emptyset$. It's worth remembering, though, that, by (3) and (4) of Theorem 6.12, we *CAN* say: For any nonempty sets S and T, we have:

$$[\ \exists S \hookrightarrow T\ ] \quad \Leftrightarrow \quad [\ \exists T \to\!\!> S\ ].$$

Another useful result says that $\mathbb{N}$ injects into any infinite set, so that $\mathbb{N}$ may be thought of as one of the "smallest" infinite set. That is:

THEOREM 6.13. *Let T be an infinite set. Then $\exists \mathbb{N} \hookrightarrow T$.*

We begin with an informal proof.

Proof. We wish to show that $\exists f$ s.t. $f : \mathbb{N} \hookrightarrow T$.

Since T is infinite, $T \neq \emptyset$, so choose $z_1 \in T$.

Since T is infinite, $T\backslash\{z_1\} \neq \emptyset$, so choose $z_2 \in T\backslash\{z_1\}$.

Since T is infinite, $T\backslash\{z_1, z_2\} \neq \emptyset$, so choose $z_3 \in T\backslash\{z_1, z_2\}$.

Since T is infinite, $T\backslash\{z_1, z_2, z_3\} \neq \emptyset$, so choose $z_4 \in T\backslash\{z_1, z_2, z_3\}$.

Continuing in this way we arrive at $z_1, z_2, z_3, \ldots \in T$ such that, for all $j, k \in \mathbb{N}$,

$$[\ i \neq j\ ] \quad \Rightarrow \quad [\ z_i \neq z_j\ ].$$

That is, the elements $z_1, z_2, z_3, \ldots$ of T are pairwise unequal.

Define $f : \mathbb{N} \to T$ by $f(j) = z_j$. We wish to show that f is one-to-one.

We leave it as an unassigned exercise to show that, because $z_1, z_2, z_3, \ldots$ are pairwise unequal, it follows that f is one-to-one. □

Recall that for any functions S and c, for any j, we often write S_j for $S(j)$ and c_j for $c(j)$. To make the preceding informal proof precise, we will need the **Axiom of Choice**:

Axiom of Choice. Let S be a function. Let $D := \mathrm{dom}[S]$ and assume, for all $j \in D$, that S_j is a nonempty set. Then there exists a function c such that $\mathrm{dom}[c] = D$ and such that, for all $j \in D$, we have $c_j \in S_j$.

The function c is sometimes called a **choice function** for S.

Recall that "$f \in D^D$" is the same as "$f : D \to D$". To make the preceding informal proof precise, we also need the **Fundamental Theorem of Dynamics**, which says that whenever you have a set D and a function $f \in D^D$, you can create a sequence of functions $\phi_0 = \mathrm{id}_D$, $\phi_1 = f$, $\phi_2 = f \circ f$, $\phi_3 = f \circ f \circ f$, *etc.*. More formally:

THEOREM 6.14. *Let D be a set, $f \in D^D$. Then $\exists! \phi : \mathbb{N}_0 \to D^D$ s.t.*

(1) $\phi_0 = \mathrm{id}_D$ *and*
(2) $\forall j \in \mathbb{N}, \quad \phi_j = f \circ \phi_{j-1}$.

Proof. Omitted. Idea: For all $j \in \mathbb{N}$, define

$$\begin{aligned} S_j \quad := \quad \{ \quad & \psi : \{0, \ldots, j\} \to D^D \quad | \\ & (\ \psi_0 = \mathrm{id}_D \ ) \ \& \ (\ \forall i \in \{1, \ldots, j\}, \psi_i = f \circ \psi_{i-1} \ ) \quad \}. \end{aligned}$$

Argue, by induction, that, $\forall j \in \mathbb{N}$, $\#[S_j] = 1$. For all $j \in \mathbb{N}$, define $\psi^j := \mathrm{ELT}(S_j)$. Argue, by induction, that, for all $j \in \mathbb{N}$, we have $\psi^{j+1} | \{1, \ldots, j\} = \psi^j$. Define $\phi_\bullet$ by $\phi_j = \psi^j_j$. □

In the statement of Theorem 6.14, we are using ϕ_0 to mean $\phi(0)$, and ϕ_1 to mean $\phi(1)$, *etc.*.

DEFINITION 6.15. *Let D be a set and let $f \in D^D$. Then, by $f_\circ$, we mean the unique function $\phi : \mathbb{N} \to D^D$ such that*

(1) $\phi_0 = \mathrm{id}_D$ *and*
(2) $\forall j \in \mathbb{N}, \quad \phi_j = f \circ \phi_{j-1}$.

For all $j \in \mathbb{N}_0$, by $f^j_\circ$, we mean $(f_\circ)_j = f_\circ(j) \in D^D$.

Informally, for any set D, for any $f : D \to D$, we have $f^0_\circ = \mathrm{id}_D$, and $f^1_\circ = f$, and $f^2_\circ = f \circ f$, and $f^3_\circ = f \circ f \circ f$, and so on.

DEFINITION 6.16. *Let D be a set and let $f : D \hookrightarrow\!\!> D$. Then, for any $j \in \mathbb{N}$, we define $f^{-j}_\circ := (f^{-1})^j_\circ$.*

Informally, for any set D, for any $f : D \hookrightarrow\!\!> D$, we have $f^{-1}_\circ = f^{-1}$, and $f^{-2}_\circ = f^{-1} \circ f^{-1}$, and $f^{-3}_\circ = f^{-1} \circ f^{-1} \circ f^{-1}$, and so on.

We can now give a more formal proof of Theorem 6.13:

Proof. We wish to show: $\exists z$ s.t. $z : \mathbb{N} \hookrightarrow T$.

Let $\mathcal{D} := \{X \in 2^T \,|\, \#X = \infty\}$ be the set of all infinite subsets of T. Note that $T \in \mathcal{D}$. Let $S := \mathrm{id}_{\mathcal{D}} : \mathcal{D} \to \mathcal{D}$. Then, $\forall X \in \mathcal{D}$, $S_X = X$.

By the Axiom of Choice, choose a function c such that $\mathrm{dom}[c] = \mathcal{D}$ and such that, $\forall X \in \mathcal{D}$, $c_X \in S_X$. Then, for all $X \in \mathcal{D}$, we have $c_X \in S_X = X$. For all $X \in \mathcal{D}$, as X is infinite, and as $\{c_X\}$ is finite, we see that $X \backslash \{c_X\} \in \mathcal{D}$. Define $f \in \mathcal{D}^{\mathcal{D}}$ by $f(X) = X \backslash \{c_X\}$.

For all $j \in \mathbb{N}$, we have $f_\circ^j(T) = f(f_\circ^{j-1}(T))$. Define $X : \mathbb{N}_0 \to \mathcal{D}$ by $X_j := f_\circ^j(T)$. Then

$(*)$ $\forall j \in \mathbb{N}$, $[X_j = f(X_{j-1})]$.

We leave it as unassigned homework to show:

$(**)$ $\forall j, k \in \mathbb{N}_0$, $[(j \leqslant k) \Rightarrow (X_k \subseteq X_j)]$.

Define $z : \mathbb{N} \to T$ by $z_j := c_{X_{j-1}}$. We wish to show that $z : \mathbb{N} \hookrightarrow T$. Since $z : \mathbb{N} \to T$, we need only show that z is one-to-one. We wish to show: $\forall j, k \in \mathbb{N}$, $([j \neq k] \Rightarrow [z_j \neq z_k])$. Let $j, k \in \mathbb{N}$ be given. We wish to show: $[j \neq k] \Rightarrow [z_j \neq z_k]$. Assume $j \neq k$. We wish to show: $z_j \neq z_k$. Let $m := \min\{j, k\}$ and $n := \max\{j, k\}$. Then $m < n$, and it suffices to show that $z_m \neq z_n$.

By $(**)$, we have $X_{n-1} \subseteq X_m$. Then $z_n = c_{X_{n-1}} \in X_{n-1} \subseteq X_m$. By $(*)$, we have $X_m = f(X_{m-1})$. Then

$$X_m \;=\; X_{m-1} \backslash \{c_{X_{m-1}}\} \;=\; X_{m-1} \backslash \{z_m\},$$

so $z_m \notin X_m$. So, since $z_n \in X_m$, we get $z_m \neq z_n$, as desired. □

Using Theorem 6.12 and Theorem 6.13, we can organize the **World of Sets** into various levels numbered 0, 1, 2, *etc.*. We picture the 0th level at the bottom, the 1st just above it, the 2nd just above the 1st, etc. Above all of these, there are infinte levels. Moreover:

(1) The 0th level contains only the empty set.
(2) The 1st level contains all sets S s.t. $\#S = 1$.
(3) The 2nd level contains all sets S s.t. $\#S = 2$.
(4) $\forall n \in \mathbb{N}_0$, the nth level contains all sets S s.t. $\#S = n$.
(5) The lowest infinite level contains $\mathbb{N}$.
(6) For any two sets S and T on the same level, $\exists S \hookrightarrow\!\!> T$.
(7) For any two sets S and T,
    if T is on a strictly higher level than S,
    then [$(\exists S \hookrightarrow T)$ and $(\nexists T \hookrightarrow S)$].

(8) For any two nonempty sets S and T,
if T is on a strictly higher level than S,
then [($\exists T \twoheadrightarrow S$) and ($\nexists S \twoheadrightarrow T$)].

So, in the World of Sets, a $\hookrightarrow\!\!>$ points horizontal, a $\twoheadrightarrow$ points down, and a $\hookrightarrow$ points up. We will eventually see that there is a level above the lowest infinite level, and, in fact, there is no top level.

Note: $\{5\}$ and $\{\{1,2,3\}\}$ are both on the 1st level. Note: $\{5,\{1,2,3\}\}$ is on the 2nd level. By Theorem 6.13, every infinite set is at or above the level of $\mathbb{N}$. Every finite set is strictly below the level of $\mathbb{N}$.

We explained why $\exists \mathbb{N} \hookrightarrow \mathbb{N}_0$ and $\exists \mathbb{N} \hookrightarrow \mathbb{Z}$. Thus $\mathbb{N}_0$ and $\mathbb{Z}$ are on the same level as $\mathbb{N}$.

DEFINITION 6.17. *Let S be a set. By S is* **countable**, *we mean: $\exists S \hookrightarrow \mathbb{N}$. By S is* **uncountable**, *we mean: $\nexists S \hookrightarrow \mathbb{N}$. By S is* **countably infinite** *or* **denumerable**, *we mean: $\exists S \hookrightarrow\!\!> \mathbb{N}$.*

The countably infinite sets are those on the same level as $\mathbb{N}$; in particular, $\mathbb{N}_0$ and $\mathbb{Z}$ are countably infinite.

A set is countably infinite iff it is both countable and infinite. The countable sets are those on or below the level of $\mathbb{N}$. The uncountable sets are those strictly above the level of $\mathbb{N}$.

In the next classes, we will place $\mathbb{Q}$ and $\mathbb{R}$ in the World of Sets.

7. Class 7 on 26 September 2017, Tu of Week 4

DEFINITION 7.1. *For all $n \in \mathbb{N}$, we define* $\displaystyle\frac{\mathbb{Z}}{n} := \left\{ \frac{m}{n} \,\middle|\, m \in \mathbb{Z} \right\}$.

So, for example, $\displaystyle\frac{\mathbb{Z}}{3} = \left\{ \ldots, \frac{-3}{3}, \frac{-2}{3}, \frac{-1}{3}, \frac{0}{3}, \frac{1}{3}, \frac{2}{3}, \frac{3}{3}, \ldots \right\}$.

Assigned HW#16.

DEFINITION 7.2. *For any set S, for any $A \subseteq S$, we define the function $\chi_A^S : S \to \{0,1\}$ by* $\chi_A^S(z) = \begin{cases} 1, & \text{if } z \in A \\ 0, & \text{if } z \notin A. \end{cases}$

Assigned HW#17 and HW#18 and HW#19.

The following result is called the **Archimedean Principle**:

THEOREM 7.3. $\forall x \in \mathbb{R}$, $\exists j \in \mathbb{N}$ *s.t.* $j > x$.

Proof. Omitted. $\square$

Assigned HW#20.

Recall: Theorem 6.12 and Theorem 6.13. By Theorem 6.13, we know: $\forall$set S, [(S is infinite) $\Leftrightarrow$ ($\exists \mathbb{N} \hookrightarrow S$)].

Recall: Definition 6.17. Using the Schroeder-Bernstein Theorem ((2) of Theorem 6.12), it's not hard to show: For any set S, we have

$$[\ S \text{ is countably infinite }] \Leftrightarrow [\ (\ S \text{ is countable}) \ \&\ (\ S \text{ is infinite})\ ].$$

Recall: Because $\mathbb{N}_0$ is "listable" as $\{0, 1, 2, 3, \ldots\}$, it follows that $\exists \mathbb{N} \hookrightarrow\!\!> \mathbb{N}_0$. Similarly, because $\mathbb{Z}$ is "listable" as $\{0, 1, -1, 2, -2, 3, -3, \ldots\}$, it follows that $\exists \mathbb{N} \hookrightarrow\!\!> \mathbb{Z}$. Thus $\mathbb{N}$ and $\mathbb{N}_0$ and $\mathbb{Z}$ are all three countably infinite.

Similarly, we can list the elements of $\mathbb{N} \times \mathbb{N}$ using two rules:

(1) For all $p, q, r, s \in \mathbb{N}$, if $p + q < r + s$, then we list (p, q) before (r, s).
(2) For all $p, q, r, s \in \mathbb{N}$, if $p + q = r + s$ and if $p < r$, then we list (p, q) before (r, s).

Following these two rules, we list $\mathbb{N} \times \mathbb{N}$ as follows:

$$\begin{array}{lcr} \{ & (1,1), & \\ & (1,2),(2,1), & \\ & (1,3),(2,2),(3,1), & \\ & (1,4),(2,3),(3,2),(4,1), \qquad \ldots & \}. \end{array}$$

DEFINITION 7.4. *A* **sequence** *is a function whose domain is* $\mathbb{N}$*.*

For any sequence a, for any $j \in \mathbb{N}$, it is traditional to denote $a(j)$ by a_j. For any sequence a, we will sometimes use $a_\bullet$ to denote a. If in a complicated situation, we have, say, sequences a, x and y, and real numbers b, c and z, it can be helpful to use $a_\bullet$, b, c, $x_\bullet$, $y_\bullet$, z so that the reader can easily remember which object is a sequence and which is a real number.

To indicate a sequence $a_\bullet$, we sometimes write $(a_1, a_2, a_3, \ldots)$, hoping that the pattern becomes clear for the reader. For example, defining $b_\bullet := (1, 4, 9, 16, \ldots)$ is the same as defining $b : \mathbb{N} \to \mathbb{R}$ by $a_j = j^2$. Defining $c := (5, 5, 5, 5, \ldots)$ is the same as defining $c : \mathbb{N} \to \mathbb{R}$ by $c_j = 5$.

For any sequence a, note that $\mathrm{im}[a] = \{a_j \mid j \in \mathbb{N}\}$, and we sometimes write $\mathrm{im}[a_\bullet] = \{a_1, a_2, \ldots\}$.

If we say "a is a sequence of real numbers", then we mean: "a is a sequence, and, for all $j \in \mathbb{N}$, $a_j \in \mathbb{R}$". Or, equivalently, "$a \in \mathbb{R}^{\mathbb{N}}$".

If we say "S is a sequence of sets", then we mean: "S is a sequence, and, for all $j \in \mathbb{N}$, S_j is a set".

If we say "W is a sequence of nonempty sets", then we mean: "W is a sequence, and, for all $j \in \mathbb{N}$, W_j is a nonempty set".

THEOREM 7.5. *Let $S_1, S_2, S_3, \ldots$ be countably infinite sets. Then $S_1 \cup S_2 \cup S_3 \cup \cdots$ is a countably infinite set.*

Colloquially, Theorem 7.5 asserts: "A countably infinite union of countably infinite sets is countably infinite." A logic purist would not appreciate the use of ellipses ("...”s) in Theorem 7.5, and would prefer the following more precise version:

Theorem. Let S be a sequence of countably infinite sets. Then $\bigcup(\mathrm{im}[S_\bullet])$ is a countably infinite set.

We begin with a proof of Theorem 7.5 that is appropriate to the level of our course, but would irritate the logic purist:

Proof. Let $U := S_1 \cup S_2 \cup \cdots$. We wish to show that U is countably infinite. That is, we wish to show that U is both countable and infinite.

Since $U \supseteq S_1$ and since S_1 is infinite, it follows that U is infinite. It remains to show that U is countable. We wish to show: $\exists U \hookrightarrow \mathbb{N}$. Then, by (3) of Theorem 6.12, it suffices to show $\exists \mathbb{N} \twoheadrightarrow U$. Since $\mathbb{N} \times \mathbb{N}$ is countable, it follows that $\exists \mathbb{N} \times \mathbb{N} \twoheadrightarrow \mathbb{N}$. Then, by transitivity of $\exists \bullet \twoheadrightarrow \bullet$, it suffices to show: $\exists \mathbb{N} \times \mathbb{N} \twoheadrightarrow U$. That is, we wish to show: $\exists \Phi$ s.t. $\Phi : \mathbb{N} \times \mathbb{N} \twoheadrightarrow U$.

Choose $f_1 : \mathbb{N} \hookrightarrow\!\!\!\!\!\twoheadrightarrow S_1$, $f_2 : \mathbb{N} \hookrightarrow\!\!\!\!\!\twoheadrightarrow S_2$, *etc.* Then, for all $j \in \mathbb{N}$, since $S_j \subseteq U$, we get $f_j : \mathbb{N} \to U$. Define $\Phi : \mathbb{N} \times \mathbb{N} \to U$ by $\Phi(j, k) = f_j(k)$. We wish to show Φ is onto U. That is, we wish to show: $\forall y \in U$, $\exists j, k \in \mathbb{N}$ s.t. $\Phi(j, k) = y$. Let $y \in U$ be given. We wish to show: $\exists j, k \in \mathbb{N}$ s.t. $\Phi(j, k) = y$.

Since $y \in U = S_1 \cup S_2 \cup \cdots$, choose $j \in \mathbb{N}$ such that $y \in S_j$. Since $f_j : \mathbb{N} \to S_j$ is onto S_j, it follows that $\mathrm{im}[f_j] = S_j$. Then $y \in S_j = \mathrm{im}[f_j]$, so choose $k \in \mathbb{N}$ such that $f_j(k) = y$. We wish to show: $\Phi(j, k) = y$.

We have $\Phi(j, k) = f_j(k) = y$, as desired. $\square$

From the purist's point of view, there are two problems with the proof given above: the use of "*etc.*" and the use of ellipses ("...s"). The second one is easily fixed, because we can

(1) replace "$U := S_1 \cup S_2 \cup \cdots$" with "$U := \bigcup(\mathrm{im}[S_\bullet])$" and
(2) replace "$y \in U = S_1 \cup S_2 \cup \cdots$" with "$y \in U = \bigcup(\mathrm{im}[S_\bullet])$".

A more difficult problem appears at the start of the third paragraph of the proof, where we say: "Choose $f_1 : \mathbb{N} \hookrightarrow\!> S_1$, $f_2 : \mathbb{N} \hookrightarrow\!> S_2$, *etc.*". We need to replace this with text that will satisfy the logic purist.

The basic problem is that, at the start of the third paragraph of the proof, we *KNOW*

$$\forall j \in \mathbb{N},\ \exists \mathbb{N} \hookrightarrow\!> S_j$$

or, in other words, we *KNOW*

$$\forall j \in \mathbb{N},\ \exists f \text{ s.t. } f : \mathbb{N} \hookrightarrow\!> S_j$$

and we would like to replace $\exists$ by "choose". This is similar to Replacement Rule [21] in the exposition handout, except that "$\exists$" is not appearing at the start of the known statement; it appears *AFTER* "$\forall j \in \mathbb{N}$". So, in this situation, we know that infinitely many objects exist, and we want to choose them all in one statement. A semi-purist would say that, the spirit of the Axiom of Choice is that we can make infinitely many choices at once, so we are, in fact, allowed to change "$\exists$" to "choose", *EXCEPT* that we have to alter the notation to keep track of the fact that the choices may well be different from one another. Since we start with "$\forall j \in \mathbb{N}$", we should change each "f" to "f_j" and we end up with

$$\forall j \in \mathbb{N},\ \text{choose } f_j \text{ s.t. } f_j : \mathbb{N} \hookrightarrow\!> S_j$$

Here then, is a proof for the semi-purist:

Proof. Let $U := \bigcup(\mathrm{im}[S_\bullet])$. We wish to show that U is countably infinite. That is, we wish to show that U is both countable and infinite.

Since $U \supseteq S_1$ and since S_1 is infinite, it follows that U is infinite. It remains to show that U is countable. We wish to show: $\exists U \hookrightarrow \mathbb{N}$. Then, by (3) of Theorem 6.12, it suffices to show $\exists \mathbb{N} \to\!> U$. Since $\mathbb{N} \times \mathbb{N}$ is countable, it follows that $\exists \mathbb{N} \times \mathbb{N} \to\!> \mathbb{N}$. Then, by transitivity of $\exists \bullet \to\!> \bullet$, it suffices to show: $\exists \mathbb{N} \times \mathbb{N} \to\!> U$. That is, we wish to show: $\exists \Phi$ s.t. $\Phi : \mathbb{N} \times \mathbb{N} \to\!> U$.

For all $j \in \mathbb{N}$, $\exists \mathbb{N} \hookrightarrow\!> S_j$. That is, for all $j \in \mathbb{N}$, $\exists f$ s.t. $f : \mathbb{N} \hookrightarrow\!> S_j$. Then, by the (semi-purist's) Axiom of Choice, for all $j \in \mathbb{N}$, choose f_j s.t. $f_j : \mathbb{N} \hookrightarrow\!> S_j$. Then, for all $j \in \mathbb{N}$, since $S_j \subseteq U$, we get $f_j : \mathbb{N} \to U$. Define $\Phi : \mathbb{N} \times \mathbb{N} \to U$ by $\Phi(j,k) = f_j(k)$. We wish to show Φ is onto U. That is, we wish to show: $\forall y \in U$, $\exists j,k \in \mathbb{N}$ s.t. $\Phi(j,k) = y$. Let $y \in U$ be given. We wish to show: $\exists j,k \in \mathbb{N}$ s.t. $\Phi(j,k) = y$.

Since $y \in U = \bigcup(\mathrm{im}[S_\bullet])$, choose $Z \in \mathrm{im}[S_\bullet]$ such that $y \in Z$. Since $Z \in \mathrm{im}[S_\bullet]$, choose $j \in \mathbb{N}$ such that $Z = S_j$. Since $f_j : \mathbb{N} \to S_j$ is onto S_j, it follows that $\mathrm{im}[f_j] = S_j$. Then $y \in Z = S_j = \mathrm{im}[f_j]$, so choose $k \in \mathbb{N}$ such that $f_j(k) = y$. We wish to show: $\Phi(j,k) = y$.

We have $\Phi(j,k) = f_j(k) = y$, as desired. $\square$

In reality, there isn't a "semi-purist's Axiom of Choice". The following is just a slight restatement of our earlier Axiom of Choice:

Axiom of Choice (slight restatement): Let S be a function and let D be a set. Assume that $\mathrm{dom}[S] = D$. Assume, for all $j \in D$, that S_j is a nonempty set. Then there exists a function c such that $\mathrm{dom}[c] = D$ and such that, for all $j \in \mathbb{N}$, we have $c_j \in S_j$.

The intuitive meaning is: If we're given D-many nonempty sets, then we can choose an element from each of them simultaneously. To be "given D-many nonempty sets" is to be given a single function S with domain D such that, for all $j \in D$, S_j is a nonempty set. To "choose an element from each of them simultaneously" is to choose a single function c with domain D such that, for all $j \in D$, $c_j \in S_j$.

It will be helpful to change S to A, c to f and to focus on the case where $D = \mathbb{N}$. Our slight restatment above therefore implies:

Axiom of Choice for $\mathbb{N}$: Let A be a function. Assume $\mathrm{dom}[A] = \mathbb{N}$. Assume, for all $j \in \mathbb{N}$, that A_j is a nonempty set. Then there exists a function f such that $\mathrm{dom}[f] = \mathbb{N}$ and such that, $\forall j \in \mathbb{N}$, $f_j \in A_j$.

This might be called the Axiom of Choice for $\mathbb{N}$-many nonempty sets. A function with domain $\mathbb{N}$ is called a sequence, $\mathbb{N}$-many objects is a sequence of objects. Using the terminology of sequences, we would state the last result as follows:

Axiom of Choice for sequences of nonempty sets: Let $A_\bullet$ be a sequence. Assume, for all $j \in \mathbb{N}$, that A_j is a nonempty set. Then there exists a sequence $f_\bullet$ such that, for all $j \in \mathbb{N}$, we have $f_j \in A_j$.

To use this in our proof, we need to create the sequence A in such a way that "$f_j \in A_j$" implies "$f_j : \mathbb{N} \hookrightarrow\!\!> S_j$. We can accomplish this by defining the function $A_\bullet$ by

$$\forall j \in \mathbb{N}, \qquad A_j \;=\; \{f \in U^{\mathbb{N}} \,|\, f : \mathbb{N} \hookrightarrow\!\!> S_j\}.$$

Looking at this, we see, for all $j \in \mathbb{N}$, that $A_j \subseteq U^{\mathbb{N}}$, so $A_j \in 2\hat{\ }(U^{\mathbb{N}})$. So, in our purist proof, we should define $A : \mathbb{N} \to 2\hat{\ }(U^{\mathbb{N}})$ by

$$A_j \;=\; \{f \in U^{\mathbb{N}} \,|\, f : \mathbb{N} \hookrightarrow\!\!> S_j\}.$$

(Remember that "$\forall j \in \mathbb{N}$" is understood here.)

This yields our purist's proof:

Proof. Let $U := \bigcup(\mathrm{im}[S_\bullet])$. We wish to show that U is countably infinite. That is, we wish to show that U is both countable and infinite.

Since $U \supseteq S_1$ and since S_1 is infinite, it follows that U is infinite. It remains to show that U is countable. We wish to show: $\exists U \hookrightarrow \mathbb{N}$. Then, by (3) of Theorem 6.12, it suffices to show $\exists \mathbb{N} \to\!\!> U$. Since $\mathbb{N} \times \mathbb{N}$ is countable, it follows that $\exists \mathbb{N} \times \mathbb{N} \to\!\!> \mathbb{N}$. Then, by transitivity of $\exists \bullet \to\!\!> \bullet$, it suffices to show: $\exists \mathbb{N} \times \mathbb{N} \to\!\!> U$. That is, we wish to show: $\exists \Phi$ s.t. $\Phi : \mathbb{N} \times \mathbb{N} \to\!\!> U$.

Define $A : \mathbb{N} \to 2\hat{\ }(U^{\mathbb{N}})$ by $A_j = \{f \in U^{\mathbb{N}} \,|\, f : \mathbb{N} \hookrightarrow\!\!> S_j\}$. By assumption, for all $j \in \mathbb{N}$, S_j is countably infinite. For all $j \in \mathbb{N}$, $\exists \mathbb{N} \hookrightarrow\!\!> S_j$. That is, for all $j \in \mathbb{N}$, $\exists f$ s.t. $f : \mathbb{N} \hookrightarrow\!\!> S_j$. That is, for all $j \in \mathbb{N}$, we have $A_j \neq \varnothing$. By the Axiom of Choice for sequences of nonempty sets, choose a sequence $f_\bullet$ such that, for all $j \in \mathbb{N}$, we have $f_j \in A_j$. Then, for all $j \in \mathbb{N}$, by definition of A_j, we see that $f_j : \mathbb{N} \hookrightarrow\!\!> S_j$. Define $\Phi : \mathbb{N} \times \mathbb{N} \to U$ by $\Phi(j,k) = f_j(k)$. We wish to show Φ is onto U. That is, we wish to show:

$$\forall y \in U, \quad \exists j,k \in \mathbb{N} \quad \text{s.t.} \quad \Phi(j,k) = y.$$

Let $y \in U$ be given. We wish to show: $\exists j,k \in \mathbb{N}$ s.t. $\Phi(j,k) = y$.

Since $y \in U = \bigcup(\mathrm{im}[S_\bullet])$, choose $Z \in \mathrm{im}[S_\bullet]$ such that $y \in Z$. Since $Z \in \mathrm{im}[S_\bullet]$, choose $j \in \mathbb{N}$ such that $Z = S_j$. As $f_j : \mathbb{N} \to S_j$ is onto S_j, it follows that $\mathrm{im}[f_j] = S_j$. Then $y \in Z = S_j = \mathrm{im}[f_j]$, so choose $k \in \mathbb{N}$ such that $f_j(k) = y$. We wish to show: $\Phi(j,k) = y$.

We have $\Phi(j,k) = f_j(k) = y$, as desired. □

We gave three proofs (original, semi-purist and purist) of Theorem 7.5. In this course, all three are acceptable.

COROLLARY 7.6. *The set $\mathbb{Q}$ is countably infinite.*

Proof. By HW#16, the sets

$$\frac{\mathbb{Z}}{1}, \quad \frac{\mathbb{Z}}{2}, \quad \frac{\mathbb{Z}}{3}, \quad \frac{\mathbb{Z}}{4}, \quad \ldots$$

are all countably infinite. So, since

$$\mathbb{Q} \quad = \quad \frac{\mathbb{Z}}{1} \cup \frac{\mathbb{Z}}{2} \cup \frac{\mathbb{Z}}{3} \cup \frac{\mathbb{Z}}{4} \cup \cdots,$$

it follows, from Theorem 7.5, that $\mathbb{Q}$ is countably infinite. □

This tells us that, in the World of Sets, $\mathbb{Q}$ belongs on the the countably infinite level, along with $\mathbb{N}$ and $\mathbb{N}_0$ and $\mathbb{Z}$ and $\mathbb{N} \times \mathbb{N}$. In the next class, we'll focus on placing $\mathbb{R}$. For now, we have a few additional observations to make about countable sets.

A subset of a countable set is countable:

PROPOSITION 7.7. *Let T be a countable set. Let $S \subseteq T$. Then S is countable.*

Proof. Since $\mathrm{id}_S : S \hookrightarrow T$, we conclude that $\exists S \hookrightarrow T$. Since T is countable, we conclude that $\exists T \hookrightarrow \mathbb{N}$. Then, by transitivity of $\exists \bullet \hookrightarrow \bullet$, it follows that $\exists S \hookrightarrow \mathbb{N}$. Then S is countable. □

Recall that $\bigcup \varnothing = \varnothing$. Also, $\varnothing : \varnothing \hookrightarrow \mathbb{N}$, so $\exists \varnothing \hookrightarrow \mathbb{N}$. That is, $\varnothing$ is countable.

REMARK 7.8. *Let S be a nonempty countable set. Then there exists $a \in S^{\mathbb{N}}$ such that $S = \{a_1, a_2, \ldots\}$.*

The purist would say "…such that $S = \mathrm{im}[a_\bullet]$".

Proof. Since S is countable, we have $\exists S \hookrightarrow \mathbb{N}$. So, since $S \neq \varnothing$, by (4) of Theorem 6.12, we have $\exists \mathbb{N} \twoheadrightarrow S$. Choose a such that $a : \mathbb{N} \twoheadrightarrow S$. Then $S = \mathrm{im}[a_\bullet]$, *i.e.*, $S = \{a_1, a_2, \ldots\}$. □

Keep in mind, in Remark 7.8, that there is no assumption that the sequence $a_\bullet$ is injective, and so $a_1, a_2, \ldots$ may have repititions. So, for example, if $S = \{2, 5, 8\}$, then $a_\bullet$ could, for example, be the sequence

$$(2, 5, 8, 2, 5, 8, 2, 5, 8, 2, 5, 8, 2, 5, 8, 2, 5, 8, \ldots).$$

LEMMA 7.9. *Let A be a countable set. Then $A \cup \mathbb{N}$ is countably infinite.*

Proof. Since $\mathbb{N} \subseteq A \cup \mathbb{N}$ and since $\mathbb{N}$ is infinite, it follows that $A \cup \mathbb{N}$ is infinite. It remains to show that $A \cup \mathbb{N}$ is countable.

Let $B := \{-1, -2, \ldots\}$. Then $x \mapsto -x : \mathbb{N} \hookrightarrow\!\!> B$, so $\exists \mathbb{N} \hookrightarrow\!\!> B$.

Since A is countable, $\exists A \hookrightarrow \mathbb{N}$. Since $\exists A \hookrightarrow \mathbb{N}$ and $\exists \mathbb{N} \hookrightarrow B$ we conclude that $\exists A \hookrightarrow B$. Choose f such that $f : A \hookrightarrow B$.

Define $g : A \cup \mathbb{N} \to \mathbb{Z}$ by

$$g(x) \quad = \quad \begin{cases} f(x), & \text{if } x \in A \\ x, & \text{if } x \notin A. \end{cases}$$

We leave it as an unassigned exercise to show that $g : A \cup \mathbb{N} \hookrightarrow \mathbb{Z}$. Thus $\exists A \cup \mathbb{N} \hookrightarrow \mathbb{Z}$. Since $\mathbb{Z}$ is countably infinite, $\exists \mathbb{Z} \hookrightarrow\!\!> \mathbb{N}$.

Since $\exists A \cup \mathbb{N} \hookrightarrow \mathbb{Z}$ and $\exists \mathbb{Z} \hookrightarrow \mathbb{N}$, we conclude, by transitivity, that $\exists A \cup \mathbb{N} \hookrightarrow \mathbb{N}$. That is, $A \cup \mathbb{N}$ is countable, as desired. $\square$

THEOREM 7.10. *Let $\mathcal{C}$ be a countable set of countable sets. Then $\bigcup \mathcal{C}$ is countable.*

The statement "Let $\mathcal{C}$ be a countable set of countable sets" means: "Let $\mathcal{C}$ be a set of sets, assume $\mathcal{C}$ is countable, and assume, for all $A \in \mathcal{C}$, that A is countable".

Colloquially, Theorem 7.10 asserts: "A countable union of countable sets is countable."

NOTE TO SELF:
Here's an easier approach. Given a sequence $S_\bullet$ of countable sets (some of which may be empty). For each $j \in \mathbb{N}$, form an injection from S_j into $\mathbb{N}$, then form an injection from $S_j \cup \{0\}$ into $\mathbb{N}_0$, then form a surjection from $\mathbb{N}_0$ onto $\{0\} \cup S_j$. Then use all these surjections to surject $\mathbb{N} \times \mathbb{N}_0$ onto $\{0\} \cup S_1 \cup S_2 \cup \cdots$. Then make a surjection $\mathbb{N}$ onto $\mathbb{N} \times \mathbb{N}_0$. Composing, get a surjection $\mathbb{N}$ onto $\{0\} \cup S_1 \cup S_2 \cup \cdots$. This gives an injection $\{0\} \cup S_1 \cup S_2 \cup \cdots$ into $\mathbb{N}$. Compose with the inclusion $S_1 \cup S_2 \cup \cdots \hookrightarrow \{0\} \cup S_1 \cup S_2 \cup \cdots$, and we get an injection $S_1 \cup S_2 \cup \cdots \hookrightarrow \mathbb{N}$. Then $S_1 \cup S_2 \cup \cdots$ is countable.
END OF NOTE TO SELF.

The proof below, of Theorem 7.10, illustrates how when we know a compound statement with "or", we can break the proof into two sections. This process is described in [16] of the exposition handout.

Proof. Let $U := \bigcup \mathcal{C}$. We wish to show: U is countable.

Either (1) $\mathcal{C} = \varnothing$ or (2) $\mathcal{C} \neq \varnothing$.

Proof in Case (1): As $\mathcal{C} = \varnothing$, we get $U = \bigcup \mathcal{C} = \bigcup \varnothing = \varnothing$. So, as $\varnothing$ is countable, we see that U is countable. *End of proof in Case (1).*

Proof in Case (2): Since $\mathcal{C}$ is nonempty and countable, by Remark 7.8, choose $A \in \mathcal{C}^{\mathbb{N}}$ such that $\mathcal{C} = \{A_1, A_2, \ldots\}$.

Then $U = \bigcup \mathcal{C} = A_1 \cup A_2 \cup \cdots$.

Note that, for all $j \in \mathbb{N}$, we have $A_j \in \mathcal{C}$, so $A_j \subseteq \bigcup \mathcal{C} = U$. Let $V := U \cup \mathbb{N}$. Then, for all $j \in \mathbb{N}$, we have $A_j \cup \mathbb{N} \subseteq U \cup \mathbb{N} \subseteq V$, so $A_j \cup \mathbb{N} \in 2^V$. Define $S : \mathbb{N} \to 2^V$ by $S_j = A_j \cup \mathbb{N}$. By Lemma 7.9, for all $j \in \mathbb{N}$, S_j is countably infinite. Then, by Theorem 7.5, we conclude that $S_1 \cup S_2 \cup \cdots$ is countable. So, since

$$\begin{aligned} S_1 \cup S_2 \cup \cdots &= (A_1 \cup \mathbb{N}) \cup (A_2 \cup \mathbb{N}) \cup \cdots \\ &= (A_1 \cup A_2 \cup \cdots) \cup \mathbb{N} = U \cup \mathbb{N} = V, \end{aligned}$$

we conclude that V is countable. So, since $U \subseteq U \cup \mathbb{N} = V$, by Proposition 7.7, that U is countable. *End of proof in Case (2).* □

There are several ellipses appearing in this proof of Theorem 7.10. Resolving them to rigor is an exercise for the interested reader.

8. Class 8 on 28 September 2017, Th of Week 4

DEFINITION 8.1. *Let X be a set. An X-**valued sequence** is a function whose domain is $\mathbb{N}$ and whose image is a subset of X.*

A **sequence of real numbers** is a sequence $a_\bullet$ such that, for all $j \in \mathbb{N}$, we have: a_j is a real numbers. That is, A **sequence of real numbers** is an $\mathbb{R}$-valued sequence. A **sequence of sets** is a sequence $A_\bullet$ such that, for all $j \in \mathbb{N}$, we have: A_j is a set. A **sequence of propositions** is a sequence $P_\bullet$ such that, for all $j \in \mathbb{N}$, we have: P_j is a proposition. *Etc.*

Let $a_\bullet$ be a sequence. Recall: $a_\bullet$ is denoted $(a_1, a_2, a_3, \ldots)$, if we can list enough terms of $a_\bullet$ that the pattern becomes apparent. For example, if $a_\bullet = (0, 1, 0, 1, 0, 1, 0, 1, \ldots)$, then, for all $j \in \mathbb{N}$, we have

$$a_j = \begin{cases} 1, & \text{if } j \text{ is even} \\ 0, & \text{if } j \text{ is odd.} \end{cases}$$

Note that, denoting the set of integers by $2\mathbb{N}$, then the characteristic function $\chi_{2\mathbb{N}}^{\mathbb{N}}$ is equal to $(0, 1, 0, 1, 0, 1, 0, 1, \ldots)$.

REMARK 8.2. *For all $m \in \{0, 1\}$, we have $1 - m \neq m$.*

Proof. Unassigned homework. □

THEOREM 8.3. *For all $z : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$, there exists $a \in \{0,1\}^{\mathbb{N}}$ such that for all $j \in \mathbb{N}$, $a \neq z_j$.*

A **bit** is an element of $\{0,1\}$. Let's use "bit-sequence" to mean "sequence of bits", *i.e.*, an element of $\{0,1\}^{\mathbb{N}}$. The theorem says: for any sequence $(z_1, z_2, \ldots)$ of bit-sequences, there exists a bit-sequence a that is not equal to any of the terms of the sequence.

We interpreted this as a game. You give me z, I give you a, you pick j and we then check whether $a \neq z_j$. If $a \neq z_j$, then I win; otherwise, you win.

We represented z as an infinite two-dimensional array of bits, where, for all $j, k \in \mathbb{N}$, the bit $z_j(k)$ is in the (j,k)-entry of the array. The kth row of the array is the bit-sequence $z_k \in \{0,1\}^{\mathbb{N}}$. You hope to make z so robust that *every* bit-sequence appears as one of the rows. We represent the bit-sequence a as a one-dimensional horizontal array of bits. I try to choose an a that is not equal to any row of the array. The theorem says that, with good play, I should win.

My strategy: I figure out $(a(1), a(2), a(3), \ldots)$ one term at a time. First, I define $a(1) := 1 - [z_1(1)]$, which guarantees that $a \neq z_1$. In other words, I look at the bit $z_1(1)$ and define $a(1)$ to be the *other* bit, so that $a(1) \neq z_1(1)$. This guarantees that my a will *not* be equal to the first row of z. You, therefore, will not win if you choose $j = 1$. Next, I define $a(2) := 1 - [z_2(2)]$. This guarantees that my a will *not* be equal to the second row of z. You, therefore, will not win if you choose $j = 2$. Next, I define $a(3) := 1 - [z_3(3)]$. This guarantees that my a will *not* be equal to the second row of z. You, therefore, will not win if you choose $j = 3$. *Etc.*

The strategy described in the last paragraph is sometimes called **Cantor diagonalization**, because we look at the diagonal entries

$$z_1(1), \quad z_2(2), \quad z_3(3), \quad z_4(4), \quad \ldots$$

in the array, and then use them to define the bit-sequence a.

We now give the formal proof of Theorem 8.3:

Proof. Let $z : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ be given. We wish to show that there exists $a \in \{0,1\}^{\mathbb{N}}$ such that for all $j \in \mathbb{N}$, $a \neq z_j$.

Define $a \in \{0,1\}^{\mathbb{N}}$ by $a(j) = 1 - [z_j(j)]$. We wish to show: for all $j \in \mathbb{N}$, $a \neq z_j$. Let $j \in \mathbb{N}$ be given. We wish to show $a \neq z_j$.

By Remark 8.2, $1 - [z_j(j)] \neq z_j(j)$. So, since $a(j) = 1 - [z_j(j)]$, we conclude that $a(j) \neq z_j(j)$. Then $a \neq z_j$, as desired. $\square$

What a cool argument! Very useful, too:

COROLLARY 8.4. $\nexists \mathbb{N} \twoheadrightarrow 2^{\mathbb{N}}$.

Proof. Assume $\exists \mathbb{N} \twoheadrightarrow 2^{\mathbb{N}}$. We aim for a contradiction.

By HW#17, $\exists \{0,1\}^{\mathbb{N}} \hookrightarrow\!\!> 2^{\mathbb{N}}$. Then, by symmetry, $\exists 2^{\mathbb{N}} \hookrightarrow\!\!> \{0,1\}^{\mathbb{N}}$. Then $\exists 2^{\mathbb{N}} \twoheadrightarrow \{0,1\}^{\mathbb{N}}$.

Since $\exists \mathbb{N} \twoheadrightarrow 2^{\mathbb{N}}$ and $\exists 2^{\mathbb{N}} \twoheadrightarrow \{0,1\}^{\mathbb{N}}$, it follows, by transitivity, that $\exists \mathbb{N} \twoheadrightarrow \{0,1\}^{\mathbb{N}}$. Choose z such that $z : \mathbb{N} \twoheadrightarrow \{0,1\}^{\mathbb{N}}$. Then z is onto $\{0,1\}^{\mathbb{N}}$. That is, we have: $\mathrm{im}[z] = \{0,1\}^{\mathbb{N}}$.

By Theorem 8.3, choose $a \in \{0,1\}^{\mathbb{N}}$ such that, for all $j \in \mathbb{N}$, $a \neq z_j$. Then $a \in \{0,1\}^{\mathbb{N}} = \mathrm{im}[z]$, so choose $j \in \mathbb{N}$ such that $a = z_j$. Then $a = z_j$ and $a \neq z_j$. Contradiction. $\square$

We can now upgrade our proofs from $\mathbb{N}$ to an arbitrary set S, both in Theorem 8.3 and in Corollary 8.4. We obtain:

THEOREM 8.5. *Let S be a set. For all $z : S \to \{0,1\}^S$, there exists $a \in \{0,1\}^S$ such that for all $j \in S$, $a \neq z_j$.*

Proof. Let $z : S \to \{0,1\}^S$ be given. We wish to show that there exists $a \in \{0,1\}^S$ such that for all $j \in S$, $a \neq z_j$.

Define $a \in \{0,1\}^S$ by $a(j) = 1 - [z_j(j)]$. We wish to show: for all $j \in S$, $a \neq z_j$. Let $j \in S$ be given. We wish to show $a \neq z_j$.

By Remark 8.2, $1 - [z_j(j)] \neq z_j(j)$. So, since $a(j) = 1 - [z_j(j)]$, we conclude that $a(j) \neq z_j(j)$. Then $a \neq z_j$, as desired. $\square$

COROLLARY 8.6. *Let S be a set. Then $\nexists S \twoheadrightarrow 2^S$.*

Proof. Assume $\exists S \twoheadrightarrow 2^S$. We aim for a contradiction.

By HW#17, $\exists \{0,1\}^S \hookrightarrow\!\!> 2^S$. Then, by symmetry, $\exists 2^S \hookrightarrow\!\!> \{0,1\}^S$. Then $\exists 2^S \twoheadrightarrow \{0,1\}^S$.

Since $\exists S \twoheadrightarrow 2^S$ and $\exists 2^S \twoheadrightarrow \{0,1\}^S$, it follows, by transitivity, that $\exists S \twoheadrightarrow \{0,1\}^S$. Choose z such that $z : S \twoheadrightarrow \{0,1\}^S$. Then z is onto $\{0,1\}^S$. That is, we have: $\mathrm{im}[z] = \{0,1\}^S$.

By Theorem 8.5, choose $a \in \{0,1\}^S$ such that, for all $j \in S$, $a \neq z_j$. Then $a \in \{0,1\}^S = \mathrm{im}[z]$, so choose $j \in S$ such that $a = z_j$. Then $a = z_j$ and $a \neq z_j$. Contradiction. $\square$

There's an interesting lesson here: The intuition beind the proof of Theorem 8.3 involved thinking of z as a two-dimensional array of bits and a as a one-dimensional array of bits. That intuition is built on our understanding that the elements of $\mathbb{N}$ are naturally ordered as $1, 2, 3, \ldots$. For an arbitrary set S, we don't have that kind of structure to its elements, so the same kind of intuition for z and a in Theorem 8.5 is harder to imagine.

Once you see the intuition behind Theorem 8.3, you may be tempted to think that there's little value in writing down a formal proof. On the other hand, once you write down such a proof, you might notice that nothing about the structure of $\mathbb{N}$ is used in the proof, and that the same proof will work for *any* set S, leading to the much better Theorem 8.5.

Formal proofs are valuable partly because they allow us to make arguments so airtight that disagreements resolve themselves quickly, making mathematics a very different subject than, say, political science. It's more than that, though. A formal proof allows us to go beyond a shallow understanding of a theorem, and to undertand, at a very deep level, what makes its proof "tick". So, for example, if a proof uses $\mathbb{N}$ over and over, but doesn't really use anything specific *about* $\mathbb{N}$, then the observant mathematician will see a possibility to generalize. By contrast, intuitive arguments often revolve around preconceived notions, like the ordering of the elements of $\mathbb{N}$.

Formality and intuition are yin and yang; neither is complete without the other. Without intuition, I can't imagine understanding most formal proofs. Without formal proofs, mathematical progress would halt. Partly this would be because disagreements would build. It's more than that, though. It would also halt because our understanding of theorems would be much more shallow.

Let S be a set. According to HW#18, in the World of Sets, we should put 2^S either at the same level or above S. According to Corollary 8.6, $\nexists S \to 2^S$ and so $\nexists S \hookrightarrow\!> 2^S$. Consequently, we cannot put 2^S at the same level as S.

The logic of the last paragraph tells us: for *any* set S, we place 2^S at a strictly higher level than S. So, while the World of Sets has a bottom set, namely $\varnothing$, it does *NOT* have a top set.

For the special case of $\mathbb{N}$, we see that $\nexists \mathbb{N} \hookrightarrow\!> 2^{\mathbb{N}}$, so $2^{\mathbb{N}}$ is *NOT* countably infinite. On the other hand $\exists \mathbb{N} \hookrightarrow 2^{\mathbb{N}}$, so $2^{\mathbb{N}}$ is infinite. Since $2^{\mathbb{N}}$ is infinite but not countably inifinite, $2^{\mathbb{N}}$ is uncountable.

So we now have many new levels in the World of Sets:

- the level of $2\hat{\ }\mathbb{N}$,
- the level of $2\hat{\ }2\hat{\ }\mathbb{N}$,
- the level of $2\hat{\ }2\hat{\ }2\hat{\ }\mathbb{N}$,
- *etc.*

Each of these levels is strictly higher than the preceding level. The first one consists of sets that are bijective with $2^{\mathbb{N}}$, and these sets are said to have **continuum cardinality** or **cc**. That is:

DEFINITION 8.7. *Let X be a set. By X* **has continuum cardinality** *or X* **has cc***, we mean: $\exists X \hookrightarrow\!> 2^{\mathbb{N}}$.*

Since $2^{\mathbb{N}}$ is uncountable, *any* set with cc is uncountable. We will eventually show that $\mathbb{R}$ has continuum cardinality, so this will show that $\mathbb{R}$ is uncountable.

It's natural to wonder if, in the World of Sets, there's any level strictly between countable and cc. That is, does there exist a set X with all four of the following properties?

(1) $\exists \mathbb{N} \hookrightarrow X$,
(2) $\nexists X \hookrightarrow \mathbb{N}$,
(3) $\exists X \hookrightarrow 2^{\mathbb{N}}$ and
(4) $\nexists 2^{\mathbb{N}} \hookrightarrow X$.

The **Continuum Hypothesis**, denoted CH asserts that no set X exists satisfying (1)-(4). *i.e.*, that there's no space between the countable and cc levels. Most mathematicians work with a set of axioms called ZFC and it is known

- both that [ZFC does not imply that CH is true]
- and that [ZFC does not imply that CH is false].

This is sometimes expressed by saying that "CH is independent of ZFC". So any mathematician is free

- either to add CH to ZFC, and work with a set of axioms that is denoted ZFC+CH,
- or to add (not CH) to ZFC, and work with a set of axioms that is denoted ZFC+(not CH).

Most of us choose to do neither, and simply to stick with ZFC. With this choice, it is unknowable whether CH is true or not.

For any infinite set S, it's natural to wonder if, in the World of Sets, there's any level strictly between S and 2^S. That is, for any set S, does there exist a set X with all four of the following properties?

(A) $\exists S \hookrightarrow X$,
(B) $\nexists X \hookrightarrow S$,
(C) $\exists X \hookrightarrow 2^S$ and
(D) $\nexists 2^S \hookrightarrow X$.

The **Generalized Continuum Hypothesis**, denoted GCH asserts, for all infinite sets S, that no set X exists satisfying (A)-(D). As with CH, GCH is independent of ZFC. As with CH, most mathematicians neither accept nor reject GCH, and consider it unknowable.

Our next goal is to show that $\mathbb{R}$ has cc, but, to get there, we need to understand $\mathbb{R}$ better, focusing on max, min, inf and sup.

DEFINITION 8.8. *Let $S \subseteq \mathbb{R}^*$ and $x \in \mathbb{R}^*$. Then:*

By $x \leqslant S$ or $S \geqslant x$, we mean: $\forall y \in S$, $x \leqslant y$.
By $x \geqslant S$ or $S \leqslant x$, we mean: $\forall y \in S$, $x \geqslant y$.

We read, "$x \leqslant S$" or "$S \geqslant x$" as "x is a lower bound for S.
We read, "$x \geqslant S$" or "$S \leqslant x$" as "x is an upper bound for S.

DEFINITION 8.9. *Let $S \subseteq \mathbb{R}^*$. Then*

$\mathrm{LB}(S) := \{x \in \mathbb{R}^* \,|\, x \leqslant S\}$ *and*
$\mathrm{UB}(S) := \{x \in \mathbb{R}^* \,|\, x \geqslant S\}$.

DEFINITION 8.10. *Let $S \subseteq \mathbb{R}^*$. Then*

$\min S := \mathrm{ELT}(S \cap [\mathrm{LB}(S)])$ *and*
$\max S := \mathrm{ELT}(S \cap [\mathrm{UB}(S)])$ *and*
$\inf S := \max(\mathrm{LB}(S))$ *and*
$\sup S := \min(\mathrm{UB}(S))$.

We read "sup" as a homonym for "soup". Sometimes you'll see "glb" to mean "inf", read "greatest lower bound. Sometimes you'll see "lub" to mean "sup", read "least upper bound.

Some examples:

S	LB	UB	min	max	inf	sup
$\{5\}$	$[-\infty, 5]$	$[5, \infty]$	5	5	5	5
$[0, 1]$	$[\infty, 0]$	$[1, \infty]$	0	1	0	1
$(0, 1)$	$[\infty, 0]$	$[1, \infty]$	☹	☹	0	1
$\{0, 1\}$	$[\infty, 0]$	$[1, \infty]$	0	1	0	1
$\{[0, 1]\}$	☹	☹	☹	☹	☹	☹
$\{(0, 1)\}$	☹	☹	☹	☹	☹	☹
$\{\{0, 1\}\}$	☹	☹	☹	☹	☹	☹
$\mathbb{R}^*$	$\{-\infty\}$	$\{\infty\}$	$-\infty$	∞	$-\infty$	∞
$\mathbb{R}$	$\{-\infty\}$	$\{\infty\}$	☹	☹	$-\infty$	∞
$\varnothing$	$\mathbb{R}^*$	$\mathbb{R}^*$	☹	☹	∞	$-\infty$
☹	☹	☹	☹	☹	☹	☹

LEMMA 8.11. *Let $S \subseteq \mathbb{R}^*$ and let $y \in \mathbb{R}^*$. Then:*

(1) $[y = \min S] \Leftrightarrow [(y \leqslant S) \& (y \in S)]$,
(2) $[y = \inf S] \Leftrightarrow [(y \leqslant S) \& (\forall x \in \mathbb{R}^*, [(x \leqslant S) \Rightarrow (x \leqslant y)])]$,
(3) $[y = \max S] \Leftrightarrow [(S \leqslant y) \& (y \in S)]$
&(4) $[y = \sup S] \Leftrightarrow [(S \leqslant y) \& (\forall z \in \mathbb{R}^*, [(S \leqslant z) \Rightarrow (y \leqslant z)])]$.

Proof. Unassigned homework. □

DEFINITION 8.12. *For all a, b, by $a =^* b$ or $b \, {}^*\!= a$, we mean:*

$$[\ b \neq ☹\ ] \quad \Rightarrow \quad [\ a = b\ ].$$

We read "$a =^* b$" or "$b \, {}^*\!= a$" as: "a is **contingent equal** to b, with the contingency on b". The meaning is that a is equal to b, *PROVIDED* b exists. If b does *NOT* exist, then a may or may not exist.

FACT 8.13. *Let $S \subseteq \mathbb{R}^*$. Then*

$$\min S \ \ {}^*\!= \ \inf S \ \neq \ ☹ \ \neq \sup S \ =^* \ \max S.$$

Proof. Omitted. □

Concerning Fact 8.13: The proof that $\min S \, {}^*\!= \inf S$ is not hard. The proof that $\sup S =^* \max S$ is not hard. The proof that $\inf S \neq$ ☹ $\neq \sup S$ is difficult, and belongs in a course on the foundations of $\mathbb{R}$, and not in our course.

That is, for any set of extended real numbers,

the inf and sup will exist and, moreover,
if the min exists, then it will equal the inf and, moreover,

if the max exists, then it will equal the sup.

FACT 8.14. *For any nonempty $S \subseteq \mathbb{N}$, we have:* $\quad \min S \neq ☹$.

Proof. Omitted. $\square$

The proof of Fact 8.14 belongs in a course on the foundations of $\mathbb{N}$, and not in our course. Sometimes Fact 8.14 is expressed by saying: "$\mathbb{N}$ is well-ordered", or "$\mathbb{N}$ is well-ordered from below".

Note that, there exist sets S such that $\max S = ☹$, *e.g.*, $S = \mathbb{N}$. So $\mathbb{N}$ is *NOT* well-ordered from above.

9. Class 9 on 3 October 2017, Tu of Week 5

We recalled Definition 8.9, Definition 8.10, Fact 8.13, Theorem 7.3, Fact 8.14 and Definition 7.4.

FACT 9.1. *For all $k \in \mathbb{N}$, we have*

$$[\ k \neq 1\ ] \quad \Rightarrow \quad [\ k - 1 \in \mathbb{N}\ ].$$

Proof. Omitted. $\square$

Next, we have the **Principle of Mathematical Induction**:

THEOREM 9.2. *Let $P_\bullet$ be a sequence of propositions. Assume P_1. Assume: $\forall j \in \mathbb{N}$, $(\, P_j \Rightarrow P_{j+1}\,)$. Then: $\forall j \in \mathbb{N}$, P_j.*

Proof. Assume $\exists j \in \mathbb{N}$ such that not P_j. We aim for a contradiction.

Let $S := \{j \in \mathbb{N} \,|\, \text{not}\, P_j\}$. Then $S \neq \varnothing$. Since P_1, we see that $1 \notin S$. By construction, we have $S \subseteq \mathbb{N}$.

Let $k := \min S$. By Fact 8.14, $k \neq ☹$. Then $k \in S \cap [\mathrm{LB}(S)]$. Then $k \in S$. Also, $k \in \mathrm{LB}(S)$, so $k \leqslant S$. Since $k \in S$, we get: not P_k. Since $k \in S$, but $1 \notin S$, we conclude that $k \neq 1$. Also, $k \in S \subseteq \mathbb{N}$. Then, by Fact 9.1, we see that $k-1 \in \mathbb{N}$. Then, by assumption, we get $P_{k-1} \Rightarrow P_j$. Since $k > k-1$, we have $k \nleqslant k-1$. So, since $k \leqslant S$, we get $k-1 \notin S$. So, as $k-1 \in \mathbb{N}$, by the definition of S, we get: P_{k-1}. So, as $P_{k-1} \Rightarrow P_k$, we get: P_k. So, since not P_k, we have a contradiction. $\square$

Let's apply the Principle of Mathematical Induction to prove:

THEOREM 9.3. *For all $j \in \mathbb{N}$, we have $1 + \cdots + j = j(j+1)/2$.*

Proof. Define a sequence $P_\bullet$ of propositions by

$$P_j \quad := \quad \left[\ 1 + \cdots + j \ = \ \frac{j(j+1)}{2}\ \right].$$

We wish to show: $\forall j \in \mathbb{N}$, P_j.

We have
$$1 = \frac{1 \cdot (1+1)}{2},$$
so P_1. By the Principle of Mathematical Induction, it now suffices to show: $\forall j \in \mathbb{N}$, $(P_j \Rightarrow P_{j+1})$. Let $j \in \mathbb{N}$ be given. We wish to show: $P_j \Rightarrow P_{j+1}$. Assume P_j. We wish to show: P_{j+1}.

We know P_j, *i.e.*, we know that
$$1 + \cdots + j = \frac{j(j+1)}{2}.$$
We wish to prove P_{j+1}, *i.e.*, we wish to prove that
$$1 + \cdots + (j+1) = \frac{(j+1)(j+2)}{2}.$$
We have
$$\begin{aligned} 1 + \cdots + (j+1) &= [1 + \cdots + j] + [j+1] \\ &= \left[\frac{j(j+1)}{2}\right] + \left[\frac{2(j+1)}{2}\right] \\ &= \frac{(j+2)(j+1)}{2} = \frac{(j+1)(j+2)}{2}. \end{aligned}$$
as desired. □

Assigned HW#21, HW#22, HW#23, HW#24, HW#25.

We now turn our attention back to showing that $\mathbb{R}$ has continuum cardinality. To do this we need to develop a theory of infinite summation.

DEFINITION 9.4. *For all* $Z \subseteq [0, \infty]$, *for all* $j \in \mathbb{N}$, *we define*
$$\mathrm{SUMS}_j(Z) \quad := \quad \{\ a_1 + \cdots + a_j \mid a : \{1, \ldots, j\} \hookrightarrow Z\ \}.$$

For example, if $Z = \{4, 7, 8\}$, then
$$\begin{aligned} \mathrm{SUMS}_1(Z) &= \{4, 7, 8\}, \\ \mathrm{SUMS}_2(Z) &= \{11, 12, 15\}, \\ \mathrm{SUMS}_3(Z) &= \{19\}, \\ \mathrm{SUMS}_4(Z) &= \varnothing, \end{aligned}$$
and, in fact, for all integers $j \geqslant 4$, we have $\mathrm{SUMS}_j(Z) = \varnothing$.

DEFINITION 9.5. *For all $Z \subseteq [0,\infty]$, we define*

$$\mathrm{SUMS}(Z) \quad := \quad [\mathrm{SUMS}_1(Z)] \ \cup [\mathrm{SUMS}_2(Z)] \ \cup \ \cdots ,$$

and we define $\operatorname{sum} Z := \sup(\mathrm{SUMS}(Z))$*.*

So, for all $Z \subseteq [0,\infty]$, we have:

$$(\ \mathrm{SUMS}(Z) \quad = \quad \bigcup \{\mathrm{SUMS}_j(Z) \,|\, j \in \mathbb{N}\}\ )$$

and ($\operatorname{sum}(Z)$ is the least upper bound of $\mathrm{SUMS}(Z)$).

If $Z = \{4,7,8\}$, then we have

$$\mathrm{SUMS}(Z) \quad = \quad \{4,7,8,11,12,15,19\}$$

and $\operatorname{sum} Z = 19$.

For another example, suppose $Z = \left\{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \ldots\right\}$. In this paragraph, we'll compute $\operatorname{sum} Z$. We explained why, for all $j \in \mathbb{N}$,

$$\mathrm{SUMS}_j(Z) \ \leqslant \ 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{j-1}} \ = \ 2 - \frac{1}{2^{j-1}} \ < \ 2.$$

Since ($\forall j \in \mathbb{N}$, $\mathrm{SUMS}_j(Z) \leqslant 2$), it follows that $\mathrm{SUMS}(Z) \leqslant 2$. Then $\sup(\mathrm{SUMS}(Z)) \leqslant 2$. For all $j \in \mathbb{N}$,

$$2 - \frac{1}{2^{j-1}} \ = \ 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{j-1}} \ \in \ \mathrm{SUMS}_j(Z) \ \subseteq \ \mathrm{SUMS}(Z).$$

Then $2-1,\ 2-\frac{1}{2},\ 2-\frac{1}{4},\ 2-\frac{1}{8},\ \ldots \ \in \ \mathrm{SUMS}(Z)$. Then

$$\left\{\ 2-1,\ 2-\frac{1}{2},\ 2-\frac{1}{4},\ 2-\frac{1}{8},\ \ldots\ \right\} \ \subseteq \ \mathrm{SUMS}(Z),$$

so

$$\sup\left\{\ 2-1,\ 2-\frac{1}{2},\ 2-\frac{1}{4},\ 2-\frac{1}{8},\ \ldots\ \right\} \ \leqslant \ \sup(\mathrm{SUMS}(Z)).$$

We explained why

$$\sup\left\{\ 2-1,\ 2-\frac{1}{2},\ 2-\frac{1}{4},\ 2-\frac{1}{8},\ \ldots\ \right\} \ = \ 2.$$

Then $2 \leqslant \sup(\mathrm{SUMS}(Z))$. So, since $\sup(\mathrm{SUMS}(Z)) \leqslant 2$, we get $\sup(\mathrm{SUMS}(Z)) = 2$. Then $\operatorname{sum} Z = \sup(\mathrm{SUMS}(Z)) = 2$.

The following fact, tells us that for any uncountable $Z \subseteq [0,\infty]$, it's easy to calculate $\operatorname{sum} Z$. So all the interesting infinite summation problems involve countable subsets of $[0,\infty]$.

FACT 9.6. *For all uncountable $Z \subseteq [0,\infty]$, we have* $\operatorname{sum} Z = \infty$*.*

Proof. Since $Z \subseteq [0, \infty]$, it follows that $Z \cap [0, \infty] = Z$. Since

$$[0, \infty] \quad = \quad \{0\} \ \cup \ (0, \infty],$$

it follows that $Z \cap [0, \infty] = (Z \cap \{0\}) \cup (Z \cap (0, \infty])$. Then

$$Z \quad = \quad Z \cap [0, \infty] \quad = \quad (Z \cap \{0\}) \ \cup \ (Z \cap (0, \infty]).$$

The set $\{0\}$ is finite, hence countable. So, since $Z \cap \{0\} \subseteq \{0\}$, we see that $Z \cap \{0\}$ is countable. By Theorem 7.10, a countable union of countable sets is countable. So, since (Z is uncountable) and since ($Z \cap \{0\}$ is countable) and since ($Z = (Z \cap \{0\}) \cup (Z \cap (0, \infty])$), we conclude that $Z \cap (0, \infty]$ is uncountable.

For all $k \in \mathbb{N}$, let $I_k := [1/k, \infty]$. Then

$$(0, \infty] = I_1 \cup I_2 \cup I_3 \cup \cdots ,$$

so

$$Z \cap (0, \infty] = (Z \cap I_1) \cup (Z \cap I_2) \cup (Z \cap I_3) \cup \cdots .$$

Then, by Theorem 7.10, choose $k \in \mathbb{N}$ such that $Z \cap I_k$ is uncountable.

Since $\operatorname{sum} Z \in \mathbb{R}^* \leqslant \infty$, it suffices to show that $\operatorname{sum} Z \geqslant \infty$. So, since $\sup\{ 1/k, 2/k, 3/k, \ldots \} = \infty$, it suffices to show that $\operatorname{sum} Z$ is an upper bound for the set $\{ 1/k, 2/k, 3/k, \ldots \}$. That is, it suffices to prove, for all $j \in \mathbb{N}$, that $\operatorname{sum} Z \geqslant j/k$.

Let $j \in \mathbb{N}$ be given. We wish to show: $\operatorname{sum} Z \geqslant j/k$.

The set $Z \cap I_k$ is uncountable, hence infinite, and it follows that $\mathrm{SUMS}_j(Z \cap I_k) \neq \varnothing$. Choose $x \in \mathrm{SUMS}_j(Z \cap I_k)$. Because

$$Z \cap I_k \quad \subseteq \quad I_k \quad \geqslant \quad 1/k,$$

it follows that $\mathrm{SUMS}_j(Z \cap I_k) \geqslant j/k$. Then $x \geqslant j/k$. We have

$$x \ \in \ \mathrm{SUMS}_j(Z \cap I_k) \ \subseteq \ \mathrm{SUMS}_j(Z) \ \subseteq \ \mathrm{SUMS}(Z).$$

Then $\{x\} \subseteq \mathrm{SUMS}(Z)$, so $\sup\{x\} \leqslant \sup(\mathrm{SUMS}(Z))$. Then

$$\operatorname{sum} Z \ = \ \sup(\mathrm{SUMS}(Z)) \ \geqslant \ \sup\{x\} \ = \ x \ \geqslant \ j/k,$$

as desired. □

So far every set that we have tried to sum has been a subset of $[0, \infty]$. When we try to sum sets with negative numbers, difficulties arise. For example, mathematicians have studied the partial sums of the series

$$1 \ - \ \frac{1}{2} \ + \ \frac{1}{3} \ - \ \frac{1}{4} \ + \ \frac{1}{5} \ - \ \frac{1}{6} \ + \ \frac{1}{7} \ + \ \cdots ,$$

and found that its partial sums tend toward $\ln 2$. It might be tempting, therefore to declare that the sum of the set

$$\left\{ 1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \frac{1}{5},\ -\frac{1}{6},\ \frac{1}{7},\ \ldots \right\}$$

is equal to $\ln 2$. The problem is that sets are unordered, and we need to set up our definitions in such a way that the sum of a set doesn't depend on the order in which its terms are presented. This turns out to be a real problem, which we explain in the next two paragraphs:

For all $j \in \mathbb{N}$, let

$$s_j := 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{2^j - 1}.$$

For all $j \in \mathbb{N}$, we have

$$\begin{array}{rcccl} 1 & > & \frac{1}{2}, & & \\ \frac{1}{2} + \frac{1}{3} & > & \frac{1}{4} + \frac{1}{4} & = & \frac{1}{2}, \\ \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} & > & \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} & = & \frac{1}{2}, \\ & \vdots & & & \\ \frac{1}{2^{j-1}} + \cdots + \frac{1}{2^j - 1} & > & \frac{1}{2^j} + \cdots + \frac{1}{2^j} & = & \frac{1}{2}. \end{array}$$

Adding these inequalities, we find, for all $j \in \mathbb{N}$, that $s_j > j/2$.

Let $Z := \{1, 1/2, 1/3, 1/4, \ldots\}$. Then: $\forall j \in \mathbb{N}$, $\operatorname{sum} Z \geqslant s_j > j/2$. Then $\operatorname{sum} Z = \infty$. Let $X := Z/2 = \{1/2, 1/4, 1/6, 1/8, \ldots\}$. It is not hard to show that $\operatorname{sum} X = (\operatorname{sum} Z)/2 = \infty/2 = \infty$. We have $1 > 1/2$, $1/3 > 1/4$, $1/5 > 1/6$, ... Let $Y := \{1, 1/3, 1/5, 1/7, \ldots\}$. Then $\operatorname{sum} Y \geqslant \operatorname{sum} X$. So, as $\operatorname{sum} X = \infty$, we get $\operatorname{sum} Y = \infty$. From these observations, we showed how we can find a reordering of

$$1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \frac{1}{5},\ -\frac{1}{6},\ \frac{1}{7},\ \ldots$$

whose partial sums tend toward ∞. We also showed how to find a different reordering whose partial sums tend toward $-\infty$. Consequently, trying to define the sum of the set

$$\left\{ 1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \frac{1}{5},\ -\frac{1}{6},\ \frac{1}{7},\ \ldots \right\}$$

is fraught, because sets are unordered, so any reordering is as good as any other. The following is the best we can do.

DEFINITION 9.7. *For all $x \in \mathbb{R}^*$, we define*

$$\begin{aligned} x^+ &:= \max\{x, 0\}, \\ x^- &:= \max\{-x, 0\} \qquad \textit{and} \\ |x| &:= \max\{x, -x\}. \end{aligned}$$

For example,

$$\begin{aligned} 5^+ &:= \max\{5, 0\} = 5, \\ 5^- &:= \max\{-5, 0\} = 0 \qquad \text{and} \\ |5| &:= \max\{5, -5\} = 5. \end{aligned}$$

Also,

$$\begin{aligned} (-3)^+ &:= \max\{-3, 0\} = 0, \\ (-3)^- &:= \max\{-(-3), 0\} = 3 \qquad \text{and} \\ |-3| &:= \max\{-3, -(-3)\} = 3. \end{aligned}$$

Unassigned HW: Show, for all $x \in \mathbb{R}^*$, that

$$\begin{aligned} x &= x^+ - x^- \qquad \text{and} \\ |x| &= x^+ + x^-. \end{aligned}$$

DEFINITION 9.8. *For all $Z \subseteq \mathbb{R}^*$, we define*

$$\begin{aligned} Z^+ &:= \{x^+ \,|\, x \in Z\}, \\ Z^- &:= \{x^- \,|\, x \in Z\} \qquad \textit{and} \\ |Z| &:= \{|x| \textit{ s.t. } x \in Z\}. \end{aligned}$$

For example if

$$Z = \left\{ 1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \frac{1}{5},\ -\frac{1}{6},\ \frac{1}{7},\ \ldots \right\},$$

then

$$\begin{aligned} Z^+ &= \left\{ 1,\ 0,\ \frac{1}{3},\ 0,\ \frac{1}{5},\ 0,\ \frac{1}{7},\ \ldots \right\}, \\ Z^- &= \left\{ 0,\ \frac{1}{2},\ 0,\ \frac{1}{4},\ 0,\ \frac{1}{6},\ 0,\ \ldots \right\} \qquad \text{and} \\ |Z| &= \left\{ 1,\ \frac{1}{2},\ \frac{1}{3},\ \frac{1}{4},\ \frac{1}{5},\ \frac{1}{6},\ \frac{1}{7},\ \ldots \right\}. \end{aligned}$$

DEFINITION 9.9. $\forall Z \in \mathbb{R}^*$, *let* $\operatorname{add} Z := [\operatorname{sum}(Z^+)] - [\operatorname{sum}(Z^-)]$.

For example, if

$$Z \quad = \quad \left\{ 1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \frac{1}{5},\ -\frac{1}{6},\ \frac{1}{7},\ \ldots \right\},$$

then

$$\begin{aligned} \mathrm{sum}\,(Z^+) &= \infty \qquad \text{and} \\ \mathrm{sum}\,(Z^-) &= \infty, \end{aligned}$$

so $\mathrm{add}\,Z = \infty - \infty = ☹$.

On the other hand, if

$$Z \quad = \quad \left\{ 1,\ -\frac{1}{2},\ \frac{1}{4},\ -\frac{1}{8},\ \frac{1}{16},\ -\frac{1}{32},\ \frac{1}{64},\ \ldots \right\},$$

then

$$\begin{aligned} Z^+ &= \left\{ 1,\ 0,\ \frac{1}{4},\ 0,\ \frac{1}{16},\ 0,\ \frac{1}{64},\ \ldots \right\} \qquad \text{and} \\ Z^- &= \left\{ 0,\ \frac{1}{2},\ 0,\ \frac{1}{8},\ 0,\ \frac{1}{32},\ 0,\ \ldots \right\}, \end{aligned}$$

so we get $\mathrm{sum}\,(Z^+) = \dfrac{4}{4-1} = \dfrac{4}{3}$ and $\mathrm{sum}\,(Z^-) = \dfrac{2}{4-1} = \dfrac{2}{3}$, and it follows that $\mathrm{add}\,Z = \dfrac{4}{3} - \dfrac{2}{3} = \dfrac{2}{3}$.

Unassigned HW: Show, for all $Z \subseteq \mathbb{R}^*$, that

$$\mathrm{sum}\,Z = \begin{cases} \mathrm{add}\,Z, & \text{if } Z \subseteq [0,\infty] \\ ☹, & \text{if } Z \not\subseteq [0,\infty]. \end{cases}$$

Consequently, we see, for all $Z \subseteq \mathbb{R}^*$, that $\mathrm{add}\,Z =^* \mathrm{sum}\,Z$.

Note: $\mathrm{SUMS}(\varnothing) = \varnothing$, so $\mathrm{sum}\,\varnothing = \sup \varnothing = -\infty$. Also,

$$\begin{aligned} \mathrm{add}\,\varnothing &= [\mathrm{sum}\,(\varnothing^+)] - [\mathrm{sum}\,(\varnothing^-)] \\ &= [\mathrm{sum}\,\varnothing] - [\mathrm{sum}\,\varnothing] \\ &= (-\infty) - (-\infty) \;=\; ☹. \end{aligned}$$

A basic limitation to studying add is that it doesn't allow for repeated terms. For example, one wants to say $7 + 7 + 7 + \cdots = \infty$, but $\{7, 7, 7, \ldots\} = \{7\}$, so $\mathrm{add}\,\{7, 7, 7, \ldots\} = \mathrm{add}\,\{7\} = 7$. To get at infinite sums with repeated terms, it helps to think, not just about how to sum a subset of $\mathbb{R}^*$, but, also, about how to add the outputs of a function z whose image is contained in $\mathbb{R}^*$. For any such z, we will next time define $\operatorname{add} z$. The definition is set up in such a way that, for any set S, if $z : S \to \mathbb{R}^*$ denotes the constant function defined by $z_j = 7$, then $\operatorname{add} Z = 7 \cdot (\#S)$. Also, the definition is set up in such a way that, for any *INJECTIVE* function z, if $\operatorname{im}[z] \subseteq \mathbb{R}^*$, then $\operatorname{add} z = \operatorname{add}(\operatorname{im}[z])$. In particular, for any $Z \subseteq \mathbb{R}^*$, we'll conclude that $\operatorname{add}(\operatorname{id}_Z) = \operatorname{add}(\operatorname{im}[\operatorname{id}_Z]) = \operatorname{add} Z$. More on all this next time.

10. Class 10 on 5 October 2017, Th of Week 5

We recalled the definitions of $\mathrm{SUMS}_j(Z)$, $\mathrm{SUMS}(Z)$, $\operatorname{sum} Z$ and $\operatorname{add} Z$. See Definition 9.4, and Definition 9.5 and Definition 9.9.

REMARK 10.1. *For all nonempty $Z \subseteq \mathbb{R}^*$, $\operatorname{add} Z =^* \operatorname{sum} Z$.*

Proof. Let a nonempty $Z \subseteq \mathbb{R}^*$ be given. We wish to prove that $\operatorname{add} Z =^* \operatorname{sum} Z$. That is, we wish to show:

$$[\ \operatorname{sum} Z \neq ☹\ ] \quad \Rightarrow \quad [\ \operatorname{add} Z = \operatorname{sum} Z\ ].$$

Assume: $\operatorname{sum} Z \neq ☹$. We wish to show: $\operatorname{add} Z = \operatorname{sum} Z$.

Since $\operatorname{sum} Z \neq ☹$, it follows that $Z \subseteq [0, \infty]$. Then $Z^+ = Z$. Then $\operatorname{sum}(Z^+) = \operatorname{sum} Z$. Also, since $\varnothing \neq Z \subseteq [0, \infty]$, we get $Z^- = \{0\}$. Then $\operatorname{sum}(Z^-) = 0$. Then

$$\operatorname{add} Z \;=\; [\operatorname{sum}(Z^+)] - [\operatorname{sum}(Z^-)] \;=\; [\operatorname{sum} Z] - 0 \;=\; \operatorname{sum} Z,$$

as desired. $\square$

DEFINITION 10.2. *For any set S, for all $z : S \to [0, \infty]$, for all $j \in \mathbb{N}$, we define*

$$\mathrm{SUMS}_j(z) \quad := \quad \{\ z(x_1) + \cdots + z(x_j) \mid x : \{1, \ldots, j\} \hookrightarrow S\ \}.$$

DEFINITION 10.3. *For any set S, for all $z : S \to [0, \infty]$, we define*

$$\mathrm{SUMS}(z) \quad := \quad [\mathrm{SUMS}_1(z)] \ \cup [\mathrm{SUMS}_2(z)] \ \cup \ \cdots,$$

and we define $\operatorname{sum} z := \sup(\mathrm{SUMS}(z))$.

Remember that, for all $z : \mathbb{N} \to [0, \infty]$, we sometimes denote z by the "infty-tuple" $(z_1, z_2, z_3, \ldots)$.

In class, we calculated $\operatorname{sum}(7, 7, 7, 7, \ldots) = \infty$.

Also, $\operatorname{sum}\left(1, 1, \frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}, \ldots\right) = 4$.

Also, sum $\left(1,0,\frac{1}{2},0,\frac{1}{4},0,\frac{1}{8},0,\frac{1}{16},0,\ldots\right)=2.$

Also, sum $\left(0,1,0,\frac{1}{2},0,\frac{1}{4},0,\frac{1}{8},0,\frac{1}{16},\ldots\right)=2.$

Also, sum $\left(1,\frac{1}{2},\frac{1}{3},\frac{1}{4},\frac{1}{5},\frac{1}{6},\frac{1}{7},\frac{1}{8},\ldots\right)=\infty.$

Also, sum $\left(1,0,\frac{1}{3},0,\frac{1}{5},0,\frac{1}{7},0,\ldots\right)=\infty.$

Also, sum $\left(0,\frac{1}{2},0,\frac{1}{4},0,\frac{1}{6},0,\frac{1}{8},\ldots\right)=\infty.$

DEFINITION 10.4. *For any set S, for all $z:S\to\mathbb{R}^*$, we define $z^+,z^-,|z|:S\to[0,\infty]$ by*

$$\begin{array}{rcl}(z^+)(x) & = & (z(x))^+,\\ (z^-)(x) & = & (z(x))^-\\ (|z|)(x) & = & |z(x)|.\end{array}$$

DEFINITION 10.5. *For any set S, for all $z:S\to\mathbb{R}^*$, we define* $\operatorname{add} z := [\operatorname{sum}(z^+)]-[\operatorname{sum}(z^-)]$.

In class, we calculated

$$\operatorname{add}\left(1,-1,\frac{1}{2},-\frac{1}{2},\frac{1}{4},-\frac{1}{4},\frac{1}{8},-\frac{1}{8},\frac{1}{16},-\frac{1}{16},\ldots\right)=2-2=0.$$

Also, sum $\left(1,-1,\frac{1}{2},-\frac{1}{2},\frac{1}{3},-\frac{1}{3},\frac{1}{4},-\frac{1}{4},\frac{1}{5},-\frac{1}{5},\ldots\right)=\infty-\infty=$ ☹.

Our goal is still to show that $\mathbb{R}$ has continuum cardinality (cc).

DEFINITION 10.6. *For any set I, I is a* **nondegenerate extended interval** *means: $\exists a,b\in\mathbb{R}^*$ such that $a<b$ and such that*

$$I \quad \in \quad \{\ (a,b),\ [a,b),\ (a,b],\ [a,b]\ \}.$$

DEFINITION 10.7. *For any set I, I is a* **degenerate extended interval** *means: $\exists a\in\mathbb{R}^*$ such that $I=[a,a]$.*

DEFINITION 10.8. *For any set I, I is an* **extended interval** *means: either (I is a nondegenerate extended interval) or (I is a degenerate extended interval).*

Note that any nondenerate extended interval is an infinite set. By contrast, any degenerate extended interval is a singleton set.

LEMMA 10.9. *For any extended interval I, for any $s, t \in I$, we have*

$$(\ s < t \ ) \quad \Rightarrow \quad (\ [s, t] \subseteq I \ ).$$

Proof. Omitted. $\square$

We turned Lemma 10.9 into a game: You choose I. Then you choose $s, t \in I$ satisfying $s < t$. We check whether $[s, t] \subseteq I$. If so, I win. Otherwise, you win.

For example, if you choose $I = [3, 8)$ and then choose $s = 3$ and $t = 7$, then, because $[3, 7] \subseteq [3, 8)$, I win.

I will happily play that game, and I will win every time, and with zero effort; I never have to make any choices!

DEFINITION 10.10. *For any set I, I is a* **nondegenerate interval** *means: (I is a nondegenerate extended interval) and ($I \subseteq \mathbb{R}$).*

DEFINITION 10.11. *For any set I, I is a* **degenerate interval** *means: (I is a degenerate extended interval) and ($I \subseteq \mathbb{R}$).*

DEFINITION 10.12. *For any set I, I is an* **interval** *means: (I is an extended interval) and ($I \subseteq \mathbb{R}$).*

DEFINITION 10.13. *For all $d \in \{0, 1, 2, \ldots, 9\}^{\mathbb{N}}$, we define*

$$(0.d_1 d_2 d_3 \cdots)_{\text{ten}} \quad := \quad \text{add} \left\{ \frac{d_1}{10}, \frac{d_2}{100}, \frac{d_3}{1000}, \frac{d_4}{10000}, \ldots \right\}.$$

Note that $(0.99999\cdots)_{\text{ten}} = 1$.

Note that $(0.099999\cdots)_{\text{ten}} = (0.100000\cdots)_{\text{ten}} = 1/10$.

Note that $(0.275399999\cdots)_{\text{ten}} = (0.275400000\cdots)_{\text{ten}} = 276/1000$.

DEFINITION 10.14. *For all $b \in \{0, 1\}^{\mathbb{N}}$, we define*

$$(0.b_1 b_2 b_3 \cdots)_{\text{two}} \quad := \quad \text{add} \left\{ \frac{b_1}{2}, \frac{b_2}{4}, \frac{b_3}{8}, \frac{b_4}{16}, \ldots \right\}.$$

Note that $(0.11111\cdots)_{\text{two}} = 1$.

Note that $(0.011111\cdots)_{\text{two}} = (0.100000\cdots)_{\text{two}} = 1/2$.

Note that

$$\begin{aligned} (0.10110011111\cdots)_{\text{two}} &= (0.10110100000\cdots)_{\text{two}} \\ &= \frac{2^5 + 2^3 + 2^2 + 1}{2^6}. \end{aligned}$$

DEFINITION 10.15. *For all $t \in \{0,1\}^{\mathbb{N}}$, we define*

$$(0.t_1t_2t_3\cdots)_{\text{three}} \quad := \quad \text{add}\left\{\frac{t_1}{3}, \frac{t_2}{9}, \frac{t_3}{27}, \frac{t_4}{81}, \ldots\right\}.$$

Note that $(0.22222\cdots)_{\text{three}} = 1$.
Note that $(0.022222\cdots)_{\text{three}} = (0.100000\cdots)_{\text{three}} = 1/2$.
Note that

$$\begin{aligned}(0.20012012122222\cdots)_{\text{three}} &= (0.20012012200000\cdots)_{\text{three}} \\ &= \frac{2\cdot 3^8 + 1\cdot 3^5 + 2\cdot 3^4 + 1\cdot 3^2 + 2\cdot 3 + 2}{3^9}.\end{aligned}$$

Define $f : \{0,\ldots,9\}^{\mathbb{N}} \to \mathbb{R}^*$ by $f(d_\bullet) = (0.d_1d_2d_3\ldots)_{\text{ten}}$. We explained why $\text{im}[f] = [0,1]$. Unfortunately, f is *NOT* one-to-one, because, for example

$$(0.099999\cdots)_{\text{ten}} \quad = \quad (0.100000\cdots)_{\text{ten}}.$$

We sought out restrictions of f that *ARE* one-to-one. We have to take some care: For example, $f|(\{0,2,5,6,9\}^{\mathbb{N}})$ is *NOT* one-to-one, because

$$(0.599999\cdots)_{\text{ten}} \quad = \quad (0.600000\cdots)_{\text{ten}}.$$

However, we did find several restrictions of f that *ARE* one-to-one:

$f|\{0,\ldots,8\}^{\mathbb{N}}$ is one-to-one.
$f|\{1,\ldots,9\}^{\mathbb{N}}$ is one-to-one.
$f|\{0,2,5,7,9\}^{\mathbb{N}}$ is one-to-one.
$f|\{0,3,6,9\}^{\mathbb{N}}$ is one-to-one.

We can exclude 0. We can exclude 9. If we include both 0 and 9, then we have to make sure that no two of the allowed digits differ by 1; the problem with $\{0,2,5,6,9\}$ was that 5 and 6 differ by 1. For any of the four restrictions described above, the image of the restriction is hard to calculate.

Now define $g : \{0,1\}^{\mathbb{N}} \to \mathbb{R}^*$ by $g(b_\bullet) = (0.b_1b_2b_3\ldots)_{\text{two}}$. We explained why $\text{im}[g] = [0,1]$. Unfortunately, g is not one-to-one, because, for example

$$(0.011111\cdots)_{\text{two}} \quad = \quad (0.100000\cdots)_{\text{two}}.$$

We don't want to restrict g to $\{0\}^{\mathbb{N}}$ or to $\{1\}^{\mathbb{N}}$, because these two sets each only have one element:

$$\begin{aligned}\{0\}^{\mathbb{N}} &= \{(0,0,0,0,0,0,0,0,0,0,\ldots)\} \\ \{1\}^{\mathbb{N}} &= \{(1,1,1,1,1,1,1,1,1,1,\ldots)\}.\end{aligned}$$

Now define $h : \{0,1,2\}^{\mathbb{N}} \to \mathbb{R}^*$ by $h(t_\bullet) = (0.t_1 t_2 t_3 \ldots)_{\text{three}}$. Unfortunately, h is not one-to-one, because, for example

$$(0.022222\cdots)_{\text{two}} \quad = \quad (0.100000\cdots)_{\text{two}}.$$

Here are three restrictions of h that *ARE* one-to-one:

$h|(\{0,1\}^{\mathbb{N}})$ is one-to-one.
$h|(\{0,2\}^{\mathbb{N}})$ is one-to-one.
$h|(\{1,2\}^{\mathbb{N}})$ is one-to-one.

The middle restriction, $h|(\{0,2\}^{\mathbb{N}})$, is one-to-one because, even though we have included both 0 and 2, the two trits 0 and 2 don't differ by 1. The images of the first and third restrictions are hard to calculate. The middle restriction is the most famous because its image is, perhaps, a bit easier to describe. By definition, the **Cantor set** is Cantor := $\mathrm{im}[h|(\{0,2\}^{\mathbb{N}})]$. For any $k \in \mathbb{N}$, the kth **Cantor set approximation** is

$$\begin{aligned} \mathrm{Cantor}_k \quad := \quad \{\ & (0.t_1t_2t_3\ldots)_{\text{three}} \mid t_\bullet \in \{0,1,2\}^{\mathbb{N}} \\ & t_1 \neq 1,\ t_2 \neq 1,\ \ldots,\ t_k \neq 1\ \}. \end{aligned}$$

We gave the geometric description of how to transition from each approximation to the next by "excluding middle thirds". The approximations form a decreasing sequence of sets, and the Cantor set is the intersection of its approximations. From this, we can see that the Cantor set is "small" in the sense that it does not contain any nondegenerate interval. On the other hand, if $\eta := h|(\{0,2\}^{\mathbb{N}})$, then $\eta : \{0,2\}^{\mathbb{N}} \hookrightarrow\!\!> \mathrm{Cantor}$. By HW#17, $\exists 2^{\mathbb{N}} \hookrightarrow\!\!> \{0,1\}^{\mathbb{N}}$. Also, $\exists\{0,1\}^{\mathbb{N}} \hookrightarrow\!\!> \{0,2\}^{\mathbb{N}}$. Also, using η, we see that $\exists\{0,2\}^{\mathbb{N}} \hookrightarrow\!\!> \mathrm{Cantor}$. Then, by transitivity, we see that $\exists 2^{\mathbb{N}} \hookrightarrow \mathrm{Cantor}$. We conclude the Cantor set is "big" in the sense that it has continuum cardinality (cc).

Since the Cantor set a subset of $\mathbb{R}$, we get $\exists \mathrm{Cantor} \hookrightarrow \mathbb{R}$, so, in the World of Sets, we must place $\mathbb{R}$ at or above the cc level. The same logic shows that $\mathbb{R}^*$ must be placed at or above the cc level. In particular, we now know that $\mathbb{R}$ and $\mathbb{R}^*$ are both uncountable.

LEMMA 10.16. $\exists[1,\infty] \hookrightarrow\!\!> [1,2]$.

Proof. By HW#19, $\exists[1,\infty] \hookrightarrow\!\!> [0,1]$. So, since $x \mapsto x+1 : [0,2] \hookrightarrow\!\!> [1,2]$, by composition, we see that $\exists[1,\infty] \hookrightarrow\!\!> [1,2]$, as desired. $\square$

LEMMA 10.17. $\exists[-\infty,-1] \hookrightarrow\!\!> [-2,-1]$.

Proof. We have $x \mapsto -x : [-\infty, -1] \hookrightarrow\!> [1, \infty]$ Also, by Lemma 10.16, we get $\exists [1, \infty] \to [1, 2]$. Aso, $x \mapsto -x : [1, 2] \hookrightarrow\!> [-2, -1]$. Therefore, by composition, we see that $\exists [-\infty, -1] \hookrightarrow\!> [-2, -1]$, as desired. □

LEMMA 10.18. $\exists \mathbb{R}^* \hookrightarrow\!> [-2, 2]$.

Proof. By Lemma 10.17, choose $\alpha : [-\infty, -1] \hookrightarrow\!> [-2, -1]$.

Let $\beta := \mathrm{id}_{(-1,1)} : (-1, 1) \to (-1, 1)$. Then $\beta : (-1, 1) \hookrightarrow\!> (-1, 1)$.

By Lemma 10.16, choose $\gamma : [1, \infty] \hookrightarrow\!> [1, 2]$.

Define $\omega : \mathbb{R}^* \to [-2, 2]$ by

$$\omega(x) \quad = \quad \begin{cases} \alpha(x), & \text{if } x \in [-\infty, -1] \\ \beta(x), & \text{if } x \in (-1, 1) \\ \gamma(x), & \text{if } x \in [1, \infty]. \end{cases}$$

Then $\omega : \mathbb{R}^* \hookrightarrow\!> [-2, 2]$, so $\exists \mathbb{R}^* \hookrightarrow\!> [-2, 2]$, as desired. □

THEOREM 10.19. *$\forall$nondegenerate extended interval I, $\exists \mathbb{R}^* \hookrightarrow I$.*

Proof. Given a nondegenerate extended interval I, want: $\exists \mathbb{R}^* \hookrightarrow I$.

Since I is a nondegenerate extended interval, it follows that I is an infinite set. Let $F := \{\infty, -\infty\}$. Then $\mathbb{R}^* \backslash F = \mathbb{R}$. Because I is infinite and F is finite, it follows that $I \backslash F$ is infinite. In particular, we have $\#(I \backslash F) \geqslant 2$. Choose $s, t \in I \backslash F$ such that $s < t$. Then $s, t \in I \backslash F \subseteq \mathbb{R}^* \backslash F = \mathbb{R}$. Also, $s, t \in I \backslash F \subseteq I$, so, by Lemma 10.9, we see that $[s, t] \subseteq I$.

Let $u := (t - s)/2$. Then $s + 2u = s + (t - s) = t$. We have

$$\begin{aligned} x \mapsto x/2 \quad &: \quad [-2, 2] \hookrightarrow\!> [-1, 1] \qquad \text{and} \\ x \mapsto xu \quad &: \quad [-1, 1] \hookrightarrow\!> [-u, u] \qquad \text{and} \\ x \mapsto x + s + u \quad &: \quad [-u, u] \hookrightarrow\!> [s, s + 2u]. \end{aligned}$$

Composing, $\exists [-2, 2] \hookrightarrow\!> [s, s + 2u]$. We have $[s, s + 2u] = [s, t] \subseteq I$, so $x \mapsto x : [s, s + 2u] \hookrightarrow I$. Then

$$\begin{aligned} \exists [-2, 2] \quad &\hookrightarrow \quad [s, s + 2u] \qquad \text{and} \\ \exists [s, s + 2u] \quad &\hookrightarrow \quad I \end{aligned}$$

Composing, we get $\exists [-2, 2] \hookrightarrow I$.

By Lemma 10.18, $\exists \mathbb{R}^* \hookrightarrow\!> [-2, 2]$. So, since $\exists [-2, 2] \hookrightarrow I$, by composing, we get $\exists \mathbb{R}^* \hookrightarrow I$, as desired. □

Recall that, in the World of Sets, $\mathbb{R}$ and $\mathbb{R}^*$ must be placed at or above the Cantor set, and the Cantor set has continuum cardinality (cc). By HW#17, $\exists\{0,1\}^{\mathbb{N}} \hookrightarrow 2^{\mathbb{N}}$, so $\{0,1\}$ has cc. Recall also that

$$b_\bullet \ \mapsto \ (b_1b_2b_3\ldots)_{\text{two}} \ : \ \{0,1\}^{\mathbb{N}} \to> [0,1].$$

Thus $[0,1]$ must be placed at or below $\{0,1\}^{\mathbb{N}}$. Because $[0,1]$ is a nondegenerate interval, by Theorem 10.19, we see that $\exists\mathbb{R}^* \hookrightarrow [0,1]$. Thus $\mathbb{R}^*$ must be placed at or below the level of $[0,1]$. Organizing our thoughts carefully, we can now show that $\mathbb{R}$ and $\mathbb{R}^*$ have cc. In fact, we'll show that *ANY* nondegenerate extended interval has cc:

THEOREM 10.20. *Let I be a nondenerate extended interval. Then I has cc.*

Proof. By Schroeder-Bernstein ((2) of Theorem 6.12), we wish to show $\exists 2^{\mathbb{N}} \hookrightarrow I$ and $\exists I \hookrightarrow 2^{\mathbb{N}}$.

By HW#17, $\exists 2^{\mathbb{N}} \hookrightarrow> \{0,1\}^{\mathbb{N}}$. Also, $\exists\{0,1\}^{\mathbb{N}} \hookrightarrow> \{0,2\}^{\mathbb{N}}$. Let C be the Cantor set. Then $t_\bullet \mapsto (0.t_1t_2t_3\ldots)_{\text{three}} : \{0,2\}^{\mathbb{N}} \hookrightarrow> C$. Since $C \subseteq \mathbb{R}^*$, we have $\exists C \hookrightarrow \mathbb{R}^*$. By Theorem 10.19, $\exists\mathbb{R}^* \hookrightarrow I$. Composing, $\exists 2^{\mathbb{N}} \hookrightarrow I$. It remains to show that $\exists I \hookrightarrow 2^{\mathbb{N}}$.

Recall that $b_\bullet \mapsto (b_1b_2b_3\ldots)_{\text{two}} : \{0,1\}^{\mathbb{N}} \to> [0,1]$. It then follows, from (3) of Theorem 6.12, that $\exists[0,1] \hookrightarrow \{0,1\}^{\mathbb{N}}$.

As $I \subseteq \mathbb{R}^*$, we get $\exists I \hookrightarrow \mathbb{R}^*$. Since $[0,1]$ is a nondegenerate interval, by Theorem 10.19, we get: $\exists\mathbb{R}^* \hookrightarrow [0,1]$. Recall: $\exists[0,1] \hookrightarrow \{0,1\}^{\mathbb{N}}$. By HW#17, $\exists\{0,1\}^{\mathbb{N}} \hookrightarrow> 2^{\mathbb{N}}$. Composing, $\exists I \hookrightarrow 2^{\mathbb{N}}$, as desired. $\square$

11. Class 11 on 10 October 2017, Tu of Week 6

Discussed Midterm 1, next class. Class will have three parts:

(1) 45 mins (approx) review,
(2) 10 min (approx) break, and
(3) 60 minute (exact) exam.

Discussed what Midterm 1 covers.

DEFINITION 11.1. *For any set $\mathcal{S}$ of sets, we define*

$$\langle \mathcal{S} \rangle_\cup \quad := \quad \left\{ \bigcup \mathcal{A} \,\middle|\, \mathcal{A} \subseteq \mathcal{Q} \right\}.$$

For any set $\mathcal{S}$ of sets, the set $\langle \mathcal{S} \rangle_\cup$ is called the **union-closure** of $\mathcal{S}$.

We calculated

$$\begin{aligned} \langle \ \{ \ \{1,2\} \ , \ \{3\} \ , \ \{4,5,6\} \ \} \ \rangle_\cup \quad = \\ \{ \ \varnothing \ , \\ \{1,2\} \ , \ \{3\} \ , \ \{4,5,6\} \ , \\ \{1,2,3\} \ , \ \{1,2,4,5,6\} \ , \ \{3,4,5,6\} \ , \\ \{1,2,3,4,5,6\} \ \}. \end{aligned}$$

Note that since $\{\{1,2\},\{3\},\{4,5,6\}\}$ has three elements, it follows that $\{\{1,2\},\{3\},\{4,5,6\}\}$ has 2^3 subsets, and each of those eight sets is a set of sets, and, for each of these eight sets of sets, we calculated the union. The answer above is the set of those eight calculated unions.

We also discussed the union-closure of: the set of bounded open intervals in $\mathbb{R}$. We found that $(1,2) \cup (3,5)$ is in that union-closure, but that $\{1\}$ and $[1,5)$ are not. The sets that are in that union closure are typically called "open subsets of $\mathbb{R}$". Real analysis in one variable focuses on using calculus to study partial functions $\mathbb{R} \dashrightarrow \mathbb{R}$. If the domain D of $f : \mathbb{R} \dashrightarrow \mathbb{R}$ is an open subset of $\mathbb{R}$, then, for all $p \in D$, there's room to move to the left or right of p, without leaving D; consequently, concepts like

the limit at p of f or
the tangent line at $(p, f(p))$ of f

are relatively easy to define. By contrast, if the domain of f were, say, $\{1\}$, then trying to study f via calculus would be hopeless. Even for a function f whose domain is $[1,5)$, the tangent line to f at the point $(1, f(1))$ would be a little sketchy, though we might get at it via one-sided limits. Anyway, the upshot is that: Using calculus to study a function $f:\mathbb{R} \dashrightarrow \mathbb{R}$ is made easier if $\mathrm{dom}[f]$ is a set that is in the union-closure of the set of bounded open intervals in $\mathbb{R}$.

We also discussed the union-closure of: the set of open disks in $\mathbb{R}^2$. The sets that are in that union closure are typically called "open subsets of $\mathbb{R}^2$". Real analysis in two variable focuses on using calculus to study partial functions $\mathbb{R}^2 \dashrightarrow \mathbb{R}$. If the domain D of $f:\mathbb{R}^2 \dashrightarrow \mathbb{R}$ is an open subset of $\mathbb{R}^2$, then, for all $p \in D$, there's room to move in all directions off of p, without leaving D; consequently, concepts like

the limit at p of f or
the tangent line at $(p, f(p))$ of f

are relatively easy to define. The upshot is that: Using calculus to study a function $f:\mathbb{R}^2 \dashrightarrow \mathbb{R}$ is made easier if $\mathrm{dom}[f]$ is a set that is in the union-closure of the set of open disks in $\mathbb{R}^2$.

We have three basic properties of union-closure:

REMARK 11.2. *The following are all true:*

(1) $\forall$*set* $\mathcal{S}$ *of sets,* $\langle \mathcal{S} \rangle_\cup \supseteq \mathcal{S}$.
(2) $\forall$*set* $\mathcal{S}$ *of sets,* $\langle \langle \mathcal{S} \rangle_\cup \rangle_\cup = \langle \mathcal{S} \rangle_\cup$.
(3) $\forall$*sets* $\mathcal{S}, \mathcal{T}$ *of sets,* $[\, (\mathcal{S} \subseteq \mathcal{T}) \Rightarrow (\langle \mathcal{S} \rangle_\cup \subseteq \langle \mathcal{T} \rangle_\cup)\,]$.

In Remark 11.2,

(1) is called the **superset property** of $\langle \bullet \rangle_\cup$,
(2) is called the **idempotence property** of $\langle \bullet \rangle_\cup$ and
(3) is called the **monotonic property** of $\langle \bullet \rangle_\cup$.

Assigned HW#26 and HW#27.

The next lemma will be called the **Exact Covering Lemma**.

LEMMA 11.3. *For any set* $\mathcal{S}$ *of sets, for any set* X*, we have*

$$(\, X \in \langle \mathcal{S} \rangle_\cup \,) \Leftrightarrow (\, \forall p \in X, \exists A \in \mathcal{S} \text{ s.t. } p \in A \subseteq X \,).$$

Proof. $\Rightarrow$ is HW#26, and $\Leftarrow$ is HW#27. $\square$

The clause "$X \in \langle \mathcal{S} \rangle_\cup$" means that X is a union of some of the sets from $\mathcal{S}$, and might be colloquially rendered by saying, "X can be exactly covered by well-chosen sets from $\mathcal{S}$". Here, "exactly" means that every point of X is covered, but, *ALSO*, that *NO* point outside of X is covered. The clause "$p \in A \subseteq X$" might be colloquially rendered by saying "A covers p without going outside X". With these colloquialisms, the Exact Covering Lemma asserts: for any collection $\mathcal{S}$ of sets, a set X can be exactly covered by sets from $\mathcal{S}$ iff each point of X can be covered without going outside X.

Single-variable real analysis focuses on functions defined on a subset of $\mathbb{R}$, typically an open subset. Two-variable real analysis focuses on functions defined on a subset of $\mathbb{R}^2$, typically an open subset. Understanding $\mathbb{R}$ and $\mathbb{R}^2$ becomes crucial and some of the basic properties of $\mathbb{R}$ and $\mathbb{R}^2$ are algebraic properties, particularly linear algebraic properties. We therefore pause to develop a bit of algebra and linear algebra.

DEFINITION 11.4. *Let S be a set and let $* : S \times S \to S$. Then*

$$\mathrm{IDS}_*^S \quad := \quad \{e \in S \,|\, \forall x \in S,\, x * e = x = e * x\}.$$

The set IDS_*^S will be called the **Identities Set** for $*$ in S.

Let $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be ordinary addition of real numbers. Then the Identities Set for $+$ on $\mathbb{R}$ is $\mathrm{IDS}_+^{\mathbb{R}} = \{0\}$.

Let $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be ordinary multiplication of real numbers. Then the Identities Set for $\cdot$ on $\mathbb{R}$ is $\mathrm{IDS}_\cdot^{\mathbb{R}} = \{1\}$.

REMARK 11.5. *Let S be a set and let $* : S \times S \to S$. Then $\#[\mathrm{IDS}_*^S] \leqslant 1$.*

Proof. We wish to show: $\forall e, \varepsilon \in \mathrm{IDS}_*^S$, $e = \varepsilon$. Let $\forall e, \varepsilon \in \mathrm{IDS}_*^S$ be given. We wish to show $e = \varepsilon$.

Since $e \in \mathrm{IDS}_*^S$, we conclude, for all $x \in S$, that $x * e = x = e * x$. Replacing x by ε, we get $\varepsilon * e = \varepsilon = e * \varepsilon$. Then $\varepsilon * e = \varepsilon$.

Since $\varepsilon \in \mathrm{IDS}_*^S$, we conclude, for all $x \in S$, that $x * \varepsilon = x = \varepsilon * x$. Replacing x by e, we get $e * \varepsilon = e = \varepsilon * e$. Then $e = \varepsilon * e$.

Then $e = \varepsilon * e = \varepsilon$, as desired. ∎

DEFINITION 11.6. *Let S be a set and let $* : S \times S \to S$. Then*

(1) *$*$ is* **commutative** *means: $\forall x, y \in S$, $x * y = y * x$, and*
(2) *$*$ is* **associative** *means: $\forall x, y, z \in S$, $(x * y) * z = x * (y * z)$.*

DEFINITION 11.7. *By $(V, \oplus, \bullet)$ is a* **real vector space**, *we mean that all of the following are true:*

(1) *V is a nonempty set;*
(2) *$\oplus : V \times V \to V$;*
(3) *$\bullet : \mathbb{R} \times V \to V$;*
(4) *$\oplus$ is commutative and associative;*

(5) $\forall x \in V$,
 (5a) $0 \bullet x \in \mathrm{IDS}_{\oplus}^{V}$ *and*
 (5b) $1 \bullet x = x$;
(6) $\forall a, b \in \mathbb{R}$, $\forall x \in V$,
 (6a) $(a + b) \bullet x = (a \bullet x) \oplus (b \bullet x)$ *and*
 (6b) $(ab) \bullet x = a \bullet (b \bullet x)$; *and*
(7) $\forall a \in \mathbb{R}$, *for all* $x, y \in V$, *we have:* $a \bullet (x \oplus y) = (a \bullet x) \oplus (a \bullet y)$.

We can define **complex vector space** in the same way, except that each $\mathbb{R}$ would be replaced by $\mathbb{C}$. This is a real analysis course, so by **vector space**, we will always mean real vector space. In a *complex* analysis course, the term "vector space" would probably mean complex vector space.

Let $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be ordinary addition of real numbers. Let $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be ordinary multiplication of real numbers. Then $(\mathbb{R}, +, \cdot)$ is a vector space.

Let $n \in \mathbb{N}$. Recall that $\mathbb{R}^n = \mathbb{R}^{\{1,\ldots,n\}}$. That is, $\mathbb{R}^n$ is equal to the set of all functions $\{1, \ldots, n\} \to \mathbb{R}$. Also, for any $x \in \mathbb{R}^n$, for any $j \in \{1, \ldots, n\}$, we sometimes denote $x(j)$ by x_j. Also, for any $x \in \mathbb{R}^n$, the function $x : \{1, \ldots, n\} \to \mathbb{R}$ is sometimes denoted $(x_1, \ldots, x_n)$.

Let $n \in \mathbb{N}$. Define $\oplus : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by $(x \oplus y)_j = x_j + y_j$. NOTE: To be more detailed, we would say "…by:

$$\forall x \in \mathbb{R}^n,\ \forall y \in \mathbb{R}^n,\ \forall j \in \{1, \ldots, n\}, \quad (x \oplus y)_j = x_j + y_j."$$

Define $\bullet : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ by $(a \bullet x)_j = a \cdot x_j$. NOTE: To be more detailed, we would say "…by:

$$\forall a \in \mathbb{R},\ \forall x \in \mathbb{R}^n,\ \forall j \in \{1, \ldots, n\}, \quad (a \bullet x)_j = a \cdot x_j."$$

Then $(\mathbb{R}^n, \oplus, \bullet)$ is a vector space.

We demonstrated how to view elements of $\mathbb{R}^2$ as arrows in a coordinate plane that start at $(0, 0)$. We showed how to think about $\oplus : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ and $\bullet : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ geometrically. In particular, $2 \bullet (3, 4) = (6, 8)$ is viewed geometrically as saying that if we double the arrow from $(0, 0)$ to $(3, 4)$, we get the arrow from $(0, 0)$ to $(6, 8)$. This idea of multiplying an arrow by a number is called "rescaling" the arrow, and the number you use is sometimes called a "scalar". That is, by "scalar" we mean number. Because this is a real analysis course, by **scalar** we mean *real* number. In a *complex* analysis course, by scalar we would likely mean complex number. When we are focused on a specific vector space $(V, \oplus, \bullet)$, an element of V is called a **vector**.

Every vector space has a unique additive identity:

REMARK 11.8. *Let $(V, \oplus, \bullet)$ be a vector space. Then $\#[\mathrm{IDS}_{\oplus}^{V}] = 1$.*

Proof. By Remark 11.5, $\#[\mathrm{IDS}_{\oplus}^{V}] \leqslant 1$. It therefore suffices to show: $\#[\mathrm{IDS}_{\oplus}^{V}] \geqslant 1$. It therefore suffices to show: $\#[\mathrm{IDS}_{\oplus}^{V}] \neq 0$. It therefore suffices to show: $\mathrm{IDS}_{\oplus}^{V} \neq \varnothing$.

By (1) of Definition 11.7, choose $x \in V$. By (5a) of Definition 11.7, we have $0 \bullet x \in \mathrm{IDS}_{\oplus}^{V}$. Then $\mathrm{IDS}_{\oplus}^{V} \neq \varnothing$, as desired. $\square$

Some notational conventions: For any vector space $(V, \oplus, \bullet)$,

(1) $(V, \oplus, \bullet)$ is usually abbreviated V;
(2) we define $0_V := \mathrm{ELT}(\mathrm{IDS}_{\oplus}^{V})$;
(3) $\forall x \in V$, we define $-x := (-1) \bullet x$;
(4) $\forall x, y \in V$, we define $x + y = x \oplus y$ and $x - y = x \oplus (-y)$; and
(5) $\forall a \in \mathbb{R}$, $\forall x \in V$, we define $ax = a \bullet x$.

Following (1) above, the vector space $(\mathbb{R}, +, \cdot)$ is denoted $\mathbb{R}$, and, for all $n \in \mathbb{N}$, the vector space $(\mathbb{R}^n, \oplus, \bullet)$ is denoted $\mathbb{R}^n$.

Because of (2) above, for any vector space V, we have: $\mathrm{IDS}_{\oplus}^{V} = \{0_V\}$.

Note that, technically, $\mathbb{R} \neq \mathbb{R}^1$, because $\mathbb{R}$ is a set of numbers, whereas $\mathbb{R}$ is the set of functions $\{1\} \to \mathbb{R}$. Nevertheless, to a vector space expert, $\mathbb{R}$ looks and feels like $\mathbb{R}^1$. The tecnical phrasing for this is: "$\mathbb{R}$ is isomorphic to $\mathbb{R}^1$." We will define "isomorphic" later.

Let $(V, \oplus, \bullet)$ be a vector space. Then $\oplus : V \times V \to V$ is often called **vector addition**. Also, $\bullet : \mathbb{R} \times V \to V$ is often called **scalar multiplication**. The **linear operations** of the vector space V are $\oplus$ and $\bullet$; that is, they are vector addition and scalar multiplication.

DEFINITION 11.9. *For all $n \in \mathbb{N}$, we define $0_n := 0_{\mathbb{R}^n}$.*

DEFINITION 11.10. *For any set S, for any a, we define the function $C_S^a : S \to \{a\}$ by $C_S^a(x) = a$.*

For any set S for any a, the function C_S^a is called the **constant function** on S with value a. Note that, for all $n \in \mathbb{N}$, we have

$$0_n \quad = \quad C^0_{\{1,\ldots,n\}} \quad \in \quad \mathbb{R}^{\{1,\ldots,n\}} \quad = \quad \mathbb{R}^n;$$

or, equivalently, $0_n = (0, \ldots, 0) \in \mathbb{R}^n$.

Let S be any set. Define $\oplus : \mathbb{R}^S \times \mathbb{R}^S \to \mathbb{R}^S$ by $(x \oplus y)_j = x_j + y_j$. Define $\bullet : \mathbb{R} \times \mathbb{R}^S \to \mathbb{R}^S$ by $(a \bullet x)_j = ax_j$. Then $(\mathbb{R}^S, \oplus, \bullet)$ is a vector space, typically denoted $\mathbb{R}^S$.

DEFINITION 11.11. *For any set S, we define $0_S := 0_{\mathbb{R}^S}$.*

Then, for any set S, we have $0_S = C_S^0 \in \mathbb{R}^{\{1,\ldots,n\}} = \mathbb{R}^n$.

Let S be a set. If $\#S = 1$, then $\mathbb{R}^S$ looks and feels like $\mathbb{R}$. Moreover, $\mathbb{R}$ looks and feels like a coordinatized line. So, if $\#S = 1$, then we can understand $\mathbb{R}^S$ intuitively, by looking at a coordinatized line. If $\#S = 2$, then $\mathbb{R}^2$ looks and feels like $\mathbb{R}^2$. Moreover, $\mathbb{R}^2$ looks and feels like a coordinatized plane. So, if $\#S = 2$, then we can understand $\mathbb{R}^S$ intuitively, by looking at a coordinatized plane. Coordinatized lines and planes can be drawn easily on a blackboard. When $\#S = 3$, things get a little harder because the intuition doesn't fit on a blackboard, but it's not too bad. When $3 < \#S < \infty$, things get harder still, but there are ways to think about $\mathbb{R}^S$. When $\#S = \infty$, we need a whole new branch of mathematics (called "functional analysis") to deal with $\mathbb{R}^S$.

In this course, we'll mainly focus on the vector spaces $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, ...

REMARK 11.12. *For any vector space V, for any $x \in V$, we have $x + (-x) = 0_V = (-x) + x$.*

Proof. Let a vector space V be given and let $x \in V$ be given. We wish to show that $x + (-x) = 0_V = (-x) + x$.

We have

$$\begin{aligned} x + (-x) &= [1 \bullet x] + [(-1) \bullet x] = [1 + (-1)] \bullet x \\ &= 0 \bullet x \in \mathrm{IDS}_{\oplus}^V = \{0_V\}, \end{aligned}$$

so $x + (-x) = 0_V$. It remains to show: $(-x) + x = 0_V$.

We have

$$\begin{aligned} (-x) + x &= [(-1) \bullet x] + [1 \bullet x] = [(-1) + 1] \bullet x \\ &= 0 \bullet x \in \mathrm{IDS}_{\oplus}^V = \{0_V\}, \end{aligned}$$

so $(-x) + x = 0_V$, as desired. ∎

Assigned HW#28, HW#29, HW#30.

DEFINITION 11.13. *For any vector space V, for any $S \subseteq V$,*

$$\langle S \rangle_{\mathrm{lin}}^V := \{0_V\} \bigcup \{ c_1 x_1 + \cdots + c_n x_n \mid n \in \mathbb{N},\ c_1, \ldots, c_n \in \mathbb{R},\ x_1, \ldots, x_n \in S \}.$$

Note, for any vector space V, that $\langle\varnothing\rangle_{\text{lin}}^V = \{0_V\}$.

When the vector space V is clear, we will often write $\langle S\rangle_{\text{lin}}$ to mean $\langle S\rangle_{\text{lin}}^V$. Recall that V abbreviates $(V, \oplus, \bullet)$.

Let V be a vector space. Then, for any $S \subseteq V$, the set $\langle S\rangle_{\text{lin}}$ is called the **span** of S. For any $n \in \mathbb{N}$, for any $c_1, \ldots, c_n \in \mathbb{R}$, for any $x_1, \ldots, x_n \in V$, the vector $c_1x_1 + \cdots + c_nx_n$ is called the **linear combination** of $x_1, \ldots, x_n$ with coefficients $c_1, \ldots, c_n$.

REMARK 11.14. *Let V be a vector space. The following are all true:*

(1) $\forall S \subseteq V$, $\langle S\rangle_{\text{lin}} \supseteq S$.
(2) $\forall S \subseteq V$, $\langle\langle S\rangle_{\text{lin}}\rangle_{\text{lin}} = \langle S\rangle_{\text{lin}}$.
(3) $\forall S, T \subseteq V$, $[\, (\, S \subseteq T\,) \;\Rightarrow\; (\,\langle S\rangle_{\text{lin}} \subseteq \langle T\rangle_{\text{lin}}\,)\,]$.

In Remark 11.14,

(1) is called the **superset property** of $\langle\bullet\rangle_{\text{lin}}$,
(2) is called the **idempotence property** of $\langle\bullet\rangle_{\text{lin}}$ and
(3) is called the **monotonic property** of $\langle\bullet\rangle_{\text{lin}}$.

FACT 11.15. *Let V be a vector space, $k \in \mathbb{N}$, $x_1, \ldots, x_k \in V$. Then:*

$$\langle\, \{x_1, \ldots, x_k\}\,\rangle_{\text{lin}} \quad = \quad \{\, c_1x_1 + \cdots + c_kx_k \mid c_1, \ldots, c_k \in \mathbb{R}\,\}.$$

Proof. Omitted. □

12. Class 12 on 17 October 2017, Tu of Week 7

We discussed several examples of "**quantified equivalence**". A few of these are detailed at the end of the exposition handout.

Here's an example of a quantified equivalence: For any set $\mathcal{S}$ of sets, for any x, we have:

$$\left[\; x \in \bigcup \mathcal{S}\;\right] \qquad \Leftrightarrow \qquad [\; \exists A \in \mathcal{S} \text{ s.t. } x \in A\;].$$

On the left-hand side of $\Leftrightarrow$, we have "$x \in \bigcup \mathcal{S}$", which exhibits no quantifiers. On the right-hand side, we have "$\exists A \in \mathcal{S}$ s.t. $x \in A$", which exhibits the quantifier $\exists$. In that kind of situation, we refer to the statement as a "quantified equivalence". In a proof, if we need to show that some object is in a union of a set of sets, we can now reinterpret that as a "Want: $\exists A \ldots$" statement, and, to try to deal with that, we have template [11] of the exposition handout. Also, in a proof, if we know that some object is in a union of a set of sets, we can now reinterpret that as a "Know: $\exists A \ldots$" statement, and by replacement rule [21] of the exposition handout, we can "Choose A. . .".

Another quantified equivalence: For any sets A and B, we have:

$$[\ A = B\ ] \qquad \Leftrightarrow \qquad [\ \forall x,\ (\,[x \in A] \Leftrightarrow [x \in B]\,)\ ].$$

Another quantified equivalence: For any set $\mathcal{S}$ of sets, for any set X,

$$[\ X \in \langle \mathcal{S} \rangle_\cup\ ] \qquad \Leftrightarrow \qquad [\ \forall z \in X,\ \exists A \in \mathcal{S} \text{ s.t. } z \in A \subseteq X\ ].$$

We worked through III.3 of Midterm 1, using quantified equivalnce as part of the motivation.

We reviewed Definition 11.13 and Remark 11.14.

Assigned HW#31.

We visualized several spans:

(1) $\langle\ \{\ (1,2)\ \}\ \rangle_{\text{lin}}$
(2) $\langle\ \{\ (1,2)\ ,\ (3,4)\ \}\ \rangle_{\text{lin}}$
(3) $\langle\ \{\ (1,0,0)\ ,\ (0,1,0)\ \}\ \rangle_{\text{lin}}$
(4) $\langle\ \{\ (1,0,0)\ ,\ (0,1,0)\ ,\ (1,1,0)\ \}\ \rangle_{\text{lin}}$
(5) $\langle\ \{\ (1,0,0)\ ,\ (0,1,0)\ ,\ (0,0,1)\ \}\ \rangle_{\text{lin}}$

DEFINITION 12.1. *For any vector space V, for any U, by U is a* **subspace** *of V, we mean: ($U \subseteq V$) and ($\langle U \rangle_{\text{lin}} = U$).*

REMARK 12.2. *For any vector space V, for any nonempty $U \subseteq V$,*

$$\begin{aligned} &[\ \ U \text{ is a subspace of } V\ \ ] \qquad \Leftrightarrow \\ &[\ \ (\forall x, y \in U,\ x + y \in U)\ \ \&\ \ (\forall a \in \mathbb{R}, \forall x \in U,\ cx \in U)\ \ ]. \end{aligned}$$

Proof. Omitted. □

Remark 12.2 asserts: a nonempty subset of a vector space is a subspace iff it respects both vector addition and scalar multiplication.

Let $(V, \oplus, \bullet)$ be a vector space and let U be a subspace of V. Then

$$(\ \ U\ ,\ \ \oplus|(U \times U)\ ,\ \ \bullet|(\mathbb{R} \times U)\ \ )$$

is a vector space, and we typically denote this vector space by U. We would say that the linear operations (vector addition and scalar multiplcation) of U are "inherited" from V.

We define $\mathbb{R}^0 := \{0\}$. Then $R^0 \subseteq \mathbb{R}$, and $\mathbb{R}^0$ is a vector space, with its linear operations inherited from $\mathbb{R}$.

Assigned HW#32

Any nonzero subspace of a vector space has continuum cardinality. So, to specify such a subspace, it's impractical to try to write out all of its elements. Instead, we hope that there's a finite set that spans the subspace and we can communicate the subspace to someone else by simply writing out the set that spans. This prompts the following definition:

DEFINITION 12.3. *Let V be a vector space and let U be a subspace of V. For any S, by*

S **spans** U *or*
S *is a* **spanning set** *for* U *or*
U *is* **spanned by** S,

we mean: both ($S \subseteq U$) and ($\langle S \rangle_{\text{lin}} = U$).

Assigned HW#33.

For example, let $V := \mathbb{R}^3$ and $U := \{(x, y, 0) \,|\, x, y \in \mathbb{R}\}$. Let $S := \{(1, 0, 0), (0, 1, 0), (3, 4, 0)\}$ and let $T := \{(1, 0, 0), (0, 1, 0)\}$. Then S and T are both spanning sets for U. Let $x := (3, 4, 0)$. Then $T = S \backslash \{x\}$. In some sense T is a better spanning set than S, because the vector x is extraneous, so it's good to omit it. We now home in on why x is extraneous. Note that $x = [3 \bullet (1, 0, 0)] + [4 \bullet (0, 1, 0)] \in \langle T \rangle_{\text{lin}}$. So since $T = S \backslash \{x\}$, we see that $x \in \langle S \backslash \{x\} \rangle_{\text{lin}}$. According to HW#33, it follows that $\langle S \rangle_{\text{lin}} = \langle S \backslash \{x\} \rangle_{\text{lin}}$, so the vector x can be omitted from S to create a new set with the same span as S.

We will say that a set of vectors is "linearly dependent" if at least one of them is extraneous in this sense:

DEFINITION 12.4. *For any vector space V, for any S, by S is* **linearly dependent** *in V, we mean:*

$$\textit{both} \quad (S \subseteq V) \quad \textit{and} \quad (\exists x \in S \textit{ s.t. } x \in \langle S \backslash \{x\} \rangle_{\text{lin}}).$$

The opposite of linearly dependent is "linearly independent":

DEFINITION 12.5. *For any vector space V, for any S, by S is* **linearly independent** *in V, we mean:*

$$\textit{both} \quad (S \subseteq V) \quad \textit{and} \quad (S \textit{ is not linearly dependent in } V).$$

Applying NOT to ($\exists x \in S$ s.t. $x \in \langle S \backslash \{x\} \rangle_{\text{lin}}$), we see that a subset S of a vector space V is linearly independent iff none of the vectors in

S is extraneous, *i.e.* iff

$$\forall x \in S, \quad x \notin S\backslash\{x\}\rangle_{\mathrm{lin}}.$$

Following this, if we want to check that a set of vectors is linearly independent, we have to look at each one and check that it cannot be omitted without changing the span. The following gives a more practical quantified equivalence for linear independence.

FACT 12.6. *Let V be a vector space Let $k \in \mathbb{N}$. Let $x_1, \ldots, x_k \in V$. Assume, for all $i, j \in \{1, \ldots, k\}$, that: $[\,(i \neq j) \Rightarrow (x_i \neq x_j)\,]$. Then:*

$$\begin{aligned} &[\ \ \{x_1, \ldots, x_k\} \text{ is linearly independent in } V\ \ ] \qquad \Leftrightarrow \\ &[\ \ \forall c_1, \ldots, c_k \in \mathbb{R}, \\ &\qquad\qquad (c_1x_1 + \cdots + c_kx_k = 0_V) \ \Rightarrow\ (c_1 = \cdots = c_k = 0)\ \ ]. \end{aligned}$$

Proof. Omitted. □

In general terms, Fact 12.6 asserts that a finite set of vectors is linearly independent iff the only linear combination of them that vanishes is the one with vanishing coefficients. The linear combination with vanishing coefficients is sometimes said to be "trivial". Then Fact 12.6 asserts that a finite set of vectors is linearly independent iff it has no nontrivial vanishing linear combination.

Note, for example, that $\{(1, 0, 0), (0, 1, 0), (3, 4, 0)\}$ has a nontrivial vanishing linear combination:

$$[3 \cdot (1, 0, 0)] \ +\ [4 \cdot (0, 1, 0)] \ +\ [(-1) \cdot (3, 4, 0)] \ =\ (0, 0, 0).$$

So, by Fact 12.6, $\{(1, 0, 0), (0, 1, 0), (3, 4, 0)\}$ is linearly dependent in $\mathbb{R}^3$. We leave it as an unassigned exercise to show that $\{(1, 0, 0), (0, 1, 0)\}$ has no nontrivial vanishing linear combination. So, by Fact 12.6, $\{(1, 0, 0), (0, 1, 0)\}$ is linearly independent in $\mathbb{R}^3$.

Assigned HW#34.

A "basis" for a vector space is a spanning set without extraneous vectors. That is:

DEFINITION 12.7. *For any vector space V, for any B, by B is a* **basis** *for V, we mean:*

$$(\ B \text{ spans } V\ ) \quad \& \quad (\ B \text{ is linearly independent in } V\ ).$$

For example, $\{(1, 0, 0), (0, 1, 0)\}$ is a basis for $\{(x, y, 0) \,|\, x, y \in \mathbb{R}\}$. Also $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is a basis for $\mathbb{R}^3$.

Assigned HW#35.

We next examine the basis $\{(1,0,0),(0,1,0),(0,0,1)\}$ for $\mathbb{R}^3$. Recall that $\mathbb{R}^3$ is, technically, equal to the set $\mathbb{R}^{\{1,2,3\}}$ consisting of all functions $\{1,2,3\} \to \mathbb{R}$. In particular, looking "under the hood" at say $(7,-9,4)$, we see that $(7,-9,4)$ is the function $\{1,2,3\} \to \mathbb{R}$ that maps 1 to 7, maps 2 to -9 and maps 3 to 4. That is, $(7,-9,4) = \begin{pmatrix} 1 \mapsto 7 \\ 2 \mapsto -9 \\ 3 \mapsto 4 \end{pmatrix}$. Similary, we have

$$(1,0,0) \quad = \quad \begin{pmatrix} 1 \mapsto 1 \\ 2 \mapsto 0 \\ 3 \mapsto 0 \end{pmatrix} \quad \in \quad \mathbb{R}^{\{1,2,3\}} \quad = \quad \mathbb{R}^3.$$

Similarly, we have

$$(0,1,0) \quad = \quad \begin{pmatrix} 1 \mapsto 0 \\ 2 \mapsto 1 \\ 3 \mapsto 0 \end{pmatrix} \quad \in \quad \mathbb{R}^{\{1,2,3\}} \quad = \quad \mathbb{R}^3.$$

Similarly, we have

$$(0,0,1) \quad = \quad \begin{pmatrix} 1 \mapsto 0 \\ 2 \mapsto 0 \\ 3 \mapsto 1 \end{pmatrix} \quad \in \quad \mathbb{R}^{\{1,2,3\}} \quad = \quad \mathbb{R}^3.$$

Recall Definition 7.2. Note that

$$\chi_{\{1\}}^{\{1,2,3\}} = \begin{pmatrix} 1 \mapsto 1 \\ 2 \mapsto 0 \\ 3 \mapsto 0 \end{pmatrix}, \ \chi_{\{2\}}^{\{1,2,3\}} = \begin{pmatrix} 1 \mapsto 0 \\ 2 \mapsto 1 \\ 3 \mapsto 0 \end{pmatrix}, \ \chi_{\{3\}}^{\{1,2,3\}} = \begin{pmatrix} 1 \mapsto 0 \\ 2 \mapsto 0 \\ 3 \mapsto 1 \end{pmatrix}.$$

Thus we have $\chi_{\{1\}}^{\{1,2,3\}} = (1,0,0)$, $\chi_{\{2\}}^{\{1,2,3\}} = (0,1,0)$, $\chi_{\{3\}}^{\{1,2,3\}} = (0,0,1)$. Let $S := \{1,2,3\}$. Then $\{(1,0,0),(0,1,0),(0,0,1)\} = \{\chi_{\{j\}}^S \,|\, j \in S\}$. This generalizes from $\mathbb{R}^3$ to give, for all $d \in \mathbb{N}$, a basis for $\mathbb{R}^d$.

DEFINITION 12.8. *Let $d \in \mathbb{N}$, $S := \{1,\ldots,d\}$. For all $j \in S$, let $\varepsilon_j := \chi_{\{j\}}^S$. Then $\{\varepsilon_1,\ldots,\varepsilon_d\}$ is the* **standard basis** *for $\mathbb{R}^d$.*

DEFINITION 12.9. *For any vector space V, we define*

$$\begin{aligned} Bas(V) &:= \{B \subseteq V \,|\, B \textit{ is a basis for } V\} \\ DIMS(V) &:= \{\#B \,|\, B \in Bas(V)\}. \end{aligned}$$

For example, let $S := \{(1,0,0),(0,1,0),(0,0,1)\}$, so that S is the standard basis for $\mathbb{R}^3$. In particular, $S \in \mathrm{Bas}\,(\mathbb{R}^3)$. Since $\#S = 3$, we get $3 \in \mathrm{DIMS}\,(\mathbb{R}^3)$. Mathematicians have studied bases of $\mathbb{R}^3$ intensely, and have verified that, for *every* $B \in \mathrm{Bas}\,(\mathbb{R}^3)$, we have $\#B = 3$. That is, $\mathrm{DIMS}\,(\mathbb{R}^3) = \{3\}$. Consequently, $\#[\mathrm{DIMS}\,(\mathbb{R}^3)] = \#\{3\} = 1$.

For example, let $U := \{(x,y,0)\,|\,x,y \in \mathbb{R}\}$, $T := \{(1,0,0),(0,1,0)\}$. Then $T \in \mathrm{Bas}\,(U)$, so, since $\#T = 2$, we get $2 \in \mathrm{DIMS}\,(U)$. Mathematicians have studied bases of U intensely, and have verified that, for *every* $B \in \mathrm{Bas}\,(U)$, we have $\#B = 2$. That is, $\mathrm{DIMS}\,(U) = \{2\}$. Consequently, $\#[\mathrm{DIMS}\,(U)] = \#\{2\} = 1$.

Mathematicians have studied bases of vector spaces intensely, and have verified that, for *every* vector space V, we have $\#[\mathrm{DIMS}\,(V)] = 1$. A single vector space of nonzero dimension will always have uncountably many bases, and yet any time we find two bases for the same vector space, it turns out that the two bases have the same number of elements. That number is called the **dimension** of the vector space:

DEFINITION 12.10. *$\forall$vector space V, $\dim V := \mathrm{ELT}(DIMS\,(V))$.*

Unassigned exercise: Show, for any vector space V, that: if $\dim V = 0$, then $V = \{0_V\}$. In particular, every 0-dimensional vector space has only one vector. By contrast, we will see that every positive-dimensional vector space has continuum cardinality.

For example,

$$\begin{aligned} \dim \mathbb{R}^3 &= \mathrm{ELT}\{3\} = 3 \qquad \text{and} \\ \dim \{(x,y,0)\,|\,x,y \in \mathbb{R}\} &= \mathrm{ELT}\{2\} = 2. \end{aligned}$$

DEFINITION 12.11. *Let V be a vector space. By V is* **finite dimensional**, *we mean $\dim V < \infty$. By V is* **infinite dimensional**, *we mean $\dim V = \infty$.*

13. Class 13 on 19 October 2017, Th of Week 7

let S be a set and let $k \in \mathbb{N}$. Recall that

$$S^k = S^{\{1,\ldots,k\}} = \{\text{functions } \{1,\ldots,k\} \to S\}.$$

Also, recall that, for all $x \in S^k$, we have

$$x \quad = \quad \begin{pmatrix} 1 \mapsto x_1 \\ \vdots \\ k \mapsto x_k \end{pmatrix} \quad = \quad (x_1, \ldots, x_k),$$

$$\text{and so} \qquad \mathrm{im}[x] \quad = \quad \{x_1, \ldots, x_k\}.$$

Let $\varepsilon := (\,(1,0)\,,\,(0,1)\,)$. Then

$$\varepsilon \quad = \quad \begin{pmatrix} 1 \mapsto (1,0) \\ 2 \mapsto (0,1) \end{pmatrix} \quad \in \quad (\mathbb{R}^2)^{\{1,2\}} \quad = \quad (\mathbb{R}^2)^2.$$

Then $\mathrm{im}[\varepsilon] = \{(1,0),(0,1)\}$, so $\mathrm{im}[\varepsilon]$ is a basis for $\mathbb{R}^2$. Note also that

$$\varepsilon_1 \quad = \quad (1,0) \quad = \quad \begin{pmatrix} 1 \mapsto 1 \\ 2 \mapsto 0 \end{pmatrix} \quad = \quad \chi_{\{1\}}^{\{1,2\}}$$

$$\text{and} \qquad \varepsilon_2 \quad = \quad (0,1) \quad = \quad \begin{pmatrix} 1 \mapsto 0 \\ 2 \mapsto 1 \end{pmatrix} \quad = \quad \chi_{\{2\}}^{\{1,2\}}.$$

DEFINITION 13.1. *For any finite dimensional vector space V, for any B, by B is an* **ordered basis** *or* **OB** *for V, we mean:*

$$(\,B \in V^{dim\,V}\,) \qquad \textit{and} \qquad (\,\textit{im}[B] \textit{ is a basis for } V\,).$$

For example, $((1,0),(0,1))$ is an ordered basis for $\mathbb{R}^2$, and $((0,1),(1,0))$ is a different ordered basis for $\mathbb{R}^2$.

DEFINITION 13.2. *Let $d \in \mathbb{N}$. For every $j \in \{1,\ldots,d\}$, define $\varepsilon_j := \chi_{\{j\}}^{\{1,\ldots,d\}}$. Then the* **standard OB** *for $\mathbb{R}^d$ is $(\varepsilon_1,\ldots,\varepsilon_d)$.*

For example, the standard ordered basis for $\mathbb{R}^2$ is $((1,0),(0,1))$.

DEFINITION 13.3. *For any finite dimensional vector space V, we define $OB(V) := \{OBs\ for\ V\}$.*

For example, $((0,1),(1,0)) \in \mathrm{OB}(\mathbb{R}^2)$.

We leave it as an unassigned exercise to show, for any finite dimensional vector space V, that, because V has a finite basis, $\mathrm{OB}(V) \neq \varnothing$.

DEFINITION 13.4. *Let V and W be vector spaces. Let $T : V \to W$. Then T is* **linear** *means:*

(1) $\forall x, y \in V,\ (T(x+y) = [T(x)] + [T(y)])$ *and*
(2) $\forall a \in \mathbb{R},\ \forall x \in V,\ (T(ax) = a \bullet [T(x)])$.

DEFINITION 13.5. *Let V and W be vector spaces. Let $T : V \to W$. Then T is an* **isomorphism** *onto W means:*

(1) *T is linear and*
(2) $T : V \hookrightarrow\!\!> W$.

When the vector space W is clear, we will sometimes just say "T is an isomorphism", instead of "T is an isomorphism onto W".

The word "isomorphism" is used in many different ways in mathematics. The "isomorphism" of Definition 13.5 is sometimes called a **vector space isomorphism** or **linear isomorphism**, to distinguish it from other kinds of isomorphisms.

LEMMA 13.6. *Let V and W be vector spaces and let $T : V \to W$. Assume that T is an isomorphism onto W. Then T^{-1} is an isomorphism onto V, i.e.,*
(1) *T^{-1} is linear and*
(2) $T^{-1} : W \hookrightarrow\!\!> V$.

Proof. Omitted. □

DEFINITION 13.7. *Let V and W be vector spaces. Then V is* **isomorphic** *to W means: $\exists T$ s.t.*
(1) *T is linear and*
(2) $T : V \hookrightarrow\!\!> W$.

That is, two vector spaces are isomorphic if there exists an isomorphism from one of them onto the other.

For example, while

$$\mathbb{R}^1 \;=\; \mathbb{R}^{\{1\}} \;=\; \{\text{functions } \{1\} \to \mathbb{R}\} \;\neq\; \mathbb{R},$$

it is nevertheless true that $\mathbb{R}^1$ "looks and feels like" $\mathbb{R}$. In more techical language, we say: $\mathbb{R}^1$ is isomorphic to $\mathbb{R}$. To describe the isomorphism, first note that, for all $t \in \mathbb{R}$, (t) is the 1-tuple whose only entry is t; that is, (t) is the function $\{1\} \to \mathbb{R}$ that is described by $1 \mapsto t$. So,

$$\forall t \in \mathbb{R}, \qquad (t) \;=\; (1 \mapsto t) \;\in\; \mathbb{R}^{\{1\}} \;=\; \mathbb{R}^1.$$

So, for all $t \in \mathbb{R}$, if we evaluate the function (t) at 1, we get t. Thus, for all $t \in \mathbb{R}$, while $(t) \neq t$, we do have $(t)_1 = t$. Unassigned exercise: Show that the function

$$t \;\mapsto\; (t) \quad : \quad \mathbb{R} \;\to \mathbb{R}^1$$

is a vector space isomorphism, and that its inverse is

$$v \;\mapsto\; v_1 \quad : \quad \mathbb{R}^1 \;\to \mathbb{R}.$$

In particular, $\mathbb{R}$ and $\mathbb{R}^1$ are isomorphic vector spaces..

For another example, let $U := \{(x, y, 0) \,|\, x, y \in \mathbb{R}\}$. Recall that U inherits vector addition and scalar multiplication from $\mathbb{R}^3$, and with these linear operations, U is a vector space. Unassigned exercise: Show that the function

$$(x, y) \;\mapsto\; (x, y, 0) \quad : \quad \mathbb{R}^2 \;\to\; U$$

is a vector space isomorphism. In particular, $\mathbb{R}^2$ and U are isomorphic vector spaces.

For another example, recall that, for any set S, $\mathbb{R}^S$ has linear operations making it into a vector space. In particular, $\mathbb{R}^{\{2,7,8\}}$ is a vector space. Unassigned exercise: Show that the function

$$(x, y, z) \;\mapsto\; \begin{pmatrix} 2 \mapsto x \\ 7 \mapsto y \\ 8 \mapsto z \end{pmatrix} \quad : \quad \mathbb{R}^3 \;\to\; \mathbb{R}^{\{2,7,8\}}$$

is a vector space isomorphism. In particular, R^3 and $\mathbb{R}^{\{2,7,8\}}$ are isomorphic vector spaces.

For *finite dimensional* vector spaces, we can connect ordered bases with isomorphisms, as follows.

DEFINITION 13.8. *Let V be a vector space, let $d := \dim V$ and let $B \in OB(V)$. Assume $d < \infty$. Then $L_B : \mathbb{R}^d \to V$ is defined by*

$$L_B(x_1, \ldots, x_d) \quad = \quad x_1 B_1 \;+\; \cdots \;+\; x_d B_d.$$

Recall that $\mathbb{R}^0 = \{0\} \subseteq \mathbb{R}$, and that $\mathbb{R}^0$ is a vector space, with linear operations inherited from $\mathbb{R}$. In Definition 13.8, if $d = 0$, then $V = \{0_V\}$, and we define $L_B : \mathbb{R}^d \to V$ by $L_B(0) = 0_V$. In theorems about L_B, we will often omit discussion of the 0-dimensional case, leaving that detail to the reader.

For example, let $U := \{(x, y, 0) \,|\, x, y \in \mathbb{R}\}$, $B := (\,(1, 0, 0)\,,\,(0, 1, 0)\,)$. Recall that $\dim U = 2$ and that $B \in \mathrm{OB}(V)$. We "calculate" the map $L_B : \mathbb{R}^2 \to U$: For all $x, y \in \mathbb{R}$, we have

$$\begin{aligned} L_B(x, y) \;&=\; xB_1 \;+\; yB_2 \\ &=\; x(1, 0, 0) \;+\; y(0, 1, 0) \\ &=\; (x, y, 0). \end{aligned}$$

Note that L_B is the isomorphism $\mathbb{R}^2 \to U$ that we described earlier. Let $C := (\,(0, 1, 0)\,,\,(1, 0, 0)\,)$. Note that B and C are different ordered bases of U, even though they give rise to the same (unordered) basis:

$$\mathrm{im}[B] \;=\; \{(1,0,0),(0,1,0)\} \;=\; \{(0,1,0),(1,0,0)\} \;=\; \mathrm{im}[C].$$

For all $x, y \in \mathbb{R}$, we have

$$\begin{aligned} L_C(x,y) &= xC_1 \;+\; yC_2 \\ &= x(0,1,0) \;+\; y(1,0,0) \\ &= (y,x,0). \end{aligned}$$

Note that L_C is slightly different from L_B. Unassigned exercise: Show that $L_C : \mathbb{R}^2 \to U$ is an isomorphism of vector spaces.

Quite generally, any time B is an ordered basis of any finite dimensional vector space, the function L_B is an isomorphism:

THEOREM 13.9. *Let V be a vector space, $d := \mathit{dim}\, V$, $B \in OB(V)$. Assume $d < \infty$. Then $L_B : \mathbb{R}^d \to V$ is a vector space isomorphism.*

Proof. Omitted. □

If we think of a vector space isomorphism as a dictionary that

- allows us to tranlsate questions about one vector space over to another, better understood, vector space, and, also,
- allows us to translate answers back,

then the importance of the isomorphism L_B of Theorem 13.9 can be expressed this way: For any $d \in \mathbb{N}$, if some d-dimensional vector space V seems very abstract and recondite, just seek out an ordered basis B for V, and then use the "dictionary" $L_B : \mathbb{R}^d \to V$; this connects the abstract V to the concrete, well-understood vector space $\mathbb{R}^d$ of d-tuples of real numbers. In particular, if $d \in \{0,1,2,3\}$, then we can "see" $\mathbb{R}^d$, and, therefore, we can "see" V. Following this thinking, we have:

COROLLARY 13.10. *Let V be a finite dimensional vector space. Then V is isomorphic to $\mathbb{R}^{\mathit{dim}\, V}$.*

Proof. Let $d := \dim V$. We wish to show that V is isomorphic to $\mathbb{R}^d$.

Since $\mathrm{OB}(V) \neq \varnothing$, choose $B \in \mathrm{OB}(V)$. Then $L_B : \mathbb{R}^d \to V$ is an isomorphism. Then $L_B^{-1} : V \to \mathbb{R}^d$ is an isomorphism. Then V is isomorphic to $\mathbb{R}^d$. □

This finishes our introduction to linear algebra. We move on to the study of norms, metrics, topology, after which we can begin analysis: limits, derivatives and integrals.

DEFINITION 13.11. *Let V be a vector space. Let $|\bullet| : V \to [0,\infty)$. Then $|\bullet|$ is a* **norm** *on V means:*

(1) ($\forall x \in V$, [$(|x| = 0) \Leftrightarrow (x = 0_V)$]),
(2) ($\forall a \in \mathbb{R}$, $\forall x \in V$, [$|ax| = |a| \cdot |x|$]) *and*
(3) ($\forall x, y \in V$, [$|x + y| \leqslant |x| + |y|$]).

In (2) of Definition 13.11, $|a|$ is the absolute value of a. In the rest of the definition, every use of $|\bullet|$ is the norm $|\bullet| : V \to [0,\infty)$.

Some terminology about Definition 13.11:

- We express (1) by saying: "$|\bullet|$ distinguishes zero".
- We express (2) by saying: "$|\bullet|$ is positive-homogeneous".
- We express (3) by saying: "$|\bullet|$ is subadditive".

Absolute value $|\bullet| : \mathbb{R} \to [0,\infty)$ is a norm on the vector space $\mathbb{R}$. It might be called the **standard norm** on $\mathbb{R}$. Note: $\forall x \in \mathbb{R}$, $|x| = \sqrt{x^2}$. We can generalize that formula to higher dimensions, as follows.

Let $d \in \mathbb{N}$. Define $|\bullet| : \mathbb{R}^d \to [0,\infty)$ by

$$|(x_1, \ldots, x_d)| \quad = \quad \sqrt{x_1^2 + \cdots + x_d^2}.$$

Then $|\bullet|$ is a norm on the vector space $\mathbb{R}^d$, and is called the **standard norm** on $\mathbb{R}^d$. Note: $\forall x_1, \ldots, x_d \in \mathbb{R}$,

$$|(x_1, \ldots, x_d)| \quad = \quad (\ |x_1|^2 \ + \ \cdots \ + \ |x_d|^2\ )^{1/2}.$$

We can generalize that formula to other exponents, as follows.

Let $d \in \mathbb{N}$ and $p \in [1,\infty)$. Define $|\bullet| : \mathbb{R}^d \to [0,\infty)$ by

$$|(x_1, \ldots, x_d)| \quad = \quad (\ |x_1|^p \ + \ \cdots \ + \ |x_d|^p\ )^{1/p}.$$

Then $|\bullet|$ is a norm on the vector space $\mathbb{R}^d$, and is called the **p-norm** or **L^p-norm** on $\mathbb{R}^d$.

Let $d \in \mathbb{N}$. Define $|\bullet| : \mathbb{R}^d \to [0,\infty)$ by

$$|(x_1, \ldots, x_d)| \quad = \quad \max\{|x_1|, \ldots, |x_d|\}.$$

Then $|\bullet|$ is a norm on the vector space $\mathbb{R}^d$, and is called the **∞-norm** or **L^∞-norm** on $\mathbb{R}^d$.

Note: $\forall d \in \mathbb{N}$, [(standard norm on $\mathbb{R}^d$) = (2-norm on $\mathbb{R}^d$)].

DEFINITION 13.12. *For any vector space V, we define*

$$\mathcal{N}(V) \quad := \quad \{\textit{norms on } V\}.$$

DEFINITION 13.13. *By $(V, |\bullet|)$ is a* **normed vector space**, *we mean: (V is a vector space) and ($|\bullet| \in \mathcal{N}(V)$).*

DEFINITION 13.14. *For any normed vector space $(V, |\bullet|)$, we define*

$$\begin{array}{lll} B_{|\bullet|} & := & \{x \in V \text{ s.t. } |x| < 1\}, \\ S_{|\bullet|} & := & \{x \in V \text{ s.t. } |x| = 1\} \qquad \textit{and} \\ \overline{B}_{|\bullet|} & := & \{x \in V \text{ s.t. } |x| \leqslant 1\}. \end{array}$$

For any normed vector space $(V, |\bullet|)$,

the set $B_{|\bullet|}$ is called the **unit ball** of $|\bullet|$,
the set $S_{|\bullet|}$ is called the **unit sphere** of $|\bullet|$ and
the set $\overline{B}_{|\bullet|}$ is called the **closed unit ball** of $|\bullet|$.

Assigned HW#36, due 31 October.

We played a game: I told you I had a secret norm $|\bullet|$ on $\mathbb{R}^2$. I drew a picture for you of $\overline{B}_{|\bullet|}$. I picked a point x in the plane slightly outside of $\overline{B}_{|\bullet|}$. I asked you to estimate $|x|$. We determined that $|x| > 1$ exactly because x is not an element of $\overline{B}_{|\bullet|}$. We dilated $\overline{B}_{|\bullet|}$ by a factor of 2, and saw that (the point x) is an element of (that dilated set). Therefore, $|x| \leqslant 2$. Continuing, we determined that $|x|$ is only slightly bigger than 1, because x is only slightly outside $\overline{B}_{|\bullet|}$. In fact, $|x|$ is the minimum of the set of all dilation factors such that

(the dilation of $\overline{B}_{|\bullet|}$) covers (the point x).

The upshot of this kind of thinking is that any norm is determined by its closed unit ball. Consequently, there's a close connection between norms and geometry. If someone shows you a norm on a finite dimensional vector space, and you would like to understand it, one approach is to try to understand, geometrically, what its closed unit ball looks like. Mathematicians have intensively studied closed unit balls of norms, and know exactly which kinds of sets arise. Suppose you would like to make your own personal norm on $\mathbb{R}^2$. Proceed as follows: Draw a subset S of $\mathbb{R}^2$ such that

(1) S contains all of its boundary points,
(2) S contains a disk about $(0, 0)$,
(3) $\forall p \in S$, we have $-p \in S$ and
(4) $\forall p, q \in S$, the line segment from p to q is contained in S.

It will then turn out that there is a unique norm $|\bullet|$ on $\mathbb{R}^2$ such that $\overline{B}_{|\bullet|} = S$. That norm is your personal norm.

Some terminology:

- We express (1) by saying: "S is closed".
- We express (2) by saying: "the origin is in the interior of S".

- We express (3) by saying: "S is symmetric through the origin".
- We express (4) by saying: "S is convex".

Let $p := 1/2$ and define a function $|\bullet| : \mathbb{R}^2 \to [0, \infty)$ by

$$|(a, b)| \quad = \quad (\ |a|^p \ + \ |b|^p\ )^{1/p}.$$

Let $S := \{x \in \mathbb{R}^2 \text{ s.t. } |x| \leqslant 1\}$. I drew the graph of S, and we verified that the resulting set is not convex. It turns out that the nonconvexity of S is driven exactly by the fact that $|\bullet|$ is *NOT* subadditive. Consequently, $|\bullet|$ is not a norm. Recall: In defining "p-norm", we required that $p \in [1, \infty]$. This explains what goes wrong when $p < 1$.

The key point, though, is that we have a geometric way to "see" the algebraic property of subadditivity; you just have to know where to look. For all $p \geqslant 1$, we "see" subadditivity in the p-norm. For all $p < 1$, we do not, so those functions are not norms. Also, following this same line of thought, HW#36 explains how to "see" the ∞-norm: $\forall p \in [1, \infty)$, we "see" the p-norm; we then take the limit as $p \to \infty$.

In the sequel, we will not be discussing p-norms, except for 2-norms, a.k.a. standard norms.

This concludes our introduction to norms. Next up are metrics. First, keep in mind that the word "metric" is used different ways. In a course on Riemannian geometry, "metric" is short for "Riemannian metric" and we will not be discussing that kind of metric in this course. Here, by "metric", we will mean "distance-metric", defined as follows:

DEFINITION 13.15. *Let X be a set. Let $d : X \times X \to [0, \infty)$. Then d is a* **metric** *or* **distance-metric** *on X means:*

(1) $(\quad \forall p, q \in X, \quad [\ (d(p, q) = 0) \ \Leftrightarrow\ (p = q)\ ] \quad)$,
(2) $(\quad \forall p, q \in X, \quad [\ d(p, q) = d(q, p)\ ] \quad)$ *and*
(3) $(\quad \forall p, q, r \in X, \quad [\ d(p, r) \leqslant (d(p, q)) + (d(q, r))\ ] \quad)$.

Some terminology about Definition 13.15:

- We express (1) by: "d distinguishes points".
- We express (2) by: "d is symmetric".
- We express (3) by: "d is satisfies the **triangle inequality**".

DEFINITION 13.16. *For any set X, we define*

$$\mathcal{M}(X) \quad := \quad \{\textit{metrics on } X\}.$$

DEFINITION 13.17. *By (X, d) is a* **metric space**, *we mean:*

$$(\ X \textit{ is a set}\ ) \qquad \textit{and} \qquad (\ d \in \mathcal{M}(X)\ ).$$

DEFINITION 13.18. *Let $(V, |\bullet|)$ be a normed vector space. Then $d_{|\bullet|} \in \mathcal{M}(V)$ is defined by $d_{|\bullet|}(x, y) = |x - y|$.*

In Definition 13.18, we call $d_{|\bullet|}$ the **induced metric** from $|\bullet|$.

Assigned HW#37, due 31 October.

According to HW#37, for any vector space V, for any $|\bullet| \in \mathcal{N}(V)$, if we measure distance in V by $d_{|\bullet|}$ and if we measure distance in $\mathbb{R}$ by $d_{\mathbb{R}}$, then the mapping $|\bullet| : V \to \mathbb{R}$ is distance-semidecreasing (*i.e.*, distance-nonincreasing). This property of norms is often useful.

We now start to move toward topology. Intuitively, a topology captures ideas of closeness, through what are called "neighborhoods" (to be defined later). Each neighborhood is a set of points that are, in some way, "close" to a given point. A collection of neighborhoods is a set of sets. One way to get a collection of neighborhoods is to find a metric space and take the set of all "balls" in the metric space. So we begin by defining what a ball in a metric space is:

DEFINITION 13.19. *Let (X, d) be a metric space. Then, for all $p \in X$, for all $r > 0$, we define*

$$B_d(p, r) \quad := \quad \{q \in X \,|\, d(p, q) < r\}.$$

In Definition 13.19, $B_d(p, r)$ is called the **ball** about p of radius r. (Sometimes, instead of "about p", we say "centered at p".) When the metric d is clear, we will often omit it from the notation, and write $B(p, r)$ in place of $B_d(p, r)$.

DEFINITION 13.20. *Let (X, d) be a metric space. Then, for all $p \in X$, we define $\mathcal{B}_d(p) := \{B(p, r) \,|\, r > 0\}$.*

In Definition 13.20, when d is clear, we omit it from $\mathcal{B}_d(p)$, and write $\mathcal{B}(p)$ instead.

For any metric space (X, d), for any $p \in X$, the set $\mathcal{B}(p)$ is the collection of all balls centered at p; it is a set of sets.

DEFINITION 13.21. *Let (X, d) be a metric space. Then we define $\mathcal{B}_d := \{B(p, r) \,|\, p \in X, r > 0\}$.*

For any metric space (X, d), the set $\mathcal{B}_d$ is the collection of all balls in X; it is a set of sets.

Assigned HW#38, due 31 October.

14. Class 14 on 24 October 2017, Tu of Week 8

Recall: $(V, |\bullet|)$ is a **normed vector space** means that V is a vector space and that $|\bullet| : V \to [0, \infty)$ satisfies:

distinguishes 0_V, positive homogeneous, subadditive.

Recall: (X, d) is a **metric space** means that X is a set and that $d : X \times X \to [0, \infty)$ satisfies:

distinguishes points, symmetric, triangle inequality.

We now start moving toward topology which is all about closeness. One way to talk about closeness is through "balls". The points that are close to a given point p might be the points that are a small distance away, *i.e.*, the points that are in a small ball about the point (with respect to some metric). Following this, we recall: Let (X, d) be a metric space. Then

- $\forall p \in X$, $\forall r > 0$, $B(p, r) := \{q \in X \,|\, d(p, q) < r\}$,
- $\forall p \in X$, $\mathcal{B}(p) := \{B(p, r) \,|\, r > 0\}$ and
- $\mathcal{B}_d := \{B(p, r) \,|\, p \in X,\, r > 0\}$.

Recalled HW#38.

Assigned HW#39.

Here are some facts about set theory that may prove useful:

FACT 14.1. *Let $\mathcal{S}$ be a set of sets. Then*

(1) $\forall X \in \mathcal{S}$, $X \subseteq \bigcup \mathcal{S}$,
(2) $\forall X \in \mathcal{S}$, $X \in \langle \mathcal{S} \rangle_\cup$,
(3) $\varnothing$, $\bigcup \mathcal{S} \in \langle \mathcal{S} \rangle_\cup$ *and*
(4) $\mathcal{S} \bigcup \{\varnothing\} \bigcup \{\bigcup \mathcal{S}\} \subseteq \langle \mathcal{S} \rangle_\cup$.

Proof. Omitted. □

Assigned HW#40.

You may find (1) of Fact 14.1 to be useful in doing HW#40.

The next lemma will be called the **Recentering Down Lemma**. It is an easy consequence of HW#38.

LEMMA 14.2. *Let (X, d) be a metric space, let $B \in \mathcal{B}_d$ and let $p \in B$. Then there exists $B_0 \in \mathcal{B}(p)$ such that $B_0 \subseteq B$.*

Lemma 14.2 states: For any a point inside any ball, we can find a new ball centered at that point that's so small that it's contained in the original ball. Of course, if the point is close to the boundary of the original ball, then the new ball will have to have a very small radius. HW#38 tells how to compute the radius of the new ball.

Proof. Since $B \in \mathcal{B}_d$, choose $q \in X$ and $s > 0$ such that $B = B(q, s)$. Let $r := s - [d(p, q)]$. By HW#38, $r > 0$ and $B(p, r) \subseteq B$. Let $B_0 := B(p, r)$. Then, as B_0 is centered at p, we see that $B_0 \in \mathcal{B}(p)$. We wish to show that $B_0 \subseteq B$.

We have $B_0 = B(p, r) \subseteq B$, as desired. □

DEFINITION 14.3. *For any set $\mathcal{S}$ of sets, we define*

(1) $\langle \mathcal{S} \rangle_\cap \quad := \quad \{ \bigcap \mathcal{A} \mid \varnothing \neq \mathcal{A} \subseteq \mathcal{S} \}$ *and*
(2) $\langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap} \quad := \quad \{ \bigcap \mathcal{A} \mid (\varnothing \neq \mathcal{A} \subseteq \mathcal{S})$ *and* $(\#\mathcal{A} < \infty) \}$.

The set $\langle \mathcal{S} \rangle_\cap$ is called the **intersection closure** of $\mathcal{S}$. The set $\langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap}$ is called the **finite intersection closure** of $\mathcal{S}$.

The operators $\langle \bullet \rangle_\cap$ and $\langle \bullet \rangle_{\mathrm{fin}\,\cap}$ both have the usual properties:

superset, idempotent, monotonic.

To say this more specifically:

REMARK 14.4. *The following are all true:*

(1A) $\forall$*set* $\mathcal{S}$ *of sets*, $\langle \mathcal{S} \rangle_\cap \supseteq \mathcal{S}$.
(1B) $\forall$*set* $\mathcal{S}$ *of sets*, $\langle \langle \mathcal{S} \rangle_\cap \rangle_\cap = \langle \mathcal{S} \rangle_\cap$.
(1C) $\forall$*sets* $\mathcal{R}, \mathcal{S}$ *of sets*, $\quad [\mathcal{R} \subseteq \mathcal{S}] \Rightarrow [\langle \mathcal{R} \rangle_\cap \subseteq \langle \mathcal{S} \rangle_\cap]$.
(2A) $\forall$*set* $\mathcal{S}$ *of sets*, $\langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap} \supseteq \mathcal{S}$.
(2B) $\forall$*set* $\mathcal{S}$ *of sets*, $\langle \langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap} \rangle_{\mathrm{fin}\,\cap} = \langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap}$.
(2C) $\forall$*sets* $\mathcal{R}, \mathcal{S}$ *of sets*, $\quad [\mathcal{R} \subseteq \mathcal{S}] \Rightarrow [\langle \mathcal{R} \rangle_{\mathrm{fin}\,\cap} \subseteq \langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap}]$.

Proof. Omitted. □

We will need a basic fact about finite nonempty sets: in any such set, you can put the elements into a finite list. More specifically:

FACT 14.5. *Let S be a set and let $k := \#S$. Assume $0 < k < \infty$. Then $\exists x_1, \ldots, x_k$ s.t. $S = \{x_1, \ldots, x_k\}$.*

Proof. Omitted. □

PROPOSITION 14.6. *Let $\mathcal{S}$ be a set of sets. We assume that, for all $U, V \in \mathcal{S}$, we have $U \cap V \in \mathcal{S}$. Then $\langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap} = \mathcal{S}$.*

Proposition 14.6 can be expressed as: "If a set of sets is closed under pairwise-intersection, then it's closed under finite intersection."

Proof. Claim: $\forall k \in \mathbb{N}$, $\forall X_1, \ldots, X_k \in \mathcal{S}$, $\quad X_1 \cap \cdots \cap X_k \in \mathcal{S}$. *Proof of claim:* For all $k \in \mathbb{N}$, let

$$P_k \;:=\; [\; \forall X_1, \ldots, X_k \in \mathcal{S}, \quad X_1 \cap \cdots \cap X_k \in \mathcal{S} \;].$$

We wish to show: $\forall k \in \mathbb{N}$, P_k. We have $\forall X_1 \in \mathcal{S}$, $X_1 \in \mathcal{S}$, and so P_1 follows. By the Principle of Mathematical Induction, it suffices to show: $\forall k \in \mathbb{N}$, $(\, P_k \;\Rightarrow\; P_{k+1}\,)$. Let $k \in \mathbb{N}$ be given. We wish to show: $P_k \Rightarrow P_{k+1}$. Assume P_k. We wish to show P_{k+1}. We know the induction assumption: $\forall X_1, \ldots, X_k \in \mathcal{S}$, $X_1 \cap \cdots \cap X_k \in \mathcal{S}$. We want to prove: $\forall X_1, \ldots, X_{k+1} \in \mathcal{S}$, $X_1 \cap \cdots \cap X_{k+1} \in \mathcal{S}$. Let $X_1, \ldots, X_{k+1} \in \mathcal{S}$ be given. We wish to prove: $X_1 \cap \cdots \cap X_{k+1} \in \mathcal{S}$.

By the induction assumption, $X_1 \cap \cdots \cap X_k \in \mathcal{S}$. Recall: $X_{k+1} \in \mathcal{S}$. Let $U := X_1 \cap \cdots \cap X_k$ and let $V := X_{k+1}$. Then $U, V \in \mathcal{S}$. Then, by hypothesis, $U \cap V \in \mathcal{S}$. Then

$$\begin{aligned} X_1 \cap \cdots \cap X_{k+1} \;&=\; (X_1 \cap \cdots \cap X_k) \cap X_{k+1} \\ &=\; U \cap V \;\in\; \mathcal{S}, \end{aligned}$$

as desired. *End of proof of claim.*

By (2A) of Remark 14.4, we have $\langle \mathcal{S} \rangle_{\text{fin}\,\cap} \supseteq \mathcal{S}$. It remains to show: $\langle \mathcal{S} \rangle_{\text{fin}\,\cap} \subseteq \mathcal{S}$. We wish to show: $\forall Z \in \langle \mathcal{S} \rangle_{\text{fin}\,\cap}$, $z \in \mathcal{S}$. Let $Z \in \langle \mathcal{S} \rangle_{\text{fin}\,\cap}$ be given. We wish to show: $Z \in \mathcal{S}$.

By definition of $\langle \bullet \rangle_{\text{fin}\,\cap}$, since $Z \in \langle \mathcal{S} \rangle_{\text{fin}\,\cap}$, choose $\mathcal{A}$ such that

$$(\; \varnothing \neq \mathcal{A} \subseteq \mathcal{S} \;) \text{ and } (\; \#\mathcal{A} < \infty \;) \text{ and } (\; \bigcap \mathcal{A} = Z \;).$$

Since $\mathcal{A} \neq \varnothing$, we get $\#\mathcal{A} > 0$. Then $0 < \#\mathcal{A} < \infty$. Let $k := \#\mathcal{A}$. Then we have $0 < k < \infty$, so, by Fact 14.5, choose $X_1, \ldots, X_k$ such that $\mathcal{A} = \{X_1, \ldots, X_k\}$. Then $X_1 \cap \cdots \cap X_k = \bigcap \mathcal{A}$. By the claim, $X_1 \cap \cdots \cap X_k \in \mathcal{S}$. Then $Z = \bigcap \mathcal{A} = X_1 \cap \cdots \cap X_k \in \mathcal{S}$, as desired. $\square$

Because intersection distributes over union, we get many "expansion formulas", like the one in the following fact.

FACT 14.7. *Let X_1, X_2, Y_1, Y_2, Y_3 be sets. Then*

$$\begin{aligned} &(X_1 \cup X_2) \;\cap\; (Y_1 \cup Y_2 \cup Y_3) \;= \\ &\qquad (X_1 \cap Y_1) \;\cup\; (X_1 \cap Y_2) \;\cup\; (X_1 \cap Y_3) \;\cup \\ &\qquad (X_2 \cap Y_1) \;\cup\; (X_2 \cap Y_2) \;\cup\; (X_2 \cap Y_3). \end{aligned}$$

Proof. Omitted, but this result is a special case of the next lemma. □

Rather than proving Fact 14.7, along with lots of similar expansion formulas, we will prove a general result:

LEMMA 14.8. *Let $\mathcal{A}$ and $\mathcal{B}$ be sets of sets. Let*

$$\mathcal{C} \quad := \quad \{\, X \cap Y \mid X \in \mathcal{A},\ Y \in \mathcal{B} \,\}.$$

Then $(\bigcup \mathcal{A}) \bigcap (\bigcup \mathcal{B}) = \bigcup \mathcal{C}$.

Proof. We wish to show: $\forall p$,

$$\left[\ p \ \in \ \left(\bigcup \mathcal{A}\right) \bigcap \left(\bigcup \mathcal{B}\right)\ \right] \quad \Leftrightarrow \quad \left[\ p \ \in \ \bigcup \mathcal{C}\ \right].$$

Let p be given. We wish to show:

$$\left[\ p \ \in \ \left(\bigcup \mathcal{A}\right) \bigcap \left(\bigcup \mathcal{B}\right)\ \right] \quad \Leftrightarrow \quad \left[\ p \ \in \ \bigcup \mathcal{C}\ \right].$$

Proof of $\Rightarrow$: Assume: $p \in (\bigcup \mathcal{A}) \bigcap (\bigcup \mathcal{B})$. Want: $p \in \bigcup \mathcal{C}$.

We have $p \in \bigcup \mathcal{A}$ and $p \in \bigcup \mathcal{B}$. As $p \in \bigcup \mathcal{A}$, choose $X \in \mathcal{A}$ such that $p \in X$. As $p \in \bigcup \mathcal{B}$, choose $Y \in \mathcal{B}$ such that $p \in Y$. Since $p \in X$ and $p \in Y$, we get $p \in X \cap Y$. Let $Z := X \cap Y$. Then, by definition of $\mathcal{C}$, $Z \in \mathcal{C}$, so $Z \subseteq \bigcup \mathcal{C}$. Then $p \in X \cap Y = Z \subseteq \bigcup \mathcal{C}$. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume: $p \in \bigcup \mathcal{C}$. Want: $p \in (\bigcup \mathcal{A}) \bigcap (\bigcup \mathcal{B})$.

As $p \in \bigcup \mathcal{C}$, choose $Z \in \mathcal{C}$ such that $p \in Z$. By definition of $\mathcal{C}$, since $Z \in \mathcal{C}$, choose $X \in \mathcal{A}$ and $Y \in \mathcal{B}$ such that $Z = X \cap Y$. As $X \in \mathcal{A}$, we get $X \subseteq \bigcup \mathcal{A}$. As $Y \in \mathcal{B}$, we get $Y \subseteq \bigcup \mathcal{B}$. We have $p \in Z = X \cap Y \subseteq X \subseteq \bigcup \mathcal{A}$ and $p \in Z = X \cap Y \subseteq Y \subseteq \bigcup \mathcal{B}$. Since $p \in \bigcup \mathcal{A}$ and $p \in \bigcup \mathcal{B}$, $p \in (\bigcup \mathcal{A}) \bigcap (\bigcup \mathcal{B})$. *End of proof of $\Leftarrow$.* □

DEFINITION 14.9. *For all $\mathcal{T}$, by $\mathcal{T}$ is a* **topology**, *we mean:*

(1) *$\mathcal{T}$ is a set of sets* *and*
(2) $\langle \mathcal{T} \rangle_{\cup} \ = \ \mathcal{T} \ = \ \langle \mathcal{T} \rangle_{\mathrm{fin}\,\cap}$.

That is, a topology is a set of sets that is closed under arbitrary union and finite intersection.

We described, intuitively what open and closed subsets of $\mathbb{R}^2$ are. (More rigorous definition later.) We described why if the domain of a function $\mathbb{R}^2 \dashrightarrow \mathbb{R}$ is not an open set, then it's hard to study that function at boundary points using, say, directional derivatives. Consequently, it's natural for analysts to be interested in open sets, and to want to define what they are in a rigorous way.

We defined $\mathcal{T} := \{\text{open subsets of } \mathbb{R}^2\}$. We drew some pictures and made the case that this set $\mathcal{T}$ is closed under arbitrary union and finite intersection. We also showed that it is *NOT* under countably infinite intersection, because a closed disk is not open, but can be written as a countably infinite decreasing intersection of open disks.

Let $\mathcal{T}$ be a topology. In particular, $\mathcal{T}$ is a set of sets, so every element of $\mathcal{T}$ is a set. We have $\bigcup \mathcal{T} \in \langle \mathcal{T} \rangle_\cup = \mathcal{T}$. Then $\bigcup \mathcal{T}$ is an element of $\mathcal{T}$ that is a superset of all the elements of $\mathcal{T}$; it's "the biggest set in $\mathcal{T}$".

DEFINITION 14.10. *Let X be a set. Then, for all $\mathcal{T}$, by $\mathcal{T}$ is a* **topology on** X*, we mean*

(1) $\mathcal{T}$ *is a topology* *and*
(2) $\bigcup \mathcal{T} = X$.

DEFINITION 14.11. *By $(X, \mathcal{T})$ is a* **topological space***, we mean*

(1) X *is a set* *and*
(2) $\mathcal{T}$ *is a topology on* X.

DEFINITION 14.12. *Let $\mathcal{T}$ be a topology and let U be a set. By U is* $\mathcal{T}$**-open***, we mean: $U \in \mathcal{T}$.*

When the topology $\mathcal{T}$ is well-understood, we may simply say **open** instead of $\mathcal{T}$-open. Also, we often abbreviate a topological space $(X, \mathcal{T})$ as X. Also, we often study sets that have a "standard topology" that we all know about, but that, in most situations, goes unnamed. In such a situation, if we say that a subset is "open", we mean that it's an element of that anonymous topology.

For example, we will soon be defining the standard topology on $\mathbb{R}^2$. After that, when we say that a subset U of $\mathbb{R}^2$ is "open", we're saying "U is an element of the standard topology on $\mathbb{R}^2$", and the reader is expected to know how that standard topology is defined.

If I want to describe a topology, *e.g.* the standard topology on $\mathbb{R}^2$, it's impractical to try to list all of its elements; the typical topology isn't countable. We faced the same problem when we wanted to try to describe a subspace of a vector space. Our solution was: For me to describe to you a subspace, I should tell you a spanning set, or, even better, a basis for the subspace. A similar situation arises with topologies. The analogue for a spanning set is called a "base" for the topology:

DEFINITION 14.13. *Let $\mathcal{T}$ be a topology. Then, for all $\mathcal{B}$, by $\mathcal{B}$ is a* **base** *for $\mathcal{T}$, we mean:*

(1) $\mathcal{B} \subseteq \mathcal{T}$ *and*

(2) $\langle \mathcal{B} \rangle_\cup = \mathcal{T}$.

Some mathematicians use “basis” instead of “base”. Be aware that, whatever you call it, it’s more like a spanning set of a subspace than a basis of a subspace. There’s *NO* assumption that each set in $\mathcal{B}$ is in any way independent of the rest. In fact, in many bases, if you select any element of the base, and then remove it, that resulting set of sets is still a base for the same topology.

In our course, the vector spaces of interest will typically be finite dimensional, so their subspaces always have a finite spanning set. By contrast, most interesting topologies don’t have a finite base. However, many interesting topologies *DO* have a countable base. So, even if you can’t list all the elements of the topology (in an infinite list), you can list the elements of a base. So, if I’m interested in a topology, and if I want to describe it to you, if I can find a countable base for it, then I can show you the first few elements of the base, and hope that you can pick up the pattern, and figure out the rest. If so, then I’ve told you my topology of interest.

Every subset of a vector space is a spanning set for some subspace. By contrast, a set of sets may or may not be a base for a topology. So one problem that faces us in this situation (that didn’t come up in linear algebra) is to know when a certain collection of sets has a possibility to be a base for a topology. So, if we’re staring at a set of sets, how can we know if it’s a base for a topology? Here’s our answer:

PROPOSITION 14.14. *Let $\mathcal{S}$ be a set of sets, and assume, for all $X, Y \in \mathcal{S}$, that $X \cap Y \in \langle \mathcal{S} \rangle_\cup$. Then $\langle \mathcal{S} \rangle_\cup$ is a topology.*

By Definition 14.13, once we know $\langle \mathcal{S} \rangle_\cup$ is a topology, it follows that $\mathcal{S}$ is a base for that topology. So a buzz phrase to summarize Proposition 14.14 is: “If every pairwise-intersection is a union, then the set of sets is a base for some topology.”

We considered the case where $\mathcal{S} = \{\text{open disks in } \mathbb{R}^2\}$, and explained why, even though an intersection of two open disks is typically *NOT* a disk, it will always be a union of open disks. Then, by Proposition 14.14, that set $\mathcal{S}$ is a base for some topology, namely $\langle \mathcal{S} \rangle_\cup$. That topology is called the “standard topology” on $\mathbb{R}^2$. More on this later.

Proof. Let $\mathcal{T} := \langle \mathcal{S} \rangle_\cup$. We wish to show that $\mathcal{T}$ is a topology. By (2) of Remark 11.2 we have $\langle \langle \mathcal{S} \rangle_\cup \rangle_\cup = \langle \mathcal{S} \rangle_\cup$. Then

$$\langle \mathcal{T} \rangle_\cup \quad = \quad \langle \langle \mathcal{S} \rangle_\cup \rangle_\cup \quad = \quad \langle \mathcal{S} \rangle_\cup \quad = \quad \mathcal{T},$$

so $\langle \mathcal{T} \rangle_\cup = \mathcal{T}$. So, by Definition 14.9, it remains to show: $\langle \mathcal{T} \rangle_{\text{fin}\,\cap} = \mathcal{T}$. Then, by Proposition 14.6, it suffices to show: $\forall U, V \in \mathcal{T}$, $U \cap V \in \mathcal{T}$. Let $U, V \in \mathcal{T}$ be given. We wish to show: $U \cap V \in \mathcal{T}$.

We have $U \in \mathcal{T} = \langle \mathcal{S} \rangle_\cup$, so we choose $\mathcal{A} \subseteq \mathcal{S}$ such that $U = \bigcup \mathcal{A}$. We have $V \in \mathcal{T} = \langle \mathcal{S} \rangle_\cup$, so we choose $\mathcal{B} \subseteq \mathcal{S}$ such that $V = \bigcup \mathcal{B}$. Let $\mathcal{C} := \{X \cap Y \mid X \in \mathcal{A},\, Y \in \mathcal{B}\}$. By Lemma 14.8, $(\bigcup \mathcal{A}) \cap (\bigcup \mathcal{B}) = \bigcup \mathcal{C}$.

Claim: $\mathcal{C} \subseteq \mathcal{T}$. *Proof of claim:* We wish to show: $\forall Z \in \mathcal{C}$, $Z \in \mathcal{T}$. Let $Z \in \mathcal{C}$ be given. We wish to show: $Z \in \mathcal{T}$.

By definition of $\mathcal{C}$, since $Z \in \mathcal{C}$, choose $X \in \mathcal{A}$ and $Y \in \mathcal{B}$ such that $Z = X \cap Y$. We have $X \in \mathcal{A} \subseteq \mathcal{C}$ and $Y \in \mathcal{B} \subseteq \mathcal{C}$. As $X, Y \in \mathcal{C}$, by hypothesis, we see that $X \cap Y \in \langle \mathcal{S} \rangle_\cup$. Then $Z = X \cap Y \in \langle \mathcal{S} \rangle_\cup = \mathcal{T}$, as desired. *End of proof of claim.*

By the claim and by (3) of Remark 11.2, we have $\langle \mathcal{C} \rangle_\cup \subseteq \langle \mathcal{T} \rangle_\cup$. By (3) of Fact 14.1, $\bigcup \mathcal{C} \in \langle \mathcal{C} \rangle_\cup$. Recall that $\langle \mathcal{T} \rangle_\cup = \mathcal{T}$. Then we have $U \cap V = (\bigcup \mathcal{A}) \cap (\bigcup \mathcal{B}) = \bigcup \mathcal{C} \in \langle \mathcal{C} \rangle_\cup \subseteq \langle \mathcal{T} \rangle_\cup = \mathcal{T}$, as desired. $\square$

In the next class, we will focus on using Proposition 14.14 to show: $\forall$set X, $\forall d \in \mathcal{M}(X)$, $\langle \mathcal{B}_d \rangle_\cup$ is a topology on X. That is, we will construct a topology on any metric space, such that set of balls in the metric space is a base for that topology. From this, every metric space will have a standard topology. So, since $\mathbb{R}^2$ has a standard metric, we can use that metric to create a standard topology on $\mathbb{R}^2$: Specifically, that topology will be the union-closure of the set of all open disks in $\mathbb{R}^2$.

15. Class 15 on 26 October 2017, Th of Week 8

Midterm 2 will be on 2 November 2017, and will cover $\leqslant$ 17 October.

Recall: Let $\mathcal{T}$ be a set of sets. Then $\mathcal{T}$ is a **topology** means:

$$\langle \mathcal{T} \rangle_\cup \;=\; \mathcal{T} \;=\; \langle \mathcal{T} \rangle_{\text{fin}\,\cap}.$$

Recall: Let $\mathcal{T}$ be a set of sets and let X be a set. Then $\mathcal{T}$ is a **topology on** X means: ($\mathcal{T}$ is a topology) and ($\bigcup \mathcal{T} = X$).

Recall: Let $\mathcal{T}$ be a topology and let $\mathcal{B} \subseteq \mathcal{T}$. Then $\mathcal{B}$ is a **base** for $\mathcal{T}$ means: $\langle \mathcal{B} \rangle_\cup = \mathcal{T}$.

Recall: Let $\mathcal{S}$ be a set of sets. Then both of the following are true:

(1) $(\forall X, Y \in \mathcal{S},\ X \cap Y \in \mathcal{S}) \Rightarrow (\langle \mathcal{S} \rangle_{\text{fin}\,\cap} = \mathcal{S})$.
(2) $(\forall X, Y \in \mathcal{S},\ X \cap Y \in \langle \mathcal{S} \rangle_{\cup}) \Rightarrow (\langle \mathcal{S} \rangle_{\cup}$ is a topology).

The buzz phrase for (1) is "if a set of sets is pairwise-intersection closed, then it's finite-intersection closed." The buzz phrase for (2) is "if every pairwise-intersection is a union, then the collection of sets forms a base for some topology." Also, note that (2) is Proposition 14.14, proved in the last class.

Recall: Let (X, d) be a metric space, let $p \in X$ and let $U, V \in \mathcal{B}(p)$. Then both of the following are true:

(1) $(U \subseteq V)$ or $(V \subseteq U)$.
(2) $U \cap V \in \{U, V\}$.

Recall the **quantified equivalence for union-closure**: Let $\mathcal{B}$ be a set of sets and let S be a set. Then: $(\, S \in \langle \mathcal{B} \rangle_{\cup} \,)$ iff

$$\forall p \in S,\ \exists A \in \mathcal{B} \text{ s.t. } p \in A \subseteq S.$$

Recall the **Recentering Down Lemma** (Lemma 14.2): Let (X, d) be a metric space, let $B \in \mathcal{B}_d$ and let $p \in B$. Then $\exists B_0 \in \mathcal{B}(p)$ s.t. $B_0 \subseteq B$.

FACT 15.1. *Let (X, d) be a metric space, let $p \in X$ and let $B \in \mathcal{B}(p)$. Then $p \in B$.*

Proof. Omitted. □

Fact 15.1 asserts: Any ball in a metric space covers its center.

FACT 15.2. *Let B and C be sets. Let $B_0 \subseteq B$ and let $C_0 \subseteq C$. Then $B_0 \cap C_0 \subseteq B \cap C$.*

Proof. Omitted. □

The next lemma asserts that, in any metric space, any intersection of balls is a union of balls; consequently, by Proposition 14.14, the set of balls is a base for some topology.

LEMMA 15.3. *Let (X, d) be a metric space. Let $B, C \in \mathcal{B}_d$. Then $B \cap C \in \langle \mathcal{B}_d \rangle_{\cup}$.*

Proof. By quantified-equivalence for union-closure, we wish to prove: $\forall p \in B \cap C$, $\exists A \in \mathcal{B}_d$ s.t. $p \in A \subseteq B \cap C$. Let $p \in B \cap C$ be given. We wish to show: $\exists A \in \mathcal{B}_d$ s.t. $p \in A \subseteq B \cap C$.

By the Recentering Down Lemma (Lemma 14.2), choose $B_0 \in \mathcal{B}(p)$ such that $B_0 \subseteq B$, and choose $C_0 \in \mathcal{B}(p)$ such that $C_0 \subseteq C$. Then $B_0 \cap C_0 \in \{B_0, C_0\} \subseteq \mathcal{B}(p) \subseteq \mathcal{B}_d$. Let $A := B_0 \cap C_0$. Then $A \in \mathcal{B}_d$, and we wish to show: $p \in A \subseteq B \cap C$.

By Fact 15.1, since $B_0 \in \mathcal{B}(p)$, it follows that $p \in B_0$. By Fact 15.1, since $C_0 \in \mathcal{B}(p)$, it follows that $p \in C_0$. Since $p \in B_0$ and $p \in C_0$, we get $p \in B_0 \cap C_0$. Then $p \in B_0 \cap C_0 = A$. It remains to show: $A \subseteq B \cap C$.

By Fact 15.2, $B_0 \cap C_0 \subseteq B \cap C$. Then $A = B_0 \cap C_0 \subseteq B \cap C$. $\square$

DEFINITION 15.4. *Let (X, d) be a metric space. Then we define $\mathcal{T}_d := \langle \mathcal{B}_d \rangle_\cup$.*

FACT 15.5. *Let (X, d) be a metric space. Then $\bigcup \mathcal{B}_d = X$.*

COROLLARY 15.6. *Let (X, d) be a metric space. Then $\bigcup \mathcal{T}_d = X$.*

COROLLARY 15.7. *Let (X, d) be a metric space. Then $\mathcal{T}_d$ is a topology on X.*

DEFINITION 15.8. *Let (X, d) be a metric space. Then $\mathcal{T}_d$ is called the* **standard topology** *on (X, d).*

We described $(V, |\bullet|) \mapsto (V, d_{|\bullet|})$ as a "functor" from the "category" of normed vector spaces to the "category" of metric spaces.

We described $(X, d) \mapsto (X, \mathcal{T}_d)$ as a "functor" from the "category" of metric spaces to the "category" of topological spaces.

DEFINITION 15.9. *For all $k \in \mathbb{N}$, we define*

$$\begin{aligned} |\bullet|_k &:= \text{ the standard norm on } \mathbb{R}^k, \\ d_k &:= d_{|\bullet|_k}, \\ \mathcal{B}_k &:= \mathcal{B}_{d_k} \qquad \text{and} \\ \mathcal{T}_k &:= \mathcal{T}_{d_k}. \end{aligned}$$

Let $k \in \mathbb{N}$. Then, for all $p \in \mathbb{R}^k$, we have

$$|p|_k \quad = \quad \sqrt{p_1^2 + \cdots + p_k^2}.$$

Also, for all $p, q \in \mathbb{R}^k$, we have

$$d_k(p, q) \quad = \quad \sqrt{(p_1 - q_1)^2 + \cdots + (p_k - q_k)^2}.$$

Also, for all $p \in \mathbb{R}^k$, for all $r > 0$, we have

$$B(p, r) \quad = \quad \{q \in \mathbb{R}^k \mid d_k(p, q) < r\}.$$

Also, $\mathcal{B}_k = \{B(p,r)\,|\,p \in \mathbb{R}^k, r > 0\}$. Also, $\mathcal{T}_k = \langle \mathcal{B}_k \rangle_\cup$.

We described some open and non-open subsets of $\mathbb{R}^2$.

DEFINITION 15.10. *Let $|\bullet|$ denote absolute value on $\mathbb{R}$. Then we define $\mathcal{B}_\mathbb{R} := \mathcal{B}_{d_{|\bullet|}}$.*

We drew a few pictures and concluded that $\mathcal{B}_\mathbb{R}$ is the set of bounded open intervals in $\mathbb{R}$. We also noticed that, for all $U, V \in \mathcal{B}_\mathbb{R}$, we have

$$U \cap V \quad \in \quad \mathcal{B}_\mathbb{R} \cup \{\varnothing\} \quad \subseteq \quad \langle \mathcal{B}_\mathbb{R} \rangle_\cup.$$

DEFINITION 15.11. *We define*

$$\begin{aligned} \mathcal{B}_\infty \quad &:= \quad \{\ (M, \infty]\ |\ M \in \mathbb{R}\ \}, \\ \mathcal{B}_{-\infty} \quad &:= \quad \{\ [-\infty, N)\ |\ N \in \mathbb{R}\ \} \qquad \textit{and} \\ \mathcal{B}_* \quad &:= \quad \mathcal{B}_{-\infty} \bigcup \mathcal{B}_\mathbb{R} \bigcup \mathcal{B}_\infty. \end{aligned}$$

We drew a few pictures, and noticed that, for all $U, V \in \mathcal{B}_*$, we have

$$U \cap V \quad \in \quad \mathcal{B}_* \cup \{\varnothing\} \quad \subseteq \quad \langle \mathcal{B}_* \rangle_\cup.$$

So, in $\mathcal{B}_*$, every pairwise-intersection is a union. It follows that $\mathcal{B}_*$ is a base for some topology.

DEFINITION 15.12. *We define $\mathcal{T}_* := \langle \mathcal{B}_* \rangle_\cup$.*

FACT 15.13. *We have: $\bigcup \mathcal{B}_* = \mathbb{R}^*$.*

COROLLARY 15.14. *We have: $\bigcup \mathcal{T}_* = \mathbb{R}^*$.*

COROLLARY 15.15. *We have: $\mathcal{T}_*$ is a topology on $\mathbb{R}^*$.*

DEFINITION 15.16. *We call $\mathcal{T}_*$ the* **standard topology** *on $\mathbb{R}^*$.*

DEFINITION 15.17. *Let $(X, \mathcal{T})$ be a topological space. We say that $(X, \mathcal{T})$ is* **metrizable** *if there exists $d \in \mathcal{M}(X)$ such that $\mathcal{T} = \mathcal{T}_d$.*

Let $\mathbb{R}^*$ be an abbreviation for $(\mathbb{R}^*, \mathcal{T}_*)$. We discussed the question: Is $\mathbb{R}^*$ metrizable? The answer is yes, but, unfortunately, there's no "standard metric" on $\mathbb{R}^*$, so it's really not so helpful to think of $\mathbb{R}^*$ as a metric space. After all, what element of $[0, \infty)$ should represent the distance from $-\infty$ to ∞? There's no obvious choice. However, there is a standard *TOPOLOGY* on $\mathbb{R}^*$, namely $\mathcal{T}_*$.

A reasonable question is: Is every topological space metrizable? In other words, is the functor $(X, d) \mapsto (X, \mathcal{T}_d) : \{\text{TSs}\} \to \{\text{MSs}\}$ surjective? The answer turns out to be no, and we next explain why.

DEFINITION 15.18. *Let X be a topological space. By X is* **Hausdorff**, *we mean: $\forall p, q \in X$, if $p \neq q$, then*

$$\exists \textit{open sets } U, V \textit{ in } X \quad \textit{s.t.} \quad [\,(p \in U)\,\&\,(q \in V)\,\&\,(U \cap V = \varnothing)\,].$$

We explained why metrizable topological spaces are Hausdorff.

DEFINITION 15.19. *Let X be a set. Then $\{\varnothing, X\}$ is a topology on X, and is called the* **indiscrete topology** *on X.*

DEFINITION 15.20. *Let $(X, \mathcal{T})$ be a topological space. We say that $(X, \mathcal{T})$ is* **indiscrete** *if $\mathcal{T}$ is the indiscrete topology on X.*

We briefly expained why, if X is a set and if $\#X \geqslant 2$, then the topological space $(X, \{\varnothing, X\})$ is *NOT* Hausdorff, and is therefore not metrizable. So not every topological space is metrizable.

In this course, we will take the point of view that a every interesting topological space is metrizable. (A value judgment!) Consequently, any non-Hausdorff topological space is uninteresting.

DEFINITION 15.21. *Let $\mathcal{A}$ and $\mathcal{B}$ be sets of sets. By ($\mathcal{A}$ is* **finer** *than $\mathcal{B}$) or by ($\mathcal{B}$ is* **coarser** *than $\mathcal{A}$), we mean: $\mathcal{A} \supseteq \mathcal{B}$.*

Let X be a set. Recall that 2^X denotes the set of all subsets of X. Among the topologies on X, the coarsest is $\{\varnothing, X\}$. There is also a finest topology on X, namely 2^X.

DEFINITION 15.22. *Let X be a set. Then 2^X is a topology on X, and is called the* **discrete topology** *on X.*

DEFINITION 15.23. *Let $(X, \mathcal{T})$ be a topological space. We say that $(X, \mathcal{T})$ is* **discrete** *if $\mathcal{T}$ is the discrete topology on X.*

So, on $\mathbb{R}^2$, we now have introduced three topologies: indiscrete, discrete and $\mathcal{T}_2$. The indiscrete topology on $\mathbb{R}^2$ is uninteresting because it's not even Hausdorff (and hence not metrizable). It turns out that the discrete topology on $\mathbb{R}^2$ is too fine to be useful, so it's also uninteresting. (More on this later.) Following Goldilocks, $\mathcal{T}_2$ is "just right".

The same situation arises on $\mathbb{R}^*$. The indiscrete topology on $\mathbb{R}^*$ is too coarse to be useful, the discrete topology on $\mathbb{R}^*$ is too fine to be useful, and the standard topology $\mathcal{T}_*$ is just right.

Let $k \in \mathbb{N}$. The same situation arises on $\mathbb{R}^k$. The indiscrete topology on $\mathbb{R}^k$ is too coarse to be useful, the discrete topology on $\mathbb{R}^k$ is too fine to be useful, and $\mathcal{T}_k$ is just right.

Recall: For any topological space $(X, \mathcal{T})$, for any $U \subseteq X$, by U is **open** or **$\mathcal{T}$-open**, we mean: $U \in \mathcal{T}$.

DEFINITION 15.24. *Let $(X, \mathcal{T})$ be a topological space. Let $C \subseteq X$. By C is* **closed** *or* **$\mathcal{T}$-closed***, we mean: $X \backslash C \in \mathcal{T}$.*

That is, a set is closed iff its complement is open.

We studied various sets in $\mathbb{R}^2$. Intuitively, a set in $\mathbb{R}^2$ is $\mathcal{T}_2$-open iff it contains *NONE* of its boundary points. Intuitively, a set in $\mathbb{R}^2$ is $\mathcal{T}_2$-closed iff it contains *ALL* of its boundary points. We verified that many sets in $\mathbb{R}^2$ are *neither* $\mathcal{T}_2$-open *nor* $\mathcal{T}_2$-closed.

So be aware: open is *NOT* the opposite of closed. So, for example, if you are asked to show that some set is closed, it is *NOT* enough to show that it's not open.

In any topological space X, both $\varnothing$ and X are "**clopen**" meaning "both closed and open". So every topological space has at least two clopen sets. We'll eventually see that some topological spaces have more than two, but, first, it will be very helpful to define and study

(1) relative norms,
(2) relative metrics and
(3) relative topologies.

DEFINITION 15.25. *Let $(V, |\bullet|)$ be a normed vector space and let S be a subspace of V. Then $|\bullet|_S := (|\bullet|)|S \in \mathcal{N}(S)$, and $|\bullet|_S$ is called the* **relative norm** *on S inherited from $(V, |\bullet|)$.*

DEFINITION 15.26. *Let (X, d) be a metric space and let $S \subseteq X$. Then $d_S := d|(S \times S) \in \mathcal{M}(S)$, and d_S is called the* **relative metric** *on S inherited from (X, d).*

DEFINITION 15.27. *Let $\mathcal{S}$ be a set of sets and let A be a set. Then the* **restriction** *of $\mathcal{S}$ to A is $\{X \cap A \,|\, X \in \mathcal{S}\}$, and is denoted either by $\mathcal{S}_A$ or by $\mathcal{S}|A$.*

FACT 15.28. *Let $\mathcal{S}$ be a set of sets and let A be a set. Then we have: $\langle \mathcal{S}|A \rangle_\cup = (\langle \mathcal{S} \rangle_\cup)|A$.*

Fact 15.28 is sometimes expressed by saying that "restriction and union-closure commute".

DEFINITION 15.29. *Let $(X, \mathcal{T})$ be a topological space. Let $A \subseteq X$. Then $\mathcal{T}_A$ is a topology on A, and $\mathcal{T}_A$ is called the* **relative topology** *on A inherited from $(X, \mathcal{T})$.*

We looked at the unit circle in $\mathbb{R}^2$. We discussed

- its relative metric inherited from $(\mathbb{R}^2, d_2)$ and
- its relative topology inherited from $(\mathbb{R}^2, \mathcal{T}_2)$.

NOTE: That relative metric is sometimes called the "chord metric", for which the distance between two points on the circle is the length of the chord connecting them. There is also the "Riemannian metric", and, for that metric, the distance between two points is the length of the *arc* connecting them. We will not study Riemannian metrics.

We looked at the unit sphere in $\mathbb{R}^3$. We discussed

- its relative metric inherited from $(\mathbb{R}^3, d_3)$ and
- its relative topology inherited from $(\mathbb{R}^3, \mathcal{T}_3)$.

NOTE: That relative metric is sometimes called the "chord metric", for which the distance between two points on the sphere is the length of the chord connecting them. There is also the "Riemannian metric", and, for that metric, the distance between two points is the length of the *arc* connecting them. We will not study Riemannian metrics.

Fix a moment in time. Let T be the function that, to any point on the surface of the earth, associates the temperature at that point at that moment. We think of the surface of the earth as a sphere in $\mathbb{R}^3$. Then T is a function whose domain is that sphere. To analyze T using techniques from calculus, we want to "do calculus" on that sphere. Even to talk about limits (much less derivatives and integrals), we will need to specify a topology on that sphere. There are many topologies, some too coarse to be useful, some too fine. Experience has shown: the "just right" topology is the relative topology inherited from $(\mathbb{R}^3, \mathcal{T}_3)$.

16. Class 16 on 31 October 2017, Tu of Week 9

Assigned HW#41.

DEFINITION 16.1. *Let X be a topological space, $q \in X$, $V \subseteq X$. By V is a* **neighborhood** *of q in X, we mean: $\exists U$ s.t. U is open in X and s.t. $q \in U \subseteq V$.*

We use **nbd** as an abbreviation for "neighborhood".

We gave some examples of neighborhoods of a point in $\mathbb{R}^2$.

DEFINITION 16.2. *For any topological space X, for any $q \in X$, we define $\mathcal{N}_X(q) := \{V \subseteq X \mid V$ is a neighborhood of q in $X\}$.*

When X is clear, we sometimes write $\mathcal{N}(q)$ for $\mathcal{N}_X(q)$.

REMARK 16.3. *Let X be a topological space, $q \in X$. Then*

(1) $\forall V \in \mathcal{N}_X(q)$, $q \in V$,
(2) $\forall V \in \mathcal{N}_X(q)$, $\forall W \subseteq X$, $\;[\,(W \supseteq V) \Rightarrow (W \in \mathcal{N}_X(q))\,]$ *and*
(3) $\forall V, W \in \mathcal{N}_X(q)$, $\;V \cap W \in \mathcal{N}_X(q)$.

Proof. Omitted. □

We next prove that an open set is a neighborhood of any of its points:

REMARK 16.4. *Let X be a topological space, let V be an open subset of X and let $q \in V$. Then $V \in \mathcal{N}_X(q)$.*

Proof. We wish to show that $\exists$open U in X such that $q \in U \subseteq V$.

Let $U := V$. By assumption, V is open in X, so U is open in X. We wish to show: $q \in U \subseteq V$. By assumption, $q \in V$. Then $q \in V = U$. It remains to show: $U \subseteq V$. We have $U = V \subseteq V$, as desired. □

COROLLARY 16.5. *Let X be a topological space and let $W \subseteq X$. Assume that W is open in X. Then, $\forall q \in W$, $\exists V \in \mathcal{N}_X(q)$ s.t. $V \subseteq W$.*

Proof. Given $q \in W$. We wish to show: $\exists V \in \mathcal{N}_X(q)$ s.t. $V \subseteq W$. Let $V := W$. Then $V = W \subseteq W$. It remains only to show that $V \in \mathcal{N}_X(q)$.

Since W is open in X and $q \in W$, it follows, from Remark 16.4, that $W \in \mathcal{N}_X(q)$. Then $V = W \in \mathcal{N}_X(q)$, as desired. □

Assigned HW#42 and HW#43.

Recall that, by Proposition 14.14, for any set $\mathcal{R}$ of sets, if

$$\text{every pairwise-}\bigcap \text{ is a } \bigcup,\ i.e.,\ \forall W, W' \in \mathcal{R},\ W \cap W' \in \langle \mathcal{R} \rangle_\cup,$$

then

$$\mathcal{R} \text{ is a base for a topology, } i.e.,\ \langle \mathcal{R} \rangle_\cup \text{ is a topology.}$$

FACT 16.6. *Let $\mathcal{S}$ and $\mathcal{T}$ be topologies. Let*

$$\mathcal{R} \quad := \quad \{\, U \times V \mid U \in \mathcal{S},\, V \in \mathcal{T} \,\}.$$

Then $\langle \mathcal{R} \rangle_\cup$ is a topology.

Proof. By Proposition 14.14, we wish to show, for all $W, W' \in \mathcal{R}$, that $W \cap W' \in \langle \mathcal{R} \rangle_\cup$. Given $W, W' \in \mathcal{R}$. We wish to show: $W \cap W' \in \langle \mathcal{R} \rangle_\cup$.

By definition of $\langle \bullet \rangle_{\mathrm{fin}\,\cap}$, we have $W \cap W' \in \langle \mathcal{R} \rangle_{\mathrm{fin}\,\cap}$. By HW#43, we get $\langle \mathcal{R} \rangle_{\mathrm{fin}\,\cap} = \mathcal{R}$. By the superset property of $\langle \bullet \rangle_\cup$, we get $\mathcal{R} \subseteq \langle \mathcal{R} \rangle_\cup$. Then $W \cap W' \in \langle \mathcal{R} \rangle_{\mathrm{fin}\,\cap} = \mathcal{R} \subseteq \langle \mathcal{R} \rangle_\cup$, as desired. □

If $(X,\mathcal{S})$ and $(Y,\mathcal{T})$ are topological spaces, then an **open rectangle** in $X \times Y$ is a subset $U \times V$ of $X \times Y$ such that $U \in \mathcal{S}$ and $V \in \mathcal{T}$. That is, an open rectangle is the Cartesian product of an open set in X by an open set in Y.

HW#43 essentially says that, for any two topological spaces X and Y, the set of open rectangles in $X \times Y$ is closed under finite intersection. It's easy to see that, if a collection of sets is closed under finite intersection, then any pairwise-intersection is a union, and so that collection of sets is a base for some topology. Putting this together, we get Fact 16.6, which asserts that for any two topological spaces X and Y, the set of open rectangles forms a base for a topology on $X \times Y$; that topology is called the **product topology**.

Under the standard bijection $\mathbb{R} \times \mathbb{R} \hookrightarrow\!\!> \mathbb{R}^2$, the standard metric on $\mathbb{R}^2$ corresponds to a metric d on $\mathbb{R} \times \mathbb{R}$. Then $\mathcal{T}_d = \langle \mathcal{B}_d \rangle_\cup$ is the topology whose base is the set $\mathcal{B}_d$ of all open d-balls. We explained why the product topology on $\mathbb{R} \times \mathbb{R}$ is equal to $\mathcal{T}_d$.

Recall: For any normed vector space $(V, |\bullet|)$, for any subspace S of V, we defined $|\bullet|_S := (|\bullet|)|S \in \mathcal{N}(S)$. Recall: For any metric space (M,d), for any $S \subseteq M$, we defined $d_S := d|(S \times S) \in \mathcal{M}(S)$. Recall: For any set $\mathcal{V}$ of sets, for any set S, we defined $\mathcal{V}|S := \{U \cap S \,|\, U \in \mathcal{V}\}$ and $\mathcal{V}_S := \mathcal{V}|S$. Recall: For any topological space $(X,\mathcal{T})$, for any $S \subseteq X$, that $\mathcal{T}_S$ is a topology on X.

We have the following "transitivity" result:

PROPOSITION 16.7. *All of the following are true:*

(1) $\forall$*normed vector space* $(V, |\bullet|)$, $\forall$*subspace* S *of* V, $\forall$*subspace* R *of* S, *we have:* $(|\bullet|_S)_R = |\bullet|_R$.
(2) $\forall$*metric space* (M,d), $\forall S \subseteq M$, $\forall R \subseteq S$, *we have:* $(d_S)_R = d_R$.
(3) $\forall$*top. space* $(X,\mathcal{T})$, $\forall S \subseteq X$, $\forall R \subseteq S$, *we have:* $(\mathcal{T}_S)_R = \mathcal{T}_R$.

Proof. Unassigned HW. □

We can generalize (3) of Proposition 16.7: $\forall$set $\mathcal{V}$ of sets, $\forall$set S, $\forall R \subseteq S$, we have: $(\mathcal{V}_S)_R = \mathcal{V}_R$. The proof is again unassigned HW.

Let $(V, |\bullet|)$ be a normed vector space and let S be a subspace of V. Then there are *TWO* possible metrics on S: First, we can restrict the norm, and then form the metric of the restricted norm. This yields $d_{|\bullet|_S}$. Second, we can form the metric of the orignal norm, and then restrict that metric. This yields $(d_{|\bullet|})_S$. These two agree:

PROPOSITION 16.8. *For any normed vector space $(V, |\bullet|)$, for any subspace S of V, we have: $d_{|\bullet|_S} = (d_{|\bullet|})_S$.*

Proof. Unassigned HW. □

Let (X, d) be a metric space and let $S \subseteq X$. Then there are *TWO* possible topologies on S: First, we can restrict the metric, and then form the topology of the restricted metric. This yields $\mathcal{T}_{d_S}$. Second, we can form the topology of the orignal metric, and then restrict that topology. This yields $(\mathcal{T}_d)_S$. These two agree:

PROPOSITION 16.9. *For any metric space (X, d), for any $S \subseteq X$, we have: $\mathcal{T}_{d_S} = (\mathcal{T}_d)_S$.*

Proof. Unassigned HW. □

Recall, for all $k \in \mathbb{N}$, that $| \bullet |_k$ denotes the standard norm on $\mathbb{R}^k$, that $d_k = d_{|\bullet|_k}$ denotes the standard metric in $\mathbb{R}^k$ and that $\mathcal{T}_k = B_{d_k}$ denotes the standard topology on $\mathbb{R}^k$.

DEFINITION 16.10. *Let $k \in \mathbb{N}$. Then:*

(1) *$\forall$subspace S of $\mathbb{R}^k$, we define $| \bullet |_S := (| \bullet |_k)_S$*
(2) *$\forall S \subseteq \mathbb{R}^k$, we define $d_S := (d_k)_S$.*
(3) *$\forall S \subseteq \mathbb{R}^k$, we define $\mathcal{T}_S := (\mathcal{T}_k)_S$.*

Next, we turn to one-dimensional topology. That is, we look to $\mathbb{R}$ and $\mathbb{R}^*$, and discuss relative structures on subsets. Recall that $\mathbb{R}$ has a standard vector space structure, where vector addition is addition of real numbers and scalr multiplication is multiplication of real numbers. On the other hand, $\mathbb{R}^*$ isn't, in any natural way, a vector space, so we can't really discuss norms on $\mathbb{R}^*$. We do have a standard norm on $\mathbb{R}$, namely absolute value. However, there are only two subspaces of the vector space $\mathbb{R}$, namely $\{0\}$ and $\mathbb{R}$, and these aren't interesting enough to make it worthwhile to restrict the absolute value norm. So we move on to metrics. We'll talk about the difficulties in metrizing $\mathbb{R}^*$ later, but, for now, let's focus $\mathbb{R}$ and its subsets:

DEFINITION 16.11. *Let $| \bullet | \in \mathcal{N}(\mathbb{R})$ denote absolute value. For all $S \subseteq \mathbb{R}$, we define $d_S := (d_{|\bullet|})_S$.*

Let $| \bullet | \in \mathcal{N}(\mathbb{R})$ continue to denote absolute value. Then $| \bullet |_{\mathbb{R}} = | \bullet |$, and so, by Definition 16.11, we get $d_{\mathbb{R}} = d_{|\bullet|}$. Then, for all $a, b \in \mathbb{R}$, we have $d_{\mathbb{R}}(a, b) = |a - b|$.

Next, let's recall the standard topology $\mathcal{T}_*$ on $\mathbb{R}^*$. Let $|\bullet| \in \mathcal{N}(\mathbb{R})$ continue to denote absolute value. We defined

$$\begin{aligned} \mathcal{B}_{\mathbb{R}} &:= \mathcal{B}_{d_{|\bullet|}} = \mathcal{B}_{d_{\mathbb{R}}}, \\ \mathcal{B}_{\infty} &:= \{ (M,\infty] \mid M \in \mathbb{R} \}, \\ \mathcal{B}_{-\infty} &:= \{ [-\infty, N) \mid N \in \mathbb{R} \}, \\ \mathcal{B}_* &:= \mathcal{B}_{-\infty} \bigcup \mathcal{B}_{\mathbb{R}} \bigcup \mathcal{B}_{\infty} \qquad \text{and} \\ \mathcal{T}_* &:= \langle \mathcal{B}_* \rangle_\cup. \end{aligned}$$

Recall that, for any metric space (X, c), we defined $\mathcal{T}_c := \langle \mathcal{B}_c \rangle_\cup$; it is the union-closure of the set of all balls in the c metric.

Let $|\bullet| \in \mathcal{N}(\mathbb{R})$ continue to denote absolute value; then $d_{\mathbb{R}} = d_{|\bullet|}$. For any metric c on $\mathbb{R}^*$, recall that $\mathcal{T}_c = \langle \mathcal{B}_c \rangle_\cup$ denotes the topology on $\mathbb{R}^*$ generated by the set of balls in the c metric. For any metric c on $\mathbb{R}^*$, recall that $c_{\mathbb{R}} = c|(\mathbb{R} \times \mathbb{R})$ denotes the restriction of c to a relative metric on $\mathbb{R}$. It turns out that there exist a metric c on $\mathbb{R}^*$ such that $\mathcal{T}_c = \mathcal{T}_*$. In fact, there are many such metrics, and none of them is, in any way, a natural choice. There also exists a metric c on $\mathbb{R}^*$ such that $c_{\mathbb{R}} = d_{\mathbb{R}}$. Once again, there are many such metrics, and none of them is, in any way, a natural choice. It is not too hard to prove that there does *NOT* exist a metric c on $\mathbb{R}^*$ such that both $\mathcal{T}_c = \mathcal{T}_*$ and $c_{\mathbb{R}} = d_{\mathbb{R}}$. From these remarks, we take the point of view that, while $(\mathbb{R}^*, \mathcal{T}_*)$ is metrizable, there is no *standard* metric on $\mathbb{R}^*$.

We now have *TWO* different natural ways to get a topology on $\mathbb{R}$. First, we can take the topology $\mathcal{T}_*$, and restrict it to $\mathbb{R}$. This yields $(\mathcal{T}_*)_{\mathbb{R}}$. Second, we can take the metric $d_{\mathbb{R}}$, and use it to create a topology. This yields $\mathcal{T}_{d_{\mathbb{R}}}$. These two agree:

PROPOSITION 16.12. *We have:* $(\mathcal{T}_*)_{\mathbb{R}} = \mathcal{T}_{d_{\mathbb{R}}}$.

Proof. Unassigned HW. □

DEFINITION 16.13. *For any subset* $S \subseteq \mathbb{R}^*$*, we define* $\mathcal{T}_S := (\mathcal{T}_*)_S$.

According to Definition 16.13, we have $\mathcal{T}_{\mathbb{R}} = (\mathcal{T}_*)_{\mathbb{R}}$. So, by Proposition 16.12, we have $\mathcal{T}_{\mathbb{R}} = (\mathcal{T}_*)_{\mathbb{R}} = \mathcal{T}_{d_{\mathbb{R}}}$.

Let $(X, \mathcal{T})$ be a topological space, let $S \subseteq X$ and let $U \subseteq S$. If we were to say that U is "open", there would be some amiguity: We might mean that U is an element of the topology $\mathcal{T}$ on X, *OR* we might mean that U is an element of the relative topology $\mathcal{T}|S$ on S. We typically distinguish these cases as follows: In the former case, we would say "U is open in X". In the latter case, we would say "U is open in S".

REMARK 16.14. *Let X be a topological space, $S \subseteq X$. Then*

$$\{\textit{open subsets of } X\} \mid S \qquad = \qquad \{\textit{open subsets of } S\}.$$

Proof. This is the definition of the relative topology. $\square$

COROLLARY 16.15. *Let X be a topological space, $S \subseteq X$. Then*

(1) $\forall$*open V in X, $\quad V \cap S$ is open in S $\qquad$ and*
(2) $\forall$*open U in S, $\exists$open V in X s.t. $U = V \cap S$.*

Proof. We leave it as unassigned HW to show that $\subseteq$ in Remark 16.14 implies (1), and that $\supseteq$ in Remark 16.14 implies (2). $\square$

We call (1) of Corollary 16.15 the "restriction result for open sets". We call (2) of Corollary 16.15 the "extension result for open sets".

Recall that a subset C of a topological space X is said to be "closed" if $X \backslash C$ is closed.

Let X be a topological space, let $S \subseteq X$ and let $C \subseteq S$. Then, by "C is closed in X", we mean: $X \backslash C$ is open in X. Also, by "C is closed in S", we mean: $S \backslash C$ is open in S.

REMARK 16.16. *Let X be a topological space, $S \subseteq X$. Then*

$$\{\textit{closed subsets of } X\} \mid S \qquad = \qquad \{\textit{closed subsets of } S\}.$$

Proof. Unassigned HW. $\square$

COROLLARY 16.17. *Let X be a topological space, $S \subseteq X$. Then*

(1) $\forall$*closed D in X, $\quad D \cap S$ is closed in S $\qquad$ and*
(2) $\forall$*closed C in S, $\exists$closed D in X s.t. $C = D \cap S$.*

Proof. We leave it as unassigned HW to show that $\subseteq$ in Remark 16.16 implies (1), and that $\supseteq$ in Remark 16.16 implies (2). $\square$

We call (1) of Corollary 16.17 the "restriction result for closed sets". We call (2) of Corollary 16.17 the "extension result for closed sets".

Let $(X, \mathcal{T})$ be a topological space, let $S \subseteq X$ and let $q \in S$. Then $\mathcal{N}_X(q)$ denotes the set of all neighborhoods of q in $(X, \mathcal{T})$, while $\mathcal{N}_S(q)$ denotes the set of all neighborhoods of q in $(S, \mathcal{T}|S)$.

REMARK 16.18. *Let X be a topological space, $S \subseteq X$, $q \in S$. Then*

$$\mathcal{N}_X(q) \mid S \qquad = \qquad \mathcal{N}_S(q).$$

Proof. Unassigned HW. $\square$

COROLLARY 16.19. *Let X be a top. space, $S \subseteq X$, $q \in S$. Then*

(1) $\forall V \in \mathcal{N}_X(q), \quad V \cap S \in \mathcal{N}_S(q)$ *and*
(2) $\forall U \in \mathcal{N}_S(q)$, $\exists V \in \mathcal{N}_X(q)$ *s.t.* $U = V \cap S$.

Proof. We leave it as unassigned HW to show that $\subseteq$ in Remark 16.18 implies (1), and that $\supseteq$ in Remark 16.18 implies (2). $\square$

We call (1) of Corollary 16.19 the "restriction result for nbds". We call (2) of Corollary 16.19 the "extension result for nbds".

Let $S := [1,2] \cup [3,4) \cup \{5\}$. We discussed why $[1,2]$, $[3,4)$ and $\{5\}$ are all clopen in S. This motivated the following two definitions.

DEFINITION 16.20. *Let S be a topological space. Then, by S is* **connected***, we mean: $\forall$clopen A in S, [$(A = \emptyset)$ or $(A = S)$].*

For example, $\mathbb{R}$ is connected. Also, for all $k \in \mathbb{N}$, $\mathbb{R}^k$ is connected. However, $[1,2] \cup [3,4) \cup \{5\}$ is not connected.

DEFINITION 16.21. *Let S be a topological space, $p \in S$. Then, by p is* **isolated** *in S, we mean: $\{p\}$ is open in S.*

For example, 5 is isolated in $[1,2] \cup [3,4) \cup \{5\}$. So, in the topological spaces that we study in this course, it can happen that a singleton set is open, and, in this case, its point is an isolated point.

Next we consider whether singleton sets may not be closed.

DEFINITION 16.22. *Let X be a topological space. Then, by X is* $\mathbf{T^1}$*, we mean: $\quad \forall p \in X$, $\{p\}$ is closed.*

When using the indiscrete topology on a set with two or more points, a singleton set is neither open nor closed, so such a topological space is *NOT* T^1. However, the indiscrete topology is typically too coarse to be of interest to us.

Typically, the topological spaces we consider are all metrizable. We explained earlier why metrizable implies Haudsdorff. The next result states that Hausdorff implies T^1. Consequently, the topological spaces of interest to us are always T^1. So, in our topological spaces, singleton sets are closed. So, throughout this course, when we encounter an isolated point, its singleton set isn't just open; it is clopen.

FACT 16.23. *Let X be a Hausdorff topological space. Then X is T^1.*

Proof. We wish to prove: $\forall p \in X$, $\{p\}$ is closed in X. Let $p \in X$ be given. We wish to prove: $\{p\}$ is closed in X. That is, we wish to prove $X \backslash \{p\}$ is open in X. Let $W := X \backslash \{p\}$. We wish to prove W is open in X. By HW#42, it suffices to show: $\forall q \in W$, $\exists V \in \mathcal{N}_X(q)$ s.t. $V \subseteq W$. Let $q \in W$ be given. We wish to show: $\exists V \in \mathcal{N}_X(q)$ s.t. $V \subseteq W$.

Since $q \in W = X \backslash \{p\}$, it follows that $p \neq q$. So, since X is Hausdorff, by Definition 15.18, choose open set U, V in X s.t. $p \in U$ and $q \in V$ and $U \cap V = \varnothing$. By Remark 16.4, $V \in \mathcal{N}_X(q)$. We wish to show: $V \subseteq W$.

Since $p \in U$, we get $X \backslash \{p\} \supseteq X \backslash U$. Since $V \subseteq X$ and $U \cap V = \varnothing$, it follows that $V \subseteq X \backslash U$. Then $V \subseteq X \backslash U \subseteq X \backslash \{p\} = W$, as desired. $\square$

17. Class 17 on 7 November 2017, Tu of Week 10

Assigned HW#46 and HW#47.

DEFINITION 17.1. *For any set Z, for any $\mathcal{S} \subseteq 2^Z$, we define*

$$\bigcap{}^Z \mathcal{S} := \begin{cases} \bigcap \mathcal{S}, & \text{if } \mathcal{S} \neq \varnothing \\ Z, & \text{if } \mathcal{S} = \varnothing. \end{cases}$$

FACT 17.2. *Let Z be a set and let $\mathcal{S} \subseteq 2^Z$. For all $A \subseteq Z$, define $A' := Z \backslash A$. Define $\mathcal{S}' := \{A' \mid A \in \mathcal{S}\}$. Then*

(1) $\left(\bigcap^Z \mathcal{S}\right)' = \bigcup(\mathcal{S}')$ *and*

(2) $\left(\bigcup \mathcal{S}\right)' = \bigcap^Z(\mathcal{S}')$.

Proof. Unassigned HW. $\square$

DEFINITION 17.3. *Let Z be a set and let $\mathcal{S} \subseteq 2^Z$. Then we define*

$$\begin{aligned} \langle \mathcal{S} \rangle_\cap^Z &:= \langle \mathcal{S} \rangle_\cap \cup \{Z\}, \qquad \text{and} \\ \langle \mathcal{S} \rangle_{\text{fin}\,\cap}^Z &:= \langle \mathcal{S} \rangle_{\text{fin}\,\cap} \cup \{Z\}. \end{aligned}$$

DEFINITION 17.4. *Let $\mathcal{S}$ be a set of sets. Then we define*

$$\langle \mathcal{S} \rangle_{\text{fin}\,\cup} := \left\{ \bigcup \mathcal{A} \mid (\varnothing \neq \mathcal{A} \subseteq \mathcal{S}) \text{ and } (\#\mathcal{A} < \infty) \right\}.$$

Assigned HW#48.

A topological space is discrete iff all of its points are isolated:

THEOREM 17.5. *Let Z be a topological space. Then*

$$[\ Z \text{ is discrete}\ ] \quad \Leftrightarrow \quad [\ \forall p \in Z,\ \{p\} \text{ is open in } Z\ ].$$

Proof. Let $\mathcal{T} := \{$open subsets of $Z\}$. We wish to show:

$$[\ \mathcal{T} = 2^Z\ ] \quad \Leftrightarrow \quad [\ \forall p \in Z,\ \{p\} \in \mathcal{T}\ ].$$

Proof of $\Rightarrow$: Assume $\mathcal{T} = 2^Z$. We wish to show: $\forall p \in Z,\ \{p\} \in \mathcal{T}$. Let $p \in Z$ be given. We wish to show: $\{p\} \in \mathcal{T}$.

Since $p \in Z$, we get $\{p\} \subseteq Z$, so $\{p\} \in 2^Z$. Then $\{p\} \in 2^Z = \mathcal{T}$, as desired. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume $\forall p \in Z,\ \{p\} \in \mathcal{T}$. We wish to show: $\mathcal{T} = 2^Z$. We have $\mathcal{T} = \{$open subsets of $Z\} \subseteq 2^Z$. It remains to show: $2^Z \subseteq \mathcal{T}$. So, by the quantified equivalence for $\subseteq$, we wish to show: $\forall S \in 2^Z$, $S \in \mathcal{T}$. Let $S \in 2^Z$ be given. We wish to show: $S \in \mathcal{T}$.

We know: $\forall p \in S, \{p\} \in \mathcal{T}$. Then $\{\{p\} \,|\, p \in S\} \subseteq \mathcal{T}$. Then $\bigcup\{\{p\} \,|\, p \in S\} \subseteq \langle \mathcal{T} \rangle_\cup$. We have $S = \bigcup\{\{p\} \,|\, p \in S\}$. Since $\mathcal{T}$ is a topology, we have $\langle \mathcal{T} \rangle_\cup = \mathcal{T}$. Then $S = \bigcup\{\{p\} \,|\, p \in S\} \subseteq \langle \mathcal{T} \rangle_\cup = \mathcal{T}$, as desired. *End of proof of $\Leftarrow$.* $\square$

DEFINITION 17.6. *For any set A, for any p, let $A_p^\times := A \backslash \{p\}$.*

DEFINITION 17.7. *For any set $\mathcal{S}$ of sets, for any p, we define $\mathcal{S}_p^{\times\times} := \{A_p^\times \,|\, A \in \mathcal{S}\}$.*

DEFINITION 17.8. *For any metric space (M, d), for any $p \in M$, for any $r > 0$, we define $B_d^\times(p, r) := (B_d(p, r))_p^\times$.*

If d is clear, we write $B(p, r)$ for $B_d(p, r)$, and $B^\times(p, r)$ for $B_d^\times(p, r)$.

DEFINITION 17.9. *For any metric space (M, d), for any $p \in M$, we define $\mathcal{B}_d^\times(p) := (\mathcal{B}_d(p))_p^{\times\times}$.*

If d is clear, we write $\mathcal{B}(p)$ for $\mathcal{B}_d(p)$, and $\mathcal{B}^\times(p)$ for $\mathcal{B}_d^\times(p)$.

DEFINITION 17.10. *For any topological space Z, for any $p \in Z$, we define $\mathcal{N}_Z^\times(p) := (\mathcal{N}_Z(p))_p^{\times\times}$.*

If Z is clear, we write $\mathcal{N}(p)$ for $\mathcal{N}_Z(p)$, and $\mathcal{N}^\times(p)$ for $\mathcal{N}_Z^\times(p)$.

Recall that a set in $\mathcal{N}_Z(p)$ is called a neighborhood of p in Z. A set in $\mathcal{N}_Z^\times(p)$ is called a **punctured neighborhood** or **pnbd** of p in Z.

REMARK 17.11. *Let Z be a topological space, let $S \subseteq Z$ and let $y \in S$. Then $(\mathcal{N}_Z^\times(y))|S = \mathcal{N}_S^\times(y)$.*

Proof. Unassigned HW. $\square$

COROLLARY 17.12. *Let Z be a top. space, $S \subseteq Z$, $y \in S$. Then*

(1) $\forall Q \in \mathcal{N}_X^\times(y), \quad Q \cap S \in \mathcal{N}_S^\times(y)$ *and*
(2) $\forall P \in \mathcal{N}_S^\times(y)$, $\exists Q \in \mathcal{N}_X^\times(y)$ *s.t.* $P = Q \cap S$.

Proof. We leave it as unassigned HW to show that $\subseteq$ in Remark 17.11 implies (1), and that $\supseteq$ in Remark 17.11 implies (2). $\square$

We call (1) of Corollary 17.12 the "restriction result for pnbds". We call (2) of Corollary 17.12 the "extension result for pnbds".

We drew a Venn diagram to illustrate the restriction and extension results for each of the following:

(1) open sets,
(2) closed sets,
(3) neighborhoods and
(4) punctured neighborhoods.

DEFINITION 17.13. *Let Z be a topological space, and let $p \in Z$. For all $\mathcal{B}$, by $\mathcal{B}$ is a* **neighborhood base** *at p in Z, we mean:*

(1) $\mathcal{B} \subseteq \mathcal{N}_Z(p)$ *and*
(2) $\forall V \in \mathcal{N}_Z(p)$, $\exists U \in \mathcal{B}$ *s.t.* $U \subseteq V$.

Any neighborhood of p covers p, so, in Definition 17.13, if you wish, you can replace "$U \subseteq V$" by "$p \in U \subseteq V$".

FACT 17.14. *Let Z be a topological space, let $Z_0 \subseteq Z$ and let $p \in Z_0$. Let $\mathcal{B}$ be a neighborhood base at p in Z. Let $\mathcal{B}_0 := \mathcal{B}|Z_0$. Then $\mathcal{B}_0$ is a neighborhood base at p in Z_0.*

Proof. Since $\mathcal{B}$ be a neighborhood base at p in Z, it follows, from (1) of Definition 17.13, that $\mathcal{B} \subseteq \mathcal{N}_Z(p)$. So, since restriction is monotonic, we get $\mathcal{B}|Z_0 \subseteq (\mathcal{N}_Z(p))|Z_0$. By Remark 16.18, $(\mathcal{N}_Z(p))|Z_0 = \mathcal{N}_{Z_0}(p)$. Then $\mathcal{B}_0 = \mathcal{B}|Z_0 \subseteq (\mathcal{N}_Z(p))|Z_0 = \mathcal{N}_{Z_0}(p)$. Following Definition 17.13 (replacing Z by Z_0, U by U_0 and V by V_0), it remains only to show: $\forall V_0 \in \mathcal{N}_{Z_0}(p)$, $\exists U_0 \in \mathcal{B}_0$ s.t. $U_0 \subseteq V_0$. Let $V_0 \in \mathcal{N}_{Z_0}(p)$ be given. We wish to show: $\exists U_0 \in \mathcal{B}_0$ s.t. $U_0 \subseteq V_0$.

By extension of neighborhhoods (see (2) of Corollary 16.19), choose $V \in \mathcal{N}_Z(p)$ s.t. $V_0 = V \cap Z_0$. Then, since $\mathcal{B}$ is a neighborhood base at p in Z, choose $U \in \mathcal{B}$ s.t. $U \subseteq V$. Since $U \in \mathcal{B}$, we get $U \cap Z_0 \in \mathcal{B}|Z_0$. Let $U_0 := U \cap Z_0$. Then $U_0 \in \mathcal{B}|Z_0 = \mathcal{B}_0$, and we wish to show: $U_0 \subseteq V_0$.

As $U \subseteq V$, $U \cap Z_0 \subseteq V \cap Z_0$. Then $U_0 = U \cap Z_0 \subseteq V \cap Z_0 = V_0$. $\square$

We assigned HW#49.

We calll Fact 17.14 the "restriction result for neighborhood bases". There is no perfect extension result for neighborhood bases; if you have a neighborhood base in a subspace, you can extend all of its neighborhoods, but, if you don't do it carefully, you won't end up with a neighborhood base in the ambient space. However if the subspace is *OPEN* in the ambient space, then, by HW#50, the neighborhood base in the subspace is, without change, a neighborhood base in the ambient space.

In the next result, when we say "$\mathcal{B}(p)$ is a neighborhood base at p in Z", it is understood that Z is given the topology $\mathcal{T}_d$. So, if you wish to be more detailed, you could say "$\mathcal{B}(p)$ is a neighborhood base at p in $(Z, \mathcal{T}_d)$".

THEOREM 17.15. *Let (Z, d) be a metric space and let $p \in Z$. Then $\mathcal{B}(p)$ is a neighborhood base at p in Z.*

Proof. By Definition 17.13 (with U replaced by B), we wish to show:

(1) $\mathcal{B}(p) \subseteq \mathcal{N}(p)$.
(2) $\forall V \in \mathcal{N}(p)$, $\exists B \in \mathcal{B}(p)$ s.t. $B \subseteq V$.

Proof of (1): We wish to show: $\forall B \in \mathcal{B}(p)$, $B \in \mathcal{N}(p)$. Let $B \in \mathcal{B}(p)$ be given. We wish to show: $B \in \mathcal{N}(p)$.

We have $B \in \mathcal{B}_d \subseteq \langle \mathcal{B}_d \rangle_\cup = \mathcal{T}_d$, so B is open in Z. Also, since $B \in \mathcal{B}(p)$, it follows that $p \in B$. Recall (Remark 16.4) that every open set is a neighborhood of each of its points. Then $B \in \mathcal{N}(p)$, as desired. *End of proof of (1).*

Proof of (2): Given $V \in \mathcal{N}(p)$. Want: $\exists B \in \mathcal{B}(p)$ s.t. $B \subseteq V$.

By definition of neighborhood, since $V \in \mathcal{N}(p)$, choose an open subset U of Z s.t. $p \in U \subseteq V$. We have $U \in \mathcal{T}_d = \langle \mathcal{B}_d \rangle_\cup$. Then $p \in U \in \langle \mathcal{B}_d \rangle_\cup$, so, by the quantified equivalence for $\langle \bullet \rangle_\cup$, choose $C \in \mathcal{B}_d$ s.t. $p \in C \subseteq U$. Since $p \in C \in \mathcal{B}_d$, by the Recentering Down Lemma (Lemma 14.2), choose $B \in \mathcal{B}(p)$ s.t. $B \subseteq C$. We wish to show: $B \subseteq V$.

We have $B \subseteq C \subseteq U \subseteq V$, as desired. *End of proof of (2).* □

Recall:

(1) $\mathcal{B}_\infty \;=\; \{\, (M, \infty] \mid M \in \mathbb{R} \,\}$,
(2) $\mathcal{B}_{-\infty} \;=\; \{\, [-\infty, N) \mid N \in \mathbb{R} \,\}$,
(3) $\mathcal{B}_\mathbb{R} \;=\; \mathcal{B}_{d_\mathbb{R}} \;=\; \{$bounded open intervals in $\mathbb{R}\}$,
(4) $\mathcal{B}_* \;=\; \mathcal{B}_\infty \;\cup\; \mathcal{B}_\mathbb{R} \;\cup\; \mathcal{B}_{-\infty}$,

(5) $\forall U, V \in \mathcal{B}_*, \quad U \cap V \;\in\; \mathcal{B}_* \cup \{\varnothing\} \;\subseteq\; \langle \mathcal{B}_* \rangle_\cup$,
(6) $\mathcal{T}_* \;=\; \langle \mathcal{B}_* \rangle_\cup$,
(7) $\mathbb{N}^* \;=\; \mathbb{N} \cup \{\infty\}$ and
(8) $\mathcal{T}_{\mathbb{N}^*} \;=\; \mathcal{T}_* \,|\, \mathbb{N}^*$.

The standard topology on $\mathbb{R}^*$ is $\mathcal{T}_*$. The standad topology on $\mathbb{N}^*$ is $\mathcal{T}_{\mathbb{N}^*}$.

Assigned HW#50.

HW#50 asserts that $\mathcal{B}_\infty$ is a neighborhood base at ∞ in $\mathbb{R}^*$.

Unassigned HW: Show: $\mathcal{B}_{-\infty}$ is a neighborhood base at $-\infty$ in $\mathbb{R}^*$.

By HW#50 and the restriction result for neighborhood bases (Fact 17.14), we conclude: $\mathcal{B}_\infty \,|\, \mathbb{N}^*$ is a neighborhood base at ∞ in $\mathbb{N}^*$. We computed

$$\mathcal{B}_\infty \,|\, \mathbb{N}^* \quad = \quad \{\ \{j, j+1, j+2, \cdots\} \,\cup\, \{\infty\} \mid j \in \mathbb{N}\}.$$

Then: $\quad (\mathcal{B}_\infty \,|\, \mathbb{N}^*)^{\times\times}_\infty \quad = \quad \{\ \{j, j+1, j+2, \cdots\} \mid j \in \mathbb{N}\}$.

Recall: For any function f, for any set S,

$$f_*(S) \;=\; \{\ f(x) \mid x \in (\mathrm{dom}[f]) \cap S\} \qquad \text{and}$$
$$f^*(S) \;=\; \{x \in \mathrm{dom}[f] \mid f(x) \in S\}.$$

The next result is “forward-image containment equivalency”; see (7) in the list of quantified equivalencies.

FACT 17.16. *For any function f, for any sets S and T,*

$$[\ f_*(S) \subseteq T\ ] \quad \Leftrightarrow \quad [\ \forall x \in dom[f],\ (\ (x \in S) \;\Rightarrow\; (f(x) \in T)\ )\ ].$$

Proof. Unassigned HW. □

We took the point of view that:

- Analysts like normed vector spaces.
- Geometers like metric spaces.
- Topologists like topological spaces.

From this perspective, Theorem 17.15 gives topologists and geometers something common. When working in a metric space, there's a natural topology to use: the union-closure of the set of balls in the metric. To approximate a point p, the geometer would use smaller and smaller balls about p, whereas the topologist would use smaller and smaller neighborhoods of p. The topologist would say “Why work with something geometric, like balls about a point p? There are lots of neighborhoods of p that are not round, but they can be very important. Why ignore them?" The savvy geometer might answer, "Why work with something non-geometric, like some amoeba-like neighborhood of p? I like geometric objects, because they have nice properties, and because there are fewer of them to track. Anyway, it doesn't really matter because the set of balls about p is a neighborhood base at p. So, for any approximation you might make with some neighborhood of p, I can make a better approximation using a ball centered at p. So anything you can do, I can do, and I don't have to keep track of so many weird-looking sets."

When working in a normed vector space, the geometer would use the norm $|\bullet|$ to create a metric d, defined by $d(p,q) = |p-q|$. To approximage a point p, the geometer would use smaller and smaller balls about p. The analyst prefers inequalities, so saying "q approximates p" would be expressed as "$|q-p| < r$ and r is small". To get better and better approximations, make r smaller and smaller. From this perspective, the next result gives a way of translating between geometric statements and analytic statements.

FACT 17.17. *Let $(V, |\bullet|)$ be a normed vector space, let $p, q \in V$ and let $r > 0$. Then*

(1) $[\ q \in B(p,r)\ ] \Leftrightarrow [\ |p-q| < r\ ]$ *and*
(2) $[\ q \in B^\times(p,r)\ ] \Leftrightarrow [\ 0 < |p-q| < r\ ]$.

Proof. Unassigned HW. $\square$

A geometer would typically prefer statements involving balls and punctured balls, *e.g.*, "$q \in B(p,r)$" or "$q \in B^\times(p,r)$". Using Fact 17.17, we can translate these into analytic statements, namely: "$|p-q| < r$" and "$0 < |p-q| < r$", respectively.

DEFINITION 17.18. *Let Y and Z be topological spaces. Let $f : Y \dashrightarrow Z$. Let $a \in Y$ and $b \in Z$. By*

near a in Y, $\quad f \to b$ in Z $\qquad$ or
$f \to b$ in Z $\quad$ near a in Y $\qquad$ or
$f(x) \to b$ in Z $\quad$ as $x \to a$ in Y,

we mean: $\forall V \in \mathcal{N}_Z(b)$, $\exists U \in \mathcal{N}_Y^\times(a)$ s.t. $f_*(U) \subseteq V$.

In Definition 17.18, the text "$\to$" is read "approaches". Sometimes Y and Z are clear, and we may omit "in Y" and "in Z", obtaining

near a, $f \to b$ or
$f \to b$ near a or
$f(x) \to b$ as $x \to a$.

In the third choice, "$f(x) \to b$ as $x \to a$", you may choose to use another variable. So, *e.g.*, "$f(t) \to b$ as $t \to a$" is also equivalent to "$f \to b$ near a". If you say, *e.g.*, "$f(s) \to b$ as $s \to a$", remember that s must be unbound at the start, and becomes unbound again at the end. If s is already bound, then you need to use another variable.

To stress that we use *punctured* neighborhoods of a, but *unpunctured* neighborhoods of b, I sometimes say:

p.near a, $f \to b$ or
$f \to b$ p.near a or
$f(x) \to b$ as $x \to\neq a$.

Here, "p.near" is an abbreviation for "punctured near". Also, the text "$x \to\neq a$" is read "x approaches a without equaling a".

We defined a function $f : [0, 4] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 3x - 5, & \text{if } x \neq 2 \\ 6, & \text{if } x = 2. \end{cases}$$

In the next class, we will prove that $f \to 1$ in $\mathbb{R}$ near 2 in $\mathbb{R}$. The proof will be based on the next lemma, which tells us that "in studying limits, it's okay to think like a geometer". More specifically, if we take the point of view that a geometer likes to work with neighborhood bases, and particularly with neighborhood bases consisiting of "geometrically nice" sets (like balls, disks, intervals), then the following will be useful:

LEMMA 17.19. *Let Y and Z be topological spaces. Let $f : Y \dashrightarrow Z$. Let $a \in Y$ and $b \in Z$. Let $\mathcal{A}$ be a neighborhood base of a in Y, and let $\mathcal{B}$ be a neighborhood base of b in Z. Then*

$$(\, f \to b \textit{ near } a\,) \quad \Leftrightarrow \quad (\,\forall B \in \mathcal{B}, \exists A \in \mathcal{A}_a^{\times\times} \textit{ s.t. } f_*(A) \subseteq B\,).$$

Proof. Next class. □

In the case of the function $f : [1, 4] \to \mathbb{R}$ defined above, we consider that $f : \mathbb{R} \dashrightarrow \mathbb{R}$, and wish to prove $f \to 1$ in $\mathbb{R}$ near 2 in $\mathbb{R}$. We use $Y = \mathbb{R}$, $Z = \mathbb{R}$, $a = 2$, $b = 1$. We use the standard metric $d_{\mathbb{R}}$ on $\mathbb{R}$, so: $\forall \delta > 0$, $B(2, \delta) = (2 - \delta, 2 + \delta)$ and $B^\times(2, \delta) = (2 - \delta, 2 + \delta) \backslash \{2\}$. Also,

$\forall\varepsilon > 0$, $B(1,\varepsilon) = (1-\varepsilon, 1+\varepsilon)$. In applying Lemma 17.19, we use

$$\begin{array}{rcccl} \mathcal{A} & = & \mathcal{B}(2) & = & \{B(2,\delta)\,|\,\delta > 0\} \quad \text{and} \\ \mathcal{B} & = & \mathcal{B}(1) & = & \{B(1,\varepsilon)\,|\,\varepsilon > 0\}. \end{array}$$

Then $\mathcal{A}_a^{\times\times} = \{B^\times(2,\delta)\,|\,\delta > 0\}$.

In the next class, we will prove: $f \to 1$ in R near 2 in $\mathbb{R}$. Equivalently, by Lemma 17.19,

$$\forall B \in \mathcal{B}, \quad \exists A \in \mathcal{A}_a^{\times\times} \quad \text{s.t.} \quad f_*(A) \subseteq B.$$

Equivalently,

$$\forall\varepsilon > 0, \quad \exists\delta > 0 \quad \text{s.t.} \quad f_*(B^\times(2,\delta)) \subseteq B(1,\delta).$$

Equivalently, by forward-image containment equivalency (Fact 17.16),

$$\begin{array}{l} \forall\varepsilon > 0, \quad \exists\delta > 0 \quad \text{s.t.} \quad \forall x \in \mathrm{dom}[f], \\ \qquad\qquad [\ x \in B^\times(2,\delta)\ ] \quad \Rightarrow \quad [\ f(x) \in B(1,\varepsilon)\ ]. \end{array}$$

Equivalently, by Fact 17.17,

$$\begin{array}{l} \forall\varepsilon > 0, \quad \exists\delta > 0 \quad \text{s.t.} \quad \forall x \in \mathrm{dom}[f], \\ \qquad\qquad [\ 0 < |\,x - 2\,| < \delta\ ] \quad \Rightarrow \quad [\ |\,[f(x)] - 1\,| < \varepsilon\ ]. \end{array}$$

NOTE: We just translated a topological statement

$$f \to 1 \text{ in } \mathbb{R} \quad \text{near 2 in } \mathbb{R}$$

into a geometric statement

$$\begin{array}{l} \forall\varepsilon > 0, \quad \exists\delta > 0 \quad \text{s.t.} \quad \forall x \in \mathrm{dom}[f], \\ \qquad\qquad [\ x \in B^\times(2,\delta)\ ] \quad \Rightarrow \quad [\ f(x) \in B(1,\varepsilon)\ ], \end{array}$$

and then into an analytic statement

$$\begin{array}{l} \forall\varepsilon > 0, \quad \exists\delta > 0 \quad \text{s.t.} \quad \forall x \in \mathrm{dom}[f], \\ \qquad\qquad [\ 0 < |\,x - 2\,| < \delta\ ] \quad \Rightarrow \quad [\ |\,[f(x)] - 1\,| < \varepsilon\ ]. \end{array}$$

This last statement is what we will prove in the next class.

18. Class 18 on 9 November 2017, Th of Week 10

We recalled Definition 17.18. Combined with Fact 17.16, we obtained the following triply quantified (with implication) formulation of the definition of the statement $f \to b$ near a.

COROLLARY 18.1. *Let Y and Z be topological spaces, $f : Y \dashrightarrow Z$, $a \in Y$, $b \in Z$. Then: $[\ f \to b$ near $a\ ]$ iff*

$$[\ \forall V \in \mathcal{N}_Z(b),\, \exists U \in \mathcal{N}_Y^\times(a) \text{ s.t.}, \forall x \in dom[f],\\ (\, x \in U\,) \ \Rightarrow\ (\, f(x) \in V\,)\ ].$$

Proof. Omitted. □

We also recalled Lemma 17.19. Combined with Fact 17.16, we obtained the following triply quantified (with implication) formulation of the definition of the statement $f \to b$ near a, given neighborhood bases of a and b.

COROLLARY 18.2. *Let Y and Z be topological spaces, $f : Y \dashrightarrow Z$, $a \in Y$, $b \in Z$. Let $\mathcal{A}$ be a neighborhood base of a in Y. Let $\mathcal{B}$ be a neighborhood base of b in Z. Then: $[\ f \to b$ near $a\ ]$ iff*

$$[\ \forall V \in \mathcal{B},\, \exists U \in \mathcal{A}_a^{\times\times} \text{ s.t.}, \forall x \in dom[f],\\ (\, x \in U\,) \ \Rightarrow\ (\, f(x) \in V\,)\ ].$$

Proof. Omitted. □

Let Y and Z be metric spaces. Recalled Theorem 17.15 to understand neighborhood bases in metric spaces. We combined this with Fact 17.17 and obtained the following analyst's triply quantified (with implication) quantified equivalence of $f \to b$ in Z near a in Y. This lead to:

COROLLARY 18.3. *Let Y and Z be metric spaces, $f : Y \dashrightarrow Z$, $a \in Y$, $b \in Z$. Then: $[\ f \to b$ near $a\ ]$ iff*

$$[\ \forall \varepsilon > 0,\, \exists \delta > 0 \text{ s.t.}, \forall x \in dom[f],\\ (\, 0 < d_Y(x,a) < \delta\,) \ \Rightarrow\ (\, d_Z(f(x),b) < \varepsilon\,)\ ].$$

Proof. Omitted. □

For normed vector spaces, the quantified equivalence reads:

COROLLARY 18.4. *Let Y, Z be normed vector spaces, $f : Y \dashrightarrow Z$, $a \in Y$, $b \in Z$. Then:* [$f \to b$ *near* a] *iff*

$$[\ \forall \varepsilon > 0,\ \exists \delta > 0\ s.t.,\ \forall x \in dom[f],$$
$$(\ 0\ <\ |\,x\ -\ a\,|_Y\ <\ \delta\ )\ \Rightarrow\ (\ |\,[f(x)]\ -\ b\,|_Z\ <\ \varepsilon\ )\ ].$$

Proof. Omitted. □

We also did a hybrid quantified equivalence in the case where the superdomain Y is a topological space, and the superimage Z is a normed vector space:

COROLLARY 18.5. *Let Y be a topological space. Let Z be a normed vector space. Let $f : Y \dashrightarrow Z$, $a \in Y$, $b \in Z$. Then:* [$f \to b$ *near* a] *iff*

$$[\ \forall \varepsilon > 0,\ \exists U \in \mathcal{N}_Y^\times(a)\ s.t.,\ \forall x \in dom[f],$$
$$(\ x \in U\ )\ \Rightarrow\ (\ |\,[f(x)]\ -\ b\,|_Z\ <\ \varepsilon\ )\ ].$$

Proof. Omitted. □

Now, let's move to the one-dimensional case, where $Y = Z = \mathbb{R}^*$.

First $\mathbb{R} = \cup\mathcal{B}_{\mathbb{R}} \in \langle \mathcal{B}_{\mathbb{R}} \rangle_\cup \subseteq \langle \mathcal{B}_* \rangle_\cup = \mathcal{T}_*$, so $\mathbb{R}$ is open in $\mathbb{R}^*$. For all $a \in \mathbb{R}$, by Theorem 17.15, we see that $\mathcal{B}(a)$ is a neighborhood base at a in $\mathbb{R}$. Then, since $\mathbb{R}$ is open in $\mathbb{R}^*$, by HW#49, for all $a \in \mathbb{R}$, we see that $\mathcal{B}(a)$ is a neighborhood base at a in $\mathbb{R}^*$. We can use this to understand limits of partial functions $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. Here's the quantified equivalence for $f \to b$ near a, when $a, b \in \mathbb{R}$:

LEMMA 18.6. *Let $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$, $a, b \in \mathbb{R}$. Then:* [$f \to b$ *near* a] *iff*

$$[\ \forall \varepsilon > 0,\ \exists \delta > 0\ s.t.,\ \forall x \in dom[f],$$
$$(\ 0\ <\ |\,x\ -\ a\,|\ <\ \delta\ )\ \Rightarrow\ (\ |\,[f(x)]\ -\ b\,|\ <\ \varepsilon\ )\ ].$$

Proof. Omitted. □

We applied this to a special case:

THEOREM 18.7. *Define $f : [0, 4] \to \mathbb{R}$ by*

$$f(x) = \begin{cases} 3x - 5, & \text{if } x \neq 2 \\ 6, & \text{if } x = 2. \end{cases}$$

Then $f \to 1$ in $\mathbb{R}^*$ near 2 in $\mathbb{R}^*$

Proof. By Lemma 18.6, we wish to show:

$$\begin{aligned}&[\ \forall\varepsilon>0,\ \exists\delta>0\text{ s.t., }\forall x\in\mathrm{dom}[f],\\&\qquad(\,0\,<\,|\,x\,-\,2\,|\,<\,\delta\,)\ \Rightarrow\ (\,|\,[f(x)]\,-\,1\,|\,<\,\varepsilon\,)\ ].\end{aligned}$$

Let $\varepsilon>0$ be given. We wish to show:

$$\begin{aligned}&[\ \exists\delta>0\text{ s.t., }\forall x\in\mathrm{dom}[f],\\&\qquad(\,0\,<\,|\,x\,-\,2\,|\,<\,\delta\,)\ \Rightarrow\ (\,|\,[f(x)]\,-\,1\,|\,<\,\varepsilon\,)\ ].\end{aligned}$$

Let $\varepsilon:=\delta/3$. We wish to show:

$$\begin{aligned}&[\ \forall x\in\mathrm{dom}[f],\\&\qquad(\,0\,<\,|\,x\,-\,2\,|\,<\,\delta\,)\ \Rightarrow\ (\,|\,[f(x)]\,-\,1\,|\,<\,\varepsilon\,)\ ].\end{aligned}$$

Let $x\in\mathrm{dom}[f]$ be given. We wish to show:

$$(\,0\,<\,|\,x\,-\,2\,|\,<\,\delta\,)\ \Rightarrow\ (\,|\,[f(x)]\,-\,1\,|\,<\,\varepsilon\,).$$

Assume $0<|x-2|<\delta$. We wish to show: $|[f(x)]-1|<\varepsilon$.

Since $0<|x-2|$, we get $x\neq 2$. Since $x\in\mathrm{dom}[f]$ and $x\neq 2$, by the definition of f, we see that $f(x)=3x-5$. Since $|x-2|<\delta$, we get $3\cdot|x-2|<3\cdot\delta$. Since $\delta=\varepsilon/3$, we get $3\cdot\delta=\varepsilon$. Then

$$\begin{aligned}|[f(x)]-1|\ &=\ |3x-5-1|\ =\ |3x-6|\ =\ |3(x-2)|\\&=\ 3\cdot|x-2|\ <\ 3\cdot\delta\ =\ \varepsilon,\end{aligned}$$

as desired. $\square$

Let $f:[0,4]\to\mathbb{R}$ be the function given in Theorem 18.7. We showed: $f\to 1$ in $\mathbb{R}^*$ near 2 in $\mathbb{R}^*$. Here are some related questions:

(1) Does $f\to 1$ in $\mathbb{R}^*$ near 2 in $\mathbb{R}$?
(2) Does $f\to 1$ in $\mathbb{R}^*$ near 2 in $[0,4]$?
(3) Does $f\to 1$ in $\mathbb{R}$ near 2 in $\mathbb{R}^*$?
(4) Does $f\to 1$ in $\mathbb{R}$ near 2 in $\mathbb{R}$?
(5) Does $f\to 1$ in $\mathbb{R}$ near 2 in $[0,4]$?
(6) Does $f\to 1$ in $[-5,7]$ near 2 in $\mathbb{R}^*$?
(7) Does $f\to 1$ in $[-5,7]$ near 2 in $\mathbb{R}$?
(8) Does $f\to 1$ in $[-5,7]$ near 2 in $[0,4]$?

The next result allows us to conclude, from Theorem 18.7, that the answer to each of these questions is yes.

We call this result **absoluteness of limits**:

THEOREM 18.8. *Let Y, Z be topological spaces, $Y_0 \subseteq Y$, $Z_0 \subseteq Z$, $f : Y_0 \dashrightarrow Z_0$, $a \in Y_0$, $b \in Z_0$. Then:* $[\, f \to b$ *in* Z_0 *near* a *in* $Y_0\,]$ *iff* $[\, f \to b$ *in* Z *near* a *in* $Y\,]$.

Proof. Proof of $\Rightarrow$*:* HW#51, due on 21 November. *End of proof of* $\Rightarrow$.

Proof of $\Leftarrow$*:* Assume: $f \to b$ in Z_0 near in Y_0. We wish to show: $f \to b$ in Z near in Y. Following Definition 17.18, we wish to show: $\forall V \in \mathcal{N}_Z(b)$, $\exists U \in \mathcal{N}_Y^\times(a)$ s.t. $f_*(U) \subseteq V$. Let $\forall V \in \mathcal{N}_Z(b)$ be given. We wish to show: $\exists U \in \mathcal{N}_Y^\times(a)$ s.t. $f_*(U) \subseteq V$.

Let $V_0 := V \cap Z_0$. Then, by (1) of Corollary 16.19, $V_0 \in \mathcal{N}_{Z_0}(b)$. So, following Definition 17.18, since $(\, f \to b$ in Z_0 near in $Y_0\,)$, choose $U_0 \in \mathcal{N}_{Y_0}^\times(a)$ such that $f_*(U_0) \subseteq V_0$. By (2) of Corollary 17.12, choose $U \in \mathcal{N}_Y^\times(a)$ such that $U_0 = U \cap Y_0$. We wish to show: $f_*(U) \subseteq V$. By Fact 17.16, we wish to show: $\forall x \in \mathrm{dom}[f]$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ f(x) \in V\ ].$$

Let $x \in \mathrm{dom}[f]$ be given. We wish to show $[x \in U] \Rightarrow [f(x) \in V]$. Assume $x \in U$. We wish to show: $f(x) \in V$.

Since $f : Y_0 \dashrightarrow Z_0$, we get $\mathrm{dom}[f] \subseteq Y_0$. Then $x \in \mathrm{dom}[f] \subseteq Y_0$. So, since $x \in U$, we get $x \in U \cap Y_0$. So, since $U \cap Y_0 = U_0$, we get $x \in U_0$. So, since $x \in \mathrm{dom}[f]$, we get $f(x) \in f_*(U_0)$. So, since $f_*(U_0) \subseteq V_0 = V \cap Z_0 \subseteq V$, we get $f(x) \in V$. *End of proof of* $\Leftarrow$. □

We next return to the one-dimensional case of functions $\mathbb{R}^* \to \mathbb{R}^*$. Then, since $\mathbb{R}$ is open in $\mathbb{R}^*$, by HW#49, for all $a \in \mathbb{R}$, we see that $\mathcal{B}(a)$ is a neighborhood base at a in $\mathbb{R}^*$. Also, by HW#50, $\mathcal{B}_\infty$ is a neighborhood base at ∞ in $\mathbb{R}^*$. Also, by an unassigned HW, $\mathcal{B}_{-\infty}$ is a neighborhood base at $-\infty$ in $\mathbb{R}^*$. So we now have a neighborhood base in $\mathbb{R}^*$ at any point of $\mathbb{R}^*$.

FACT 18.9. *Let $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$, $b \in \mathbb{R}$. Then:* $[\ f \to b$ *near* $\infty\ ]$ *iff*

$$\begin{aligned} &[\ \forall M \in \mathbb{R}, \exists \delta > 0 \ s.t., \forall x \in dom[f], \\ &\qquad (\, M < x < \infty\,) \ \Rightarrow\ (\, |\, [f(x)] - b\,| < \varepsilon\,)\ ]. \end{aligned}$$

Proof. Omitted. □

FACT 18.10. *Let $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$, $a \in \mathbb{R}$. Then:* $[\, f \to -\infty$ *near* $a\,]$ *iff*

$$\begin{aligned} &[\ \forall N \in \mathbb{R}, \exists \delta > 0 \ s.t., \forall x \in dom[f], \\ &\qquad (\, 0 < |\, x - a\,| < \delta\,) \ \Rightarrow\ (\, f(x) < N\,)\ ]. \end{aligned}$$

Proof. Omitted. □

FACT 18.11. *Let* $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. *Then:* $[\, f \to -\infty$ *near* $\infty\,]$ *iff*

$$[\ \forall N \in \mathbb{R},\, \exists M \in \mathbb{R}\ s.t.,\, \forall x \in dom[f],$$
$$(\, M \ < \ x \ < \ \infty\,) \ \Rightarrow \ (\, f(x) \ < \ N\,)\ ].$$

Proof. Omitted. □

FACT 18.12. *Let* $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. *Then:* $[\, f \to \infty$ *near* $\infty\,]$ *iff*

$$[\ \forall M \in \mathbb{R},\, \exists L \in \mathbb{R}\ s.t.,\, \forall x \in dom[f],$$
$$(\, L \ < \ x \ < \ \infty\,) \ \Rightarrow \ (\, f(x) \ > \ M\,)\ ].$$

Proof. Omitted. □

FACT 18.13. *Let* $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. *Then:* $[\, f \to \infty$ *near* $-\infty\,]$ *iff*

$$[\ \forall M \in \mathbb{R},\, \exists N \in \mathbb{R}\ s.t.,\, \forall x \in dom[f],$$
$$(\, -\infty \ < \ x \ < \ N\,) \ \Rightarrow \ (\, f(x) \ > \ M\,)\ ].$$

Proof. Omitted. □

DEFINITION 18.14. *We define* $\mathbb{Z}^* := \mathbb{Z} \cup \{\infty, -\infty\}$.

DEFINITION 18.15. *Let* $a, b \in \mathbb{R}^*$. *If* $a < b$, *then we define*

- $(a..b) := (a, b) \cap \mathbb{Z}^*$,
- $[a..b) := [a, b) \cap \mathbb{Z}^*$ *and*
- $(a..b] := (a, b] \cap \mathbb{Z}^*$.

If $a \leqslant b$, *then we define*

- $[a..b] := [a, b] \cap \mathbb{Z}^*$.

Then, for example,

$$(5..\infty] \ = \ \{5, 6, 7, \ldots\} \cup \{\infty\} \qquad \text{and}$$
$$(5..\infty) \ = \ \{5, 6, 7, \ldots\}.$$

Recall that a sequence is a function whose domain is $\mathbb{N}$. Let X be a topological space. An element of $\mathbb{R}^{\mathbb{N}}$ is a "sequence of real numbers". We wish to discuss limits of sequences of real numbers, as follows. For every $s \in \mathbb{R}^{\mathbb{N}}$, we consider $s : \mathbb{N}^* \dashrightarrow \mathbb{R}^*$ and try to find $b \in \mathbb{R}^*$ such that $s \to b$ in $\mathbb{R}^*$ near ∞ in $\mathbb{N}^*$. Recall that the topology on $\mathbb{N}^*$ is the relative topology inherited from $\mathcal{T}_*$ on $\mathbb{R}^*$. Because we are taking limits "near ∞ in $\mathbb{N}^*$", we want to have an easily described neighborhood base at ∞. We also want to describe the corresponding set of punctured neighborhoods of ∞ in $\mathbb{N}^*$.

Since $\mathcal{B}_\infty$ is a neighborhood base of ∞ in $\mathbb{R}$, it follows, from Fact 17.14, that $\mathcal{B}_\infty|\mathbb{N}^*$ is a neighborhood base of ∞ in $\mathbb{N}^*$. It is therefore worthwhile to calculate both $\mathcal{B}_\infty|\mathbb{N}^*$ and the corresponding set $(\mathcal{B}_\infty|\mathbb{N}^*)_\infty^{\times\times}$ of punctured neighborhoods of ∞ in $\mathbb{N}^*$. We calculated:

$$\begin{aligned} \mathcal{B}_\infty|\mathbb{N}^* &= \{\, [K..\infty] \mid K \in \mathbb{N} \,\} \qquad \text{and} \\ (\mathcal{B}_\infty|\mathbb{N}^*)_\infty^{\times\times} &= \{\, [K..\infty) \mid K \in \mathbb{N} \,\}. \end{aligned}$$

FACT 18.16. *Let $s : \mathbb{N}^* \dashrightarrow \mathbb{R}^*$. Then: $[\, s \to \infty$ near $\infty \,]$ iff*

$$\begin{aligned} &[\, \forall M \in \mathbb{R}, \exists K \in \mathbb{N} \text{ s.t.}, \forall j \in \mathit{dom}[s], \\ &\qquad (\, K \leqslant j < \infty \,) \Rightarrow (\, s_j > M \,) \,]. \end{aligned}$$

Proof. Omitted. □

FACT 18.17. *Let $s : \mathbb{N}^* \dashrightarrow \mathbb{R}^*$, $b \in \mathbb{R}$. Then: $[\, s \to b$ near $\infty \,]$ iff*

$$\begin{aligned} &[\, \forall \varepsilon > 0, \exists K \in \mathbb{N} \text{ s.t.}, \forall j \in \mathit{dom}[s], \\ &\qquad (\, K \leqslant j < \infty \,) \Rightarrow (\, |\, s_j - b \,| < \varepsilon \,) \,]. \end{aligned}$$

Proof. Omitted. □

FACT 18.18. *Let $s : \mathbb{N}^* \dashrightarrow \mathbb{R}^*$. Then: $[\, s \to -\infty$ near $\infty \,]$ iff*

$$\begin{aligned} &[\, \forall N \in \mathbb{R}, \exists K \in \mathbb{N} \text{ s.t.}, \forall j \in \mathit{dom}[s], \\ &\qquad (\, K \leqslant j < \infty \,) \Rightarrow (\, s_j < N \,) \,]. \end{aligned}$$

Proof. Omitted. □

19. Class 19 on 14 November 2017, Tu of Week 11

Assigned HW#51 and HW#52 and HW#53.

Recalled Definition 18.15 and Definition 18.14.
For any set S, for any $k \in \mathbb{N}$, we have $S^k = S^{[1..k]}$.
We recalled:

$$\begin{aligned} \mathcal{B}_\infty|\mathbb{N}^* &= \{\, [K..\infty] \mid K \in \mathbb{N} \,\} \qquad \text{and} \\ (\mathcal{B}_\infty|\mathbb{N}^*)_\infty^{\times\times} &= \{\, [K..\infty) \mid K \in \mathbb{N} \,\}. \end{aligned}$$

We recalled the quantified equivalence for $f \to b$ near ∞, in the situation where $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$ and $b \in \mathbb{R}$:

FACT 19.1. *Let* $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$, $b \in \mathbb{R}$. *Then:* $[\, f \to b$ *near* $\infty\,]$ *iff*

$$[\ \forall \varepsilon > 0,\ \exists M \in \mathbb{R}\ s.t.,\ \forall x \in dom[f],$$
$$(\ \infty\ >\ x\ >\ M\ )\ \Rightarrow\ (\ |\, [f(x)]\ -\ b\,|\ <\ \varepsilon\ )\ ].$$

Proof. Omitted. □

We developed the quantified equivalence for $f \to b$ near ∞, in the situation where $f : \mathbb{R} \dashrightarrow \mathbb{R}^*$ and $b \in \mathbb{R}$. Note that there's no need to "puncture" the input interval, because the domain of f does not include ∞. That is, there's no need to say "$\infty > x$":

FACT 19.2. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}^*$, $b \in \mathbb{R}$. *Then:* $[\, f \to b$ *near* $\infty\,]$ *iff*

$$[\ \forall \varepsilon > 0,\ \exists M \in \mathbb{R}\ s.t.,\ \forall x \in dom[f],$$
$$(\ x\ >\ M\ )\ \Rightarrow\ (\ |\, [f(x)]\ -\ b\,|\ <\ \varepsilon\ )\ ].$$

Proof. Omitted. □

We recalled the quantified equivalence for $s \to b$ near ∞, in the situation where $s : \mathbb{N}^* \dashrightarrow \mathbb{R}^*$ and $b \in \mathbb{R}$:

FACT 19.3. *Let* $s : \mathbb{N}^* \dashrightarrow \mathbb{R}^*$, $b \in \mathbb{R}$. *Then:* $[\, s \to b$ *near* $\infty\,]$ *iff*

$$[\ \forall \varepsilon > 0,\ \exists K \in \mathbb{N}\ s.t.,\ \forall j \in dom[s],$$
$$(\ \infty\ >\ j\ \geqslant\ K\ )\ \Rightarrow\ (\ |\, s_j\ -\ b\,|\ <\ \varepsilon\ )\ ].$$

Proof. Omitted. □

We developed the quantified equivalence for $s \to b$ near ∞, in the situation where $s : \mathbb{N} \dashrightarrow \mathbb{R}^*$ and $b \in \mathbb{R}$. Note that there's no need to "puncture" the input interval, because the domain of s does not include ∞. That is, there's no need to say "$\infty > j$":

FACT 19.4. *Let* $s : \mathbb{N} \dashrightarrow \mathbb{R}^*$, $b \in \mathbb{R}$. *Then:* $[\, s \to b$ *near* $\infty\,]$ *iff*

$$[\ \forall \varepsilon > 0,\ \exists K \in \mathbb{N}\ s.t.,\ \forall j \in dom[s],$$
$$(\ j\ \geqslant\ K\ )\ \Rightarrow\ (\ |\, s_j\ -\ b\,|\ <\ \varepsilon\ )\ ].$$

Proof. Omitted. □

DEFINITION 19.5. *Let* Z *be a topological space,* $s : \mathbb{N}^* \dashrightarrow Z$, $b \in Z$. *Then* $s_\bullet \to b$ *in* Z *means:* $s \to b$ *in* Z *near* ∞ *in* $\mathbb{N}^*$.

If Z is clear, we sometimes omit "in Z", and simply say "$s_\bullet \to b$".

In the situation where $Z = \mathbb{R}^*$, where $s \in \mathbb{R}^{\mathbb{N}}$ and where $b \in \mathbb{R}$, the quantified equivalence for $s_\bullet \to b$ is given by Fact 19.4. Bearing in mind that, in this situation, $\mathrm{dom}[s] = \mathbb{N}$, we get:

FACT 19.6. *Let $s \in \mathbb{R}^{\mathbb{N}}$, $b \in \mathbb{R}$. Then:* $[\, s_\bullet \to b \,]$ *iff*

$$[\ \forall \varepsilon > 0,\ \exists K \in \mathbb{N}\ s.t.,\ \forall j \in \mathbb{N},$$
$$(\ j \geqslant K\ )\ \Rightarrow\ (\ |\,s_j - b\,| < \varepsilon\ )\ ].$$

Proof. Omitted. □

In the situation where $Z = \mathbb{R}^*$, where $s \in \mathbb{R}^{\mathbb{N}}$ and where $b = \infty$, the quantified equivalence for $s_\bullet \to b$ reads:

FACT 19.7. *Let $s \in \mathbb{R}^{\mathbb{N}}$. Then:* $[\, s_\bullet \to \infty \,]$ *iff*

$$[\ \forall M \in \mathbb{R},\ \exists K \in \mathbb{N}\ s.t.,\ \forall j \in \mathbb{N},$$
$$(\ j \geqslant K\ )\ \Rightarrow\ (\ s_j > M\ )\ ].$$

Proof. Omitted. □

In the situation where $Z = \mathbb{R}^*$, where $s \in \mathbb{R}^{\mathbb{N}}$ and where $b = -\infty$, the quantified equivalence for $s_\bullet \to b$ reads:

FACT 19.8. *Let $s \in \mathbb{R}^{\mathbb{N}}$. Then:* $[\, s_\bullet \to -\infty \,]$ *iff*

$$[\ \forall N \in \mathbb{R},\ \exists K \in \mathbb{N}\ s.t.,\ \forall j \in \mathbb{N},$$
$$(\ j \geqslant K\ )\ \Rightarrow\ (\ s_j < N\ )\ ].$$

Proof. Omitted. □

We have absoluteness of limits of sequences:

THEOREM 19.9. *Let Z be a topological space and let $Z_0 \subseteq Z$. Let $s : \mathbb{N}^* \dashrightarrow Z_0$ and let $b \in Z_0$. Then: [$s_\bullet \to b$ in Z_0] $\Leftrightarrow$ [$s_\bullet \to b$ in Z].*

Proof. This is Theorem 18.8 (with Y replaced by $\mathbb{N}^*$ and f by s). □

Assigned HW#54.

DEFINITION 19.10. *For any set S, for any $f, g : S \dashrightarrow \mathbb{R}^*$, we define $f+g, fg, f-g, f/g : S \dashrightarrow \mathbb{R}^*$ by*

$$\begin{aligned}(f+g)(x) &= [f(x)] + [g(x)],\\ (fg)(x) &= [f(x)][g(x)],\\ (f-g)(x) &= [f(x)] - [g(x)] \qquad \text{and}\\ (f/g)(x) &= [f(x)]/[g(x)].\end{aligned}$$

We sometimes write $\dfrac{f}{g}$ for f/g.

DEFINITION 19.11. *For any set S, for any $a \in \mathbb{R}^*$, for any $g : S \dashrightarrow \mathbb{R}^*$, we define $ag, a/g, g/a : S \dashrightarrow \mathbb{R}^*$ by*

$$\begin{aligned}(ag)(x) &= a[g(x)],\\ (a/g)(x) &= a/[g(x)] \qquad \text{and}\\ (g/a)(x) &= [g(x)]/a.\end{aligned}$$

We sometimes write $\dfrac{a}{g}$ for a/g and we sometimes write $\dfrac{g}{a}$ for g/a.

DEFINITION 19.12. *For any set S, for any vector space V, for any $f, g : S \dashrightarrow V$, we define $f+g : S \dashrightarrow V$ by*

$$(f+g)(x) = [f(x)] + [g(x)].$$

DEFINITION 19.13. *For any set S, for any vector space V, for any $f : S \dashrightarrow \mathbb{R}$, for any $g : S \dashrightarrow V$, we define $fg : S \dashrightarrow V$ by*

$$(fg)(x) = [f(x)][g(x)].$$

DEFINITION 19.14. *For any set S, for any vector space V, for any $a \in \mathbb{R}$, for any $g : S \dashrightarrow V$, we define $ag, g/a : S \dashrightarrow V$ by*

$$\begin{aligned}(ag)(x) &= a[g(x)] \qquad \text{and}\\ (g/a)(x) &= [1/a][g(x)].\end{aligned}$$

We sometimes write $\dfrac{g}{a}$ for g/a.

Two sequences $a, \alpha \in (\mathbb{R}\backslash\{0\})^{\mathbb{N}}$ are **asymptotic** if $(a/\alpha)_\bullet \to 1$. For sequences s, σ, t, τ of positive real numbers, if s is asymptotic to σ and t is asymptotic to τ, then $s+t$ is asymptotic to $\sigma+\tau$:

FACT 19.15. *Let $s, \sigma, t, \tau \in (0, \infty)^{\mathbb{N}}$. Assume $(s/\sigma)_\bullet \to 1$, $(t/\tau)_\bullet \to 1$. Then $([s+t]/[\sigma+\tau])_\bullet \to 1$.*

The next proof was omitted from class.

Proof. We wish to show: $\forall \varepsilon > 0$, $\exists K \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad \left[\ \left|\frac{s_j + t_j}{\sigma_j + \tau_j} - 1\right| < \varepsilon\ \right].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists K \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad \left[\ \left|\frac{s_j + t_j}{\sigma_j + \tau_j} - 1\right| < \varepsilon\ \right].$$

Since $(s/\sigma)_\bullet \to 1$, choose $L \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant L\ ] \quad \Rightarrow \quad \left[\ \left|\frac{s_j}{\sigma_j} - 1\right| < \varepsilon\ \right].$$

Since $(t/\tau)_\bullet \to 1$, choose $M \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant M\ ] \quad \Rightarrow \quad \left[\ \left|\frac{t_j}{\tau_j} - 1\right| < \varepsilon\ \right].$$

Let $K := \max\{L, M\}$. We wish to show: $\forall j \in \mathbb{N}$,

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad \left[\ \left|\frac{s_j + t_j}{\sigma_j + \tau_j} - 1\right| < \varepsilon\ \right].$$

Let $j \in \mathbb{N}$ be given. We wish to show:

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad \left[\ \left|\frac{s_j + t_j}{\sigma_j + \tau_j} - 1\right| < \varepsilon\ \right].$$

Assume $j \geqslant K$. We wish to show: $\left|\dfrac{s_j + t_j}{\sigma_j + \tau_j} - 1\right| < \varepsilon$. We wish to show: $1 - \varepsilon < \dfrac{s_j + t_j}{\sigma_j + \tau_j} < 1 + \varepsilon$. Since $\sigma_j + \tau_j > 0$, we wish to show:

$$(1 - \varepsilon)(\sigma_j + \tau_j) \quad < \quad s_j + t_j \quad < \quad (1 + \varepsilon)(\sigma_j + \tau_j).$$

Since $j \geqslant K \geqslant L$, we get $\left|\dfrac{s_j}{\sigma_j} - 1\right| < \varepsilon$. Then $1 - \varepsilon < \dfrac{s_j}{\sigma_j} < 1 + \varepsilon$, so, since $\sigma_j > 0$, it follows that

$$(1 - \varepsilon)\sigma_j \quad < \quad s_j \quad < \quad (1 + \varepsilon)\sigma_j.$$

Since $j \geqslant K \geqslant M$, we get $\left|\dfrac{t_j}{\tau_j} - 1\right| < \varepsilon$. Then $1 - \varepsilon < \dfrac{t_j}{\tau_j} < 1 + \varepsilon$, so, since $\tau_j > 0$, it follows that

$$(1 - \varepsilon)\tau_j \quad < \quad t_j \quad < \quad (1 + \varepsilon)\tau_j.$$

Adding
$$(1-\varepsilon)\sigma_j \quad < \quad s_j \quad < \quad (1+\varepsilon)\sigma_j$$
to
$$(1-\varepsilon)\tau_j \quad < \quad t_j \quad < \quad (1+\varepsilon)\tau_j,$$
we get
$$(1-\varepsilon)(\sigma_j+\tau_j) \quad < \quad s_j+t_j \quad < \quad (1+\varepsilon)(\sigma_j+\tau_j),$$
as desired. $\square$

Assigned HW#55.

The **summation change formula** asserts: for any vector space Z, for any $u, u_1, v, v_1 \in Z$,
$$[u_1 \;+\; v_1] \;-\; [u \;+\; v] \;=\; [u_1 \;-\; u] \;+\; [v_1 \;-\; v].$$
In particular, this formula is true when $Z = \mathbb{R}$.

THEOREM 19.16. *Let $f, g : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. Let $a, u, v \in \mathbb{R}$. Assume: ($f \to u$ near a) and ($g \to v$ near a). Then: $f + g \to u + v$ near a.*

Proof. We wish to show: $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[f+g]$,
$$[\ 0 < |x-a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$
Let $\varepsilon > 0$ be given. We wish to show: $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[f+g]$,
$$[\ 0 < |x-a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$
Let $\lambda := \varepsilon/2$. Since $f \to u$ near a, choose $\beta > 0$ s.t., $\forall x \in \mathrm{dom}[f]$,
$$[\ 0 < |x-a| < \beta\ ] \quad \Rightarrow \quad [\ |\,[f(x)] - u\,| < \lambda\ ].$$
Since $g \to u$ near a, choose $\gamma > 0$ s.t., $\forall x \in \mathrm{dom}[g]$,
$$[\ 0 < |x-a| < \gamma\ ] \quad \Rightarrow \quad [\ |\,[g(x)] - v\,| < \lambda\ ].$$
Let $\delta := \min\{\beta, \gamma\}$. We wish to show: $\forall x \in \mathrm{dom}[f+g]$,
$$[\ 0 < |x-a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$
Let $x \in \mathrm{dom}[f+g]$ be given. We wish to show:
$$[\ 0 < |x-a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$
Assume $0 < |x-a| < \delta$. We wish to show: $|\,[(f+g)(x)] - [u+v]\,| < \varepsilon$.

Since $x \in \mathrm{dom}[f+g] \subseteq \mathrm{dom}[f]$ and since $0 < |x-a| < \delta \leqslant \beta$, by choice of β, we get $|\,[f(x)]-u]\,| < \lambda$. Since $x \in \mathrm{dom}[f+g] \subseteq \mathrm{dom}[g]$ and since $0 < |x-a| < \delta \leqslant \gamma$, by choice of γ, we get $|\,[g(x)]-v]\,| < \lambda$.

Let $u_1 := f(x)$ and $v_1 := g(x)$. Then $|u_1 - u| < \lambda$ and $|v_1 - v| < \lambda$. By the summation change formula,

$$[u_1 \ + \ v_1] \ - \ [u \ + \ v] \ = \ [u_1 \ - \ u] \ + \ [v_1 \ - \ v].$$

We have $(f + g)(x) = [f(x)] + [g(x)] = u_1 + v_1$. Then

$$\begin{aligned} |\,[(f+g)(x)] - [u+v]\,| &= |\,[u_1+v_1] - [u+v]\,| \\ &= |\,[u_1-u] + [v_1-v]\,| \\ &\leqslant |u_1-u| + |v_1-v| \\ &< \lambda+\lambda \ = \ 2\lambda \ = \ 2\cdot[\varepsilon/2] \ = \ \varepsilon, \end{aligned}$$

as desired. □

Next, we "upgrade" the preceding result to a result about partial functions from any topological space to any normed vector space:

THEOREM 19.17. *Let Y be a topological space and let Z be a normed vector space. Let $f, g : Y \dashrightarrow Z$. Let $a \in Y$, $u, v \in Z$. Assume that: ($f \to u$ near a) and ($g \to v$ near a). Then: $f + g \to u + v$ near a.*

Proof. We wish to show: $\forall \varepsilon > 0$, $\exists \Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[f + g]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[f+g]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$

Let $\lambda := \varepsilon/2$.

Since $f \to u$ near a, choose $B \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[f]$,

$$[\ x \in B\ ] \quad \Rightarrow \quad [\ |\,[f(x)] - u\,| < \lambda\ ].$$

Since $g \to u$ near a, choose $\Gamma \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ x \in \Gamma\ ] \quad \Rightarrow \quad [\ |\,[g(x)] - v\,| < \lambda\ ].$$

Let $\Delta := B \cap \Gamma$. Since $B, \Gamma \in \mathcal{N}^\times(a)$, it follows that $\Delta \in \mathcal{N}^\times(a)$. We wish to show: $\forall x \in \mathrm{dom}[f + g]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$

Let $x \in \mathrm{dom}[f + g]$ be given. We wish to show:

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(f+g)(x)] - [u+v]\,| < \varepsilon\ ].$$

Assume $x \in \Delta$. We wish to show: $|\,[(f+g)(x)] - [u+v]\,| < \varepsilon$.

Since $x \in \mathrm{dom}[f+g] \subseteq \mathrm{dom}[f]$ and since $x \in \Delta \subseteq B$, by choice of B, we get $|\,[f(x)] - u]\,| < \lambda$. Since $x \in \mathrm{dom}[f+g] \subseteq \mathrm{dom}[g]$ and since $x \in \Delta \subseteq \Gamma$, by choice of Γ, we get $|\,[g(x)] - v]\,| < \lambda$.

Let $u_1 := f(x)$ and $v_1 := g(x)$. Then $|u_1 - u| < \lambda$ and $|v_1 - v| < \lambda$. By the summation change formula,

$$[u_1 \ + \ v_1] \ - \ [u \ + \ v] \ = \ [u_1 \ - \ u] \ + \ [v_1 \ - \ v].$$

We have $(f+g)(x) = [f(x)] + [g(x)] = u_1 + v_1$. Then

$$\begin{aligned} |\,[(f+g)(x)] - [u+v]\,| &= |\,[u_1+v_1] - [u+v]\,| \\ &= |\,[u_1 - u] + [v_1 - v]\,| \\ &\leqslant |u_1 - u| + |v_1 - v| \\ &< \lambda + \lambda \ = \ 2\lambda \ = \ 2 \cdot [\varepsilon/2] \ = \ \varepsilon, \end{aligned}$$

as desired. □

Note that, in the proof of Theorem 19.17, we "think like a topologist" on the input side, using neighborhoods like B, Γ, Δ. At the same time, we "think like an analyst on the output side, using inequalities like $|[f(x)] - u| < \lambda$, $\ |[g(x)] - v| < \lambda$, $\ |[(f+g)(x)] - [u+v]| < \varepsilon$. This is in contrast to the proof of Theorem 19.16, where we're thinking like an analyst both on the input side and on the output side.

HW#52 asserts: Let $f, g : \mathbb{R}^* \to \mathbb{R}^*$, $a \in \mathbb{R}$. Assume $f \to \infty$ near a and $g \to \infty$ near a. Then $f + g \to \infty$ near a.

There is another similar statement: Let $f, g : \mathbb{R}^* \to \mathbb{R}^*$. Assume $f \to \infty$ near ∞ and $g \to \infty$ near ∞. Then $f + g \to \infty$ near ∞.

There is another similar statement: Let $f, g : \mathbb{R}^* \to \mathbb{R}^*$. Assume $f \to \infty$ near $-\infty$ and $g \to \infty$ near $-\infty$. Then $f + g \to \infty$ near $-\infty$.

HW#52 and the other two statements are all three true, though UNPROVED. If we "think like an analyst", then we cannot avoid having three different proofs, but we'll certainly notice that the proofs are similar, and we'll wonder if there might be some way to handle all three statements at once. The solution is to "think like a topologist" on the input side, while "thinking like an analyst" on the output side, and "upgrade" the statement to HW#52 to:

> Let Y be a topogical space. Let $f, g : Y \to \mathbb{R}^*$. Let $a \in Y$. Assume that $f \to \infty$ near a, and, also, that $g \to \infty$ near a. Then $f + g \to \infty$ near a.

This statement is UNPROVED. However, you might try "upgrading" your solution to prove it. Perhaps in your solution, you have $\beta, \gamma > 0$ and $\delta = \min\{\beta, \gamma\}$, as in Theorem 19.16. Then, in your upgraded proof, you'll have $B, \Gamma \in \mathcal{N}^\times(a)$ and $\Delta = B \cap \Gamma$, as in Theorem 19.17.

Here's another similar statement:

> Let Y be a topogical space. Let $f, g : Y \to \mathbb{R}^*$. Let $a \in Y$. Assume that $f \to -\infty$ near a and, also, that $g \to 7$ near a. Then $f + g \to -\infty$ near a.

This statement is UNPROVED. You might try to adapt your previous upgrade of HW#57 to prove it, but it will be unsatisfying, becuase there are many other similar statements, and it's natural to wonder whether they can all be subsumed into one result. To accomplish this, we need the concept of **contingent approaches**, denoted $\to^*$:

DEFINITION 19.18. *Let Y and Z be topological spaces, $a \in Y$, $f : Y \dashrightarrow Z$. Then, for all b, by $f \to^* b$ in Z near a in Y, we mean:*

either $[\ b = \frown\ ]$
or $[\ (\ b \in Z\ )$ *and* $(\ f \to b$ *in* Z *near* a *in* $Y\ )\ ]$.

That is, ($f \to^* b$ in Z near a in Y) means either $b = \frown$ or we can "remove the contingency" and change $\to^*$ to $\to$.

When Y and Z are clear, we somtimes omit the text "in Y" and "in Z".

We now present a very general result about adding functions f and g, when their superdomain is a topological space Y and their superimage is $\mathbb{R}^*$:

THEOREM 19.19. *Let Y be a topological space, $f, g : Y \dashrightarrow \mathbb{R}^*$, $a \in Y$. Then, for all b, c, we have:*

if $[\ (\ f \to^* b$ *near* $a\ )$ *and* $(\ g \to^* c$ *near* $a\ )\ ]$,
then $[\ f + g \to^* b + c$ *near* $a\ ]$.

Theorem 19.19 is UNPROVED. However, we'll talk about how one might prove it: Structuring the proof as usualy, eventually both b and c are bound. Then one of the following must be true:

(1) $b = -\infty$ or
(2) $b \in \mathbb{R}$ or
(3) $b = \infty$ or
(4) $b = \frown$.

Also, one of the following must be true:

(A) $c = -\infty$ or
(B) $c \in \mathbb{R}$ or
(C) $c = \infty$ or
(D) $c = ☹$.

This gives us 16 cases:

1A, 1B, 1C, 1D, 2A, 2B, 2C, 2D,
3A, 3B, 3C, 3D, 4A, 4B, 4C, 4D.

Note that Case 2B is proved in Theorem 19.16 (with u replaced by b and v by c). Case 3C is UNPROVED, but our upgrade of HW#52 would prove it. In the nine cases

1C, 3A,
4A, 4B, 4C,
1D, 2D, 3D and
4D,

we have $b + c = ☹$; those cases are "easy". The remaining cases are

1A, 1B, 2A, 2C, 3B

and are all UNPROVED. They make good problems for homeworks and exams.

Now we change gears and move from studing sums of functions to products of functions. We begin with the **product change formula**:

REMARK 19.20. *Let $u, u_1, v, v_1 \in \mathbb{R}$. Then*

$$u_1 v_1 - uv \;=\; (u_1 - u)v \;+\; u(v_1 - v) \;+\; (u_1 - u)(v_1 - v).$$

Proof. Omitted. □

We showed how a rectangle, partitioned into four subrectangles, gives motivation for Remark 19.20. To prove Remark 19.20, one simply expands the RHS, and then simplifies.

We can "upgrade" Remark 19.20 to vector spaces:

REMARK 19.21. *Let V be a vector space. Let $b, b_1 \in \mathbb{R}$ and let $v, v_1 \in V$. Then*

$$b_1 v_1 - bv \;=\; (b_1 - b)v \;+\; b(v_1 - v) \;+\; (b_1 - b)(v_1 - v).$$

Proof. Omitted. □

To prove Remark 19.21, one, again, simply expands the RHS, and then simplifies.

We now show that the limit of the product is the product of the limits, for partial functions $\mathbb{R}^* \dashrightarrow \mathbb{R}^*$.

THEOREM 19.22. *Let $f, g : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. Let $a, u, v \in \mathbb{R}$. Assume: ($f \to u$ near a) and ($g \to v$ near a). Then: $fg \to uv$ near a.*

Proof. We wish to show: $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[fg]$,

$$[\ 0 < |x - a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[fg]$,

$$[\ 0 < |x - a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Let $\lambda \ := \ \min\left\{\ 1\ ,\ \dfrac{\varepsilon}{|b| + |v| + 1}\ \right\}.$

Since $f \to u$ near a, choose $\beta > 0$ s.t., $\forall x \in \mathrm{dom}[f]$,

$$[\ 0 < |x - a| < \beta\ ] \quad \Rightarrow \quad [\ |\,[f(x)] - u\,| < \lambda\ ].$$

Since $g \to u$ near a, choose $\gamma > 0$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ 0 < |x - a| < \gamma\ ] \quad \Rightarrow \quad [\ |\,[g(x)] - v\,| < \lambda\ ].$$

Let $\delta := \min\{\beta, \gamma\}$. We wish to show: $\forall x \in \mathrm{dom}[fg]$,

$$[\ 0 < |x - a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Let $x \in \mathrm{dom}[fg]$ be given. We wish to show:

$$[\ 0 < |x - a| < \delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Assume $0 < |x - a| < \delta$. We wish to show: $|\,[(fg)(x)] - uv\,| < \varepsilon$.

Since $x \in \mathrm{dom}[fg] \subseteq \mathrm{dom}[f]$ and since $0 < |x-a| < \delta \leqslant \beta$, by choice of β, we get $|\,[f(x)] - u]\,| < \lambda$. Since $x \in \mathrm{dom}[fg] \subseteq \mathrm{dom}[g]$ and since $0 < |x - a| < \delta \leqslant \gamma$, by choice of γ, we get $|\,[g(x)] - v]\,| < \lambda$.

Let $u_1 := f(x)$ and $v_1 := g(x)$. Then $|u_1 - u| < \lambda$ and $|v_1 - v| < \lambda$. Then $|u_1 - u| \cdot |v| \leqslant \lambda \cdot |v|$ and $|u| \cdot |v_1 - v| \leqslant |u| \cdot \lambda$. By definition of λ, we have $\lambda \leqslant 1$, so $\lambda^2 \leqslant \lambda$. Then $|u_1 - u| \cdot |v_1 - v| < \lambda^2 \leqslant \lambda$. Because

$$\begin{aligned}
&|u_1 - u| \cdot |v| \leqslant \lambda \cdot |v|, \\
&|u| \cdot |v_1 - v| \leqslant |u| \cdot \lambda \qquad \text{and} \\
&|u_1 - u| \cdot |v_1 - v| < \lambda,
\end{aligned}$$

we conclude:

$$\begin{array}{rl} & |u_1 - u| \cdot |v| + |u| \cdot |v_1 - v| + |u_1 - u| \cdot |v_1 - v| \\ < & \lambda \cdot |v| \ + \ |u| \cdot \lambda \ + \ \lambda. \end{array}$$

By the product change formula,

$$u_1 v_1 - uv \ = \ (u_1 - u)v \ + \ u(v_1 - v) \ + \ (u_1 - u)(v_1 - v).$$

By definition of λ, we have $\lambda \leqslant \dfrac{\varepsilon}{|u| + |v| + 1}$, so $(|u| + |v| + 1) \cdot \lambda \ \leqslant \ \varepsilon$. We have $(fg)(x) = [f(x)][g(x)] = u_1 v_1$. Then

$$\begin{array}{rcl} |\, [(fg)(x)] - uv \,| & = & |\, u_1 v_1 - uv \,| \\ & = & |\, (u_1 - u)v \ + \ u(v_1 - v) \ + \ (u_1 - u)(v_1 - v) \,| \\ & \leqslant & |u_1 - u| \cdot |v| + |u| \cdot |v_1 - v| + |u_1 - u| \cdot |v_1 - v| \\ & < & \lambda \cdot |v| \ + \ |u| \cdot \lambda \ + \ \lambda \\ & = & (|u| + |v| + 1) \cdot \lambda \ \leqslant \ \varepsilon, \end{array}$$

as desired. □

We now "upgrade":

THEOREM 19.23. *Let Y be a topological space and let Z be a normed vector space. Let $f : Y \dashrightarrow \mathbb{R}$ and $g : Y \dashrightarrow Z$. Let $a \in Y$ and $b \in \mathbb{R}$ and $v \in Z$. Assume that: ($f \to b$ near a) and ($g \to v$ near a). Then we have: $fg \to bv$ near a.*

Proof. We wish to show: $\forall \varepsilon > 0$, $\Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[fg]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\, [(fg)(x)] - bv \,| < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[fg]$,

$$[\ 0 < |x - a| < \delta\ ] \quad \Rightarrow \quad [\ |\, [(fg)(x)] - bv \,| < \varepsilon\ ].$$

Let $\lambda \ := \ \min\left\{ 1 \ , \ \dfrac{\varepsilon}{|b| + |v| + 1} \right\}$.

Since $f \to b$ near a, choose $B \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[f]$,

$$[\ x \in B\ ] \quad \Rightarrow \quad [\ |\, [f(x)] - b \,| < \lambda\ ].$$

Since $g \to b$ near a, choose $\Gamma \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ x \in \Gamma\ ] \quad \Rightarrow \quad [\ |\, [g(x)] - v \,| < \lambda\ ].$$

Let $\Delta := B \cap \Gamma$. We wish to show: $\forall x \in \mathrm{dom}[fg]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\, [(fg)(x)] - bv \,| < \varepsilon\ ].$$

Let $x \in \mathrm{dom}[fg]$ be given. We wish to show:

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - bv\,| < \varepsilon\ ].$$

Assume $x \in \Delta$. We wish to show: $|\,[(fg)(x)] - bv\,| < \varepsilon$.

Since $x \in \mathrm{dom}[fg] \subseteq \mathrm{dom}[f]$ and since $x \in \Delta \subseteq B$, by choice of B, we get $|\,[f(x)] - b]\,| < \lambda$. Since $x \in \mathrm{dom}[fg] \subseteq \mathrm{dom}[g]$ and since $x \in \Delta \subseteq \Gamma$, by choice of Γ, we get $|\,[g(x)] - v]\,| < \lambda$.

Let $b_1 := f(x)$ and $v_1 := g(x)$. Then $|b_1 - b| < \lambda$ and $|v_1 - v| < \lambda$. Then $|b_1 - b| \cdot |v| \leqslant \lambda \cdot |v|$ and $|b| \cdot |v_1 - v| \leqslant |b| \cdot \lambda$. By definition of λ, we have $\lambda \leqslant 1$, so $\lambda^2 \leqslant \lambda$. Then $|b_1 - b| \cdot |v_1 - v| < \lambda^2 \leqslant \lambda$. Because

$$\begin{array}{ll} |b_1 - b| \cdot |v| \leqslant \lambda \cdot |v|, & \\ |b| \cdot |v_1 - v| \leqslant |b| \cdot \lambda & \text{and} \\ |b_1 - b| \cdot |v_1 - v| < \lambda, & \end{array}$$

we conclude:

$$\begin{aligned} & |b_1 - b| \cdot |v| + |b| \cdot |v_1 - v| + |b_1 - b| \cdot |v_1 - v| \\ < \ & \lambda \cdot |v| \ + \ |b| \cdot \lambda \ + \ \lambda. \end{aligned}$$

By the product change formula,

$$b_1 v_1 - bv \ = \ (b_1 - b)v \ + \ b(v_1 - v) \ + \ (b_1 - b)(v_1 - v).$$

By definition of λ, we have $\lambda \leqslant \dfrac{\varepsilon}{|b| + |v| + 1}$, so $(|b| + |v| + 1) \cdot \lambda \ \leqslant \ \varepsilon$. We have $(fg)(x) = [f(x)][g(x)] = b_1 v_1$. Then

$$\begin{aligned} |\,[(fg)(x)] - bv\,| \ &= \ |\,b_1 v_1 - bv\,| \\ &= \ |\,(b_1 - b)v \ + \ b(v_1 - v) \ + \ (b_1 - b)(v_1 - v)\,| \\ &\leqslant \ |b_1 - b| \cdot |v| + |b| \cdot |v_1 - v| + |b_1 - b| \cdot |v_1 - v| \\ &< \ \lambda \cdot |v| \ + \ |b| \cdot \lambda \ + \ \lambda \\ &= \ (|b| + |v| + 1) \cdot \lambda \ \leqslant \ \varepsilon, \end{aligned}$$

as desired. □

Here's slightly different "upgrade" of Theorem 19.22:

THEOREM 19.24. *Let Y be a topological space. Let $f, g : Y \dashrightarrow \mathbb{R}^*$. Let $a, u, v \in \mathbb{R}$. Assume: ($f \to u$ near a) and ($g \to v$ near a). Then: $fg \to uv$ near a.*

Proof. We wish to show: $\forall \varepsilon > 0$, $\exists \Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[fg]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[fg]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Let $\lambda \ := \ \min\left\{\ 1\ ,\ \dfrac{\varepsilon}{|b| + |v| + 1}\ \right\}$.

Since $f \to u$ near a, choose $B \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[f]$,

$$[\ x \in B\ ] \quad \Rightarrow \quad [\ |\,[f(x)] - u\,| < \lambda\ ].$$

Since $g \to u$ near a, choose $\Gamma \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ x \in \Gamma\ ] \quad \Rightarrow \quad [\ |\,[g(x)] - v\,| < \lambda\ ].$$

Let $\Delta := B \cap \Gamma$. We wish to show: $\forall x \in \mathrm{dom}[fg]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Let $x \in \mathrm{dom}[fg]$ be given. We wish to show:

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[(fg)(x)] - uv\,| < \varepsilon\ ].$$

Assume $x \in \Delta$. We wish to show: $|\,[(fg)(x)] - uv\,| < \varepsilon$.

Since $x \in \mathrm{dom}[fg] \subseteq \mathrm{dom}[f]$ and since $x \in \Delta \subseteq B$, by choice of B, we get $|\,[f(x)] - u]\,| < \lambda$. Since $x \in \mathrm{dom}[fg] \subseteq \mathrm{dom}[g]$ and since $x \in \Delta \subseteq \Gamma$, by choice of Γ, we get $|\,[g(x)] - v]\,| < \lambda$.

Let $u_1 := f(x)$ and $v_1 := g(x)$. Then $|u_1 - u| < \lambda$ and $|v_1 - v| < \lambda$. Then $|u_1 - u| \cdot |v| \leqslant \lambda \cdot |v|$ and $|u| \cdot |v_1 - v| \leqslant |u| \cdot \lambda$. By definition of λ, we have $\lambda \leqslant 1$, so $\lambda^2 \leqslant \lambda$. Then $|u_1 - u| \cdot |v_1 - v| < \lambda^2 \leqslant \lambda$. Because

$$\begin{array}{ll} |u_1 - u| \cdot |v| \leqslant \lambda \cdot |v|, & \\ |u| \cdot |v_1 - v| \leqslant |u| \cdot \lambda & \qquad \text{and} \\ |u_1 - u| \cdot |v_1 - v| < \lambda, & \end{array}$$

we conclude:

$$\begin{aligned} & |u_1 - u| \cdot |v| + |u| \cdot |v_1 - v| + |u_1 - u| \cdot |v_1 - v| \\ < \ \ & \lambda \cdot |v| \ + \ |u| \cdot \lambda \ + \ \lambda. \end{aligned}$$

By the product change formula,

$$u_1 v_1 - uv \ = \ (u_1 - u)v \ + \ u(v_1 - v) \ + \ (u_1 - u)(v_1 - v).$$

By definition of λ, we have $\lambda \leqslant \dfrac{\varepsilon}{|u| + |v| + 1}$, so $(|u| + |v| + 1) \cdot \lambda \leqslant \varepsilon$. We have $(fg)(x) = [f(x)][g(x)] = u_1 v_1$. Then

$$\begin{aligned}
|\,[(fg)(x)] - uv\,| &= |\,u_1 v_1 - uv\,| \\
&= |\,(u_1 - u)v \,+\, u(v_1 - v) \,+\, (u_1 - u)(v_1 - v)\,| \\
&\leqslant |u_1 - u| \cdot |v| + |u| \cdot |v_1 - v| + |u_1 - u| \cdot |v_1 - v| \\
&< \lambda \cdot |v| \,+\, |u| \cdot \lambda \,+\, \lambda \\
&= (|u| + |v| + 1) \cdot \lambda \,\leqslant\, \varepsilon,
\end{aligned}$$

as desired. □

20. Class 20 on 16 November 2017, Th of Week 11

FACT 20.1. *For all $a, b \in \mathbb{R}^*$, we have*

(1) $a - b = a + (-b)$ *and*
(2) $a/b = a \cdot (1/b)$.

Proof. Unassigned HW. □

FACT 20.2. *For all $a, b, c, d \in \mathbb{R}^*$, we have*

$$\frac{a}{b} \;+\; \frac{c}{d} \quad =^* \quad \frac{ad \;+\; bc}{bd}.$$

Proof. Unassigned HW. □

Note that

$$\frac{3}{\infty} \;+\; \frac{\infty}{2} \quad = \quad 0 \;+\; \infty \quad = \quad \infty$$

and that

$$\frac{3 \cdot 2 \;+\; \infty \cdot \infty}{\infty \cdot 2} \quad = \quad \frac{6 + \infty}{\infty} \quad = \quad \frac{\infty}{\infty} \quad = \quad ☹,$$

so

$$\frac{3}{\infty} \;+\; \frac{\infty}{2} \quad \neq \quad \frac{3 \cdot 2 \;+\; \infty \cdot \infty}{\infty \cdot 2}.$$

We recalled Theorem 19.22, but replaced u by b and v by c.

We recalled Theorem 19.23, but replaced v by c.

HW#56 (due Tuesday 28 November): Let $f, g : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$ and let $a \in \mathbb{R}$. Assume: ($f \to \infty$ near a) and ($g \to \infty$ near a). Show: $fg \to \infty$ near a.

We have two variants of HW#56, as follows.

Unassigned HW: Let $f, g : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. Assume: ($f \to \infty$ near ∞) and ($g \to \infty$ near ∞). Show: $fg \to \infty$ near ∞.

Unassigned HW: Let $f, g : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$. Assume: ($f \to \infty$ near $-\infty$) and ($g \to \infty$ near $-\infty$). Show: $fg \to \infty$ near $-\infty$.

If you "think like an analyst", then HW#56 and the two variants all require different proofs. However, if you can "think like a topologist on the input side", then there's a common generalization of all three results:

THEOREM 20.3. *Let Y be a topological space, let $f, g : Y \dashrightarrow \mathbb{R}^*$ and let $a \in Y$. Assume: ($f \to \infty$ near a) and ($g \to \infty$ near a). Then: $fg \to \infty$ near a.*

Theorem 20.3 is UNPROVED, but is almost proved, in the sense that, to prove it, you can take your solution to HW#56, and "upgrade" from (β, γ and $\delta = \min\{\beta, \gamma\}$) to (B, Γ and $\Delta = B \cap \Gamma$).

Here's a variant of Theorem 20.3:

Unassigned HW: Let Y be a topological space, let $f, g : Y \dashrightarrow \mathbb{R}^*$ and let $a \in Y$. Assume: ($f \to \infty$ near a) and ($g \to -3$ near a). Show: $fg \to -\infty$ near a.

We now seek a common generalization of this last unassigned HW and Theorem 20.3. The following is that common generalization:

THEOREM 20.4. *Llet Y be a topological space, let $f, g : Y \dashrightarrow \mathbb{R}^*$ and let $a \in Y$. Then, for all b, c, we have:*

$$[\ (\ f \to^* b \textit{ near } a\ ) \quad \& \quad (\ g \to^* c \textit{ near } a\ )\ ]$$
$$\Rightarrow \qquad [\ fg \to^* bc \textit{ near } a\ ].$$

Partial proof: Let b, c be given. We wish to show

$$[\ (\ f \to^* b \text{ near } a\ ) \quad \& \quad (\ g \to^* c \text{ near } a\ )\ ]$$
$$\Rightarrow \qquad [\ fg \to^* bc \text{ near } a\ ].$$

Assume that $f \to^* b$ near a and that $g \to^* c$ near a. We wish to show that $fg \to^* bc$ near a. That is, we wish to show: either [$bc = ☹$] or

[both ($bc \in \mathbb{R}^*$) and ($fg \to bc$ near a)].

Since $f \to^* b$ near a, it follows either that $b \in \mathbb{R}^*$ or that $b = ☹$. Then one of the following must be true:

(1) $b = -\infty$,
(2) $b \in (-\infty, 0)$,
(3) $b = 0$,
(4) $b \in (0, \infty)$,
(5) $b = \infty$ or
(6) $b = ☹$.

Since $g \to^* c$ near a, it follows either that $c \in \mathbb{R}^*$ or that $c = ☹$. Then one of the following must be true:

(A) $c = -\infty$,
(B) $c \in (-\infty, 0)$,
(C) $c = 0$,
(D) $c \in (0, \infty)$,
(E) $c = \infty$ or
(F) $c = ☹$.

We now have 36 cases to consider, from 1A to 6F.

Case 5E is a consequence of Theorem 20.3, which is "almost proved".

In Cases 2B, 2C, 2D, 3B, 3C, 3D, 4B, 4C and 4D, we have $b \in \mathbb{R}$ and $c \in \mathbb{R}$, so, by Theorem 19.24 (with u replaced by b and v by c), we get what we want.

The following cases are all "easy" because, in each of them, $bc = ☹$:

1C, 5C, 3A, 3E, 1F, 2F, 3F, 4F, 5F, 6A, 6B, 6C, 6D, 6E, 6F

This almost finishes 5E, and completely finishes 24 of the 36 cases. The other 11 cases are

1A, 1B, 1D, 1E, 2A, 2E, 4A, 4E, 5A, 5B, 5D

and they are all unproved. *End of partial proof.*

Assigned HW#57 (due Tuesday 28 November).

LEMMA 20.5. *Let Y be a topological space, let $g : Y \dashrightarrow \mathbb{R}^*$ and let $a \in Y$. Then, for all c, we have:*

$$(\, g \ \to^* \ c \textit{ near } a \,) \quad \Rightarrow \quad (\, 1/g \ \to^* \ 1/c \textit{ near } a \,).$$

Partial proof: Let c be given. We wish to show:

$$(\, g \ \to^* \ c \text{ near } a \,) \quad \Rightarrow \quad (\, 1/g \ \to^* \ 1/c \text{ near } a \,).$$

Assume that $g \ \to^* \ c$ near a. We wish to show: $1/g \ \to^* \ 1/c$ near a. That is, we wish to show: either [$1/c = ☹$] or

[($1/c \in \mathbb{R}^*$) and ($1/g \ \to \ 1/c$ near a)].

One of the following must be true:

(1) $c = ☹$ or
(2) $c \neq ☹$.

Proof in Case 1: It suffices to show that $1/c = ☹$.

Since $c = ☹$, it follows that $1/c = ☹$. *End of proof in Case 1.*

Proof in Case 2: Since $g \;\to^*\; c$ near a, and $c \neq ☹$, we conclude that

$$[\ (\ c \in \mathbb{R}^*\ )\ \text{and}\ (\ g \;\to\; c\ \text{near}\ a\ )\ ].$$

Since $c \in \mathbb{R}^*$, one of the following must be true:

(A) $c = -\infty$,
(B) $c = 0$,
(C) $c = \infty$ or
(D) $c \in \mathbb{R}\backslash\{0\}$.

Case A is UNPROVED.

Case B is "easy" because, in that case, $1/c = ☹$.

Case C follows from HW#57.

Proof in Case D: It suffices to show that

$$[\ (\ 1/c \in \mathbb{R}^*\ )\ \text{and}\ (\ 1/g \;\to\; 1/c\ \text{near}\ a\ )\ ].$$

As $c \in \mathbb{R}\backslash\{0\}$, we get $1/c \in \mathbb{R}^*$. It remains to prove: $1/g \;\to\; 1/c$ near a. We wish to show: $\forall \varepsilon > 0$, $\exists \Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \text{dom}[1/g]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad \left[\ \left| \left[\left(\frac{1}{g}\right)(x) \right] - \frac{1}{c} \right| \ <\ \varepsilon\ \right].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \text{dom}[1/g]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad \left[\ \left| \left[\left(\frac{1}{g}\right)(x) \right] - \frac{1}{c} \right| \ <\ \varepsilon\ \right].$$

Let $b := |c|$. Since $c \in \mathbb{R}\backslash\{0\}$, we see that $b > 0$ and $c^2 > 0$. Let $\lambda := \min\{\, b/2\,,\, c^2\varepsilon/2\,\}$. Then $\lambda > 0$ and $\lambda \leqslant b/2$ and $2\lambda \leqslant c^2\varepsilon$. Recalling that $1/g \;\to\; 1/c$ near a, choose $\Delta \in \mathcal{N}^\times(a)$ s.t., $\forall x \in \text{dom}[g]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad [\ |\,[g(x)] - c\,| \ <\ \lambda\ ].$$

We wish to show: $\forall x \in \text{dom}[1/g]$,

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad \left[\ \left| \left[\left(\frac{1}{g}\right)(x) \right] - \frac{1}{c} \right| \ <\ \varepsilon\ \right].$$

Let $x \in \text{dom}[1/g]$ be given. We wish to show:

$$[\ x \in \Delta\ ] \quad \Rightarrow \quad \left[\ \left| \left[\left(\frac{1}{g}\right)(x) \right] - \frac{1}{c} \right| \ <\ \varepsilon\ \right].$$

Assume $x \in \Delta$. We wish to show: $\left| \left[\left(\frac{1}{g} \right) (x) \right] - \frac{1}{c} \right| < \varepsilon$.

We have $x \in \mathrm{dom}[1/g] \subseteq \mathrm{dom}[g]$ and $x \in \Delta$, so, by the choice of Δ, we conclude that $|[g(x)] - c| < \lambda$. Let $c_1 := g(x)$. Then $|c_1 - c| < \lambda$, and we wish to show: $\left| \frac{1}{c_1} - \frac{1}{c} \right| < \varepsilon$.

Since $|\bullet|$ is distance-semidecreasing, we get $|\,|c_1| - |c|\,| \leqslant |\,c_1 - c\,|$. Recall that $b = |c|$. Then $b^2 = |c|^2 = c^2$. Let $b_1 := |c_1|$. Then

$$|\,b_1 - b\,| = |\,|c_1| - |c|\,| \leqslant |\,c_1 - c\,| < \lambda,$$

so $|b_1 - b| < \lambda$, so $b - \lambda < b_1 < b + \lambda$. Since $\lambda \leqslant b/2$, we conclude that $b - \lambda \geqslant b - (b/2) = b/2$. Then $b_1 > b - \lambda \geqslant b/2$. Also, we have $|c - c_1| = |c_1 - c| < \lambda$. Since $b_1 > b/2$ and $b > 0$ and $0 \leqslant |c - c_1| < \lambda$, we get $\frac{|c - c_1|}{b_1 \cdot b} < \frac{\lambda}{(b/2) \cdot b}$. Recall: $2\lambda \leqslant c^2\varepsilon$ and $b^2 = c^2 > 0$. Then

$$\begin{aligned} \left| \frac{1}{c_1} - \frac{1}{c} \right| &= \left| \frac{c - c_1}{c_1 c} \right| = \frac{|c - c_1|}{|c_1| \cdot |c|} = \frac{|c - c_1|}{b_1 \cdot b} \\ &< \frac{\lambda}{(b/2) \cdot b} = \frac{2\lambda}{b^2} \leqslant \frac{c^2 \varepsilon}{c^2} = \varepsilon, \end{aligned}$$

as desired. *End of proof in Case D. End of proof in Case 2. End of partial proof.*

THEOREM 20.6. *Let Y be a topological space. Let $f, g : Y \dashrightarrow \mathbb{R}^*$. Then, for all b, c, we have:*

$$[\ (\ f \to^* b \textit{ near } a\ )\ \&\ (\ g \to^* c \textit{ near } a\ )\ ] \quad \Rightarrow \quad [\ f/g \to^* b/c \textit{ near } a\ ].$$

Proof. Let b and c be given. We wish to prove:

$$[\ (\ f \to^* b \text{ near } a\ )\ \&\ (\ g \to^* c \text{ near } a\ )\ ] \quad \Rightarrow \quad [\ f/g \to^* b/c \text{ near } a\ ].$$

Assume: $(\ f \to^* b$ near $a\ )\ \&\ (\ g \to^* c$ near $a\ )$. We wish to prove:

$f/g \to^* b/c$ near a.

Since $g \to^* c$ near a, by Lemma 20.5, we see that $1/g \to^* 1/c$ near a. So, since $f \to^* b$ near a, by Theorem 20.4, we get $f \cdot (1/g) \to^* b \cdot (1/c)$. So, as $f/g = f \cdot (1/g)$ and $b/c = b \cdot (1/c)$, we get $f/g \to^* b/c$ near a. $\square$

DEFINITION 20.7. *Let P be a set and let f and g be functions. Then $f = g$ on P means: $\forall x \in P$, $f(x) = g(x)$. Also, $f =^* g$ on P means: $\forall x \in P$, $f(x) =^* g(x)$.*

NOTE TO SELF: Let's change to "$f =^{**} g$ on P" next year.

THEOREM 20.8. *Let Y and Z be topological spaces, let $a \in Y$, $b \in Z$ and let $f, g : Y \dashrightarrow Z$. Assume: $\exists P \in \mathcal{N}^\times(a)$ s.t. $f =^* g$ on P. Assume that $f \to b$ near a. Then $g \to b$ near a.*

Proof. We wish to show: $\forall V \in \mathcal{N}(b)$, $\exists U \in \mathcal{N}^\times(a)$ s.t. $g_*(U) \subseteq V$. Let $V \in \mathcal{N}(b)$ be given. We wish to show: $\exists U \in \mathcal{N}^\times(a)$ s.t. $g_*(U) \subseteq V$.

Choose $P \in \mathcal{N}^\times(a)$ s.t. $f =^* g$ on P. Since $f \to b$ near a, choose $Q \in \mathcal{N}^\times(a)$ s.t. $f_*(Q) \subseteq V$. Let $U := P \cap Q$. Then $U \in \mathcal{N}^\times(a)$, and we wish to show: $g_*(U) \subseteq V$. We wish to show: $\forall x \in \mathrm{dom}[g]$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ g(x) \in V\ ].$$

Let $x \in \mathrm{dom}[g]$ be given. We wish to show:

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ g(x) \in V\ ].$$

Assume: $x \in U$. We wish to show: $g(x) \in V$.

Since $x \in \mathrm{dom}[g]$, we have $g(x) \neq ☹$. Since $x \in U = P \cap Q \subseteq P$, and since $f =^* g$ on P, we get $f(x) =^* g(x)$. So, since $g(x) \neq ☹$, we get $f(x) = g(x)$. Since $f(x) = g(x) \neq ☹$, we get $x \in \mathrm{dom}[f]$. Since $x \in \mathrm{dom}[f]$ and $x \in U = P \cap Q \subseteq Q$, we get $f(x) \in f_*(Q)$. Then $g(x) = f(x) \in f_*(Q) \subseteq V$, as desired. □

21. Class 21 on 21 November 2017, Tu of Week 12

DEFINITION 21.1. *Let Y be a set. Let $f, g : Y \dashrightarrow \mathbb{R}^*$. Let $P \subseteq Y$. Then $f \leqslant g$ on P means: $\forall x \in P$, $☹ \neq f(x) \leqslant g(x) \neq ☹$.*

The following is the **Squeeze Theorem**.

THEOREM 21.2. *Let Y be a topological space, $f, g, h : Y \dashrightarrow \mathbb{R}^*$. Let $a \in Y$ and let $b \in \mathbb{R}^*$. Assume: $\exists P \in \mathcal{N}_Y^\times(a)$ s.t. $f \leqslant g \leqslant h$ on P. Assume: ($f \to b$ near a) and ($h \to b$ near a). Then: $g \to b$ near a.*

Partial proof: One of the following must be true:

(1) $b = \infty$ or
(2) $b = -\infty$ or
(3) $b \in \mathbb{R}$.

Case 1 follows from HW#58.

Case 2 is UNPROVED.

Proof in Case 3: Want: $\forall\varepsilon > 0$, $\exists U \in \mathcal{N}^\times(a)$ s.t. $\forall x \in \mathrm{dom}[g]$,

$$[\; x \in U \;] \quad \Rightarrow \quad [\; |\,[\,g(x)\,] - b\,| < \varepsilon \;].$$

Let $\varepsilon > 0$ be given. Want: $\exists U \in \mathcal{N}^\times(a)$ s.t. $\forall x \in \mathrm{dom}[g]$,

$$[\; x \in U \;] \quad \Rightarrow \quad [\; |\,[\,g(x)\,] - b\,| < \varepsilon \;].$$

Choose $P \in \mathcal{N}^\times(a)$ s.t.

$$f \leqslant g \qquad \text{on } P.$$

Since $f \to b$ near a, choose $Q \in \mathcal{N}^\times(a)$ s.t. $\forall x \in \mathrm{dom}[f]$,

$$[\; x \in Q \;] \quad \Rightarrow \quad [\; |\,[\,f(x)\,] - b\,| < \varepsilon \;].$$

Since $h \to b$ near a, choose $R \in \mathcal{N}^\times(a)$ s.t. $\forall x \in \mathrm{dom}[h]$,

$$[\; x \in R \;] \quad \Rightarrow \quad [\; |\,[\,h(x)\,] - b\,| < \varepsilon \;].$$

Let $U := P \cap Q \cap R$. Then $U \in \mathcal{N}^\times(a)$. We wish to show: $\forall x \in \mathrm{dom}[g]$,

$$[\; x \in U \;] \quad \Rightarrow \quad [\; |\,[\,g(x)\,] - b\,| < \varepsilon \;].$$

Let $x \in \mathrm{dom}[g]$ be given. We wish to show:

$$[\; x \in U \;] \quad \Rightarrow \quad [\; |\,[\,g(x)\,] - b\,| < \varepsilon \;].$$

Assume $x \in U$. We wish to show: $|\,[\,g(x)\,] - b\,| < \varepsilon$.

By the choice of P, we have

$$P \subseteq \mathrm{dom}[f] \quad \text{and} \quad \text{and} \quad P \subseteq \mathrm{dom}[g] \quad \text{and} \quad P \subseteq \mathrm{dom}[h].$$

Also, since $x \in U \subseteq P$, by choice of P, we have: $f(x) \leqslant g(x) \leqslant h(x)$.

Since $x \in U \subseteq P \subseteq \mathrm{dom}[f]$ and since $x \in U \subseteq Q$, by choice of Q, we see that $|\,[\,f(x)\,] - b\,| < \varepsilon$. Then $b - \varepsilon < f(x) < b + \varepsilon$.

Since $x \in U \subseteq P \subseteq \mathrm{dom}[h]$ and since $x \in U \subseteq R$, by choice of R, we see that $|\,[\,h(x)\,] - b\,| < \varepsilon$. Then $b - \varepsilon < h(x) < b + \varepsilon$.

Since $b - \varepsilon < f(x) \leqslant g(x)$, we see that $b - \varepsilon < g(x)$.

Since $g(x) \leqslant h(x) < b + \varepsilon$, we see that $g(x) < b + \varepsilon$.

Then $b - \varepsilon < g(x) < b + \varepsilon$. Then $|\,[\,g(x)\,] - b\,| < \varepsilon$, as desired.

End of proof in Case 3. End of partial proof.

DEFINITION 21.3. *Let Y and Z be topological spaces, $f : Y \dashrightarrow Z$. Then, for all a, by f is (Y, Z)-***continuous** *at a we mean:*

$$(\; a \in \mathit{dom}[f] \;) \quad \textit{and} \quad (\; f \to f(a) \textit{ in } Z \textit{ near } a \textit{ in } Y \;).$$

We sometimes simplify and say "continuous" instead of "(Y, Z)-continuous", provided Y and Z are clear. This is very common.

Let $(Y, \mathcal{T})$ and $(Z, \mathcal{U})$ be topological spaces. In cases where it's important to track topologies, we can use "$((Y, \mathcal{T}), (Z, \mathcal{U}))$-continuous" or "$(\mathcal{T}, \mathcal{U})$-continuous" instead of "(Y, Z)-continuous". This is rare.

The next result is **Absoluteness of continuity**:

THEOREM 21.4. *Let Y and Z be topological spaces, let $Y_0 \subseteq Y$ and let $Z_0 \subseteq Z$. Let $f : Y_0 \dashrightarrow Z_0$ and let $a \in dom[f]$. Then:*

$$[\, f \textit{ is } (Y_0, Z_0)\textit{-continuous at } a \,] \;\Leftrightarrow\; [\, f \textit{ is } (Y, Z)\textit{-continuous at } a \,].$$

Proof. By Theorem 18.8 (with b replaced by $f(a)$), we have

$$\begin{aligned} & [\, f \to f(a) \text{ in } Z_0 \text{ near } a \text{ in } Y_0 \,] \\ \Leftrightarrow\; & [\, f \to f(a) \text{ in } Z \text{ near } a \text{ in } Z \,]. \end{aligned}$$

Then

$$\begin{aligned} & [\, f \text{ is } (Y_0, Z_0)\text{-continuous at } a \,] \\ \Leftrightarrow\; & [\, f \to f(a) \text{ in } Z_0 \text{ near } a \text{ in } Y_0 \,] \\ \Leftrightarrow\; & [\, f \to f(a) \text{ in } Z \text{ near } a \text{ in } Z \,] \\ \Leftrightarrow\; & [\, f \text{ is } (Y, Z)\text{-continuous at } a \,], \end{aligned}$$

as desired. $\square$

REMARK 21.5. *Let Y and Z be topological spaces, let $f : Y \dashrightarrow Z$ and let $a \in dom[f]$. Then:*

$$\begin{aligned} & [\quad f \textit{ is continuous at } a \quad] \\ \Leftrightarrow\; & [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}(a) \textit{ s.t. } f_*(U) \subseteq V \quad]. \end{aligned}$$

Proof. We have

$$\begin{aligned}
& [\quad f \text{ is continuous at } a \quad] \\
\Leftrightarrow\ & [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}^\times(a) \text{ s.t. } f_*(U) \subseteq V \quad] \\
\Leftrightarrow\ & [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}^\times(a) \text{ s.t., } \forall x \in \mathrm{dom}[f], \\
& \qquad\qquad (\, x \in U\,) \Rightarrow (\, f(x) \in V\,) \quad] \\
\Leftrightarrow\ & [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}(a) \text{ s.t., } \forall x \in \mathrm{dom}[f], \\
& \qquad\qquad (\, x \in U_a^\times\,) \Rightarrow (\, f(x) \in V\,) \quad] \\
\Leftrightarrow\ & [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}(a) \text{ s.t., } \forall x \in \mathrm{dom}[f], \\
& \qquad\qquad (\, x \in U\,) \Rightarrow (\, f(x) \in V\,) \quad] \\
\Leftrightarrow\ & [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}(a) \text{ s.t. } f_*(U) \subseteq V \quad],
\end{aligned}$$

as desired. □

The main content of Remark 21.5 is that, to check continuity, one may choose to use *UNPUNCTURED* neighborhoods on the input side. The only somewhat hard step in the proof is

$$\begin{aligned}
& [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}(a) \text{ s.t., } \forall x \in \mathrm{dom}[f], \\
& \qquad\qquad (\, x \in U_a^\times\,) \Rightarrow (\, f(x) \in V\,) \quad] \\
\Leftrightarrow\ & [\quad \forall V \in \mathcal{N}(f(a)),\, \exists U \in \mathcal{N}(a) \text{ s.t., } \forall x \in \mathrm{dom}[f], \\
& \qquad\qquad (\, x \in U\,) \Rightarrow (\, f(x) \in V\,) \quad].
\end{aligned}$$

We explained this in terms of playing two games. In both games, one player chooses V, the other chooses U and then the first one chooses x. In the first game, to win, the x-chooser must pick $x \in U_a^\times$. In the second game, to win, the x-chooser must pick $x \in U$. Note that $U \backslash U_a^\times = \{a\}$. So, in the second game, the x-chooser has the possibility of choosing $x = a$, but in the first, that leads to an immediate loss. It may seem that, in switching from the first game to the second, the x-chooser gets a small benefit, but that is illusory because $V \in \mathcal{N}(f(a))$, and so $f(a) \in V$, and so, even in the second game, to win, the x-chooser must not pick $x = a$. This informal discussion can be converted into a proof.

To check continuity, given neighborhood bases (on input and output sides), one may choose to use only basic open neighborhoods:

FACT 21.6. *Let Y and Z be topological spaces, let $f : Y \dashrightarrow Z$ and let $a \in dom[f]$. Let $\mathcal{A}$ be a neighborhood base at a in Y. Let $\mathcal{B}$ be a* *neighborhood base at $f(a)$ in Z. Then:*

$$\begin{aligned} & [\quad f \textit{ is continuous at } a \quad] \\ \Leftrightarrow \quad & [\quad \forall B \in \mathcal{B}, \exists A \in \mathcal{A} \textit{ s.t. } f_*(A) \subseteq B \quad]. \end{aligned}$$

Proof. Proof of $\Leftarrow$*:* Unassigned HW. *End of proof of* $\Leftarrow$*.*

Proof of $\Rightarrow$*:* Assume: $\forall B \in \mathcal{B}$, $\exists A \in \mathcal{A}$ s.t. $f_*(A) \subseteq B$. We wish to show: f is continuous at a. We wish to show: $\forall V \in \mathcal{N}(f(a))$, $\exists U \in \mathcal{N}(a)$ s.t. $f_*(U) \subseteq V$. Let $V \in \mathcal{N}(f(a))$ be given. We wish to show: $\exists U \in \mathcal{N}(a)$ s.t. $f_*(U) \subseteq V$.

Since $\mathcal{B}$ is a neighborhood base at $f(a)$ in Z and since $V \in \mathcal{N}(f(a))$, choose $B \in \mathcal{B}$ s.t. $f(a) \in B \subseteq V$. Then, by assumption, choose $A \in \mathcal{A}$ s.t. $f_*(A) \subseteq B$. As $\mathcal{A}$ is a neighborhood base of a in Y, it follows that $\mathcal{A} \subseteq \mathcal{N}(a)$. So, as $A \in \mathcal{A}$, we get $A \in \mathcal{N}(a)$. Let $U := A$. Then $U \in \mathcal{N}(a)$, and we wish to show: $f_*(U) \subseteq V$.

We have $f_*(U) = f_*(A) \subseteq B \subseteq V$. *End of proof of* $\Rightarrow$*.* □

For any metric space M, if the metric is anonymous, but, at some point, we need it, we will denote it by d_M, or, if M is clear, by d.

Since, for any point p in a metric space $\mathcal{B}(p)$ is a neighborhood base at p, it follows, from Fact 21.6, that we can get a geometric quantified equivalence for continuity in metric spaces:

REMARK 21.7. *Let Y and Z be metric spaces, let $f : Y \dashrightarrow Z$ and let $a \in dom[f]$. Then*

$$\begin{aligned} & [\quad f \textit{ is continuous at } a \quad] \\ \Leftrightarrow \quad & [\quad \forall \varepsilon > 0, \exists \delta > 0 \textit{ s.t. } f_*(B(a,\delta)) \;\subseteq\; B(f(a),\varepsilon) \quad]. \\ \Leftrightarrow \quad & [\quad \forall \varepsilon > 0, \exists \delta > 0 \textit{ s.t.}, \forall x \in dom[f], \\ & \qquad\qquad (\; d_Y(x,a) \;<\; \delta \;) \Rightarrow (\; d_Z(f(x), f(a)) \;<\; \varepsilon \;) \quad]. \end{aligned}$$

Proof. Omitted. □

For any normed vector space V, if the norm is anonymous, but, at some point, we need to use it, we will denote it by $|\bullet|_V$, or, if V is clear, by $|\bullet|$.

Remark 21.7 yields an analytic quantified equivalence for continuity in normed vector spaces:

REMARK 21.8. *Let Y, Z be normed vector spaces, let $f : Y \dashrightarrow Z$ and let $a \in dom[f]$. Then*

$$\begin{aligned} &[\quad f \text{ is continuous at } a \quad] \\ \Leftrightarrow\ &[\quad \forall \varepsilon > 0, \exists \delta > 0 \text{ s.t.}, \forall x \in dom[f], \\ &\qquad (\ |x - a|_Y < \delta\ ) \Rightarrow (\ |[f(x)] - [f(a)]|_Z < \varepsilon\ ) \quad]. \end{aligned}$$

Remark 21.5 and Remark 21.7 and Remark 21.8 give quantified equivalencies for continuity of maps

from topological spaces to topological spaces,
from metric spaces to metric spaces and
from normed vector spaces to normed vector spaces,

respectively. There are also "mixed" situations, like maps

from normed vector spaces to topological spaces.

We won't quantify all of them, but we give one such quantification in Remark 21.9 below. Students are expected to be able to figure out the others, as needed.

REMARK 21.9. *Let Y be a normed vector space and let Z be a topological space. Let $f : Y \dashrightarrow Z$ and let $a \in dom[f]$. Then*

$$\begin{aligned} &[\quad f \text{ is continuous at } a \quad] \\ \Leftrightarrow\ &[\quad \forall V \in \mathcal{N}(f(a)), \exists \delta > 0 \text{ s.t.}, \forall x \in dom[f], \\ &\qquad (\ |x - a| < \delta\ ) \Rightarrow (\ f(x) \in V\ ) \quad]. \end{aligned}$$

Recall that, for any set S, for any a, $C_S^a : S \to \{a\}$ is the **constant fuction** on S with value a, defined by $C_S^a(x) = a$.

FACT 21.10. *Let Y and Z be topological spaces and let $f : Y \dashrightarrow Z$. Let $a \in Y$ and $b \in Z$. Then $C_X^b \to b$ near a.*

Proof. Unassigned HW. □

Assigned HW#59 and HW#60.

THEOREM 21.11. *Let Y be a topological space, let $f, g : Y \dashrightarrow \mathbb{R}^*$ and let $a \in Y$. Assume that both f and g are continuous at a. Let $b := f(a)$ and let $c := g(a)$. Then*

(1) $(\ b + c \neq ☹\ ) \Rightarrow (\ f + g$ *is continuous at* $a\ )$,
(2) $(\ bc \neq ☹\ ) \Rightarrow (\ fg$ *is continuous at* $a\ )$,
(3) $(\ b - c \neq ☹\ ) \Rightarrow (\ f - g$ *is continuous at* $a\ )$ *and*

(4) $(\, b/c \neq ☹ \,) \;\Rightarrow\; (\, f/g$ *is continuous at* $a\,)$.

Proof. Proof of (2),(3),(4): Unassigned HW. *End of proof of (2),(3),(4).*

Proof of (1): Assume: $b + c \neq ☹$. Want: $f + g$ is continuous at a.

By definition of continuity, and by definition of b and c, we have both $f \to b$ near a and $g \to c$ near a. Then, by Theorem 19.19, we get $f + g \to^* b + c$ near a. So, since $b + c \neq ☹$, we get $f + g \to b + c$ near a. So, as $(f + g)(a) = [f(a)] + [g(a)] = b + c$, we see, from the definition of continuity, that $f + g$ is continuous at a. *End of proof of (1).* □

THEOREM 21.12. *Let* Y *be a topological space, let* $a \in Y$*, let* $b \in \mathbb{R}^*$ *and let* $g : Y \dashrightarrow \mathbb{R}^*$*. Assume that* g *is continuous at* a*. Assume that* $(bg)(a) \neq ☹$*. Then* bg *is continuous at* a*.*

Proof. Omitted. □

THEOREM 21.13. *Let* Y *be a topological space, let* Z *be a normed vector space, let* $f, g : Y \dashrightarrow Z$ *and let* $a \in Y$*. Assume* f *and* g *are both continuous at* a*. Then* $f + g$ *is continuous at* a*.*

Proof. Omitted. □

THEOREM 21.14. *Let* Y *be a topological space, let* Z *be a normed vector space, let* $a \in Y$*, let* $g : Y \dashrightarrow \mathbb{R}^*$ *and let* $b \in \mathbb{R}$*. Assume that* g *is continuous at* a*. Then* bg *is continuous at* a*.*

Proof. Omitted. □

DEFINITION 21.15. *Let* f *be a function. Then, for all non-☹* a, b*, the function* $\mathrm{adj}_a^b f : (\mathit{dom}[f]) \cup \{a\} \to (\mathit{im}[f]) \cup \{b\}$ *is defined by*

$$(\mathrm{adj}_a^b f)(x) = \begin{cases} f(x), & \text{if } x \neq a \\ b, & \text{if } x = a. \end{cases}$$

The function $\mathrm{adj}_a^b f$ is the same as f, except that the value at a is adjusted from $f(a)$ to b.

For example, define $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ by

$$f(x) \;=\; \frac{x}{x} \quad \text{and} \quad g(x) = \begin{cases} 1, & \text{if } x \neq 0 \\ 2, & \text{if } x = 0. \end{cases}$$

Note that f and g are both discontinuous (*i.e.*, *NOT* continuous) at 0, and that $g = \mathrm{adj}_0^2 f$. In general, given a function with a discontinuity at 0, like f, we would like, if possible, to adjust the function in such a way as to remove the discontinuity. The function g is an adjustment of f, but that particular adjustment doesn't do anyting to remove the discontinuity. We can however, adjust both f and g in such a way that the discontinuity goes away. Specifically, $\mathrm{adj}_0^1 f = \mathrm{adj}_0^1 g = C_{\mathbb{R}}^1$, and the constant function $C_{\mathbb{R}}^1$ is continuous at 0.

The next result say that if a function has a limit near a, then the limit is unaffected by an adjustment at a.

REMARK 21.16. *let Y and Z be topological spaces, let $f : Y \dashrightarrow Z$, let $a \in Y$ and let $b, c \in Z$. Let $g := \mathrm{adj}_a^b f$. Then*

$$[\, f \to c \textit{ near } a \,] \quad \Leftrightarrow \quad [\, g \to c \textit{ near } a \,].$$

Proof. Let $\mathcal{T} :=$ {open sets in Y}. Then $\mathcal{T}$ is a topology on Y, so $\bigcup \mathcal{T} = Y$. Then $Y \in \langle \mathcal{T} \rangle_\cup = \mathcal{T}$. That is, Y is open in Y. By Remark 16.4 any open set is a neighborhood of each of its points, so $Y \in \mathcal{N}(a)$. Let $P := Y_a^\times$. Then $P \in \mathcal{N}^\times(a)$. By definition of g, we have: $f = g$ on P. Then, as $f \to b$ near a, by Theorem 20.8, we get $[\, f \to c$ near $a \,] \Leftrightarrow [\, g \to c$ near $a \,]$, as desired. □

The next result say that if a function has a limit near a, then, even if it's discontinuous at a, the discontinuity is "removable". We will call this the **Discontinuity Removal Theorem**.

REMARK 21.17. *let Y and Z be topological spaces, let $f : Y \dashrightarrow Z$, let $a \in Y$ and let $b \in Z$. Then:*

$$[\ f \to b \textit{ near } a\ ] \quad \Leftrightarrow \quad [\ \mathrm{adj}_a^b f \textit{ is continuous at } a\ ].$$

Proof. Let $g := \mathrm{adj}_a^b f$. Then $g(a) = b$.

By Remark 21.16, $[\, f \to b$ near $a \,] \Leftrightarrow [\, g \to b$ near $a \,]$. Then

$$\begin{aligned} & [\ f \to b \text{ near } a\ ] \\ \Leftrightarrow\ & [\ g \to b \text{ near } a\ ] \\ \Leftrightarrow\ & [\ g \to g(a) \text{ near } a\ ] \\ \Leftrightarrow\ & [\ g \text{ is continuous at } a\ ] \\ \Leftrightarrow\ & [\ \mathrm{adj}_a^b f \text{ is continuous at } a\ ], \end{aligned}$$

as desired. □

DEFINITION 21.18. *let Y and Z be topological spaces, $f : Y \dashrightarrow Z$. Then, for any set S, by f is* **continuous on** *S, we mean:*

$$(\ S \subseteq \mathit{dom}[f]\ ) \quad \textit{and} \quad (\ \forall a \in S,\ f \textit{ is continuous at } a\ ).$$

By (f is **continuous** *) or (f is* **everywhere continuous** *), we mean: f is continuous on* $\mathrm{dom}[f]$.

We have a quantified equivalence for everywhere continuity:

THEOREM 21.19. *Let Y, Z be topological spaces, $f : Y \to Z$. Then*

$$[\ f \text{ is continuous}\ ] \quad \Leftrightarrow \quad [\ \forall \text{open } V \text{ in } Z,\ f^*(V) \text{ is open in } Y\ ].$$

Note that, in Theorem 21.19, we have $f : Y \to Z$, not $f : Y \dashrightarrow Z$.

Theorem 21.19 asserts: f is everywhere continuous iff "the preimage of any open set is open".

Proof. Proof of $\Leftarrow$: Assume that f is continuous. We wish to show: $\forall$open V in Z, $f^*(V)$ is open in Y. Let an open V in Z be given. We wish to show: $f^*(V)$ is open in Y. By HW#42 (with X replaced by Y, W by $f^*(V)$), we want: $\forall a \in f^*(V), \exists U \in \mathcal{N}(a)$ s.t. $U \subseteq f^*(V)$. Let $a \in f^*(V)$ be given. We wish to show: $\exists U \in \mathcal{N}(a)$ s.t. $U \subseteq f^*(V)$.

Since $a \in f^*(V)$, we get $f(a) \in V$. Also V is open in Z. Then, by Remark 16.4, $V \in \mathcal{N}(f(a))$. So, as f is continuous at a, by Remark 21.5, choose $U \in \mathcal{N}(a)$ s.t. $f_*(U) \subseteq V$. We wish to show: $U \subseteq f^*(V)$.

As $f_*(U) \subseteq V$, $f^*(f_*(U)) \subseteq f^*(V)$. Then $U \subseteq f^*(f_*(U)) \subseteq f^*(V)$, as desired. *End of proof of $\Leftarrow$.*

Proof of $\Rightarrow$: Assume: $\forall$open V in Z, $f^*(V)$ is open in Y. We wish to show: f is continuous. Since $f : Y \to Z$, we have $Y = \mathrm{dom}[f]$. We therefore wish to show: $\forall a \in Y$, f is continuous at a. Let $a \in Y$ be given. We want: f is continuous at a. Then, by Remark 21.5 (with V replaced by W), want: $\forall W \in \mathcal{N}(f(a))$, $\exists U \in \mathcal{N}(a)$ s.t. $f_*(U) \subseteq W$. Let $W \in \mathcal{N}(f(a))$ be given. We want: $\exists U \in \mathcal{N}(a)$ s.t. $f_*(U) \subseteq W$.

By definition of $\mathcal{N}(f(a))$, as $W \in \mathcal{N}(f(a))$, choose an open set V in Z such that $f(a) \in V \subseteq W$. Then, by our assumption, $f^*(V)$ is open in Y. Since $f(a) \in V$, we get $a \in f^*(V)$. Then, by Remark 16.4, we have $f^*(V) \in \mathcal{N}(a)$. Let $U := f^*(V)$. We wish to show: $f_*(U) \subseteq W$.

We have $f_*(U) = f_*(f^*(V)) \subseteq V$, as desired. *End of proof of $\Rightarrow$.* $\square$

22. Class 22 on 28 November 2017, Tu of Week 13

THEOREM 22.1. *Let X and Y and Z be topological spaces. Let $f : X \dashrightarrow Y$ and $g : Y \dashrightarrow Z$. Let $a \in X$. Assume that f is continuous at a and that g is continuous at $f(a)$. Then $g \circ f$ is continuous at a.*

Proof. Since f is continuous at a, we see that $f \to f(a)$ near a. We wish to show: $g \circ f \to (g \circ f)(a)$ near a.

Let $b := f(a)$. Then $f \to b$ near a. Also, $g(b) = g(f(a)) = (g \circ f)(a)$. We therefore wish to show: $g \circ f \to g(b)$ near a.

Since $f \to b$ near a and since g is continuous at b, we conclude, from HW#60, that $g \circ f \to g(b)$ near a, as desired. $\square$

THEOREM 22.2. *Let X and Y and Z all be topological spaces. Let $f : X \dashrightarrow Y$ and $g : Y \dashrightarrow Z$. Assume that f and g are both continuous. Then $g \circ f$ is continuous.*

Proof. Unassigned HW. $\square$

THEOREM 22.3. *Let X and Y be topological spaces, $f : X \dashrightarrow Y$. Let $S \subseteq dom[f]$ and let $a \in S$. Assume that f is continuous at a. Then $f|S$ is continuous at a.*

Proof. We wish to show: $f|S \to (f|S)(a)$ in Y near a in X. Note that $S = \mathrm{dom}[f|S]$. We wish to show: $\forall V \in \mathcal{N}_Y(f(a))$, $\exists U \in \mathcal{N}_X(a)$ s.t., $\forall x \in S$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ f(x) \in V\ ].$$

Let $V \in \mathcal{N}_Y(f(a))$ be given. We wish to show: $\exists U \in \mathcal{N}_X(a)$ s.t., $\forall x \in S$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ f(x) \in V\ ].$$

Since f is continuous at a, we know that $f \to f(a)$ in Y near a in X. So, since $V \in \mathcal{N}_Y(f(a))$, choose $U \in \mathcal{N}_X(a)$ s.t., $\forall x \in \mathrm{dom}[f]$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ f(x) \in V\ ].$$

We wish to show: $\forall x \in S$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ f(x) \in V\ ].$$

Let $x \in S$ be given. We wish to show:

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ f(x) \in V\ ].$$

Assume $x \in U$. We wish to show: $f(x) \in V$.

As $x \in S \subseteq \mathrm{dom}[f]$ and $x \in U$, by choice of U, we have $f(x) \in V$. $\square$

COROLLARY 22.4. *Let X and Y be topological spaces, $f : X \dashrightarrow Y$. Let $S \subseteq dom[f]$. Assume that f is continuous. Then $f|S$ is continuous.*

Proof. Unassigned HW. $\square$

DEFINITION 22.5. *Let s be a sequence and let $K \in \mathbb{N}$. Then the K-**tail** of s is the sequence $(s_K, s_{K+1}, s_{K+2}, \ldots)$.*

That is, the K-tail of s is the sequence $i \mapsto s_{K+i-1} : \mathbb{N} \to \mathrm{im}[s]$.

Sometimes we are sloppy and refer to the set $\{s_K, s_{K+1}, s_{K+2}, \ldots\}$ as the K-tail of s. That is, in conversation, we might sometimes confuse the K-tail with its image. In these notes, however, we will try to be accurate in our terminology.

Assigned HW#61.

DEFINITION 22.6. *Let Y, Z be metric spaces, $f : Y \dashrightarrow Z$. Then*

(1) *f is **uniformly continuous** or (Y, Z)**-uniformly continuous** means: $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall w, x \in dom[f]$,*

$$[\ d_Y(w, x) < \delta\ ] \quad \Rightarrow \quad [\ d_Z(f(w), f(x)) < \varepsilon\ ].$$

(2) *$\forall K \geqslant 0$, f is K**-Lipschitz** or (Y, Z, K)**-Lipschitz** means: $\forall w, x \in dom[f]$,*

$$[\ d_Z(f(w), f(x)) \ \leqslant\ K \cdot [d_Y(w, x)]\ ].$$

(3) *f is **Lipschitz** or (Y, Z)**-Lipschitz** means:*

$$\exists K \geqslant 0 \quad s.t. \quad f \text{ is } K\text{-Lipschitz.}$$

For all three concepts in Definition 22.6, we have absoluteness:

FACT 22.7. *Let Y, Z be MSs, $Y_0 \subseteq Y$, $Z_0 \subseteq Z$, $f : Y_0 \dashrightarrow Z_0$. Then*

(1) *(f is (Y_0, Z_0)-continuous) $\Leftrightarrow$ (f is (Y, Z)-continuous),*
(2) *(f is (Y_0, Z_0)-uniformly continuous) $\Leftrightarrow$ (f is (Y, Z)-uniformly continuous),*
(3) *$\forall K \geqslant 0$, [(f is (Y_0, Z_0, K)-Lipschitz) $\Leftrightarrow$ (f is (Y, Z, K)-Lipschitz)] and*
(4) *(f is (Y_0, Z_0)-Lipschitz) $\Leftrightarrow$ (f is (Y, Z)-Lipschitz).*

Proof. Omitted. □

A function between metric spaces is 0-Lipschitz iff it is constant. A function between metric spaces is called **distance-semiincreasing** or **distance-nondecreasing** if it is 1-Lipschitz.

Let Y, Z be metric spaces, $f : Y \dashrightarrow Z$, $K \geqslant 0$. We observed that the function f is K-Lipschitz iff: $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall w, x \in \mathrm{dom}[f]$,

$$[\ w \neq x\ ] \quad \Rightarrow \quad [\ d_Z(f(w), f(x)) \ \leqslant\ K \cdot [d_Y(w, x)]\ ].$$

Or: f is K-Lipschitz iff: $\forall\varepsilon > 0$, $\exists\delta > 0$ s.t., $\forall w, x \in \mathrm{dom}[f]$,

$$[\ w \neq x\ ] \quad \Rightarrow \quad \left[\ \frac{d_Z(f(w), f(x))}{d_Y(w, x)}\ \leqslant\ K\ \right].$$

Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$, $K \geqslant 0$. Then the function f is K-Lipschitz iff: $\forall\varepsilon > 0$, $\exists\delta > 0$ s.t., $\forall w, x \in \mathrm{dom}[f]$,

$$[\ w \neq x\ ] \quad \Rightarrow \quad \left[\ \left|\frac{[f(w)] - [f(x)]}{w - x}\right|\ \leqslant\ K\ \right].$$

Or: f is K-Lipschitz iff: $\forall\varepsilon > 0$, $\exists\delta > 0$ s.t., $\forall w, x \in \mathrm{dom}[f]$,

$$[\ w \neq x\ ] \quad \Rightarrow \quad \left[\ -K\ \leqslant\ \frac{[f(w)] - [f(x)]}{w - x}\ \leqslant\ K\ \right].$$

Recall that a **secant line** of f is a line that passes through at least two points of the graph of f. A **secant slope** of f is a real number that is the slope of some secant line of f. Then: f is K-Lipschitz iff:

$$-K \quad \leqslant \quad \{\text{secant slopes of } f\} \quad \leqslant \quad K.$$

Let $m, b \in \mathbb{R}$. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = mx + b$. (This kind of function is called "linear" in high school courses and in freshman calculus courses, but is more properly called an "affine" function.) The graph of f is a line, and all of its secant lines are that same line. The set of all secant slopes of f is $\{m\}$. Thus f is $|m|$-Lipschitz.

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 1/(1 + x^2)$. The graph of f is a "fat-tailed bell-curve". We observed that the set of secant slopes of f is both bounded above and bounded below. Then f is Lipschitz. When we want to refer to this function anonymously, we may use: "the function $1/(1 + (\bullet)^2) : \mathbb{R} \to \mathbb{R}$".

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. We observed that the set of secant slopes of f is neither bounded above nor bounded below. Then f is *NOT* Lipschitz. When we want to refer to this function anonymously, we may use: "the function $(\bullet)^2 : \mathbb{R} \to \mathbb{R}$".

Let $f : [0, \infty) \to [0, \infty)$ be defined by $f(x) = \sqrt{x}$. We observed that the set of secant slopes of f is neither bounded above nor bounded below. Then f is *NOT* Lipschitz. When we want to refer to this function anonymously, we may use: "the function $\sqrt{\bullet} : [0, \infty) \to [0, \infty)$".

FACT 22.8. *Let Y, Z be metric spaces, $f : Y \dashrightarrow Z$. Then:*

(1) *(f is Lipschitz) $\Rightarrow$ (f is uniformly continuous), and*
(2) *(f is unformly continuous) $\Rightarrow$ (f is continuous).*

Proof. Omitted. □

Here is a summary of our class discussion about Fact 22.8:

To prove (1), the δ-strategy is: choose a Lipschitz constant $K > 0$ for f and let $\delta := \varepsilon/K$.

The function f is uniformly continuous iff: $\forall\varepsilon > 0$, $\exists\delta > 0$ s.t., $\forall w \in \mathrm{dom}[f]$, $\forall x \in \mathrm{dom}[f]$,

$$[\ d_Y(w,x) < \delta\ ] \quad \Rightarrow \quad [\ d_Z(f(w), f(x)) < \varepsilon\ ].$$

The function f is continuous iff: $\forall w \in \mathrm{dom}[f]$, $\forall\varepsilon > 0$, $\exists\delta > 0$ s.t., $\forall x \in \mathrm{dom}[f]$,

$$[\ d_Y(w,x) < \delta\ ] \quad \Rightarrow \quad [\ d_Z(f(w), f(x)) < \varepsilon\ ].$$

We can commute the universally quantified clauses "$\forall w \in \mathrm{dom}[f]$" and "$\forall\varepsilon > 0$". This shows that f is continuous iff: $\forall\varepsilon > 0$, $\forall w \in \mathrm{dom}[f]$, $\exists\delta > 0$ s.t., $\forall x \in \mathrm{dom}[f]$,

$$[\ d_Y(w,x) < \delta\ ] \quad \Rightarrow \quad [\ d_Z(f(w), f(x)) < \varepsilon\ ].$$

To prove (2), say we have a δ-strategy for uniform continuity. In this strategy, δ depends only on ε. Say we then change the game to continuity, so that δ is allowed to depend both on ε and on w. This only makes life easier on the δ-picker, who can continue with the same strategy as before. That is: any δ-strategy that works for uniform continuity will work for continuity.

This concludes the summary of our class discussion about Fact 22.8.

We leave it as an unassigned exercise to show: $(\bullet)^2 : \mathbb{R} \to \mathbb{R}$ is continuous. In HW#63, you'll show that this function is *NOT* uniformly continuous. Thus the converse to (2) of Fact 22.8 does *NOT* hold.

In Corollary 22.10 below, we will show that $\sqrt{\bullet} : [0,\infty) \to [0,\infty)$ is uniformly continuous. We observed above that this function is *NOT* Lipschitz. Thus the converse of (1) of Fact 22.8 does *NOT* hold.

LEMMA 22.9. *Let $w, x \geqslant 0$ and let $\delta > 0$. Assume that $|w - x| < \delta$. Then $|\sqrt{w} - \sqrt{x}| < \sqrt{\delta}$.*

Proof. Let $a := \min\{w,x\}$, and let $b := \max\{w,x\}$. Then we have $a \leqslant b$ and $|w - x| = b - a$. Then $b - a < \delta$. Also, we have $\sqrt{a} \leqslant \sqrt{b}$ and and $|\sqrt{w} - \sqrt{x}| = \sqrt{b} - \sqrt{a}$. We wish to show: $\sqrt{b} - \sqrt{a} < \sqrt{\delta}$. Assume that $\sqrt{b} - \sqrt{a} \geqslant \sqrt{\delta}$. We aim for a contradiction.

We have $\sqrt{b}+\sqrt{a} \geqslant \sqrt{\delta}+2\sqrt{a} \geqslant \sqrt{\delta}$. Since $\sqrt{b}-\sqrt{a} \geqslant \sqrt{\delta} \geqslant 0$ and $\sqrt{b}+\sqrt{a} \geqslant \sqrt{\delta} \geqslant 0$, we get $[\sqrt{b}-\sqrt{a}][\sqrt{b}+\sqrt{a}] \geqslant [\sqrt{\delta}][\sqrt{\delta}]$. Then $b-a = [\sqrt{b}-\sqrt{a}][\sqrt{b}+\sqrt{a}] \geqslant [\sqrt{\delta}][\sqrt{\delta}] = \delta$, so $b-a \geqslant \delta$, so $\delta \leqslant b-a$. Recall: $b-a < \delta$. Then $\delta \leqslant b-a < \delta$, so $\delta < \delta$. Contradiction. □

COROLLARY 22.10. *The function* $\sqrt{\bullet} : [0,\infty) \to [0,\infty)$ *is uniformly continuous.*

Proof. We wish to show: $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall w, x \in [0,\infty)$,

$$[\ |w-x| < \delta\ ] \quad \Rightarrow \quad [\ |\sqrt{w}-\sqrt{x}| < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \delta > 0$ s.t., $\forall w, x \in [0,\infty)$,

$$[\ |w-x| < \delta\ ] \quad \Rightarrow \quad [\ |\sqrt{w}-\sqrt{x}| < \varepsilon\ ].$$

Let $\delta := \varepsilon^2$. We wish to show: $\forall w, x \in [0,\infty)$,

$$[\ |w-x| < \delta\ ] \quad \Rightarrow \quad [\ |\sqrt{w}-\sqrt{x}| < \varepsilon\ ].$$

Let $w, x \in [0,\infty)$ be given. We wish to show:

$$[\ |w-x| < \delta\ ] \quad \Rightarrow \quad [\ |\sqrt{w}-\sqrt{x}| < \varepsilon\ ].$$

Assume $|w-x| < \delta$. We wish to show: $|\sqrt{w}-\sqrt{x}| < \varepsilon$.

As $w, x \geqslant 0$, $\delta > 0$ and $|w-x| < \delta$, by Lemma 22.9, $|\sqrt{w}-\sqrt{x}| < \sqrt{\delta}$. Also, $\sqrt{\delta} = \sqrt{\varepsilon^2} = \varepsilon$. Then $|\sqrt{w}-\sqrt{x}| < \sqrt{\delta} = \varepsilon$, as desired. □

Assigned HW#62, HW#63 and HW#64.

DEFINITION 22.11. *For any function* f*, for any set* $\mathcal{S}$ *of sets,*

$$\begin{aligned} f_{**}(\mathcal{S}) &:= \{f_*(A) \,|\, A \in \mathcal{S}\} \qquad \textit{and} \\ f^{**}(\mathcal{S}) &:= \{f^*(A) \,|\, A \in \mathcal{S}\}. \end{aligned}$$

We have a "**functoriality property**": For any functions f and g, for any set A, $(g \circ f)_*(A) = g_*(f_*(A))$ and $(g \circ f)^*(A) = f^*(g^*(A))$. Consequently, for any functions f and g, for any set $\mathcal{S}$ of sets, we have: $(g \circ f)_{**}(\mathcal{S}) = g_{**}(f_{**}(\mathcal{S}))$ and $(g \circ f)^{**}(\mathcal{S}) = f^{**}(g^{**}(\mathcal{S}))$.

REMARK 22.12. *Let* f *be a one-to-one function,* $g := f^{-1}$*. Then*

(1) $\forall$*set* A*, [(* $f_*(A) = g^*(A)$ *) and (* $f^*(A) = g_*(A)$ *)]* *and*
(2) $\forall$*set* $\mathcal{S}$ *of sets, [(* $f_{**}(\mathcal{S}) = g^{**}(\mathcal{S})$ *) and (* $f^{**}(\mathcal{S}) = g_{**}(\mathcal{S})$ *)].*

Proof. Omitted. □

FACT 22.13. *Let* $(Y, \mathcal{T})$ *and* $(Z, \mathcal{U})$ *be topological spaces,* $f : Y \to Z$*. Then:* $[\ f \textit{ is continuous}\ ] \quad \Leftrightarrow \quad [\ f^{**}(\mathcal{U}) \subseteq \mathcal{T}\ ].$

Proof. This follows from Theorem 21.19. $\square$

In Fact 22.13, we need "$f : Y \to Z$" and *NOT* "$f : Y \dashrightarrow Z$".

In Fact 22.13, if Y or Z or $\mathcal{T}$ or $\mathcal{U}$ is unclear, we might replace "continuous" by any of

- (Y, Z)-continuous,
- $((Y, \mathcal{T}), (Z, \mathcal{U}))$-continuous or
- $(\mathcal{T}, \mathcal{U})$-continuous.

DEFINITION 22.14. *Let Y, Z be topological spaces, $f : Y \dashrightarrow Z$. Then f is* **open** *or (Y, Z)*-**open** *means:*

$$\forall \text{open } U \text{ in } Y,\ f_*(U) \text{ is open in } Z.$$

FACT 22.15. *Let $(Y, \mathcal{T})$ and $(Z, \mathcal{U})$ be topological spaces, $f : Y \dashrightarrow Z$. Then:* $[\ f \text{ is open}\ ] \quad \Leftrightarrow \quad [\ f_{**}(\mathcal{T}) \subseteq \mathcal{U}\ ]$.

Proof. This follows from Definition 22.14. $\square$

Fact 22.15, holds even for *PARTIAL* functions $f : Y \dashrightarrow Z$.

In Fact 22.15, if Y or Z or $\mathcal{T}$ or $\mathcal{U}$ is unclear, we might replace "open" by any of

- (Y, Z)-open,
- $((Y, \mathcal{T}), (Z, \mathcal{U}))$-open or
- $(\mathcal{T}, \mathcal{U})$-open.

REMARK 22.16. *Let Y, Z be topological spaces, $f : Y \hookrightarrow\!\!> Z$. Then:*

(1) *([f is open] iff [f^{-1} is continuous]) and*
(2) *([f is continuous] iff [f^{-1} is open]).*

Proof. This follows from Fact 22.15 and Remark 22.12. $\square$

DEFINITION 22.17. *Let Y and Z be topological spaces. Then*

(1) *$\forall f$, by f is a* **homeomorphism** *from Y onto Z, we mean:*
$$f : Y \hookrightarrow\!\!> Z \text{ and } f \text{ is continuous and } f \text{ is open;} \quad \text{and}$$
(2) *by Y and Z are* **homeomorphic**, *we mean:*
$$\exists f \text{ s.t. } f \text{ is a homeomorphism from } Y \text{ onto } Z.$$

When Y and Z are clear, we might simply say that f is a homeomorphism, and leave it to the reader to know the domain and image of f.

Instead of "Y and Z are homeomorphic", we might say "Y is homeomorphic to Z" or, equivalently, "Z is homeomorphic to Y".

REMARK 22.18. *Let $(Y, \mathcal{T})$ and $(Z, \mathcal{U})$ be topological spaces and let $f : Y \hookrightarrow\!\!> Z$. Then we have:*

$$\begin{aligned}
[\ f \textit{ is a homeomorphism from } Y \textit{ onto } Z\ ] \quad &\Leftrightarrow \quad [\ f_{**}(\mathcal{T}) = \mathcal{U}\ ] \\
&\Leftrightarrow \quad [\ f^{**}(\mathcal{U}) = \mathcal{T}\ ].
\end{aligned}$$

Proof. Let $g := f^{-1}$. By functoriality, $[f_{**}(\mathcal{T}) = \mathcal{U}] \Leftrightarrow [\mathcal{T} = g_{**}(\mathcal{U})]$, so, by Remark 22.12, we conclude that $[f_{**}(\mathcal{T}) = \mathcal{U}] \Leftrightarrow [\mathcal{T} = f^{**}(\mathcal{U})]$. It remains to show:

$$[\ f \text{ is a homeomorphism from } Y \text{ onto } Z\ ] \quad \Leftrightarrow \quad [\ f_{**}(\mathcal{T}) = \mathcal{U}\ ].$$

By functoriality, $[f_{**}(\mathcal{T}) \subseteq \mathcal{U}] \Leftrightarrow [\mathcal{T} \subseteq g_{**}(\mathcal{U})]$, so, by Remark 22.12, we conclude that $[f^{**}(\mathcal{U}) \subseteq \mathcal{T}] \Leftrightarrow [\mathcal{U} \subseteq f_{**}(\mathcal{T})]$. Then, by Fact 22.13, $[f \text{ is continuous}] \Leftrightarrow [\mathcal{U} \subseteq f_{**}(\mathcal{T})]$.

By Fact 22.15, $[f \text{ is open}] \Leftrightarrow [f_{**}(\mathcal{T}) \subseteq \mathcal{U}]$. Then

$$\begin{aligned}
&[\ f \text{ is a homeomorphism from } Y \text{ onto } Z\ ] \\
\Leftrightarrow\ &[\ (\ f \text{ is continuous}\ )\ \&\ (\ f \text{ is open}\ )\ ] \\
\Leftrightarrow\ &[\ (\ \mathcal{U} \subseteq f_{**}(\mathcal{T})\ )\ \&\ (\ f_{**}(\mathcal{T}) \subseteq \mathcal{U}\ )\ ] \\
\Leftrightarrow\ &[\ f_{**}(\mathcal{T}) = \mathcal{U}\ ].
\end{aligned}$$

as desired. □

In Remark 22.18, we need "$f : Y \hookrightarrow\!\!> Z$". In Remark 22.18, we might replace "homeomorphism from Y onto Z" by any of

- (Y, Z)-homeomorphism,
- $((Y, \mathcal{T}), (Z, \mathcal{U}))$-homeomorphism or
- $(\mathcal{T}, \mathcal{U})$-homeomorphism.

A homeomorphism between topological spaces can be thought of as a dictionary that allows us to translate any topological information about one of the spaces into topological information about the other. So, for example, if two topological spaces are homeomorphic and one of them is metrizable, then the other one must be as well. To a topologist, when two topological spaces are homeomorphic they are "essentially the same".

Assigned HW#65.

The basic idea in HW#65 is that, when we $[1, 2)$ and $[2, 3)$ together, we are performing a continuous operation. However, when we tear $[5, 7)$

apart, we are doing something discontinuous. Gluing is continuous; tearing apart is discontinuous.

Reviewing Definition 22.17, one might wonder: Is any continuous bijection is automatically open? We can answer this "No" as follows. Let f be the function in HW#65. By (a) of HW#65, f is a bijection. By (b) of HW#65, f is continuous. By (c) of HW#65 and by (1) of Remark 22.16, f is not open.

We now look at some important homeomorphisms.

REMARK 22.19. *Let* $C^\times := \{(u,v) \in \mathbb{R}^2 \,|\, (u^2 + v^2 = 1) \& (v \neq 1)\}$. *Then* $C^\times$ *is homeomorphic to* $\mathbb{R}$.

Proof. Define $\sigma : C^\times \to \mathbb{R}$ and $\tau : \mathbb{R} \to C^\times$ by

$$\sigma(u,v) \quad = \quad \frac{u}{1-v} \quad \text{ and } \quad \tau(x) \quad = \quad \left(\frac{2x}{x^2+1}, \frac{x^2-1}{x^2+1} \right).$$

It suffices to show that σ is a homeomorphism.

We leave it as unassigned homework to show:

(A) σ is continuous,
(B) τ is continuous,
(C) $\sigma \circ \tau = \mathrm{id}_{\mathbb{R}}$ and
(D) $\tau \circ \sigma = \mathrm{id}_{C^\times}$.

From (C) and (D), we see that $\sigma : C^\times \hookrightarrow\!\!> \mathbb{R}$ and that $\tau = \sigma^{-1}$. Then, from (B) and from (1) of Remark 22.16, we see that σ is open. Then, by (A), σ is a homeomorphism. □

The mapping σ in the proof of Remark 22.19 is called **stereographic projection**.

REMARK 22.20. *Let* I *be the interval* $(-1,1)$. *Then* I *is homeomorphic to* $\mathbb{R}$.

Proof. Define $f : I \to \mathbb{R}$ and $g : \mathbb{R} \to I$ by

$$f(t) \quad = \quad \frac{t}{\sqrt{1-t^2}} \quad \text{ and } \quad g(x) \quad = \quad \frac{x}{\sqrt{1+x^2}}.$$

It suffices to show that f is a homeomorphism.

We leave it as unassigned homework to show:

(A) f is continuous,
(B) g is continuous,
(C) $f \circ g = \mathrm{id}_{\mathbb{R}}$ and

(D) $g \circ f = \mathrm{id}_I$.

From (C) and (D), we see that $f : I \hookrightarrow\!\!> \mathbb{R}$ and that $g = f^{-1}$. Then, from (B) and from (1) of Remark 22.16, we see that f is open. Then, by (A), f is a homeomorphism. $\square$

REMARK 22.21. *Let I^* be the interval $[-1, 1]$. Then I^* is homeomorphic to $\mathbb{R}^*$.*

Proof. Let I be the interval $(-1, 1)$.

Define $f : I \to \mathbb{R}$ and $g : \mathbb{R} \to I$ by

$$f(t) \;=\; \frac{t}{\sqrt{1-t^2}} \quad \text{and} \quad g(x) \;=\; \frac{x}{\sqrt{1+x^2}}.$$

Define $f^* : I^* \to \mathbb{R}^*$ by

$$f^*(t) \;=\; \begin{cases} -\infty, & \text{if } t = -1 \\ f(t), & \text{if } t \in I \\ \infty, & \text{if } t = 1, \end{cases}$$

and define $g^* : \mathbb{R}^* \to I^*$ by

$$g^*(x) \;=\; \begin{cases} -1, & \text{if } x = -\infty \\ g(x), & \text{if } x \in \mathbb{R} \\ 1, & \text{if } x = \infty. \end{cases}$$

It suffices to show that f^* is a homeomorphism.

We leave it as unassigned homework to show:

(A) f^* is continuous,
(B) g^* is continuous,
(C) $f^* \circ g^* = \mathrm{id}_{\mathbb{R}}$ and
(D) $g^* \circ f^* = \mathrm{id}_I$.

From (C) and (D), we see that $f^* : I \hookrightarrow\!\!> \mathbb{R}$ and that $g^* = (f^*)^{-1}$. Then, from (B) and from (1) of Remark 22.16, we see that f^* is open. Then, by (A), f^* is a homeomorphism. $\square$

Let $I^* := [-1, 1]$, as in Remark 22.21. Recall that the standard metric on I^* is $d_{I^*} = (d_{\mathbb{R}})_{I^*} = d_{\mathbb{R}}|(I^* \times I^*)$. Recall that the standard topology on I^* is $\mathcal{T}_{I^*} = \mathcal{T}_{\mathbb{R}^*}|I^*$. Recall that $\mathcal{T}_{d_{I^*}} = \mathcal{T}_{I^*}$. Then I^* is metrizable. So, by Remark 22.21, $\mathbb{R}^*$ is metrizable.

One might ask the question: What is the distance from $-\infty$ to ∞? The functions f^* and g^* (in the proof of Remark 22.21) yield a dictionary between $\mathbb{R}^*$ and I^*. In this dictionary,

- $-\infty \in \mathbb{R}^*$ corresponds to $-1 \in I^*$ and
- $\infty \in \mathbb{R}^*$ corresponds to $1 \in I^*$.

So, using this dictionary to translate d_{I^*} to a metric d on $\mathbb{R}^*$, we would find that the d-distance from $-\infty$ to ∞ is equal to the d_{I^*}-distance from -1 to 1, which is 2.

Uing arctan, we can construct a homemorphism between $\mathbb{R}^*$ and $[-\pi/2, \pi/2]$, in which

- $-\infty \in \mathbb{R}^*$ corresponds to $-\pi/2$ and
- $\infty \in \mathbb{R}^*$ corresponds to $\pi/2$.

This yields a metric on $\mathbb{R}^*$ s.t. the distance from $-\infty$ to ∞ is π.

While $\mathbb{R}^*$ is a metrizable topological space, we don't consider it to be a metric space, because there isn't any *STANDARD* metric on $\mathbb{R}^*$. There are many metrics, and none of them is considered canonical.

23. Class 23 on 30 November 2017, Th of Week 13

DEFINITION 23.1. *Let X be a metric space and S a set. By S is* **bounded** *in X, we mean: $\exists B \in \mathcal{B}_{d_X}$ s.t. $S \subseteq B$.*

When X is clear, we may simply say "bounded", instead of "bounded in X".

DEFINITION 23.2. *Let $S \subseteq \mathbb{R}^*$.*

By S **has a maximum***, we mean:* $\max S \neq ☹$.
By S **has a minimum***, we mean:* $\min S \neq ☹$.

REMARK 23.3. *Let $S \subseteq \mathbb{R}$. Assume that S has a maximum and a minimum. Then S is bounded.*

Proof. Let $a := \min S$ and $b := \max S$. Then $a, b \in S \subseteq \mathbb{R}$ and $a \leqslant S \leqslant b$. Let $c := (a + b)/2$ and let $r := (b - a + 1)/2$. Then $S \subseteq [a, b] \subseteq (c - r, c + r) = B(c, r) \in \mathcal{B}_{\mathbb{R}}$. Then S is bounded in $\mathbb{R}$. $\square$

The converse of Remark 23.3 is not true: Let S be the interval $(-1, 1)$. Then S is bounded in $\mathbb{R}$, but S has neither maximum nor minimum.

DEFINITION 23.4. *Let X be a metric space and let f be a function. By f is X-***bounded***, we mean im$[f]$ is bounded in X.*

If X is clear, we simply say "bounded", instead of "X-bounded".

DEFINITION 23.5. *Let f be a function. Assume that im$[f] \subseteq \mathbb{R}^*$.*

By f **has a maximum***, we mean: im$[f]$ has a maximum.*
By f **has a minimum***, we mean: im$[f]$ has a minimum.*

REMARK 23.6. *Let f be a function. Assume that im$[f] \subseteq \mathbb{R}$. Assume that f has a maximum and a minimum. Then f is bounded.*

Proof. Let $S := \mathrm{im}[f]$. Then S has a maximum and a minimum, so, by Remark 23.3, S is bounded. Then f is bounded. □

The converse of Remark 23.6 is not true: Let $S := (-1, 1)$, $f := \mathrm{id}_S$. Then f is bounded, but f has neither maximum nor minimum.

Let $C^\times := \{(u, v) \in \mathbb{R}^2 \,|\, (u^2 + v^2 = 1)\&(v \neq 1)\}$. Let $\sigma : C^\times \to \mathbb{R}$ be stereographic projection. Then σ has neither maximum nor minimum. So not all max-min problems have a solution. (Max-min problems are typically called "optimization problems".) Let $X := \{(x, 0) \,|\, x \in \mathbb{R}\}$ be the horizontal axis in $\mathbb{R}^2$. Since $\sigma : C^\times \to \mathbb{R}$ is a homeomorphism, and since $x \mapsto (x, 0) : \mathbb{R} \to X$ is a homeomoprhism, it follows that $C^\times$ is homeomorphic to X. Since $C^\times$ is bounded in $\mathbb{R}^2$, while X is not, boundedness does not appear to be a topological concept; it is geometric by nature. However, there is something similar to boundedness that topologists study. To understand it, look at the full circle $C := \{(u, v) \in \mathbb{R}^2 \,|\, u^2 + v^2 = 1\}$, instead of the punctured circle $C^\times$. It is difficult to prove, but it turns out that, if a subset S of $\mathbb{R}^2$ is homeomorphic to C, then S is bounded. So, in some sense, C is *SO* bounded that it even appears bounded to a topologist. This kind of "superboundedness" has come to be called "compactness". In these notes we will focus on a related concept called "sequential compactness", to be defined later. Whatever the technical definitions of compact and sequentially compact, keep in mind that, sequential compactness is easier to define and to study, Also, in this course, we only care about metrizable topological spaces, and, in a metrizable topological space, sequential compactness is the same as compactness. To get at all the relevant ideas, we first need to upgrade our understanding of sequences.

DEFINITION 23.7. *Let $s \in (\mathbb{R}^*)^{\mathbb{N}}$. Then*

- *s is* **semiincreasing** *means: $\forall j \in \mathbb{N}$, $s_j \leqslant s_{j+1}$.*
- *s is* **semidecreasing** *means: $\forall j \in \mathbb{N}$, $s_j \geqslant s_{j+1}$.*
- *s is* **strictly increasing** *means: $\forall j \in \mathbb{N}$, $s_j < s_{j+1}$.*
- *s is* **strictly decreasing** *means: $\forall j \in \mathbb{N}$, $s_j > s_{j+1}$.*

FACT 23.8. *Let $s \in (\mathbb{R}^*)^{\mathbb{N}}$. Then all of the following are true:*

(1) *s is semiincreasing iff:* $\forall i, j \in \mathbb{N}$, $[(i \leqslant j) \Rightarrow (s_i \leqslant s_j)]$.
(2) *s is semidecreasing iff:* $\forall i, j \in \mathbb{N}$, $[(i \leqslant j) \Rightarrow (s_i \geqslant s_j)]$.
(3) *s is strictly increasing iff:* $\forall i, j \in \mathbb{N}$, $[(i < j) \Rightarrow (s_i < s_j)]$.
(4) *s is strictly decreasing iff:* $\forall i, j \in \mathbb{N}$, $[(i < j) \Rightarrow (s_i > s_j)]$.

Proof. Omitted. □

Let s be a sequence and $\ell \in \mathbb{N}^{\mathbb{N}}$. Then $\ell : \mathbb{N} \to \mathbb{N}$ and $s : \mathbb{N} \to \operatorname{im}[s]$, so $s \circ \ell : \mathbb{N} \to \operatorname{im}[s]$. In particular, $(s \circ \ell)_\bullet$ is a sequence. Also, for all $j \in \mathbb{N}$, we have: $(s \circ \ell)_j \;=\; (s \circ \ell)(j) \;=\; s(\ell(j)) \;=\; s_{\ell_j}$.

We computed $(2, 4, 6, 8, \ldots) \circ (3, 4, 3, 4, 3, 4, \ldots) \;=\; (6, 8, 6, 8, 6, 8, \ldots)$. Precomposition of $(2, 4, 6, 8, \ldots)$ with $(3, 4, 3, 4, 3, 4, \ldots)$ creates a new sequence out of $(2, 4, 6, 8, \ldots)$ by writing down the 3rd term, then 4th, then 3rd, then 4th, then 3rd, then 4th, *etc.* That is, we write down 6, then 8, then 6, then 8, then 6, then 8, *etc.*, yielding $(6, 8, 6, 8, 6, 8, \ldots)$.

DEFINITION 23.9. *Let s and t be sequences. By $t_\bullet$ is a* **subsequence** *of $s_\bullet$, we mean: $\exists$strictly increasing $\ell \in \mathbb{N}^{\mathbb{N}}$ s.t. $t_\bullet = (s \circ \ell)_\bullet$.*

In our preceding example, $(6, 8, 6, 8, 6, 8, \ldots)$ is *NOT* a subsequence of $(2, 4, 6, 8, \ldots)$ because $(3, 4, 3, 4, 3, 4, \ldots)$ is not strictly increasing. By contrast, $(8, 12, 16, 20, 24, 28, 32, \ldots)$ *IS* a subsequence of $(2, 4, 6, 8, \ldots)$ because it is obtained by precomposing $(2, 4, 6, 8, \ldots)$ with the strictly increasing sequence $(4, 6, 8, 10, 12, 14, 16, \ldots)$.

Intuitively, you can think of a subsequence of $(s_1, s_2, s_3, \ldots)$ as being obtained by circling infinitely many of the terms of $(s_1, s_2, s_3, \ldots)$ and then discarding all the others, leaving only the circled terms, in their original order. Then erase all the circles.

For any sequence s, for any $K \in \mathbb{N}$, (the K-tail of $s_\bullet$) is equal to

$$(s_K, s_{K+1}, s_{K+2}, \ldots) \quad = \quad s_\bullet \;\circ\; .(K, K+1, K+2, \ldots),$$

so (the K-tail of $s_\bullet$) is a subsequence of $s_\bullet$.

REMARK 23.10. *Let $\ell, m \in \mathbb{N}^{\mathbb{N}}$. Then $(\ell \circ m)_\bullet \in \mathbb{N}^{\mathbb{N}}$ and*

$$(\ell_\bullet, m_\bullet \textit{ are strictly increasing}) \Rightarrow ((\ell \circ m)_\bullet \textit{ is strictly increasing}).$$

Proof. Since $m : \mathbb{N} \to \mathbb{N}$ and $\ell : \mathbb{N} \to \mathbb{N}$, it follows that $\ell \circ m : \mathbb{N} \to \mathbb{N}$, *i.e.*, that $(\ell \circ m)\bullet \in \mathbb{N}^{\mathbb{N}}$. It remains to show:

$$(\ell_\bullet, m_\bullet \text{ are strictly increasing}) \Rightarrow ((\ell \circ m)_\bullet \text{ is strictly increasing}).$$

Assume $\ell_\bullet, m_\bullet$ are strictly increasing. We wish to show: $(\ell \circ m)_\bullet$ is strictly increasing. By (3) of Fact 23.8, we wish to show:

$$\forall i, j \in \mathbb{N}, \qquad [(i < j) \quad \Rightarrow \quad ((\ell \circ m)_i \ < \ (\ell \circ m)_j)].$$

Let $i, j \in \mathbb{N}$ be given. We wish to show: $(i < j) \Rightarrow ((\ell \circ m)_i < (\ell \circ m)_j)$. Assume $i < j$. We wish to show: $(\ell \circ m)_i < (\ell \circ m)_j$.

As m is strictly increasing and $i < j$, by (3) of Fact 23.8, we have $m_i < m_j$. So, as ℓ is stricly increasing, by (3) of Fact 23.8, we have $\ell_{m_i} < \ell_{m_j}$. That is, $(\ell \circ m)_i < (\ell \circ m)_j$, as desired. □

Similar arguments show:

(A) $\forall \ell, m \in \mathbb{N}^\mathbb{N}$, if ℓ and m are both strictly decreasing, then $\ell \circ m$ is strictly increasing.
(B) $\forall \ell, m \in \mathbb{N}^\mathbb{N}$, if ℓ and m are both semidecreasing, then $\ell \circ m$ is semiincreasing.
(C) $\forall \ell, m \in \mathbb{N}^\mathbb{N}$, if ℓ and m are both semiincreasing, then $\ell \circ m$ is semiincreasing.
(D) $\forall \ell, m \in \mathbb{N}^\mathbb{N}$, if ℓ is semiincreasing and m is semidecreasing, then $\ell \circ m$ is semidecreasing.

For example, to prove (A), we note that, because ℓ and m both reverse strict inequalities, it follows that $\ell \circ m$ will preserve strict inequalities. (If you turn an inequality around twice, it's the same as not turning it around at all.)

COROLLARY 23.11. *Let s be a sequence, let t be a subsequence of $s_\bullet$ and let u be a subsequence of $t_\bullet$. Then $u_\bullet$ is a subsequence of $s_\bullet$.*

Proof. Since $t_\bullet$ is a subsequence of $s_\bullet$, choose a strictly increasing $\ell \in \mathbb{N}^\mathbb{N}$ such that $t_\bullet = (s \circ \ell)_\bullet$. Since $u_\bullet$ is a subsequence of $t_\bullet$, choose a strictly increasing $m \in \mathbb{N}^\mathbb{N}$ such that $u_\bullet = (t \circ m)_\bullet$. Let $n := \ell \circ m$. By Remark 23.10, $n_\bullet \in \mathbb{N}^\mathbb{N}$ and $n_\bullet$ is strictly increasing. We have

$$u \ = \ t \circ m \ = \ (s \circ \ell) \circ m \ = \ s \circ (\ell \circ m) \ = \ s \circ n.$$

That is, $u_\bullet = (s \circ n)_\bullet$. Then $u_\bullet$ is a subsequence of $s_\bullet$. □

In class, we gave an informal proof of the following result. Below, we give a formal proof, using the Principle of Mathematical Induction.

REMARK 23.12. *Let $\ell \in \mathbb{N}^\mathbb{N}$ and assume that $\ell_\bullet$ is strictly increasing. Then, $\forall j \in \mathbb{N}$, $\ell_j \in [j..\infty)$.*

Proof. For all $j \in \mathbb{N}$, let $P_j := [\ell_j \in [j..\infty)]$. Want: $\forall j \in \mathbb{N}$, P_j. We have $\ell_1 \in \mathbb{N} = [1..\infty)$. Then P_1 holds. By the Principle of Mathematical Induction, we wish to show: $\forall j \in \mathbb{N}$, $(P_j \Rightarrow P_{j+1})$. Let $j \in \mathbb{N}$ be given. We wish to show: $P_j \Rightarrow P_{j+1}$. Assume P_j. Want: P_{j+1}. Let $k := j+1$. We wish to show: P_k. Know: $\ell_j \in [j..\infty)$. Want: $\ell_k \in [k..\infty)$.

Since $\ell_j \in [j..\infty)$, it follows that $\ell_j \geqslant j$. Since $\ell_\bullet$ is strictly increasing, it follows that $\ell_{j+1} > \ell_j$. Then $\ell_k = \ell_{j+1} > \ell_j \geqslant j$. Then $\ell_k > j$. So, since $\ell_k, j \in \mathbb{N}$, we get $\ell_k \geqslant j+1$. Then $\ell_k \geqslant j+1 = k$, so $\ell_k \geqslant k$. So, since $\ell_k \in \mathbb{N}$, we get $\ell_k \in [k..\infty)$, as desired. $\square$

The next result says, for example, that the 500th term of a subsequence must be, for some $k \in [500..\infty)$, the kth term of the original sequence. More generally, each term of the subsequence is a "later" term of the sequence. More precisely:

COROLLARY 23.13. *Let s be a sequence and let t be a subsequence of $s_\bullet$. Let $j \in \mathbb{N}$. Then $\exists k \in [j..\infty)$ s.t. $t_j = s_k$.*

Proof. Choose a strictly increasing $\ell \in \mathbb{N}^{\mathbb{N}}$ such that $t_\bullet = (s \circ \ell)_\bullet$. By Remark 23.12, $\ell_j \in [j..\infty)$. Let $k := \ell_j$. We wish to show: $t_j = s_k$.

We have $t_j = (s \circ \ell)_j = s_{\ell_j} = s_k$, as desired. $\square$

It's a consequnce of Corollary 23.13 that the image of a subsequence is a subset of the image of the sequence. Consequently, for any bounded sequence in a metric space, all of its subsequences are bounded as well.

Another consequence of Corollary 23.13 is:

THEOREM 23.14. *Let X be a topological space, $s \in X^{\mathbb{N}}$, $z \in X$. Let t be a subsequence of $s_\bullet$. Assume: $s_\bullet \to z$ in X. Then $t_\bullet \to z$ in X.*

Proof. We wish to show: $\forall U \in \mathcal{N}_X(z)$, $\exists I \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant I\ ] \quad \Rightarrow \quad [\ t_j \in U\ ].$$

Let $U \in \mathcal{N}_X(z)$ be given. We wish to show: $\exists I \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant I\ ] \quad \Rightarrow \quad [\ t_j \in U\ ].$$

Since $s_\bullet \to z$ in X and $U \in \mathcal{N}_X(z)$, choose $I \in \mathbb{N}$ such that, $\forall k \in \mathbb{N}$,

$$[\ k \geqslant I\ ] \quad \Rightarrow \quad [\ s_k \in U\ ].$$

We wish to show: $\forall j \in \mathbb{N}$,

$$[\ j \geqslant I\ ] \quad \Rightarrow \quad [\ t_j \in U\ ].$$

Let $j \in \mathbb{N}$ be given. We wish to show:

$$[\ j \geqslant I\ ] \quad \Rightarrow \quad [\ t_j \in U\ ].$$

Assume that $j \geqslant I$. We wish to show: $t_j \in U$.

By Corollary 23.13, choose $k \in [j..\infty)$ s.t. $t_j = s_k$. Since $k \in \mathbb{N}$ and $k \geqslant j \geqslant I$, by the choice of U, we conclude that $s_k \in U$. Then $t_j = s_k \in U$, as desired. □

DEFINITION 23.15. *Let X be a topological space and let $s \in X^{\mathbb{N}}$. By $s_\bullet$ is* **convergent**, *we mean: $\exists z \in X$ s.t. $s_\bullet \to z$ in X.*

COROLLARY 23.16. *Let X be a topological space, let $s \in X^{\mathbb{N}}$ and let t be a subsequence of $s_\bullet$. Assume that $s_\bullet$ is convergent in X. Then $t_\bullet$ is convergent in X.*

Proof. Since $s_\bullet$ is convergent in X, choose $z \in X$ s.t. $s_\bullet \to z$ in X. By Theorem 23.14, we have $t_\bullet \to z$ in X. Then $t_\bullet$ is convergent in X. □

The converse of Corollary 23.16 is false: Let $s_\bullet = (-1, 1, -1, 1, \ldots)$ and let $t_\bullet = (1, 1, 1, 1, \ldots)$. Then $t_\bullet$ is convergent in $\mathbb{R}$, but $s_\bullet$ is not.

DEFINITION 23.17. *Let X be a topological space, $s \in X^{\mathbb{N}}$. Then by $s_\bullet$ is* **subconvergent** *in X, we mean: there exists a subsequence t of $s_\bullet$ such that $t_\bullet$ is convergent in X.*

Note that any sequence a subsequence of itself, so convergent implies subconvergent. The converse fails: $(-1, 1, -1, 1, \ldots)$ is subconvergent in $\mathbb{R}$, but is not convergent in $\mathbb{R}$.

The sequence $(1, 2, 3, 4, \ldots)$ is not subconvergent in $\mathbb{R}$, and is therefore not convergent in $\mathbb{R}$. However, $(1, 2, 3, 4, \ldots)$ is convergent in $\mathbb{R}^*$, and is therefore subconvergent in $\mathbb{R}^*$.

DEFINITION 23.18. *Let X be a topological space. By X is* **sequentially compact**, *we mean: $\forall s \in X^{\mathbb{N}}$, $s_\bullet$ is subconvergent in X.*

That is, a topological space is sequentially compact if every sequence has a convergent subsequence.

Because $(1, 2, 3, 4, \ldots)$ is not subconvergent in $\mathbb{R}$, we see that $\mathbb{R}$ is not sequentially compact. Let $C := \{(u, v) \in \mathbb{R}^2 \mid u^2 + v^2 = 1\}$. We eventually wish to show that C *IS* sequentially compact, but that requires a few preliminary results.

DEFINITION 23.19. *Let X be a metric space. By X is* **proper**, *we mean: $\forall$bounded $s \in X^{\mathbb{N}}$, $s_\bullet$ is subconvergent.*

Here is a brief description of the upcoming program in this course:

We will eventually show that $\mathbb{R}$ is proper. So, as we saw earlier, in $\mathbb{R}$, there are sequences with no convergent subsequences, but you can't find a *BOUNDED* sequence with no convergent subsequence.

We will eventually show that $\mathbb{R}^2$ is proper. We will eventually show that every closed bounded subset of a proper metric space is sequentially compact. This will show that C, being a closed bounded subset of $\mathbb{R}^2$, is sequentially compact. Let D be some other subset of $\mathbb{R}^2$ that is homeomorphic to C. Then D is sequentially compact. We will eventually show that every sequentially compact subset of a metric space is closed and bounded in the metric space. Then D is necessarily closed and bounded in $\mathbb{R}^2$. The upshot of this discussion: We will see that C "looks bounded" even to a topologist, because there is no homeomorphic image of it (in *ANY* metric space) that is unbounded.

This concludes our description of the upcoming program.

It will be helpful, going forward to have some dynamical ideas, and the basis of dynamics is iterated maps. That is, given a function f, we can compose it with itself repeatedly, and we develop the following notation to handle the basic bookkeeping of that.

DEFINITION 23.20. *Let f be a function. We define $f_\circ^0 := \mathrm{id}_{dom[f]}$. We define $f_\circ^1 := f$. For all $j \in [2..\infty)$, we define $f_\circ^j := f \circ \cdots \circ f$, the composition of j copies of f. For all $j \in \mathbb{N}_0$, the function $f_\circ^j$ is called the jth* **composition power** *of f.*

The logic purist would not like the informality of "$f_\circ^j = f \circ \cdots \circ f$".

To be more formal, we could rewrite Definition 23.20, as follows: Let f be a function. For all $x \in \mathrm{dom}[f]$, $\forall j \in \mathbb{N}$, let

$$\begin{aligned} \mathrm{ORBS}_f^j(x) \quad &:= \quad \{\, y \in ((\mathrm{dom}[f]) \cup (\mathrm{im}[f]))^{[0..j]} \mid \\ &\qquad\qquad (y_0 = x) \quad \& \quad (\forall i \in [1..j],\, y_i = f(y_{i-1}))\,\}, \\ \mathrm{ENDORBS}_f^j(x) \quad &:= \quad \{\, y_j \in \mathrm{im}[f] \mid y \in \mathrm{ORBS}_f^j(x)\,\}. \end{aligned}$$

For all $x \in \mathrm{dom}[f]$, $\forall j \in \mathbb{N}$, an element of $\mathrm{ORBS}_f^j(x)$ is called a j-length **orbit** of x under f, and, informally, is a tuple $(y_0, \ldots, y_j)$ s.t.

$$y_0 = x, \qquad y_1 = f(y_0), \qquad y_2 = f(y_1), \qquad \cdots \qquad y_j = f(y_{j-1}).$$

It is not hard to show, $\forall x \in \mathrm{dom}[f]$, $\forall j \in \mathbb{N}$, that $\mathrm{ORBS}_f^j(x)$ has at most one element; consequently $\mathrm{ENDORBS}_f^j(x)$ also has at most one element. For all $j \in \mathbb{N}_0$, we define $f_\circ^j : \mathrm{dom}[f] \dashrightarrow \mathrm{im}[f]$ by

$$f_\circ^j(x) \quad = \quad \begin{cases} x, & \text{if } j = 0 \\ \mathrm{ELT}(\mathrm{ENDORBS}_f^j(x)), & \text{if } j \in \mathbb{N}. \end{cases}$$

For an example, define $f : [0, 100] \to \mathbb{R}$ by $f(x) = x + 1$. Then

- $f_\circ^0(0) = 0$,
- $f_\circ^1(0) = 1$,
- $f_\circ^2(0) = 2$,
- $f_\circ^3(0) = 3$,
- $\vdots$
- $f_\circ^{100}(0) = 100$,
- $f_\circ^{101}(0) = 101$ and
- $f_\circ^{102}(0) = ☹$.

We leave it as an unassigned exercise to show: For any set P, for any $f : P \to P$, for any $j \in \mathbb{N}$, we have $f_\circ^j : P \to P$.

We also leave as an unassigned exercise: For any function f, for any x, for any $j \in \mathbb{N}_0$, we have $f(f_\circ^j(x)) = f_\circ^{j+1}(x)$.

In class, we gave an informal proof of the following result. Below, we give a formal proof, using composition powers.

FACT 23.21. *Let $P \subseteq \mathbb{N}$. Assume $\#P = \infty$. Then $\exists \ell \in P^{\mathbb{N}}$ s.t. $\ell_\bullet$ is strictly increasing.*

Proof. As $\#P = \infty$, we get $P \neq \varnothing$. So, by the Well-Ordering of $\mathbb{N}$, P has a minimum. Let $m := \min P$. Then $m \in P \subseteq \mathbb{N}$.

For all $k \in \mathbb{N}$, because $\#P = \infty$ and $\#[1..k] = k < \infty$, it follows that $\#(P \backslash [1..k]) = \infty$, and so $P \backslash [1..k] \neq \varnothing$. So, by the Well-Ordering of $\mathbb{N}$, we see: $\forall k \in \mathbb{N}$, $P \backslash [1..k]$ has a minimum. Define $f : P \to P$ by $f(k) = \min(P \backslash [1..k])$. Then, for all $k \in P$, we have $f(k) \in P \backslash [1..k]$.

As $f : P \to P$, we see, for all $j \in \mathbb{N}_0$, that $f_\circ^j : P \to P$. Define $\ell \in P^{\mathbb{N}}$ by $\ell_j = f_\circ^j(m)$. Want: $\ell_\bullet$ is strictly increasing. We wish to show: $\forall j \in \mathbb{N}$, $\ell_j < \ell_{j+1}$. Let $j \in \mathbb{N}$ be given. We wish to show: $\ell_j < \ell_{j+1}$.

Let $k := \ell_j$. Then $f(k) = f(\ell_j) = f(f_\circ^j(m)) = f_\circ^{j+1}(m) = \ell_{j+1}$. Also, as $f(k) \in P \backslash [1..k] \subseteq \mathbb{N} \backslash [1..k] = (k..\infty)$, we get $f(k) > k$. Then $\ell_j = k < f(k) = \ell_{j+1}$, as desired. $\square$

LEMMA 23.22. *Let $s \in \mathbb{R}^{\mathbb{N}}$. Assume that $s_\bullet$ is semiincreasing and $\mathbb{R}$-bounded. Then $s_\bullet$ is convergent in $\mathbb{R}$.*

Proof. Let $A := \{s_1, s_2, s_3, \ldots\}$ and let $y := \sup A$. Then $A = \mathrm{im}[s]$. As $s_\bullet$ is $\mathbb{R}$-bounded, it follows that A is bounded in $\mathbb{R}$. Choose $M \in \mathbb{R}$ such that $A \leqslant M$. Then $\sup A \leqslant M$. We have

$$-\infty \ < \ s_1 \ \in \ A \ \leqslant \ \sup A \ = \ y,$$

so $-\infty < y$. Also $y = \sup A \leqslant M < \infty$, so $y < \infty$. Then $-\infty < y < \infty$, so $y \in \mathbb{R}$. It suffices to show: $s_\bullet \to y$ in $\mathbb{R}$.

We wish to show: $\forall \varepsilon > 0$, $\exists K \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad [\ |s_j - y| < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists K \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad [\ |s_j - y| < \varepsilon\ ].$$

We have $y - \varepsilon < y = \sup A$, so $y - \varepsilon < \sup A$. Then $\mathrm{NOT}(A \leqslant y - \varepsilon)$. Choose $z \in A$ s.t. $z > y - \varepsilon$. Since $z \in A = \{s_1, s_2, s_3, \ldots\}$, choose $K \in \mathbb{N}$ s.t. $z = s_K$. We wish to show: $\forall j \in \mathbb{N}$,

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad [\ |s_j - y| < \varepsilon\ ].$$

Let $j \in \mathbb{N}$ be given. We wish to show:

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad [\ |s_j - y| < \varepsilon\ ].$$

Assume: $j \geqslant K$. Want: $|s_j - y| < \varepsilon$. Want: $y - \varepsilon < s_j < y + \varepsilon$.

Since $K \leqslant j$, by (1) of Fact 23.8, we see that $s_K \leqslant s_j$. We have $y - \varepsilon < z = s_K \leqslant s_j$, so $y - \varepsilon < s_j$. It remains to show: $s_j < y + \varepsilon$.

We have $s_j \in A \leqslant \sup A = y < y + \varepsilon$, so $s_j < y + \varepsilon$, as desired. □

For any $s \in \mathbb{R}^{\mathbb{N}}$, we define $-s \in \mathbb{R}^{\mathbb{N}}$ by $(-s)_j = -s_j$. We leave it as an unassigned exercise to show: for any $s \in \mathbb{R}^{\mathbb{N}}$, if $s_\bullet$ is semidecreasing, then $(-s)_\bullet$ is semiincreasing. We leave it as an unassigned exercise to show: for any $s \in \mathbb{R}^{\mathbb{N}}$, if $s_\bullet$ is $\mathbb{R}$-bounded, then $(-s)_\bullet$ is $\mathbb{R}$-bounded. We leave it as an unassigned exercise to show: for any $s \in \mathbb{R}^{\mathbb{N}}$, if $(-s)_\bullet$ is convergent in $\mathbb{R}$, then $s_\bullet$ is convergent in $\mathbb{R}$.

DEFINITION 23.23. *Let $s \in (\mathbb{R}^*)^{\mathbb{N}}$. Then*

- $s_\bullet$ *is* **semimonotone** *means:*
  ($s_\bullet$ is semiincreasing) or ($s_\bullet$ is semidecreasing).
- $s_\bullet$ *is* **strictly monotone** *means:*
  ($s_\bullet$ is strictly increasing) or ($s_\bullet$ is strictly decreasing).

LEMMA 23.24. *Let $s \in \mathbb{R}^{\mathbb{N}}$. Assume that $s_\bullet$ is semimonotone and $\mathbb{R}$-bounded. Then $s_\bullet$ is convergent in $\mathbb{R}$.*

Proof. As $s_\bullet$ is semimonotone, we conclude:

(1) $s_\bullet$ is semiincreasing or
(2) $s_\bullet$ is semidecreasing.

Proof in Case 1: By Lemma 23.22, $s_\bullet$ is convergent in $\mathbb{R}$, as desired. *End of proof in Case 1.*

Proof in Case 2: Since $(-s)_\bullet$ is semiincreasing and $\mathbb{R}$-bounded, by Lemma 23.22, $(-s)_\bullet$ is convergent in $\mathbb{R}$. Then $s_\bullet$ is convergent in $\mathbb{R}$, as desired. *End of proof in Case 2.* □

We told a story that indicated how to prove the following result, but delayed the formal proof until the next class.

THEOREM 23.25. *Let $s \in \mathbb{R}^{\mathbb{N}}$. Then there exists a subsequence t of $s_\bullet$ such that $t_\bullet$ is semimonotone.*

Proof. Next class. See Theorem 24.12. □

24. Class 24 on 5 December 2017, Tu of Week 14

Assigned HW#66.

DEFINITION 24.1. *Let Y be a topological space. Let $A \subseteq Y$. Then*

$$\begin{aligned}
\mathrm{Cl}_Y A &:= \bigcap \{\, C \text{ closed in } Y \mid C \supseteq A \,\}, \\
\mathrm{Int}_Y A &:= \bigcup \{\, W \text{ open in } Y \mid W \subseteq A \,\}, \\
\partial_Y A &:= [\mathrm{Cl}_Y A] \setminus [\mathrm{Int}_Y A], \\
\mathrm{Isol}_Y A &:= \{\, z \in Y \mid \exists V \in \mathcal{N}_Y(z) \text{ s.t. } V \cap A = \{z\} \,\}, \qquad \text{and} \\
\mathrm{LP}_Y A &:= [\mathrm{Cl}_Y A] \setminus [\mathrm{Isol}_Y A].
\end{aligned}$$

The sets $\mathrm{Cl}_Y A$, $\mathrm{Int}_Y A$, $\partial_Y A$, $\mathrm{Isol}_Y A$ and $\mathrm{LP}_Y A$ are called the **closure**, **interior**, **boundary**, **isolated set** and **limit point set** of A, respectively. When Y is clear, it can be omitted and we have: $\mathrm{Cl}\, A$, $\mathrm{Int}\, A$, ∂A, $\mathrm{Isol}\, A$ and $\mathrm{LP}\, A$.

We drew a neither-open-nor-closed subset A of $\mathbb{R}^2$ that had three isolated points. We went through and described $\mathrm{Cl}_{\mathbb{R}^2}(A)$, $\mathrm{Int}_{\mathbb{R}^2}(A)$, $\partial_{\mathbb{R}^2}(A)$, $\mathrm{Isol}_{\mathbb{R}^2}(A)$ and $\mathrm{LP}_{\mathbb{R}^2}(A)$.

We did another example: Let $A := [1,2) \cup \{3\}$. Then $A \subseteq \mathbb{R}$ and

$$\begin{aligned} \mathrm{Cl}_{\mathbb{R}} A &= [1,2] \cup \{3\}, \\ \mathrm{Int}_{\mathbb{R}} A &= (1,2) \\ \partial_{\mathbb{R}} A &= \{1,2,3\} \\ \mathrm{Isol}_{\mathbb{R}} A &= \{3\} \qquad \text{and} \\ \mathrm{LP}_{\mathbb{R}} A &= [1,2]. \end{aligned}$$

For any topological space Y, the operator $\mathrm{Cl}_Y : 2^Y \to 2^y$ has the superset property, the monotonicity property and the idempotence property. That is, for any topological space Y, we have:

$$\begin{aligned} (\ \forall A \subseteq Y, &\quad [\ \mathrm{Cl}\, A \supseteq A\ ]\ ), \\ (\ \forall A, B \subseteq Y, &\quad [\ (A \subseteq B) \Rightarrow (\mathrm{Cl}\, A \subseteq \mathrm{Cl}\, B)\ ]\ ) \qquad \text{and} \\ (\ \forall A \subseteq Y, &\quad [\ \mathrm{Cl}(\mathrm{Cl}\, A) \supseteq \mathrm{Cl}\, A\ ]\ ). \end{aligned}$$

Also, for any topological space Y, for any $A \subseteq Y$, we have:

$$(\ A \text{ is closed in } Y\ ) \qquad \text{iff} \qquad (\ \mathrm{Cl}\, A = A\ ).$$

FACT 24.2. *Let Y be a topological space, let $A \subseteq Y$ and let C be a closed subset of Y. Assume that $A \subseteq C$. Then $\mathrm{Cl}\, A \subseteq C$.*

Proof. By monotonicity, $\mathrm{Cl}\, A \subseteq \mathrm{Cl}\, C$. Since C is closed in Y, we have $\mathrm{Cl}\, C = C$. Then $\mathrm{Cl}\, A \subseteq \mathrm{Cl}\, C = C$, as desired. □

Let Y be a topological space, $A \subseteq Y$. An element of $\mathrm{Cl}_Y A$ is called a **closure point** of A in Y. An element of $\mathrm{Int}_Y A$ is called an **interior point** of A in Y. An element of $\partial_Y A$ is called an **boundary point** of A in Y. An element of $\mathrm{Isol}_Y A$ is called an **isolated point** of A in Y. Note: $\forall z \in Y$, we have:

$$\begin{aligned} &(\ z \text{ is an isolated point of } A \text{ in } Y\ ) \qquad \text{iff} \\ &(\ [\ z \in A\ ] \text{ and } [\ \{z\} \text{ is open in } A\ ]\ ) \qquad \text{iff} \\ &(\ [\ z \in A\ ] \text{ and } [\ z \text{ is isolated in } A\ ]\ ). \end{aligned}$$

An element of $\mathrm{LP}_Y A$ is called a **limit point** of A in Y.

Assigned HW#67.

We have the following quantified equivalences, one for closure points, the other for limit points:

FACT 24.3. *Let Y be a topological space, $A \subseteq Y$ and $z \in Y$. Then*

(1) $(z \in \mathrm{Cl}\, A) \Leftrightarrow (\forall V \in \mathcal{N}(z),\ V \cap A \neq \varnothing)$ *and*

(2) $(z \in \mathrm{LP}\, A) \;\Leftrightarrow\; (\forall V \in \mathcal{N}^\times(z),\, V \cap A \neq \varnothing)$.

Proof. Proof of (1): Proof of $\Rightarrow$*:* Homework#67. *End of proof of* $\Rightarrow$*.*

Proof of $\Leftarrow$*:* Assume: $\forall V \in \mathcal{N}(z)$, $V \cap A \neq \varnothing$. We wish to show: $z \in \mathrm{Cl}\, A$. That is, we wish to show:

$$z \quad \in \quad \bigcap \{\, C \text{ closed in } Y \mid C \supseteq A \,\}.$$

Let $\mathcal{S} := \{C \text{ closed in } Y \mid C \supseteq A\}$. We wish to show: $\forall C \in \mathcal{S}$, $z \in C$. Let $C \in \mathcal{S}$ be given. We wish to show $z \in C$. Assume that $z \notin C$. We aim for a contradiction.

Since $C \in \mathcal{S}$, it follows both that C is closed in Y and that $C \supseteq S$. Let $V := Y \backslash C$. Since C is closed in Y, it follows that V is open in Y. So, since $z \in V$, it follows, from Remark 16.4, that $V \in \mathcal{N}(z)$. Then, by assumption, $V \cap A \neq \varnothing$.

We have $A \subseteq C$, so $V \cap A \subseteq V \cap C$. Since $V = Y \backslash C$, it follows that $V \cap C = \varnothing$. Then $V \cap A \subseteq V \cap C = \varnothing$, so $V \cap A = \varnothing$. Contradiction. *End of proof of* $\Leftarrow$*. End of proof of (1).*

Proof of (2): Unassigned homework. *End of proof of (2).* □

There are sequential versions of closure, interior, boundary, isolated set and limit point set. We will only need sequential closure:

DEFINITION 24.4. *Let Y be a topological space and let $A \subseteq Y$. Then* $\mathrm{sCl}_Y A := \{z \in Y \mid \exists s \in A^{\mathbb{N}} \text{ s.t. } s_\bullet \to z \text{ in } Y\}$.

Let Y be a topological space and let $A \subseteq Y$. Then $\mathrm{sCl}_Y A$ is called the **sequential closure** of A in Y. An element of $\mathrm{sCl}_Y A$ is called a **sequential closure point** of A in Y. When Y is clear, we may omit the subscript and use $\mathrm{sCl}\, A$.

For any topological space Y, the operator $\mathrm{sCl}_Y : 2^Y \to 2^Y$ has the superset property and the monotonicity property. That is, for any topological space Y, we have:

$$(\ \forall A \subseteq Y, \qquad [\ \mathrm{Cl}\, A \supseteq A\ ]\ ) \qquad \text{and}$$
$$(\ \forall A, B \subseteq Y, \qquad [\ (A \subseteq B) \;\Rightarrow\; (\mathrm{Cl}\, A \subseteq \mathrm{Cl}\, B)\ ]\ ).$$

WARNING: There are strange topological spaces Y in which a subset $A \subseteq Y$ exists satisfying $\mathrm{Cl}_Y(\mathrm{Cl}_Y A) \neq \mathrm{Cl}_Y A$. That is, for certain Y, the operator $\mathrm{Cl}_Y : 2^Y \to 2^Y$ is *NOT* idempotent. However, we will eventually see, for any *METRIZABLE* topological space Y, that $\mathrm{sCl}_Y = \mathrm{Cl}_Y$. So, for any metrizable topological space Y, the idempotence of Cl_Y implies that of sCl_Y.

We drew a neither-open-nor-closed subset A of $\mathbb{R}^2$ that had three isolated points. We verified that $\mathrm{sCl}_{\mathbb{R}^2}(A) = \mathrm{Cl}_{\mathbb{R}^2}(A)$.

Assigned HW#68.

DEFINITION 24.5. *Let Y be a set, $d \in \mathcal{M}(Y)$, $s \in Y^{\mathbb{N}}$, $z \in Y$. Then $d(s_\bullet, z) \in [0,\infty)^{\mathbb{N}}$ is defined by $(d(s_\bullet, z))_j = d(s_j, z)$.*

Assigned HW#69 and HW#70.

The next definition saves us a bit of writing:

DEFINITION 24.6. *Let Y be a topological space and let f be a function. Assume that dom$[f] \subseteq Y$. Then $\mathrm{LPD}_Y\, f := \mathrm{LP}_Y(\textit{dom}[f])$.*

As usual, we may omit the subscript, and write $\mathrm{LPD}\, f$ instead of $\mathrm{LPD}_Y\, f$, provided Y is clear.

We graphed a function $f : \mathbb{R} \dashrightarrow \mathbb{R}$ s.t. $\mathrm{dom}[f] = [1,2) \cup \{3\}$. We observed that $3 \notin \mathrm{LPD}_{\mathbb{R}}\, f$, and argued

$$\begin{array}{rl} \text{both that} & f \;\to\; 8 \;\text{ near }\; 3 \\ \text{and that} & f \;\to\; 267 \;\text{ near }\; 3. \end{array}$$

In general, we have:

PROPOSITION 24.7. *Let Y, Z be topological spaces, $f : Y \dashrightarrow Z$. Let $a \in Y \backslash (\mathrm{LPD}_Y[f])$. Then: $\forall b \in Z$, $f \to b$ in Z near a in Y.*

Proof. Unassigned HW. □

According to Proposition 24.7, we have extreme NON-uniqueness of limits whenever, on the input side, we try to approach a non-limit point. On the other hand, what possible reason could anyone have in being interested in a limit at a non-limit point? By contrast:

THEOREM 24.8. *Let Y be a topological space, Z a Hausdorff topological space. Let $f : Y \dashrightarrow Z$. Let $a \in \mathrm{LPD}_Y f$, $b, c \in Z$. Assume:*

(1) *$f \to b$ in Z near a in Y and*
(2) *$f \to c$ in Z near a in Y.*

Then $b = c$.

Proof. Assume that $b \neq c$. We aim for a contradiction.

Since Z is Hausdorff, $b, c \in Z$ and $b \neq c$, choose $V \in \mathcal{N}(b)$, $W \in \mathcal{N}(c)$ such that $V \cap W = \varnothing$. Since $f \to b$ near a and $V \in \mathcal{N}(b)$, choose

$P \in \mathcal{N}^\times(a)$ s.t. $f_*(P) \subseteq V$. Since $f \to c$ near a and $W \in \mathcal{N}(c)$, choose $Q \in \mathcal{N}^\times(a)$ s.t. $f_*(Q) \subseteq W$. As $P, Q \in \mathcal{N}^\times(a)$, we get $P \cap Q \in \mathcal{N}^\times(a)$.

Let $D := \mathrm{dom}[f]$. Then $a \in \mathrm{LPD}\, f = \mathrm{LP}\, D$. So, as $P \cap Q \in \mathcal{N}^\times(a)$, we conclude, from $\Rightarrow$ of (2) of Fact 24.3 (with A replaced by D and V by $P \cap Q$), that $P \cap Q \cap D \neq \varnothing$. Choose $x \in P \cap Q \cap D$.

Since $x \in D = \mathrm{dom}[f]$ and $x \in P$, it follows that $f(x) \in f_*(P)$. Since $x \in D = \mathrm{dom}[f]$ and $x \in Q$, it follows that $f(x) \in f_*(Q)$. Since $f(x) \in f_*(P) \subseteq V$ and $f(x) \in f_*(Q) \subseteq W$, it follows that $f(x) \in V \cap W$. Then $V \cap W \neq \varnothing$. Contradiction. ☐

Theorem 24.8, has an *input* hypothesis: $a \in \mathrm{LPD}_Y\, f$. This is a tame hypothesis. After all, what possible reason could anyone have in being interested in a limit at a non-limit point? It also has an *output* hypothesis: Z is Hausdorff. This is also tame. After all, what possible interest could anyone have in a non-Hausdorff topological space?

COROLLARY 24.9. *Let Z be a Hausdorff topological space. Let $s \in Z^{\mathbb{N}}$ and let $b, c \in Z$. Assume:*

(1) *$s_\bullet \to b$ in Z and*
(2) *$s_\bullet \to c$ in Z.*

Then $b = c$.

Proof. Let $Y := \mathbb{N}^*$ and let $a := \infty$. We have $s : Y \dashrightarrow Z$ and

(1') $s \to b$ in Z near a in Y and
(2') $s \to c$ in Z near a in Y.

Since $s \in Z^{\mathbb{N}}$, we have $\mathrm{dom}[s] = \mathbb{N}$. We have

$$\mathrm{LPD}_Y[s] \;=\; \mathrm{LP}_Y(\mathrm{dom}[s]) \;=\; \mathrm{LP}_{\mathbb{N}^*}(\mathbb{N}) \;=\; \{\infty\}.$$

Then $a = \infty \in \{\infty\} = \mathrm{LPD}_Y[s]$.

Then, by Theorem 24.8 (with f replaced by s), $b = c$, as desired. ☐

LEMMA 24.10. *Let Y be a metric space, let $s \in Y^{\mathbb{N}}$ and let t be a subsequence of $s_\bullet$. Assume that $s_\bullet$ is Y-bounded. Then $t_\bullet$ is Y-bounded.*

Proof. As $s_\bullet$ is Y-bounded, we see that $\mathrm{im}[s_\bullet]$ is a bounded subset of Y. As $t_\bullet$ is a subsequence of $s_\bullet$, we see that $\mathrm{im}[t_\bullet] \subseteq \mathrm{im}[s_\bullet]$. It follows that $\mathrm{im}[t_\bullet]$ is a bounded subset of Y. Then $t_\bullet$ is Y-bounded. ☐

LEMMA 24.11. *Let $s \in (\mathbb{R}^2)^{\mathbb{N}}$. Define $p, q : \mathbb{R}^2 \to \mathbb{R}$ by $p(x, y) = x$ and $q(x, y) = y$. Then*

$$\begin{aligned} & [\ s_\bullet \textit{ is } \mathbb{R}^2\textit{-bounded}\ ] \\ \Leftrightarrow\ & [\ ((p \circ s)_\bullet \textit{ is } \mathbb{R}\textit{-bounded})\ \&\ ((q \circ s)_\bullet \textit{ is } \mathbb{R}\textit{-bounded})\ ]. \end{aligned}$$

Proof. Unassigned homework. □

We now pick up where we left off at the end of the last class:

THEOREM 24.12. *Let $s \in \mathbb{R}^{\mathbb{N}}$. Then there exists a subsequence t of $s_\bullet$ such that $t_\bullet$ is semimonotone.*

Proof. Let $P := \{j \in \mathbb{N} \mid s_j \geqslant \{s_{j+1}, s_{j+2}, s_{j+3}, \ldots\}\}$. Then

(A) $\forall j \in P$, $\forall k \in (j..\infty)$, $s_j \geqslant s_k$ and

(B) $\forall j \in \mathbb{N}\backslash P$, $\exists k \in (j..\infty)$ s.t. $s_j < s_k$.

One of the following must be true:

(1) $\#P = \infty$ or

(2) $\#P < \infty$.

Proof in Case 1: By Fact 23.21, choose $\ell \in P^{\mathbb{N}}$ s.t. $\ell_\bullet$ is strictly increasing. Since $P \subseteq \mathbb{N}$, we get $P^{\mathbb{N}} \subseteq \mathbb{N}^{\mathbb{N}}$. Then $\ell_\bullet \in P^{\mathbb{N}} \subseteq \mathbb{N}^{\mathbb{N}}$. Let $t := (s \circ \ell)_\bullet$. Then $t_\bullet$ is a subsequence of $s_\bullet$. We wish to show that $t_\bullet$ is semimonotone. We will show that $t_\bullet$ is semidecreasing. We wish to show: $\forall i \in \mathbb{N}$, $t_i \geqslant t_{i+1}$. Let $i \in \mathbb{N}$ be given. Want: $t_i \geqslant t_{i+1}$.

Since $\ell_\bullet$ is strictly increasing, it follows that $\ell_i < \ell_{i+1}$. Let $j := \ell_i$ and $k := \ell_{i+1}$. Then $j < k$. Since $\ell_\bullet \in P^{\mathbb{N}}$, we get $j, k \in P$. Since $j < k$ and since $j, k \in P \subseteq \mathbb{N}$, we conclude that $k \in (j..\infty)$. Then, by (A), we see that $s_j \geqslant s_k$. Then $t_i = s_{\ell_i} = s_j \geqslant s_k = s_{\ell_{i+1}} = t_{i+1}$, as desired. *End of proof in Case 1.*

Proof in Case 2: Since $P \subseteq \mathbb{N} \subseteq \mathbb{R}$ and $\#P < \infty$, we see that P has a maximum. Let $m := \max P$. Then $P \subseteq [1..m]$. Then $\mathbb{N}\backslash P \supseteq \mathbb{N}\backslash[1..m]$. For all $j \in (m..\infty)$, let $Q_j := \{k \in (j..\infty) \mid s_j < s_k\}$. For all $j \in (m..\infty)$, we have $j \in \mathbb{N}\backslash[1..m] \subseteq \mathbb{N}\backslash P$, so, by (B), we see that $Q_j \neq \varnothing$. Then, for all $j \in (m..\infty)$, we have $\varnothing \neq Q_j \subseteq \mathbb{N}$, so, by Well-Ordering of $\mathbb{N}$, Q_j has a minimum. Define $f : (m..\infty) \to (m..\infty)$ by $f(j) = \min Q_j$. As $f : (m..\infty) \to (m..\infty)$, we see: $\forall i \in \mathbb{N}$, $f_\circ^{i-1} : (m..\infty) \to (m..\infty)$. Define $\ell \in (m..\infty)^{\mathbb{N}}$ by $\ell_i = f_\circ^{i-1}(m+1)$. Since $(m..\infty) \subseteq \mathbb{N}$, we get $(m..\infty)^{\mathbb{N}} \subseteq \mathbb{N}^{\mathbb{N}}$. Then $\ell_\bullet \in (m..\infty)^{\mathbb{N}} \subseteq \mathbb{N}^{\mathbb{N}}$. For all $i \in \mathbb{N}$, we have

$$\ell_{i+1} \ = \ f_\circ^i(m+1) \ = \ f(f_\circ^{i-1}(m)) \ = \ f(\ell_i).$$

Claim: $\ell_\bullet$ is strictly increasing. *Proof of Claim:* We wish to show: $\forall i \in \mathbb{N}$, $\ell_i < \ell_{i+1}$. Let $i \in \mathbb{N}$ be given. We wish to show: $\ell_i < \ell_{i+1}$. Let $j := \ell_i$ and let $k := \ell_{i+1}$. We wish to show: $j < k$.

By definition of Q_j, we have $Q_j \subseteq (j..\infty)$. Then

$$k \;=\; \ell_{i+1} \;=\; f(\ell_i) \;=\; f(j) \;=\; \min Q_j \;\in\; Q_j \;\subseteq\; (j..\infty).$$

Then $k \in (j..\infty)$, so $j < k$, as desired. *End of proof of Claim.*

As $\ell_\bullet \in \mathbb{N}^{\mathbb{N}}$, it follows, from the Claim, that $(s \circ \ell)_\bullet$ is a subsequence of $s_\bullet$. Let $t := (s \circ \ell)_\bullet$. We wish to show: $t_\bullet$ is semimonotone. We will show: $t_\bullet$ is strictly increasing. We wish to show: $\forall i \in \mathbb{N}$, $t_i < t_{i+1}$. Let $i \in \mathbb{N}$ be given. We wish to show: $t_i < t_{i+1}$.

Let $j := \ell_i$ and $k := \ell_{i+1}$. Then

$$k \;=\; \ell_{i+1} \;=\; f(\ell_i) \;=\; f(j).$$

Then $k = f(j) = \min Q_j \in Q_j$. So, by definition of Q_j, we have $s_j < s_k$. Then $t_i = s_{\ell_i} = s_j < s_k = s_{\ell_{i+1}} = t_{i+1}$. *End of proof in Case 2.* □

COROLLARY 24.13. *The topological space* $\mathbb{R}$ *is proper.*

Proof. We wish to show, for any $\mathbb{R}$-bounded $s \in \mathbb{R}^{\mathbb{N}}$, that $s_\bullet$ is subconvergent in $\mathbb{R}$. Let an $\mathbb{R}$-bounded $s \in \mathbb{R}^{\mathbb{N}}$ be given. We wish to show that $s_\bullet$ is subconvergent in $\mathbb{R}$.

By Theorem 24.12, choose a subsequence t of $s_\bullet$ such that $t_\bullet$ is semimontone. Since $t_\bullet$ is a subsequence of $s_\bullet$ and $s_\bullet$ is $\mathbb{R}$-bounded, by Lemma 24.10, $t_\bullet$ is $\mathbb{R}$-bounded. As $t_\bullet$ is semimonotone and $\mathbb{R}$-bounded, by Lemma 23.24 (with s replaced by t), $t_\bullet$ is convergent in $\mathbb{R}$. So, since $t_\bullet$ is a subsequence of $s_\bullet$, we see that $s_\bullet$ is subconvergent in $\mathbb{R}$. □

COROLLARY 24.14. *The topological space* $\mathbb{R}^2$ *is proper.*

Proof. We wish to show, for any $\mathbb{R}^2$-bounded $s \in (\mathbb{R}^2)^{\mathbb{N}}$, that $s_\bullet$ is subconvergent in $\mathbb{R}^2$. Let an $\mathbb{R}^2$-bounded $s \in (\mathbb{R}^2)^{\mathbb{N}}$ be given. We wish to show that $s_\bullet$ is subconvergent in $\mathbb{R}^2$.

Define $p, q : \mathbb{R}^2 \to \mathbb{R}$ by $p(x, y) = x$ and $q(x, y) = y$. Because $s_\bullet$ is $\mathbb{R}^2$-bounded, it follows, from $\Rightarrow$ of Lemma 24.11, that $(p \circ s)_\bullet$ and $(q \circ s)_\bullet$ are both $\mathbb{R}$-bounded. By Corollary 24.13, choose a strictly increasing $\ell \in \mathbb{N}^{\mathbb{N}}$ such that $(p \circ s \circ \ell)_\bullet$ is convergent in $\mathbb{R}$. As $(q \circ s)_\bullet$ is $\mathbb{R}$-bounded and $(q \circ s \circ \ell)_\bullet$ is a subsequence of $(q \circ s)_\bullet$, by Lemma 24.10, we see that $(q \circ s \circ \ell)_\bullet$ is $\mathbb{R}$-bounded. By Corollary 24.13, choose a strictly increasing $m \in \mathbb{N}^{\mathbb{N}}$ such that $(q \circ s \circ \ell \circ m)_\bullet$ is convergent in $\mathbb{R}$.

As $(s\circ\ell\circ m)_\bullet$ is a subsequence of $(s\circ\ell)_\bullet$ and $(s\circ\ell)_\bullet$ is a subsequence of $s_\bullet$, we see, by Corollary 23.11, that $(s\circ\ell\circ m)_\bullet$ is a subsequence of $s_\bullet$. It therefore suffices to show that $(s\circ\ell\circ m)_\bullet$ is convergent in $\mathbb{R}^2$.

Since $(p\circ s\circ\ell)_\bullet$ is convergent in $\mathbb{R}$ and $(p\circ s\circ\ell\circ m)_\bullet$ is a subsequence of $(p\circ s\circ\ell)_\bullet$, by Corollary 23.16, we see that $(p\circ s\circ\ell\circ m)_\bullet$ is convergent in $\mathbb{R}$. Since $(p\circ s\circ\ell\circ m)_\bullet$ and $(q\circ s\circ\ell\circ m)_\bullet$ are both convergent in $\mathbb{R}$, it follows, from $\Leftarrow$ of HW#66, that $(s\circ\ell\circ m)_\bullet$ is convergent in $\mathbb{R}^2$. $\square$

25. Class 25 on 7 December 2017, Th of Week 14

THEOREM 25.1. *Let Y be a metric space and let $A\subseteq Y$. Assume that A is sequentially compact. Then A is closed and bounded in Y.*

Proof. We wish to show:

(1) A is closed in Y and
(2) A is bounded in Y.

Proof of (1): We wish to show: $A = \mathrm{Cl}_Y A$. By the superset property, $A\subseteq \mathrm{Cl}_Y A$. We wish to show: $\mathrm{Cl}_Y A\subseteq A$. By HW#70, $\mathrm{Cl}_Y A = \mathrm{sCl}_Y A$. We wish to show: $\mathrm{sCl}_Y A\subseteq A$. We wish to show: $\forall z\in \mathrm{sCl}_Y A$, $z\in A$. Let $z\in \mathrm{sCl}_Y A$ be given. We wish to show: $z\in A$.

Choose $s\in A^{\mathbb{N}}$ s.t. $s_\bullet\to z$ in Y. Because A is sequentially compact, we know that $s_\bullet$ is subconvergent in A. Choose a subsequence t of $s_\bullet$ s.t. $t_\bullet$ is convergent in A. Choose $x\in A$ s.t. $t_\bullet\to x$ in A. By Theorem 19.9, $t_\bullet\to x$ in Y. Since $t_\bullet$ is a subsequence of $s_\bullet$ and since $s_\bullet\to z$ in Y, it follows, from Theorem 23.14, that $t_\bullet\to z$ in Y. By HW#41, any metric space is Hausdorff, so Y is Hausdorff. So, since both $t_\bullet\to x$ in Y and $t_\bullet\to z$ in Y, it follows, from Corollary 24.9, that $x=z$. Then $z=x\in A$, as desired. *End of proof of (1).*

Proof of (2): Assume A is unbounded in Y. Want: Contradiction.

Since A is unbounded in Y, it follows that $A\neq\varnothing$. Choose $p\in A$. Since A is unbounded in Y, it follows, for all $k\in\mathbb{N}$, that $A\not\subseteq B(p,k)$, and so $A\backslash[B(p,k)]\neq\varnothing$. By the Axiom of Choice, choose $s\in A^{\mathbb{N}}$ s.t. $\forall k\in\mathbb{N}$, $s_k\notin B(p,k)$.

Since A is sequentially compact, choose a subsequence t of $s_\bullet$ s.t. $t_\bullet$ is convergent in A. Choose $x\in A$ s.t. $t_\bullet\to x$ in A. Let d denote the metric on Y. Choose $I\in\mathbb{N}$ s.t., $\forall j\in\mathbb{N}$,

$$[\ j\geqslant I\ ]\quad\Rightarrow\quad[\ d(t_j,x)<1\ ].$$

Let $M := d(x,p)$, and, by the Archimedean Principle, choose $j \in \mathbb{N}$ s.t. $j \geqslant \max\{I, 1+M\}$. Then $j \geqslant I$ and $j \geqslant 1+M$. By Corollary 23.13, choose $k \in [j..\infty)$ s.t. $t_j = s_k$. Then $d(t_j,p) = d(s_k,p)$. Since $j \geqslant I$, by the choice of I, we have $d(t_j,x) < 1$. Then

$$d(t_j,p) \;\leqslant\; [d(t_j,x)] \,+\, [d(x,p)] \;<\; 1 \,+\, M.$$

Since $k \in [j..\infty)$, it follows that $j \leqslant k$. By the choice of $s_\bullet$, we have $s_k \notin B(p,k)$. Then $d(s_k,p) \geqslant k$. Then

$$j \;\leqslant\; k \leqslant\; d(s_k,p) \;=\; d(t_j,p) \;<\; 1+M \;\leqslant\; j,$$

so $j < j$. Contradiction. *End of proof of (2).* $\square$

COROLLARY 25.2. *Let Y be a metrizable topological space, $A \subseteq Y$. Assume that A is sequentially compact. Then A is closed in Y.*

Proof. Choose a metric d on Y s.t. $\mathcal{T}_d$ is the topology on Y. Then, by Theorem 25.1, A is closed and bounded in the metric space (Y,d). Then A is closed in the topological space Y. $\square$

THEOREM 25.3. *Let Y be a proper metric space. Then:*

$$[\text{ } A \textit{ is sequentially compact }] \;\Leftrightarrow\; [\text{ } A \textit{ is closed and bounded in } Y \text{ }].$$

Proof. Proof of $\Rightarrow$*:* Follows from Theorem 25.1. *End of proof of* $\Rightarrow$*.*

Proof of $\Rightarrow$*:* Assume that A is closed and bounded in Y. We wish to show that A is sequentially compact in Y. That is, we wish to show: $\forall s \in A^{\mathbb{N}}$, $s_\bullet$ is subconvergent in A. Let $s \in A^{\mathbb{N}}$ be given. We wish to show: $s_\bullet$ is subconvergent in A.

Since $s_\bullet \in A^{\mathbb{N}}$, we see that $\mathrm{im}[s_\bullet] \subseteq A$. So, as A is bounded in Y, we conclude that $\mathrm{im}[s_\bullet]$ is bounded in Y. Then $s_\bullet$ is bounded in Y. So, since Y is proper, $s_\bullet$ is subconvergent in Y. Choose a subsequence t of $s_\bullet$ s.t. $t_\bullet$ is convergent in Y. Choose $z \in Y$ s.t. $t_\bullet \to z$ in Y. As $t_\bullet$ is a subsequence of $s_\bullet$ and $s_\bullet \in A^{\mathbb{N}}$, we conclude that $t_\bullet \in A^{\mathbb{N}}$. So, since $t_\bullet \to z$ in Y, it follows that $z \in \mathrm{sCl}_Y(A)$. By HW#70, $\mathrm{sCl}_Y(A) = \mathrm{Cl}_Y(A)$. Since A is closed in Y, we get $\mathrm{Cl}_Y(A) = A$. Then $z \in \mathrm{sCl}_Y(A) = \mathrm{Cl}_Y(A) = A$. Since $t_\bullet \to z$ in Y and $t_\bullet$ in $A^{\mathbb{N}}$ and $z \in A$, it follows, from Theorem 19.9, that $t_\bullet \to z$ in A.

So, since $t_\bullet$ is a subsequence of $s_\bullet$, we conclude that $s_\bullet$ is subconvergent in A, as desired. *End of proof of* $\Rightarrow$*.* $\square$

COROLLARY 25.4. *Let $C := \{(u,v) \in \mathbb{R}^2 \,|\, u^2 + v^2 = 1\}$. Then C is sequentially compact.*

Proof. By Corollary 24.14, $\mathbb{R}^2$ is proper. So, since C is closed and bounded in $\mathbb{R}^2$, by Theorem 25.3, C is sequentially compact. $\square$

LEMMA 25.5. *Let Y and Z be topological spaces and let $g : Y \to Z$. Let $s \in Y^{\mathbb{N}}$ and $b \in Y$. Assume that $s_\bullet \to b$ in Y. Assume that g is continuous at x. Then $(g \circ s)_\bullet \to g(b)$ in Z.*

Proof. True by HW#60 (with X replaced by $\mathbb{N}^*$ and a by ∞). $\square$

THEOREM 25.6. *Let K and Z be topological spaces. Let $f : K \to Z$. Assume that K is sequentially compact and that f is continuous. Then im$[f]$ is sequentially compact.*

Proof. Let $Y := \mathrm{im}[f]$. We wish to show: Y is sequentially compact. We wish to show: $\forall s \in Y^{\mathbb{N}}$, $s_\bullet$ is subconvergent in Y. Let $s \in Y^{\mathbb{N}}$ be given. We wish to show that $s_\bullet$ is subconvergent in Y.

By assumption, f is (K, Z)-continuous. So, by $\Leftarrow$ of (1) of Fact 22.7, f is (K, Y)-continuous.

For all $j \in \mathbb{N}$, since $s_j \in Y = \mathrm{im}[f]$, there exists $r \in K$ s.t. $f(r) = s_j$. By the Axiom of Choice, choose $r \in K^{\mathbb{N}}$ s.t., for all $j \in \mathbb{N}$, $f(r_j) = s_j$. For all $j \in \mathbb{N}$, we have $(f \circ r)_j = f(r_j) = s_j$. Then $(f \circ r)_\bullet = s_\bullet$. Since K is sequentially compact and $r \in K^{\mathbb{N}}$, we conclude that $r_\bullet$ is subconvergent in K. Choose a subsequence q of $r_\bullet$ s.t. q is convergent in K. Choose $x \in K$ s.t. $q_\bullet \to x$ in K. Since $q_\bullet \to x$ in K and since f is (K, Y)-continuous at x, by Lemma 25.5 (with g replaced by f and b by x), we see that $(f \circ q)_\bullet \to f(x)$ in Y. $\square$

COROLLARY 25.7. *Let $C := \{(u, v) \in \mathbb{R}^2 \,|\, u^2 + v^2 = 1\}$. Let Z be a metric space. Let $f : C \to Z$ be continuous. Then im$[f]$ is closed and bounded in Z.*

Proof. By Corollary 25.4, C is sequentially compact. Then, by Theorem 25.6, $\mathrm{im}[f]$ is sequentially compact. Then, by Theorem 25.3, $\mathrm{im}[f]$ is closed and bounded in Z. $\square$

DEFINITION 25.8. *Let $A \subseteq \mathbb{R}$. Then A is* **bounded above** *in $\mathbb{R}$ means: $\exists M \in \mathbb{R}$ s.t. $A \leqslant M$. Also, A is* **bounded below** *in $\mathbb{R}$ means: $\exists N \in \mathbb{R}$ s.t. $N \leqslant A$.*

We often omit "in $\mathbb{R}$" if it will cause no confusion.

Recall that, in any metric space, "bounded" means "contained in a ball". In $\mathbb{R}$, a ball is the same as a bounded open interval, so "bounded in $\mathbb{R}$" means "contained in a bounded open interval".

REMARK 25.9. *Let $A \subseteq \mathbb{R}$. Then:*

$$[\ A \text{ is bounded in } \mathbb{R}\ ] \ \Leftrightarrow\ [\ A \text{ is bounded above and below in } \mathbb{R}\ ].$$

Proof. Unassigned HW. ☐

THEOREM 25.10. *Let $r, s, t \in (\mathbb{R}^*)^{\mathbb{N}}$. Let $b \in \mathbb{R}^*$. Assume, $\forall j \in \mathbb{N}$, $r_j \leqslant s_j \leqslant t_j$. Assume: ($r_\bullet \to b$) and ($t_\bullet \to b$). Then: $s_\bullet \to b$.*

Proof. Let $Y := \mathbb{N}^*$ and let $a := \infty$. Let $P := \mathbb{N}$. Then $P \in \mathcal{N}_Y^\times(a)$ and, by assumption, $r_\bullet \leqslant s_\bullet \leqslant t_\bullet$ on P. Then, by Theorem 21.2 (with f replaced by $r_\bullet$, g by $s_\bullet$ and h by $t_\bullet$), we get $s_\bullet \to b$, as desired. ☐

THEOREM 25.11. *Let $A \subseteq \mathbb{R}$ be nonempty and bounded above. Then* $\sup A \in \mathrm{Cl}_{\mathbb{R}} A$.

Proof. Let $y := \sup A$. Then $y \in \mathbb{R}^*$ and we wish to show: $y \in \mathrm{Cl}_{\mathbb{R}} A$. Since $\mathbb{R}$ is a metric space, by HW#70, we get $\mathrm{Cl}_{\mathbb{R}} A = \mathrm{sCl}_{\mathbb{R}} A$, so it suffices to show: $y \in \mathrm{sCl}_{\mathbb{R}} A$. Want: $\exists s \in A^{\mathbb{N}}$ s.t. $s_\bullet \to y$ in $\mathbb{R}$.

Since A is bounded above, choose $M \in \mathbb{R}$ s.t. $A \leqslant M$. Then

$$\sup A = \min(\mathrm{UB}(M)) \leqslant \mathrm{UB}(M) \qquad \text{and} \qquad M \in \mathrm{UB}(M),$$

so $\sup A \leqslant M$. Then $y = \sup A \leqslant M < \infty$. Then $y < \infty$. Since A is nonempty, choose $b \in A$. We have $y = \sup A = \min(\mathrm{UB}(A)) \in \mathrm{UB}(A)$, so $A \leqslant y$. Then $-\infty < b \in A \leqslant y$. Then $-\infty < y$.

Since $-\infty < y < \infty$ and $y \in \mathbb{R}^*$, we see that $y \in \mathbb{R}$. For all $j \in \mathbb{N}$, $y - (1/j) < y = \sup A = \min(\mathrm{UB}(A))$, so $y - (1/j) \notin \mathrm{UB}(A)$, so $\mathrm{NOT}(A \leqslant y - (1/j))$, so $\exists s \in A$ s.t. $s > y - (1/j)$. So, by the Axiom of Choice, choose $s \in A^{\mathbb{N}}$ s.t., $\forall j \in \mathbb{N}$, $s_j > y-(1/j)$. Want: $s_\bullet \to y$ in $\mathbb{R}$.

Define $r, t \in \mathbb{R}^{\mathbb{N}}$ by $r_j = y-(1/j)$ and $t_j = y+(1/j)$. Then $r_\bullet \to y$ and $t_\bullet \to y$, so, by Theorem 25.10, it suffices to show: $\forall j \in \mathbb{N}$, $r_j \leqslant s_j \leqslant t_j$. Let $j \in \mathbb{N}$ be given. We wish to show: $r_j \leqslant s_j \leqslant t_j$.

By choice of $s_\bullet$, we have $s_j > y - (1/j)$. Then $r_j = y - (1/j) < s_j$. Then $r_j < s_j$, so $r_j \leqslant s_j$. It remains to show: $s_j \leqslant t_j$.

Since $s_\bullet \in A^{\mathbb{N}}$, we get $s_j \in A$. Then $s_j \in A \leqslant y < y + (1/j) = t_j$. Then $s_j < t_j$, so $s_j \leqslant t_j$, as desired. ☐

COROLLARY 25.12. *Let $A \subseteq \mathbb{R}$ be nonempty, closed in $\mathbb{R}$ and bounded above in $\mathbb{R}$. Then A has a maximum.*

Proof. We wish to show: $\max A \neq$ ☹.

Let $y := \sup A$. By Theorem 25.11, we have $y \in \mathrm{Cl}_{\mathbb{R}} A$. Then $y \neq ☹$. It therefore suffices to show: $\max A = y$. That is, we wish to show $y \in A \cap [\mathrm{UB}(A)]$.

Since A is closed in $\mathbb{R}$, it follows that $\mathrm{Cl}_{\mathbb{R}} A = A$. Then $y \in \mathrm{Cl}_{\mathbb{R}} A = A$. We wish to show: $y \in \mathrm{UB}(A)$.

We have $y = \sup A = \min(\mathrm{UB}(A))$. Then $y \in \mathrm{UB}(A)$, as desired. $\square$

COROLLARY 25.13. *Let $B \subseteq \mathbb{R}$ be nonempty and bounded below. Then* $\inf B \in \mathrm{Cl}_{\mathbb{R}} B$.

Sketch of proof, details left to the reader: Let $A := -B$. Since B is nonempty and bounded below, it follows that A is nonempty and bounded above. Then, by Theorem 25.11, we have $\sup A \in \mathrm{Cl}_{\mathbb{R}} A$. Then $\inf B = -(\sup A) \in -(\mathrm{Cl}_{\mathbb{R}} A) = \mathrm{Cl}_{\mathbb{R}} B$. *End of sketch of proof.*

COROLLARY 25.14. *Let $B \subseteq \mathbb{R}$ be nonempty, closed in $\mathbb{R}$ and bounded below in $\mathbb{R}$. Then B has a minimum.*

Sketch of proof, details left to the reader: Let $A := -B$. Since B is nonempty, closed and bounded below, it follows that A is nonempty, closed and bounded above. Then, by Theorem 25.12, A has a maximum. Then $-A$ has a minimum. Since $-A = -(-B) = B$, we conclude that B has a minimum, as desired. *End of sketch of proof.*

THEOREM 25.15. *Let $S \subseteq \mathbb{R}$ be nonempty, closed and bounded in $\mathbb{R}$. Then S has a maximum and a minimum.*

Proof. By $\Rightarrow$ of Remark 25.9, we see that S is bounded above and below in $\mathbb{R}$. Then, by Corollary 25.12 and Corollary 25.14, S has a maximum and a minimum. $\square$

The next result is called the **Extreme Value Theorem**.

THEOREM 25.16. *Let K be a nonempty sequentially compact topological space and let $f : K \to \mathbb{R}$ be continuous. Then f has a maximum and a minimum.*

Proof. Since K is nonempty and $\mathrm{dom}[f] = K$, it follows that $\mathrm{im}[f]$ is nonempty. By Theorem 25.6, $\mathrm{im}[f]$ is sequentially compact. Let $S := \mathrm{im}[f]$. Then S is nonempty and sequentially compact. By Definition 23.5, we wish to show: S has a maximum and a minimum.

Since S is sequentially compact, it follows, from $\Rightarrow$ of Theorem 25.3, that S is closed and bounded in $\mathbb{R}$. So, since S is also nonempty, by Theorem 25.15, we see that S has a maximum and a minimum. $\square$

DEFINITION 25.17. *Let Y, Z be topological spaces, $f : Y \dashrightarrow Z$. Let $a \in Y$. Then*

$$(Y,Z)\text{-}\operatorname*{LIMS}_a f \quad := \quad \{\, b \in Z \mid f \to b \text{ in } Z \text{ near } a \text{ in } Y \,\} \qquad \textit{and}$$
$$(Y,Z)\text{-}\lim_a f \quad := \quad \operatorname{ELT}(\ (Y,Z)\text{-}\operatorname*{LIMS}_a f\ ).$$

When Y and Z are clear, we often omit "(Y,Z)-" and write $\operatorname*{LIMS}_a f$ and $\lim_a f$. However, some confusion can arise, so care is required:

Define $f : \mathbb{R}\backslash\{0\} \to (0,\infty)$ by $f(x) = 1/x^2$. Then $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and $f : \mathbb{R} \dashrightarrow \mathbb{R}^*$. We have

$$(\mathbb{R},\mathbb{R})\text{-}\operatorname*{LIMS}_0 f = \varnothing \qquad \text{and} \qquad (\mathbb{R},\mathbb{R}^*)\text{-}\operatorname*{LIMS}_0 f = \{\infty\}.$$

Then $(\mathbb{R},\mathbb{R})\text{-}\lim_0 f = ☹$ and $(\mathbb{R},\mathbb{R}^*)\text{-}\lim_0 f = \infty$.

We have alternate notation for $\lim_a f$:

- At any point where the variable x is unbound, we can use $\lim_{x\to a} f(x)$ to denote $\lim_a f$. Then x is temporarily bound, from the text "$\lim_{x\to a}$" to the text "$f(x)$", and becomes free afterward.
- At any point where the variable t is unbound, we can use $\lim_{t\to a} f(t)$ to denote $\lim_a f$. Then t is temporarily bound, from the text "$\lim_{t\to a}$" to the text "$f(t)$", and becomes free afterward.
- *Etc.*

NOTE TO SELF: Next year, using Theorem 20.8 we will show that if $f = g$ on a pnbd of a, then

$$\text{both} \quad (\ \operatorname*{LIMS}_a f \subseteq \operatorname*{LIMS}_a g\ ) \quad \text{and} \quad (\ \operatorname*{LIMS}_a g \subseteq \operatorname*{LIMS}_a f\ ),$$

and so $\operatorname*{LIMS}_a f = \operatorname*{LIMS}_a g$, and so $\lim_a f = \lim_a g$.

NOTE TO SELF: We did prove this, see Theorem 41.18

REMARK 25.18. *Let Y, Z be topological spaces, $f : Y \dashrightarrow Z$. Let $a \in Y$, $b \in Z$. Then: ($\lim_a f = b$) $\Rightarrow$ ($f \to b$ near a).*

Proof. Assume $\lim_a f = b$. We wish to show: $f \to b$ near a.

Since $b \in Z$, we know that $b \neq ☹$. So, since $\lim_a f = b$, we see that $\operatorname*{LIMS}_a f = \{b\}$. Then $b \in \operatorname*{LIMS}_a f$. Then $f \to b$ near a. $\square$

REMARK 25.19. *Let Y be a topological space. Let Z be a Hausdorff topological space. Let $f : Y \dashrightarrow Z$. Let $a \in \mathrm{LPD}_Y f$. Let $b \in Z$. Then:* $(\ \lim\limits_a f = b\ ) \Leftrightarrow (\ f \to b$ *near* $a\ )$.

Proof. Proof of $\Rightarrow$*:* This follows from Remark 25.18. *End of proof of* $\Rightarrow$*.*

Proof of $\Leftarrow$*:* Assume $f \to b$ near a. We wish to show: $\lim\limits_a f = b$.

By Theorem 24.8, $\#[\mathop{\mathrm{LIMS}}\limits_a f] \leqslant 1$. Since $b \in Z$ and $f \to b$ near a, we see that $b \in \mathop{\mathrm{LIMS}}\limits_a f$. Then $\{b\} = \mathop{\mathrm{LIMS}}\limits_a f$. Then

$$\lim_a f \quad = \quad \mathrm{ELT}(\mathop{\mathrm{LIMS}}_a f) \quad = \quad \mathrm{ELT}(\{b\}) \quad = \quad b,$$

as desired. *End of proof of* $\Leftarrow$*.* $\square$

DEFINITION 25.20. *Let V be a vector space, S a set, $f : V \dashrightarrow S$, $p \in V$. Then $f(p + \bullet) : V \dashrightarrow S$ is defined by $(f(p + \bullet))(h) = f(p + h)$.*

The function $f(p+\bullet)$ is called the "**horizontal translate** of f by p". We explained how the graph of $f(p + \bullet)$ is obtained from the graph of f by a horizontal translation.

DEFINITION 25.21. *Let V, W be vector spaces. Let $f : V \dashrightarrow W$. Let $p \in dom[f]$. Then $f_p^T : V \dashrightarrow W$ is defined by*

$$f_p^T(h) \quad = \quad [f(p + h] \;-\; [f(p)].$$

The function f_p^T is called the "**double translate** of f by p". We explained how the graph of f_p^T is obtained from the graph of f by two translations, one horizontal, one vertical.

DEFINITION 25.22. *Let W be a vector space. Let $f : \mathbb{R} \dashrightarrow W$. Let $p \in dom[f]$. Then $SS_f^p : \mathbb{R} \dashrightarrow W$ is defined by*

$$(SS_f^p)(h) \quad = \quad \frac{[f(p+h)] \;-\; [f(p)]}{h}.$$

The function SS_f^p is called the "**secant slope function** of f at p".

DEFINITION 25.23. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$. Then $f' : \mathbb{R} \dashrightarrow W$ is defined by $f'(p) = \lim\limits_0 SS_f^p$.*

The function f' is called the "**derivative** of f".

In Definition 25.23, by $\lim\limits_0$, we mean $(\mathbb{R}, W)$-$\lim\limits_0$. In the special case where $W = \mathbb{R}$, we mean $(\mathbb{R}, \mathbb{R})$-$\lim\limits_0$, and *NOT* $(\mathbb{R}^*, \mathbb{R}^*)$-$\lim\limits_0$. Thus, for any function $f : \mathbb{R} \dashrightarrow \mathbb{R}$, for any p, we have $f'(p) \notin \{\infty, -\infty\}$.

REMARK 25.24. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$. Then $dom[f'] \subseteq dom[f]$.*

Proof. We wish to show: $\forall p \in \mathrm{dom}[f']$, $p \in \mathrm{dom}[f]$. Let $p \in \mathrm{dom}[f']$ be given. We wish to show: $p \in \mathrm{dom}[f]$.

Since $p \in \mathrm{dom}[f']$, we get $f'(p) \neq ☹$. That is, $\lim\limits_0 SS_f^p \neq ☹$. Then $SS_f^p \neq ☹$. Then, by Definition 25.22, we get $p \in \mathrm{dom}[f]$. □

Recall: $\forall$set S, $\forall b \neq ☹$, $C_S^b : S \to \{b\}$ is defined by $C_S^b(x) = b$.

Recall: $\forall$set S, $\mathrm{id}_S : S \to S$ is defined by $\mathrm{id}_S(x) = x$.

We explained: For any normed vector space W, for any $y \in W$, we have $(C_{\mathbb{R}}^y)' = C_{\mathbb{R}}^{0_W}$. That is, the derivative of a constant is zero.

We explained: $(\mathrm{id}_{\mathbb{R}})' = C_{\mathbb{R}}^1$, *i.e.*, the derivative of the identity is one.

26. Class 26 on 12 December 2017, Tu of Week 15

Recall, for any topological space Y, for any $A \subseteq Y$, that

$$\begin{aligned} \mathrm{Cl}\, A &= \{\, z \in Y \mid \forall U \in \mathcal{N}(z),\ U \cap A \neq \varnothing \,\} \qquad \text{and} \\ \mathrm{LP}\, A &= \{\, z \in Y \mid \forall U \in \mathcal{N}^\times(z),\ U \cap A \neq \varnothing \,\}. \end{aligned}$$

Recall that a topological space is said to be T_1 if all of its singleton sets are closed. In this course, metrizable is a "tame" hypothesis, in the sense that every topological space we care about is metrizable. Recall that, for topological spaces,

$$\text{metrizable} \Rightarrow \text{Hausdorff} \Rightarrow T_1.$$

Then Hausdorff and T_1 are even tamer than metrizable.

FACT 26.1. *Let Y be a T_1 topological space, let $p, z \in Y$ and let $U \in \mathcal{N}^\times(z)$. Then $U_p^\times \in \mathcal{N}^\times(z)$.*

Proof. One of the following is true:

(1) $p \notin U$ or
(2) $p \in U$.

Case 1: As $p \notin U$, $U_\times^p = U$. Then $U_p^\times = U \in \mathcal{N}^\times(z)$. *End of Case 1.*

Case 2: Choose $V \in \mathcal{N}(z)$ such that $V_z^\times = U$. Then $z \notin V_z^\times = U$. So, since $p \in U$, we conclude that $p \neq z$. Since $W \in \mathcal{N}(z)$, choose an open subset W of Y such that $z \in W \subseteq V$. Since $p \neq z \in W$, we get $z \in W_p^\times$. Since $W \subseteq V$, we get $W_p^\times \subseteq V_p^\times$. Then $z \in W_p^\times \subseteq V_p^\times$.

Since Y is T_1, we conclude that $\{p\}$ is closed in Y. Then $Y \backslash \{p\}$ is open in Y. So, since W is also open in Y, it follows that $W \cap (Y \backslash \{p\})$ is open in Y. So, since $W_p^\times = W\backslash\{p\} = W \cap (Y\backslash\{p\})$, we see that $W_p^\times$ is open in Y. So, since $z \in W_p^\times \subseteq V_p^\times$, it follows that $V_p^\times \in \mathcal{N}(z)$. Then $(V_p^\times)_z^\times \in \mathcal{N}^\times(z)$. Then $U_p^\times = (V_z^\times)_p^\times = V\backslash\{p,z\} = (V_p^\times)_z^\times \in \mathcal{N}^\times(z)$, as desired. *End of Case 2.* □

According to Remark 21.16, a one point adjustment of a function doesn't affect its limits. There's an analogous result for limit points and sets. It states that a one point adjustment of a set doesn't affect its limit points:

COROLLARY 26.2. *Let Y be a T_1 topological space, let $p \in Y$ and let $A \subseteq Y$. Then* $\mathrm{LP}(A_p^\times) = \mathrm{LP}\,A$.

Proof. As $A_p^\times \subseteq A$, we get $\mathrm{LP}(A_p^\times) \subseteq \mathrm{LP}\,A$. Want: $\mathrm{LP}\,A \subseteq \mathrm{LP}(A_p^\times)$. We wish to show: $\forall z \in \mathrm{LP}\,A$, $z \in \mathrm{LP}(A_p^\times)$. Let $z \in \mathrm{LP}\,A$ be given. We wish to show: $z \in \mathrm{LP}(A_p^\times)$. We wish to show: $\forall U \in \mathcal{N}^\times(z)$, $U \cap A_p^\times \neq \varnothing$. Let $U \in \mathcal{N}^\times(z)$ be given. We wish to show: $U \cap A_p^\times \neq \varnothing$.

By Fact 26.1, $U_p^\times \in \mathcal{N}^\times(z)$. So, since $z \in \mathrm{LP}\,A$, we get $U_p^\times \cap A \neq \varnothing$. Then $U \cap A_p^\times = (U \cap A)\backslash\{p\} = U_p^\times \cap A \neq \varnothing$, as desired. □

DEFINITION 26.3. *$\forall$vector space V, $\forall A \subseteq V$, $\forall p \in V$, we define*

$$\begin{aligned} A+p &:= \{\, x+p \mid x \in A \,\} \qquad \textit{and} \\ A-p &:= \{\, x-p \mid x \in A \,\}. \end{aligned}$$

REMARK 26.4. *Let V be a normed vector space, let $A \subseteq V$ and let $p \in V$. Then* $\mathrm{LP}_V(A-p) = (\mathrm{LP}_V A) - p$.

Proof. Unassigned HW. □

REMARK 26.5. *Let V be a vector space and let S be a set. Let $f : V \dashrightarrow S$ and let $p \in V$. Then:*

(1) $\forall h \in V, \quad (f(p+\bullet))(h) = f(p+h) \qquad$ *and*
(2) $dom[f(p+\bullet)] = (dom[f]) - p$.

Proof. Unassigned HW. □

REMARK 26.6. *Let V and W be vector spaces. Let $f : V \dashrightarrow W$ and let $p \in dom[f]$. Then*

(1) $\forall h \in V, \quad f_p^T(h) = [f(p+h)] - [f(p)]$,
(2) $f_p^T(0_V) = 0_W \qquad$ *and*
(3) $dom[f_p^T] = (dom[f]) - p$.

Proof. Unassigned HW. □

REMARK 26.7. *Let W be a vector space. Let $f : \mathbb{R} \dashrightarrow W$ and let $p \in dom[f]$. Then*

(1) $\forall h \in \mathbb{R}, \quad SS_f^p(h) = [f_p^T(h)]/h,$
(2) $dom[SS_f^p] = ((dom[f]) - p)_0^\times$ *and*
(3) $\mathrm{LPD}_\mathbb{R}[SS_f^p] = (LPD_\mathbb{R} f) - p.$

Proof. Unassigned HW. □

REMARK 26.8. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$. Then*

(1) $\forall p \in \mathbb{R}, \quad f'(p) = \lim\limits_0 SS_f^p = \lim\limits_{h \to 0} \dfrac{[f(p+h)] - [f(p)]}{h}$ *and*
(2) $dom[f'] \subseteq dom[f]$.

Proof. Unassigned HW. □

Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = |x|$. Then

$$\mathrm{dom}[f'] \quad = \quad \mathbb{R} \backslash \{0\} \quad \subsetneq \quad \mathbb{R} \quad = \quad \mathrm{dom}[f].$$

So, in (2) of Remark 26.8, we cannot replace "$\subseteq$" by "$=$".

For any normed vector space W, for any $q \in W$, $(C_\mathbb{R}^q)' = C_\mathbb{R}^{0_W}$.

We have: $(\mathrm{id}_\mathbb{R})' = C_\mathbb{R}^1$.

Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = (x-1)^2 + 2$. We graphed f. Let $A := \{1+1, 1+(1/2), 1+(1/3), 1+(1/4), \ldots\}$ and let $B := A \cup \{1\}$. Let $g := f|B : B \to \mathbb{R}$. We graphed g. The graph of g has only countably many points and is a subset of the graph of f. We explained why $g'(1) = 0$. We noted: $\forall x \in \mathbb{R} \backslash \{1\}$, $g'(x) = ☹$.

DEFINITION 26.9. *For any normed vector space W, for any $f : \mathbb{R} \dashrightarrow W$, for any p, by f is* **differentiable** *at p, we mean: $p \in dom[f']$.*

In this course, we make the convention that $0^0 = 1$.

FACT 26.10. *Let $p \in \mathbb{R}$. Then, $\forall k \in \mathbb{N}$, $\exists$continuous $g : \mathbb{R} \to \mathbb{R}$ s.t.*

$$\forall h \in \mathbb{R}, \qquad (p+h)^k \quad = \quad p^k \; + \; kp^{k-1}h \; + \; [g(h)]\,h^2.$$

Proof. For all $k \in \mathbb{N}$, let $P_k := \Big[\ \exists$continuous $g : \mathbb{R} \to \mathbb{R}$ s.t.

$$\forall h \in \mathbb{R}, \qquad (p+h)^k \quad = \quad p^k \; + \; kp^{k-1}h \; + \; [g(h)]\,h^2 \ \Big].$$

We wish to show: $\forall k \in \mathbb{N}$, P_k. We leave it as an unassigned exercise to show: P_1. By the Principle of Mathematical Induction, we wish to show: $\forall k \in \mathbb{N}$, $(P_k \Rightarrow P_{k+1})$. Let $k \in \mathbb{N}$ be given. We wish to show: $P_k \Rightarrow P_{k+1}$. Assume P_k. We wish to show: P_{k+1}.

By P_k, choose a continuous $g : \mathbb{R} \to \mathbb{R}$ s.t.

$$\forall h \in \mathbb{R}, \qquad (p+h)^k \quad = \quad p^k \ + \ kp^{k-1}h \ + \ [g(h)]\, h^2.$$

We wish to show: $\exists$continuous $\gamma : \mathbb{R} \to \mathbb{R}$ s.t.

$$\forall h \in \mathbb{R}, \qquad (p+h)^{k+1} \quad = \quad p^{k+1} \ + \ (k+1)p^k h \ + \ [\gamma(h)]\, h^2.$$

Define $\gamma : \mathbb{R} \to \mathbb{R}$ by

$$\gamma(h) \quad = \quad p[g(h)] \ + \ kp^{k-1} \ + \ h\,[g(h)].$$

Then $\gamma : \mathbb{R} \to \mathbb{R}$ is continuous, and we wish to show:

$$\forall h \in \mathbb{R}, \qquad (p+h)^{k+1} \quad = \quad p^{k+1} \ + \ (k+1)p^k h \ + \ [\gamma(h)]\, h^2.$$

Let $h \in \mathbb{R}$ be given. We wish to show:

$$(p+h)^{k+1} \quad = \quad p^{k+1} \ + \ (k+1)p^k h \ + \ [\gamma(h)]\, h^2.$$

We have

$$(p+h)^k \quad = \quad p^k \ + \ kp^{k-1}h \ + \ [g(h)]\, h^2.$$

Multiplying this by $p+h$, we get

$$\begin{aligned}
(p+h)^{k+1} &= (\ p \ + \ h\ )(\ p^k \ + \ kp^{k-1}h \ + \ [g(h)]\, h^2\ ) \\
&= p\ (\ p^k \ + \ kp^{k-1}h \ + \ [g(h)]\, h^2\ ) \quad + \\
&\qquad h\ (\ p^k \ + \ kp^{k-1}h \ + \ [g(h)]\, h^2\ ) \\
&= p^{k+1} \ + \qquad kp^k h \ + \ p\,[g(h)]\, h^2 \quad + \\
&\qquad\qquad p^k h \ + \ kp^{k-1}h^2 \ + \ h\,[g(h)]\, h^2 \\
&= p^{k+1} \ + \ (k+1)p^k h \ + \ [\gamma(h)]\, h^2,
\end{aligned}$$

as desired. $\square$

COROLLARY 26.11. *Let $k \in \mathbb{N}$. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^k$. let $p \in \mathbb{R}$. Then $f'(p) = kp^{k-1}$.*

Proof. We wish to show: $\lim\limits_{0} SS_f^p = kp^{k-1}$.

By (3) of Remark 26.7, we have $\mathrm{LPD}_{\mathbb{R}}(SS_f^p) = (\mathrm{LPD}_{\mathbb{R}} f) - p$. Also, we have $\mathrm{LPD}_{\mathbb{R}} f = \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[f]) = \mathrm{LP}_{\mathbb{R}}\mathbb{R} = \mathbb{R}$. Then

$$0 \ = \ p - p \ \in \ \mathbb{R} - p \ = \ (\mathrm{LPD}_{\mathbb{R}} f) - p \ = \ \mathrm{LPD}_{\mathbb{R}}(SS_f^p).$$

Then, by Remark 25.19, it suffices to show: $SS_f^p \to kp^{k-1}$ near 0.

By Fact 26.10, choose a continuous $g : \mathbb{R} \to \mathbb{R}$ s.t.

$$\forall h \in \mathbb{R}, \qquad (p+h)^k \quad = \quad p^k \ + \ kp^{k-1}h \ + \ [g(h)]\, h^2.$$

For all $h \in \mathbb{R}$, we have $f(p+h) = (p+h)^k$. Also, we have $f(p) = p^k$. Let $z := kp^{k-1}$. Then

$$\forall h \in \mathbb{R}, \qquad f(p+h) \quad = \quad [f(p)] \; + \; zh \; + \; [g(h)]\, h^2.$$

We wish to prove: $SS_f^p \to z$ near 0.

For all $h \in \mathbb{R}$, we have

$$SS_f^p(h) \;=\; \frac{[f(p+h)] \,-\, [f(p)]}{h} \;=\; \frac{zh \,+\, [g(h)]h^2}{h}.$$

Then, $\forall h \in \mathbb{R}_0^\times$, $SS_f^p(h) = z + [g(h)]h$. Define $i := \mathrm{id}_\mathbb{R} : \mathbb{R} \to \mathbb{R}$. Then $\forall h \in \mathbb{R}$, $i(h) = h$. Then $i(0) = 0$. Also, the function $i : \mathbb{R} \to \mathbb{R}$ is distance-semidecreasing, hence 1-Lipschitz, hence uniformly continuous, hence continuous (see Fact 22.8). For all $h \in \mathbb{R}_0^\times$, we have

$$SS_f^p(h) \;= z \,+\, [g(h)]h \;= z \,+\, [g(h)][i(h)] \;=\; (C_\mathbb{R}^z \,+\, gi)(h).$$

That is, $SS_f^p = C_\mathbb{R}^z + gi$ on $\mathbb{R}_0^\times$.

As $g : \mathbb{R} \to \mathbb{R}$ is continuous, we get: $g \to g(0)$ near 0. As $i : \mathbb{R} \to \mathbb{R}$ is continuous, we get: $i \to i(0)$ near 0. Then, since $C_\mathbb{R}^z \to z$ near 0, we get: $C_\mathbb{R}^z + gi \to z + [g(0)][i(0)]$ near 0. So, since

$$z \,+\, [g(0)]\,[i(0)] \;=\; z \,+\, [g(0)] \cdot 0 \;=\; z \,+\, 0 \;=\; z,$$

we get: $C_\mathbb{R}^z + gi \to z$ near 0. So, since $SS_f^p = C_\mathbb{R}^z + gi$ on $\mathbb{R}_0^\times$, it follows, from Theorem 20.8, that $SS_f^p \to z$ near 0, as desired. $\square$

Recall: $\forall$function f, $\forall$non-☹ a, b, $\forall x$, we have:

$$(\mathrm{adj}_a^b\, f)(x) \quad = \quad \begin{cases} f(x), & \text{if } x \neq a \\ b, & \text{if } x = a. \end{cases}$$

We can use this "adjustment" operator to remove discontinuities, as in $\Rightarrow$ of Remark 21.17. When f is differentiable at a point p, the secant slope function SS_f^p has a removable discontinuity at 0, and, in the next definition, we remove it by an adjustment that introduces the tangent slope $f'(p)$ at 0.

DEFINITION 26.12. *For any normed vector space W, for any $f : \mathbb{R} \dashrightarrow W$, for any $p \in dom[f']$, we define*

$$STS_f^p \quad := \quad \mathrm{adj}_0^{f'(p)}(SS_p^f).$$

Keep in mind that, if f is not differentiable at p, then $STS_f^p = ☹$.

In Definition 26.12, the function STS_f^p is called the **secant tangent slope function** of f at p. Its values include all the secant slopes of f

at p together with the one tangent slope of f at p. The introduction of this one tangent slope removes the discontinuity at 0, and so we have:

REMARK 26.13. *Let W be a normed vector space, let $f : \mathbb{R} \dashrightarrow W$ and let $p \in dom[f']$. Then STS_f^p is continuous at 0.*

Proof. Since $p \in \mathrm{dom}[f']$, we get: $SS_f^p \to f'(p)$ near 0. Then, by $\Rightarrow$ of Remark 21.17 (with f replaced by SS_f^p, a by 0 and b by $f'(p)$), we see that STS_f^p is continuous at 0, as desired. $\square$

FACT 26.14. *Let W be a normed vector space, let $f : \mathbb{R} \dashrightarrow W$ and let $p \in dom[f']$. Then $(\mathrm{id}_{\mathbb{R}}) \cdot (STS_f^p) = f_p^T$.*

Proof. Since $\mathbb{R}$ is a common superdomain for $(\mathrm{id}_{\mathbb{R}}) \cdot (STS_f^p)$ and f_p^T, it suffices to show: $\forall h \in \mathbb{R}$, $((\mathrm{id}_{\mathbb{R}}) \cdot (STS_f^p))(h) = f_p^T(h)$. Let $h \in \mathbb{R}$ be given. We wish to show: $((\mathrm{id}_{\mathbb{R}}) \cdot (STS_f^p))(h) = f_p^T(h)$. We have

$$((\mathrm{id}_{\mathbb{R}}) \cdot (STS_f^p))(h) \;=\; [\mathrm{id}_{\mathbb{R}}(h)] \cdot [(STS_f^p)(h)] \;=\; h \cdot [(STS_f^p)(h)].$$

Also, we have $f_p^T(h) = [f(p+h)] - [f(p)]$. We therefore wish to show: $h \cdot [(STS_f^p)(h)] = [f(p+h)] - [f(p)]$.

One of the following must be true:

(1) $h \neq 0$ or
(2) $h = 0$.

Case 1: Since $h \neq 0$, we get $STS_f^p(h) = SS_f^p(h)$. Then

$$\begin{aligned} h \cdot [(STS_f^p)(h)] &= h \cdot [(SS_f^p)(h)] \\ &= h \cdot \left[\frac{[f(p+h)] - [f(p)]}{h} \right] \\ &= [f(p+h)] - [f(p)], \end{aligned}$$

as desired. *End of case 1.*

Case 2: Since $h = 0$, we get $STS_f^p(h) = f'(p)$. Then

$$\begin{aligned} h \cdot [(STS_f^p)(h)] &= 0 \cdot [f'(p)] \;=\; 0 \;=\; [f(p)] - [f(p)] \\ &= [f(p+0)] - [f(p)] \;=\; [f(p+h)] - [f(p)], \end{aligned}$$

as desired. *End of case 2.* $\square$

FACT 26.15. *Let V be a normed vector space, let Z be a topological space, let $f : V \dashrightarrow Z$ and let $p \in V$. Then:*

$$[\; f \text{ is continuous at } p \;] \quad \Leftrightarrow \quad [\; f(p + \bullet) \text{ is continuous at } 0_V \;].$$

Proof. Proof of $\Leftarrow$*:* Unassigned HW. *End of proof of* $\Leftarrow$*.*

Proof of $\Rightarrow$*:* Let $g := f(p + \bullet)$. We wish to show: g is continuous at 0_V. We wish to show: $\forall E \in \mathcal{N}(g(0_V))$, $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ g(x) \in E\ ].$$

Let $E \in \mathcal{N}(g(0_V))$ be given. We wish to show: $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ g(x) \in E\ ].$$

We have $g(0_V) = (f(p + \bullet))(0_V) = f(p)$, so $g(0_V) = f(p)$. It follows that $E \in \mathcal{N}(g(0_V)) = \mathcal{N}(f(p))$. So, since f is continuous at p, choose $\delta > 0$ s.t.,$\forall w \in \mathrm{dom}[f]$,

$$[\ |w - p|_V < \delta\ ] \quad \Rightarrow \quad [\ f(w) \in E\ ].$$

We wish to show: $\forall x \in \mathrm{dom}[g]$,

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ g(x) \in E\ ].$$

Let $x \in \mathrm{dom}[g]$ be given. We wish to show:

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ g(x) \in E\ ].$$

Assume $|x|_V < \delta$. We wish to show $g(x) \in E$.

By (2) of Remark 26.5, $\mathrm{dom}[f_p^T] = (\mathrm{dom}[f]) - p$. Then

$$x \ \in \ \mathrm{dom}[g] \ = \ \mathrm{dom}[f(p + \bullet)] \ = \ (\mathrm{dom}[f]) \ - \ p,$$

and so $x + p \in \mathrm{dom}[f]$. Let $w := x + p$. Then $w \in \mathrm{dom}[f]$. We have $|w - p|_V = |x|_V < \delta$. So, by choice of δ, we get $f(w) \in E$. Then $g(x) = (f(p + \bullet))(x) = f(p + x) = f(w) \in E$. *End of proof of* $\Rightarrow$*.* $\square$

FACT 26.16. *Let V, W be a normed vector spaces. Let $f : V \dashrightarrow W$ and let $p \in V$. Then:*

$$[\ f \textit{ is continuous at } p\ ] \quad \Leftrightarrow \quad [\ f_p^T \textit{ is continuous at } 0_V\ ].$$

Proof. Proof of $\Leftarrow$*:* Unassigned HW. *End of proof of* $\Leftarrow$*.*

Proof of $\Rightarrow$*:* By (2) of Remark 26.6, $f_T^p(0_V) = 0_W$. Let $g := f_p^T$. Then $g(0_V) = 0_W$. We wish to show: g is continuous at 0_V. We wish to show: $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ |g(x)|_W\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \delta > 0$ s.t., $\forall x \in \mathrm{dom}[g]$,

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ |g(x)|_W < \varepsilon\ ].$$

So, since f is continuous at p, choose $\delta > 0$ s.t.,$\forall w \in \mathrm{dom}[f]$,

$$[\ |w - p|_V < \delta\ ] \quad \Rightarrow \quad [\ |[f(w)] - [f(p)]|_W < \varepsilon\ ].$$

We wish to show: $\forall x \in \mathrm{dom}[g]$,

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ |g(x)|_W < \varepsilon\ ].$$

Let $x \in \mathrm{dom}[g]$ be given. We wish to show:

$$[\ |x|_V < \delta\ ] \quad \Rightarrow \quad [\ |g(x)|_W < \varepsilon\ ].$$

Assume $|x|_V < \delta$. We wish to show $|g(x)|_W < \varepsilon$.

By (3) of Remark 26.6, $\mathrm{dom}[f_p^T] = (\mathrm{dom}[f]) - p$. Then

$$x \ \in\ \mathrm{dom}[g] \ =\ \mathrm{dom}[f_p^T] \ =\ (\mathrm{dom}[f]) \ -\ p,$$

and so $x+p \in \mathrm{dom}[f]$. Let $w := x+p$. Then $w \in \mathrm{dom}[f]$. Also, we have $|w - p|_V = |x|_V < \delta$. So, by choice of δ, we get $|[f(w)] - [f(p)]|_W < \varepsilon$. We have $g(x) = f_p^T(x) = [f(x+p)] - [f(x)] = [f(w)] - [f(p)]$. Then $|g(x)|_W = |[f(w)] - [f(p)]|_W < \varepsilon$. *End of proof of* $\Rightarrow$. □

THEOREM 26.17. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$, and $p \in \mathbb{R}$. Assume f is differentiable at p. Then f is continuous at p.*

Proof. By Definition 26.9, we see that $p \in \mathrm{dom}[f']$. By $\Leftarrow$ of Fact 26.16, it suffices to show that f_p^T is continuous at 0. By Fact 26.14, we have $(\mathrm{id}_{\mathbb{R}}) \cdot (STS_f^p) = f_p^T$. By Remark 26.13, STS_f^p is continuous at 0. Then, by Fact 19.22, it suffices to show that $\mathrm{id}_{\mathbb{R}}$ is continuous at 0.

The map $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ is distance-semidecreasing, hence 1-Lipschitz, hence uniformly continuous, hence continuous (see Fact 22.8). It follows that $\mathrm{id}_{\mathbb{R}}$ is continuous at 0, as desired. □

Let $+, / : \mathbb{R}^* \times \mathbb{R}^* \dashrightarrow \mathbb{R}^*$ denote addition and division, respectively. Then: $\forall a, b, h$, $(a+b)/h = (a/h) + (b/h)$. From this we see:

FACT 26.18. *Let W be a vector space, let $f, g : \mathbb{R} \dashrightarrow W$ and let $p \in \mathbb{R}$. Then $SS_{f+g}^p = [SS_f^p] + [SS_g^p]$.*

Proof. We wish to show: $\forall h$, $(SS_{f+g}^p)(h) = [(SS_f^p)(h)] + [(SS_g^p)(h)]$. Let h be given. We wish to show: $(SS_{f+g}^p)(h) = [(SS_f^p)(h)] + [(SS_g^p)(h)]$.

We have

$$\begin{aligned}(SS^p_{f+g})(h) &= \frac{[(f+g)(p+h)]-[(f+g)(p)]}{h}\\ &= \frac{[f(p+h)]+[g(p+h)]-[f(p)]-[g(p)]}{h}\\ &= \frac{[f(p+h)]-[f(p)]+[g(p+h)]-[g(p)]}{h}\\ &= \left[\frac{[f(p+h)]-[f(p)]}{h}\right]+\left[\frac{[g(p+h)]-[g(p)]}{h}\right]\\ &= [(SS^p_f)(h)]+[(SS^p_g)(h)],\end{aligned}$$

as desired. □

The preceding result, Additivity of Secant Slope, is a precalculus precursor of the next result.

The next result is called **Additivity of Limit**.

THEOREM 26.19. *Let W be a normed vector space, $f, g : \mathbb{R} \dashrightarrow W$ and $p \in \mathrm{LPD}_{\mathbb{R}}(f+g)$. Then $(f+g)'(p) =^* [f'(p)] + [g'(p)]$.*

Proof. Let $z := [f'(p)] + [g'(p)]$. We wish to show:

$$[\ z \neq ☹\ ] \quad \Rightarrow \quad [\ (f+g)'(p) = z\ ].$$

Assume: $z \neq ☹$. We wish to show: $(f+g)'(p) = z$. Then, by Definition 25.23, we wish to show: $\lim\limits_0 (S^p_{f+g}) = z$.

By (3) of Remark 26.7, $\mathrm{LPD}_{\mathbb{R}}(SS^p_{f+g}) = [\mathrm{LPD}_{\mathbb{R}}(f+g)] - p$. By assumption $p \in \mathrm{LPD}_{\mathbb{R}}(f+g)$. Then

$$0 = p - p \in [\mathrm{LPD}_{\mathbb{R}}(f+g)] - p = \mathrm{LPD}_{\mathbb{R}}(SS^p_{f+g}).$$

Then, by $\Leftarrow$ of Remark 25.19, it suffices to show: $S^p_{f+g} \to z$ near 0.

Since $[f'(p)] + [g'(p)] = z \neq ☹$, it follows that $f'(p) \neq ☹ \neq g'(p)$. So, since $\lim\limits_0 (SS^p_f) = f'(p)$ and $\lim\limits_0 (SS^p_g) = g'(p)$, by Remark 25.18,

$$(\ SS^p_f \to f'(p) \text{ near } 0\ ) \qquad \text{and} \qquad (\ SS^p_g \to g'(p) \text{ near } 0\ ).$$

Then, by Theorem 19.17, $[SS^p_f] + [SS^p_g] \to [f'(p)] + [g'(p)]$ near 0. By Fact 26.18, $SS^p_{f+g} = [SS^p_f] + [SS^p_g]$. So, as $z = [f'(p)] + [g'(p)]$, we get: $S^p_{f+g} \to z$ near 0, as desired. □

We noted that, if $f = C^1_{(-\infty,0]}$ and $g = C^2_{[0,\infty)}$, then $f + g = C^3_{\{0\}}$, and we get $(f+g)'(0) = ☹$, $f'(0) = 0$ and $g'(0) = 0$, and it follows that

$(f+g)'(0) \neq [f'(0)] + [g'(0)]$. So the assumption, in Theorem 26.19, that $p \in \mathrm{LPD}_{\mathbb{R}}(f+g)$ cannot be omitted.

We noted that, if $f, g : \mathbb{R} \to \mathbb{R}$ are defined by $f(x) = |x| + 1$ and $g(x) = -|x| - 1$, then $f + g = C^0_{\mathbb{R}}$, and we get $(f+g)'(0) = 0$, $f'(0) = ☹$ and $g'(0) = ☹$, and it follows that $(f+g)'(0) \neq [f'(0)] + [g'(0)]$. So, in Theorem 26.19, we cannot change $=^*$ to $=$.

For all $a, a_1, b, b_1 \in \mathbb{R}$,

$$a_1 b_1 \ - \ ab \ = \ (a_1 - a)\, b \ + \ a\,(b_1 - b) \ + \ (a_1 - a)\,(b_1 - b);$$

if we use $\triangle(ab)$ to abbreviate $a_1 b_1 - ab$, and $\triangle a$ to abbreviate $a_1 - a$, and $\triangle b$ to abbreviate $b_1 - b$, then we have

$$\triangle(ab) \ = (\triangle a)\, b \ + \ a\,(\triangle b) \ + \ (\triangle a)\,(\triangle b).$$

FACT 26.20. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$, Then, for all p, h,*

(1) $(fg)^T_p(h) = [f^T_p(h)][g(p)] + [f(p)][g^T_p(h)] + [f^T_p(h)][g^T_p(h)]$ *and*
(2) $SS^p_{fg}(h) = [SS^p_f(h)][g(p)] + [f(p)][SS^p_g(h)] + [SS^p_f(h)][SS^p_g(h)][h]$.

Proof. Let p and h be given. We wish to prove (1) and (2).

One of the following must be true:

(A) ([$p \notin \mathrm{dom}[f]$] or [$p + h \notin \mathrm{dom}[f]$] or [$h \notin \mathbb{R}$]) or
(B) ([$p \in \mathrm{dom}[f]$] and [$p + h \in \mathrm{dom}[f]$] and [$h \in \mathbb{R}$]).

Case A: We have both

$$(fg)^T_p(h) = ☹ = [f^T_p(h)][g(p)] + [f(p)][g^T_p(h)] + [f^T_p(h)][g^T_p(h)]$$

and

$$SS^p_{fg}(h) = ☹ = [SS^p_f(h)][g(p)] + [f(p)][SS^p_g(h)] + [SS^p_f(h)][SS^p_g(h)][h],$$

as desired. *End of Case A.*

Case B: Proof of (1): Let

$$a_1 := f(p+h), \quad b_1 := g(p+h), \quad a := f(p), \quad b := g(p).$$

By Definition 25.21, $a_1 - a = f^T_p(h)$ and $b_1 - b = g^T_p(h)$. So, since

$$a_1 b_1 \ - \ ab \ = \ (a_1 - a)b \ + \ a(b_1 - b) \ + \ (a_1 - a)(b_1 - b),$$

we get

$$(fg)^T_p(h) \ = \ [f^T_p(h)][g(p)] \ + \ [f(p)][g^T_p(h)] \ + \ [f^T_p(h)][g^T_p(h)],$$

as desired. *End of proof of (1).*

Proof of (2): Dividing (1) of Fact 26.20 by h, we get

$$\begin{aligned}\frac{(fg)_p^T(h)}{h} \quad &= \quad \left[\frac{f_p^T(h)}{h}\right][g(p)] \;+\; [f(p)]\left[\frac{g_p^T(h)}{h}\right] \\ &\qquad + \left[\frac{f_p^T(h)}{h}\right]\left[\frac{g_p^T(h)}{h}\right][h],\end{aligned}$$

and, by (1) of Remark 26.7, it follows that

$$\begin{aligned}SS_{fg}^p(h) \quad &= \quad [SS_f^p(h)][g(p)] \;+\; [f(p)][SS_g^p(h)] \\ &\qquad + [SS_f^p(h)][SS_g^p(h)][h],\end{aligned}$$

as desired. *End of proof of (2). End of Case B.* □

COROLLARY 26.21. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$, Then, for all p,*

$$SS_{fg}^p \;=\; [SS_f^p]\,[C_{\mathbb{R}}^{g(p)}] \;+\; [C_{\mathbb{R}}^{f(p)}]\,[SS_g^p] \;+\; [SS_f^p]\,[SS_g^p]\,[\mathrm{id}_{\mathbb{R}}].$$

Proof. Let $\Psi := [SS_f^p]\,[C_{\mathbb{R}}^{g(p)}] + [C_{\mathbb{R}}^{f(p)}]\,[SS_g^p] + [SS_f^p]\,[SS_g^p]\,[\mathrm{id}_{\mathbb{R}}]$. We wish to show that $SS_{fg}^p = \Psi$. Since $\mathbb{R}$ is a common superdomain for both SS_{fg}^p and Ψ, We wish to show: $\forall h \in \mathbb{R}$, $(SS_{fg}^p)(h) = \Psi(h)$. Let $h \in \mathbb{R}$ be given. We wish to show: $(SS_{fg}^p)(h) = \Psi(h)$.

By (2) of Fact 26.20,

$$SS_{fg}^p(h) = [SS_f^p(h)][g(p)] + [f(p)][SS_g^p(h)] + [SS_f^p(h)][SS_g^p(h)][h].$$

By definition of Ψ, we have

$$\Psi(h) = [SS_f^p(h)][g(p)] + [f(p)][SS_g^p(h)] + [SS_f^p(h)][SS_g^p(h)][h].$$

Then $(SS_{fg}^p)(h) = \Psi(h)$, as desired. □

The preceding result, the Product Rule for Secant Slopes, is a precalculus precursor of the next result.

The next result is called the **Product Rule**.

THEOREM 26.22. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ and $p \in \mathrm{LPD}_{\mathbb{R}}(fg)$. Then $(fg)'(p) =^* [f'(p)]\,[g(p)] + [f(p)]\,[g'(p)]$.*

Proof. Let $z := [f'(p)]\,[g(p)] + [f(p)]\,[g'(p)]$. We wish to show:

$$[\; z \neq ☹ \;] \quad \Rightarrow \quad [\; (fg)'(p) \;=\; z \;].$$

Assume: $z \neq ☹$. We wish to show: $(fg)'(p) = z$. Then, by Definition 25.23, we wish to show: $\lim\limits_0 (S_{fg}^p) = z$.

By (3) of Remark 26.7, $\mathrm{LPD}_{\mathbb{R}}(SS^p_{fg}) = [\mathrm{LPD}_{\mathbb{R}}(fg)] - p$. By assumption $p \in \mathrm{LPD}_{\mathbb{R}}(fg)$. Then

$$0 \;=\; p - p \;\in\; [\mathrm{LPD}_{\mathbb{R}}(fg)] - p \;=\; \mathrm{LPD}_{\mathbb{R}}(SS^p_{fg}).$$

Then, by $\Leftarrow$ of Remark 25.19, it suffices to show: $S^p_{fg} \to z$ near 0.

Since $[f'(p)]\,[g(p)] \,+\, [f(p)]\,[g'(p)] \,=\, z \,\neq\, ☹$, it follows that

$$f'(p) \qquad \neq \qquad ☹ \qquad \neq \qquad g'(p).$$

So, since $\lim\limits_0 (SS^p_f) = f'(p)$ and $\lim\limits_0 (SS^p_g) = g'(p)$, by Remark 25.18,

$$(\; SS^p_f \to f'(p) \text{ near } 0 \;) \qquad \text{and} \qquad (\; SS^p_g \to g'(p) \text{ near } 0 \;).$$

By Corollary 26.21, we have

$$SS^p_{fg} \quad = \quad [SS^p_f]\,[C^{g(p)}_{\mathbb{R}}] \,+\, [C^{f(p)}_{\mathbb{R}}]\,[SS^p_g] \,+\, [SS^p_f]\,[SS^p_g]\,[\mathrm{id}_{\mathbb{R}}].$$

By Theorem 19.17 and Theorem 19.23, we see that

$$\begin{aligned} & [SS^p_f]\,[C^{g(p)}_{\mathbb{R}}] \,+\, [C^{f(p)}_{\mathbb{R}}]\,[SS^p_g] \,+\, [SS^p_f]\,[SS^p_g]\,[\mathrm{id}_{\mathbb{R}}] \\ \to\;\; & [f'(p)]\,[g(p)] \,+\, [f(p)]\,[g'(p)] \,+\, [f'(p)]\,[g'(p)]\,[0] \end{aligned}$$

near 0. Also, we have

$$\begin{aligned} & [f'(p)]\,[g(p)] \,+\, [f(p)]\,[g'(p)] \,+\, [f'(p)]\,[g'(p)]\,[0] \\ =\;\; & [f'(p)]\,[g(p)] \,+\, [f(p)]\,[g'(p)] \;=\; z\,. \end{aligned}$$

Then $S^p_{fg} \to z$ near 0, as desired. ☐

FACT 26.23. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$, let $p \in dom[f]$, and let $h \in \mathbb{R}$. Let $k := f^T_p(h)$. Let $q := f(p)$. Assume that $q \in dom[g']$. Then $(SS^p_{g\circ f})(h) \;=\; [(STS^q_g(k))]\,[(SS^p_f)(h)]$.*

Proof. Next semester. ☐

COROLLARY 26.24. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$. Then, for all p,*

$$SS^p_{g\circ f} \quad =^* \quad [\,(\,STS^{f(p)}_g\,) \circ (\,f^T_p\,)\,] \;\cdot\; [\,SS^p_f\,] \qquad \textit{on } \mathbb{R}.$$

Proof. Next semester. ☐

By Definition 20.7, the conclusion means: $\forall p$, $\forall h \in \mathbb{R}$,

$$(\; SS^p_{g\circ f}\;)(\;h\;) \quad =^* \quad \Big(\,[\,(\,STS^{f(p)}_g\,) \circ (\,f^T_p\,)\,] \;\cdot\; [\,SS^p_f\,]\,\Big)\;(\;h\;).$$

The preceding result, the Chain Rule for Secant Slopes, is a precursor of the next result. We indicated proofs of both of them, but will give detailed arguments next semester.

The next result is called the **Chain Rule**.

THEOREM 26.25. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$. Let $p \in \mathrm{LPD}_{\mathbb{R}}(g \circ f)$. Then $(g \circ f)'(p) \;=^*\; [f'(g(p))]\,[g'(p)]$.*

Proof. Next semester. $\square$

27. Class 1 on 16 January 2018, Tu of Week 1

Assigned HW#1-1 and HW#1-2.

THEOREM 27.1. *Let K be a sequentially compact metric space, let Y be a metric space and let $f : K \to Y$ be continuous. Then $f : K \to Y$ is uniformly continuous.*

Proof. For all $p \in K$, we define $\overline{p} := f(p)$; then $\overline{p} \in Y$. For all $p \in K^{\mathbb{N}}$, we define $\overline{p} := f \circ p$; then $\overline{p} \in Y^{\mathbb{N}}$.

Assume f is not uniformly continuous. We aim for a contradiction. By HW#1-2, choose $\varepsilon > 0$ and $s, t \in K^{\mathbb{N}}$ and $u \in K$ s.t.

$$\begin{aligned} (\quad & [\ s_\bullet \to u \text{ in } K\ ] \quad \text{and} \quad [\ t_\bullet \to u \text{ in } K\ ] \quad \text{and} \\ & [\ \forall j \in \mathbb{N}, \qquad d_Y(\overline{s}_j, \overline{t}_j) \geqslant \varepsilon\ ] \qquad). \end{aligned}$$

Since f is continuous, it follows that f is continuous at u. So, since $s_\bullet \to u$ in K, by Lemma 25.5, we see that $\overline{s}_\bullet \to \overline{u}$ in Y. Also, since $t_\bullet \to u$ in K, by Lemma 25.5, we see that $\overline{t}_\bullet \to \overline{u}$ in Y.

Since $\overline{s}_\bullet \to \overline{u}$ in Y, choose $L \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant L\ ] \quad \Rightarrow \quad [\ d_Y(\overline{s}_j, \overline{u}) < \varepsilon/2\ ].$$

Since $\overline{t}_\bullet \to \overline{u}$ in Y, choose $M \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant M\ ] \quad \Rightarrow \quad [\ d_Y(\overline{t}_j, \overline{u}) < \varepsilon/2\ ].$$

Let $j := \max\{L, M\}$. Then $j \in \mathbb{N}$. Since $j \geqslant L$, we get $d_Y(\overline{s}_j, \overline{u}) < \varepsilon/2$. Also, since $j \geqslant M$, we get $d_Y(\overline{t}_j, \overline{u}) < \varepsilon/2$. Also, by choice of s and t, we conclude that $d_Y(\overline{s}_j, \overline{t}_j) \geqslant \varepsilon$. By the triangle inequality, we have $d_Y(\overline{s}_j, \overline{t}_j) \leqslant [d(\overline{s}_j, \overline{u})] + [d(\overline{u}, \overline{t}_j)]$. Then

$$\begin{aligned} \varepsilon \;\;\leqslant\;\; d_Y(\overline{s}_j, \overline{t}_j) \;\;&\leqslant\;\; [d(\overline{s}_j, \overline{u})] + [d(\overline{u}, \overline{t}_j)] \\ &<\;\; [\varepsilon/2] + [\varepsilon/2] \;\;=\;\; \varepsilon, \end{aligned}$$

so $\varepsilon < \varepsilon$. Contradiction. $\square$

DEFINITION 27.2. *Let Z be a set and let $D : Z \times Z \to [0, \infty]$. Then D is an* **extended metric** *means:*

(1) $\forall p, q \in Z, \quad (\ [\ D(p, q) = 0\ ] \;\Leftrightarrow\; [\ p = q\ ]\ ),$

(2) $\forall p, q \in Z, \quad (\ D(p,q) \ = \ D(q,p) \ ) \qquad and$
(3) $\forall p, q, r \in Z, \quad (\ D(p,r) \ \leqslant \ [\ D(p,q) \ ] \ + \ [\ D(q,r) \ ] \ ).$

Note: an extended metric is a metric iff its image is contained in $\mathbb{R}$.

DEFINITION 27.3. *For any set Z, $\mathcal{M}^*(Z) := \{$extended metrics on $Z\}$.*

DEFINITION 27.4. *Let Z be a set, $D \in \mathcal{M}^*(Z)$, $p \in Z^{\mathbb{N}}$ and $q \in Z$. Then $p_\bullet \to q$ in (Z, D) means: $\forall \varepsilon > 0$, $\exists K \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,*

$$[\ j \geqslant K \ ] \quad \Rightarrow \quad [\ D(p_j, q) \ < \ \varepsilon \ ].$$

DEFINITION 27.5. *Let X be a set, let (Y, d) be a metric space and let $Z := Y^X$. Then $D_X^d \in \mathcal{M}^*(X)$ is defined by*

$$D_X^d(p,q) \quad = \quad \sup \ \{ \ d\,(\,p(x)\,,\,q(x)\,) \ | \ x \in X \ \}.$$

Let X be a set and let (Y, d) be a metric space. We call D_X^d the **uniform metric** of X and d. Let $Z := Y^X$. For any $p \in Z^{\mathbb{N}}$, for any $q \in Z$, by $p_\bullet$ **converges uniformly** to q with respect to X and d, we mean: $p_\bullet \to q$ in (Z, D_X^d).

We considered the special case where $X = \mathbb{R}$ and $Y = \mathbb{R}$ and $d = d_{\mathbb{R}}$. Let $Z := \mathbb{R}^{\mathbb{R}}$ and let $\varepsilon > 0$. We drew the graph of an example function $f \in Z$. We drew the graph of $f + C_{\mathbb{R}}^\varepsilon$ and $f - C_{\mathbb{R}}^\varepsilon$, and we shaded the region between those two graphs. Call the closed shaded region R. We made the point that, for any $p \in Z$, we have:

$$[\ D_X^d(p, f) \ \leqslant \ \varepsilon \ ] \quad \Leftrightarrow \quad [\text{ the graph of } p \text{ is in } R \ ].$$

We again considered the special case where $X = \mathbb{R}$ and $Y = \mathbb{R}$ and $d = d_{\mathbb{R}}$. Let $Z := \mathbb{R}^{\mathbb{R}}$ and let $f := \mathrm{adj}_0^1(C_{\mathbb{R}}^0)$. For all $x \in \mathbb{R}$, we have

$$f(x) \quad = \quad \begin{cases} 0, & \text{if } x \neq 0 \\ 1, & \text{if } x = 0. \end{cases}$$

We drew the graph of f; note that $f : \mathbb{R} \to \mathbb{R}$ is discontinuous at 0. Define $p \in Z^{\mathbb{N}}$ by: $\forall j \in \mathbb{N}$, $\forall x \in \mathbb{R}$,

$$p_j(x) = \begin{cases} 0, & \text{if } x \leqslant -1/j \\ 1 + jx, & \text{if } -1/j \leqslant x \leqslant 0 \\ 1 - jx, & \text{if } 0 \leqslant x \leqslant 1/j \\ 0, & \text{if } 1/j \leqslant x. \end{cases}$$

We drew the graphs of p_1 and p_2 and p_3 and p_4; note that, for all $j \in \mathbb{N}$, $p_j : \mathbb{R} \to \mathbb{R}$ is continuous. We observed: $\forall x \in \mathbb{R}$, $p_j(x) \to f(x)$. This observation is expressed by saying "$p_\bullet$ converges pointwise to f". Thus a pointwise limit of contiuous functions is not necessarily continuous. We drew the graphs of $f + C_{\mathbb{R}}^{1/3}$ and $f - C_{\mathbb{R}}^{1/3}$ and observed:

$$\text{NOT} [\ p_\bullet \to f \text{ in } (Z, D_X^d)\ ].$$

That is, $p_\bullet$ does NOT converge uniformly to f. The next result asserts that a uniform limit of continuous functions is continuous.

THEOREM 27.6. *Let X be a topological space and let (Y, d) be a metric space. Let $Z := Y^X$ and let $D := D_X^d$. Let $p \in Z^{\mathbb{N}}$ and let $f \in Z$. Assume that $p_\bullet \to f$ in (Z, D). Assume, for all $j \in \mathbb{N}$, that $p_j : X \to Y$ is continuous. Then $f : X \to Y$ is continuous.*

Proof. We wish to show: $\forall a \in X$, f is continuous at a. Let $a \in X$ be given. We wish to show: f is continuous at a. We wish to show: $\forall \varepsilon > 0$, $\exists U \in \mathcal{N}_X(a)$ s.t., $\forall x \in X$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ d\,(\,f(x)\,,\,f(a)\,)\ <\ \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists U \in \mathcal{N}_X(a)$ s.t., $\forall x \in X$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ d\,(\,f(x)\,,\,f(a)\,)\ <\ \varepsilon\ ].$$

Since $p_\bullet \to f$ in (Z, D), choose $K \in \mathbb{N}$ s.t., $\forall j \in \mathbb{N}$,

$$[\ j \geqslant K\ ] \quad \Rightarrow \quad [\ D\,(\,p_j\,,\,f\,)\ <\ \varepsilon/3\ ].$$

By assumption, $p_K : \mathbb{R} \to \mathbb{R}$ is continuous. Also, $D(p_k, f) < \varepsilon/3$. Let $q := p_K$. Then $q : \mathbb{R} \to \mathbb{R}$ is continuous; it follows that q is continuous at a. Also, $D(q, f) < \varepsilon/3$; it follows, for all $x \in X$, that $d(q(x), f(x)) < \varepsilon/3$. In particular, we have $d(q(a), f(a)) < \varepsilon/3$. Since q is continuous at a, choose $U \in \mathcal{N}_X(a)$ s.t., $\forall x \in X$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ d\,(\,q(x)\,,\,q(a)\,)\ <\ \varepsilon/3\ ].$$

We wish to show: $\forall x \in X$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ d\,(\,f(x)\,,\,f(a)\,)\ <\ \varepsilon\ ].$$

Let $x \in X$ be given. We wish to show:

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ d\,(\,f(x)\,,\,f(a)\,)\ <\ \varepsilon\ ].$$

Assume: $x \in U$. We wish to show: $d(f(x), f(a)) < \varepsilon$.

Recall: $d(f(a), q(a)) < \varepsilon/3$. Then we have: $d(q(a), f(a)) < \varepsilon/3$. Since $x \in X$, we get $d(f(x), q(x)) < \varepsilon/3$. Since $x \in U$, by choice of U, we see that $d(q(x), q(a)) < \varepsilon/3$. By the triangle inequality, we have

$$d(f(x), f(a)) \leqslant [d(f(x), q(x))] + [d(q(x), q(a))] + [d(q(a), f(a))].$$

Then $d(f(x), f(a)) < [\varepsilon/3] + [\varepsilon/3] + [\varepsilon/3] = \varepsilon$, as desired. □

DEFINITION 27.7. *Let X be a metric space, let $A \subseteq X$ and let $\varepsilon > 0$. Then A is ε-**crowded** means: $\forall p, q \in A$, we have $d_X(p, q) < \varepsilon$.*

DEFINITION 27.8. *Let X be a metric space and let $s \in X^{\mathbb{N}}$. Then $s_\bullet$ is **Cauchy** means: $\forall \varepsilon > 0$, $\exists K \in \mathbb{N}$ s.t., $\forall i, j \in \mathbb{N}$,*

$$[\ i, j \geqslant K\ ] \quad \Rightarrow \quad [\ d_X(s_i, s_j) < \varepsilon\ ].$$

Note: a sequence in a metric space is Cauchy iff, for every $\varepsilon > 0$, it has an ε-crowded tail.

Assigned HW#1-3.

Recall: For any sequence s, for any subsequence t of $s_\bullet$, for any $j \in \mathbb{N}$,

$$\exists j \in [j..\infty) \quad \text{s.t.} \quad t_j = s_k.$$

Assigned HW#1-4, HW#1-5.

By HW#1-3, we see that convergent implies Cauchy. The converse does not hold. For example: Let $X := \mathbb{Q}$ (with the standard metric on $\mathbb{Q}$, *i.e.*, with the relative metric on $\mathbb{Q}$ inherited from $d_{\mathbb{R}}$ on $\mathbb{R}$). Let $s := (3, 3.1, 3.14, 3.141, 3.1415, 3.14159, \ldots)$ be the sequence of decimal approximations to π. Since $s_\bullet$ is convergent in $\mathbb{R}$, it follows, from HW#1-3, that $s_\bullet$ is Cauchy in $\mathbb{R}$. So, since $s_\bullet \in \mathbb{Q}^{\mathbb{N}}$, we conclude that $s_\bullet$ is Cauchy in $\mathbb{Q}$. However, $s_\bullet$ is not convergent in $\mathbb{Q}$.

By HW#1-5, we see that Cauchy implies bounded. The converse does not hold. For example: Let $X := \mathbb{R}$ (with its standard metric $d_{\mathbb{R}}$). Let $s := (-1, 1, -1, 1, -1, 1, -1, 1, \ldots)$. Then $s_\bullet$ is bounded. However, as no tail of $s_\bullet$ is 2-crowded, we see that $s_\bullet$ is not Cauchy in $\mathbb{R}$.

DEFINITION 27.9. *Let X be a metric space. Then X is **complete** means: $\forall s \in X^{\mathbb{N}}$, $[\,(\ s_\bullet$ is Cauchy in $X\ ) \Rightarrow (\ s_\bullet$ is convergent in $X\ )\,]$.*

By considering the decimal approximations to π, we observed that Cauchy does not imply convergent in $\mathbb{Q}$. Therefore, $\mathbb{Q}$ is not complete.

In the next class, we will use HW#1-4 and HW#1-5 to show that every proper metric space is complete. We will also show, by example, that the converse is not true.

In Corollary 24.13, we showed that $\mathbb{R}$ is proper; it follows that $\mathbb{R}$ is complete. In Corollary 24.14, we showed that $\mathbb{R}^2$ is proper; it follows that $\mathbb{R}^2$ is complete. Similar arguments show, for all $k \in \mathbb{N}$, that $\mathbb{R}^k$ is proper and, therefore, complete.

28. Class 2 on 18 January 2018, Th of Week 1

Recall that, for any metric space X, for any $s \in X^{\mathbb{N}}$, by $s_\bullet$ is **Cauchy**, we mean: $\forall \varepsilon > 0$, $\exists K \in \mathbb{N}$ s.t., $\forall i, j \in \mathbb{N}$,

$$[\ i, j \geqslant K\ ] \quad \Rightarrow \quad [\ d(s_i, s_j) < \varepsilon\ ].$$

That is, we mean: $\forall \varepsilon > 0$, there is an ε-crowded tail of $s_\bullet$.

By HW#1-4, Cauchy and subconvergent implies convergent.

By HW#1-5, Cauchy implies bounded.

Recall: For any metric space X,

(1) X is **proper** means: $\forall$bounded $s \in X^{\mathbb{N}}$, $s_\bullet$ is subconvergent;
(2) X is **complete** means: $\forall$Cauchy $s \in X^{\mathbb{N}}$, $s_\bullet$ is convergent.

Recall that we showed that $\mathbb{R}$ and $\mathbb{R}^2$ are both proper. Recall that we showed that $\mathbb{Q}$ is not complete.

We next show that proper implies complete:

THEOREM 28.1. *Let X be a metric space. Assume that X is proper. Then X is complete.*

Proof. We wish to show: $\forall$Cauchy $s \in X^{\mathbb{N}}$, $s_\bullet$ is convergent. Let a Cauchy $s \in X^{\mathbb{N}}$ be given. We wish to show: $s_\bullet$ is convergent.

Since $s_\bullet$ is Cauchy, by HW#1-5, we conclude that $s_\bullet$ is bounded. So, as X is proper, we see that $s_\bullet$ is subconvergent. Then $s_\bullet$ is Cauchy and subconvergent. So, by HW#1-4, we conclude that $s_\bullet$ is convergent. $\square$

By Corollary 24.13, $\mathbb{R}$ is proper. Then, by Theorem 28.1, $\mathbb{R}$ is complete. By Corollary 24.14, $\mathbb{R}^2$ is proper. Then, by Theorem 28.1, $\mathbb{R}^2$ is complete. We will argue, below, that: $\forall k \in \mathbb{N}$, $\mathbb{R}^k$ is proper and complete.

While proper implies complete, the converse is not true. There are many important infinite dimensional normed vector spaces that are complete, but not proper. In this course, we focus on finite dimensional vector spaces, and it's not hard to construct a metric δ on $\mathbb{R}$ such that $(\mathbb{R},\delta)$ is complete, but not proper. We proceed as follows.

Recall, for all $a,b\in\mathbb{R}$, that $d_{\mathbb{R}}(a,b)=|a-b|$. We define $\delta\in\mathcal{M}(\mathbb{R})$ by $\delta(a,b)=\min\{|a-b|,1\}$. Then $B_\delta(0,2)=\mathbb{R}$. Unassigned HW: Show

(1) $\forall s\in\mathbb{R}^{\mathbb{N}}$, [($s_\bullet$ is δ-Cauchy) iff ($s_\bullet$ is $d_{\mathbb{R}}$-Cauchy)] and
(2) $\forall s\in\mathbb{R}^{\mathbb{N}}$, [($s_\bullet$ is δ-convergent) iff ($s_\bullet$ is $d_{\mathbb{R}}$-convergent)].

Since $(\mathbb{R},d_{\mathbb{R}})$ is complete, we know:

(3) $\forall s\in\mathbb{R}^{\mathbb{N}}$, [($s_\bullet$ is $d_{\mathbb{R}}$-Cauchy) iff ($s_\bullet$ is $d_{\mathbb{R}}$-convergent)].

By (1), (2) and (3), we conclude:

- $\forall s\in\mathbb{R}^{\mathbb{N}}$, [($s_\bullet$ is δ-Cauchy) iff ($s_\bullet$ is δ-convergent)].

That is, $(\mathbb{R},\delta)$ is complete. It remains to show that $(\mathbb{R},\delta)$ is not proper. Let $s:=(1,2,3,4,\ldots)\in\mathbb{R}^{\mathbb{N}}$. Then $\mathrm{im}[s_\bullet]\in\mathbb{R}=B_\delta(0,2)$, so $s_\bullet$ is bounded in $(\mathbb{R},\delta)$. However, no subsequence of $s_\bullet$ is convergent in $(\mathbb{R},d_{\mathbb{R}})$, so, by (2), no subsequence of $s_\bullet$ is convergent in $(\mathbb{R},\delta)$. That is, $s_\bullet$ is not subconvergent in $(\mathbb{R},\delta)$. Since $s_\bullet$ is bounded and not subconvergent in $(\mathbb{R},\delta)$, it follows that $(\mathbb{R},\delta)$ is not proper, as desired.

Let X and Y be sets, let $\mathcal{T}$ be a topology on X and let $\mathcal{U}$ be a topology on Y. Let

$$\mathcal{B}:=\{V\times W\,|\,(V\in\mathcal{T})\text{ and }(W\in\mathcal{U})\}.$$

Recall that the **product topology** of $\mathcal{T}$ and $\mathcal{U}$ is $\langle\mathcal{B}\rangle_\cup$; it is a topology on $X\times Y$. There are product constructions for metric spaces and normed vector spaces as well:

For any two topological spaces X and Y, the product topology of $\mathcal{T}_X$ and $\mathcal{T}_Y$ is the standard topology on $X\times Y$.

DEFINITION 28.2. *Let X and Y be sets, let $b\in\mathcal{M}(X)$ and let $c\in\mathcal{M}(Y)$. Then the* **product metric** *of b and c is the metric $d\in\mathcal{M}(X\times Y)$ defined by*

$$d(\,(a,b)\,,\,(x,y)\,)\;=\;\sqrt{[d_X(a,x)]^2\,+\,[d_Y(b,y)]^2}.$$

For any two metric spaces X and Y, the product metric of d_X and d_Y is the standard metric on $X\times Y$.

DEFINITION 28.3. *Let V and W be sets, let $|\bullet|\in\mathcal{N}(V)$ and let $\|\bullet\|\in\mathcal{N}(W)$. Then the* **product norm** *of $|\bullet|$ and $\|\bullet\|$ is the norm $|||\bullet|||\in\mathcal{N}(V\times W)$ defined by*

$$|||\,(v,w)\,|||\;=\;\sqrt{|v|^2\,+\,\|w\|^2}.$$

For any two normed vector spaces V and W, the product norm of $|\bullet|_V$ and $|\bullet|_W$ is the standard norm on $V \times W$.

FACT 28.4. *Let X and Y be metric spaces. Let d be the product metric of d_X and d_Y. Then $\mathcal{T}_d$ is the product topology of $\mathcal{T}_{d_X}$ and $\mathcal{T}_{d_Y}$.*

Proof. Omitted. □

FACT 28.5. *Let V and W be normed vector spaces, and let $|\bullet|$ be the product norm of $|\bullet|_V$ and $|\bullet|_W$. Then $d_{|\bullet|}$ is the product metric of $d_{|\bullet|_V}$ and $d_{|\bullet|_W}$.*

Proof. Omitted. □

THEOREM 28.6. *Let X and Y be metric spaces. Then*

(1) *(X and Y are both proper) implies ($X \times Y$ is proper) and*
(2) *(X and Y are both complete) implies ($X \times Y$ is complete).*

Proof. Proof of (1): Assume: X and Y are both proper. We wish to show: $X \times Y$ is proper. Let $Z := X \times Y$. We wish to show: Z is proper. We wish to show: $\forall s \in Z^{\mathbb{N}}$,

$$(s_\bullet \text{ bounded in } Z) \quad \Rightarrow \quad (s_\bullet \text{ is subconvergent in } Z).$$

Let $s \in Z^{\mathbb{N}}$ be given. We wish to show:

$$(s_\bullet \text{ bounded in } Z) \quad \Rightarrow \quad (s_\bullet \text{ is subconvergent in } Z).$$

Assume: $s_\bullet$ is bounded in Z. Want: $s_\bullet$ is subconvergent in Z.

Let $p : Z \to X$ and $q : Z \to Y$ be the coordinate projections, defined by $p(x, y) = x$ and $q(x, y) = y$. Since $s_\bullet$ is bounded in Z, it follows that $(p \circ s)_\bullet$ is bounded in X. So, since X is complete, $(p \circ s)_\bullet$ is subconvergent in X. Choose a strictly increasing $\ell \in \mathbb{N}^{\mathbb{N}}$ s.t. $(p \circ s \circ \ell)_\bullet$ is convergent in X. Since $s_\bullet$ is bounded in Z, it follows that $(q \circ s)_\bullet$ is bounded in Y, so $(q \circ s \circ \ell)_\bullet$ is bounded in Y. So, since Y is complete, $(q \circ s \circ \ell)_\bullet$ is subconvergent in Y. Choose a strictly increasing $m \in \mathbb{N}^{\mathbb{N}}$ s.t. $(q \circ s \circ \ell \circ m)_\bullet$ is convergent in Y.

Since $(p \circ s \circ \ell)_\bullet$ is convergent in X, it follows that $(p \circ s \circ \ell \circ m)_\bullet$ is convergent in X. So, since $(q \circ s \circ \ell \circ m)_\bullet$ is convergent in Y, we conclude that $(s \circ \ell \circ m)_\bullet$ is convergent in Z. Then $s_\bullet$ is subconvergent in Z, as desired. *End of proof of (1).*

Proof of (2): Unassigned homework. *End of proof of (2).* □

COROLLARY 28.7. *For all $k \in \mathbb{N}$, $\mathbb{R}^k$ is proper and complete.*

Proof. Following Theorem 28.6, this is a simple proof by induction, which we omit. $\square$

Let f be a function. Recall: $\forall$ set A,

$$\begin{array}{rcll} f_*(A) & := & \{f(x) \,|\, x \in A \cap (\mathrm{dom}[f])\} & \text{and} \\ f^*(A) & := & \{x \in \mathrm{dom}[f] \,|\, f(x) \in A\}. & \end{array}$$

Recall: $\forall$ set $\mathcal{S}$ of sets,

$$\begin{array}{rcll} f_{**}(\mathcal{S}) & := & \{f_*(A) \,|\, A \in \mathcal{S}\} & \text{and} \\ f^{**}(\mathcal{S}) & := & \{f^*(A) \,|\, A \in \mathcal{S}\}. & \end{array}$$

FACT 28.8. *Let f be a function and let $\mathcal{S}$ be a set of sets. Then*

(1) $f^*(\bigcup \mathcal{S}) = \bigcup [f^{**}(\mathcal{S})]$,
(2) $f^*(\bigcap \mathcal{S}) = \bigcap [f^{**}(\mathcal{S})]$,
(3) $f_*(\bigcup \mathcal{S}) = \bigcup [f_{**}(\mathcal{S})]$,

Proof. Omitted. $\square$

We gave heuristic explanations of (1), (2) and (3) of Fact 28.8. We also gave a heuristic explanation of why there exist a function f and sets A and B s.t. $f_*(A \cap B) \neq [f_*(A)] \cap [f_*(B)]$.

FACT 28.9. *Let f be a function and let A and B be sets. Then $f^*(A \backslash B) = [f^*(A)] \backslash [f^*(B)]$.*

Proof. Omitted. $\square$

We gave a heuristic explanations of Fact 28.9. We also gave a heuristic explanation of why there exist a function f and sets A and B s.t. $f_*(A \backslash B) \neq [f_*(A)] \backslash [f_*(B)]$.

The previous two facts can be summarized by saying: "Preimage behaves well with respect to all three set theoretic operations (union, intersection and difference). Forward image only behaves well with respect to union."

Recall, by Theorem 21.19, that a map between topological spaces is continuous iff the preimage of any open set is open:

THEOREM 28.10. *Let $(X, \mathcal{T})$ and $(Y, \mathcal{U})$ be topological spaces, and let $f : X \to Y$. Then:* $[\ f \textit{ is continuous}\ ] \quad \Leftrightarrow \quad [\ f^{**}(\mathcal{U}) \subseteq \mathcal{T}\ ]$.

Proof. This follows from Theorem 21.19. $\square$

From this we can show that a map between topological spaces is continuous iff the preimage of any closed set is closed:

THEOREM 28.11. *Let X, Y be topological spaces. Let $f : X \to Y$. Let $\mathcal{C}$ denote the set of all closed sets in X. Let $\mathcal{D}$ denote the set of all closed sets in Y. Then:* $\quad$ [*f is continuous*] $\quad\Leftrightarrow\quad$ [$f^{**}(\mathcal{D}) \subseteq \mathcal{C}$].

Proof. Proof of $\Leftarrow$: Unassigned HW. *End of proof of $\Leftarrow$.*

Proof of $\Rightarrow$: Assume: f is continuous. We wish to show: $f^{**}(\mathcal{D}) \subseteq \mathcal{C}$.

We wish to show: $\forall A \in f^{**}(\mathcal{D})$, $A \in \mathcal{C}$.

Let $A \in f^{**}(\mathcal{D})$ be given. We wish to show: $A \in \mathcal{C}$.

Since $A \in f^{**}(\mathcal{D})$, choose $B \in \mathcal{D}$ s.t. $A = f^*(B)$. Since $B \in \mathcal{D}$, we see that B is closed in Y. Then $Y\backslash B$ is open in Y. Then, by Theorem 28.10, $f^*(Y\backslash B)$ is open in X. Then $X\backslash[f^*(Y\backslash B)]$ is closed in X. That is, $X\backslash[f^*(Y\backslash B)] \in \mathcal{C}$. Since $f^*(Y) = X$, by Fact 28.9, we conclude that $f^*(Y\backslash B) = X\backslash[f^*(B)]$. Then $X\backslash[f^*(Y\backslash B)] = f^*(B)$.

Then $A = f^*(B) = X\backslash[f^*(Y\backslash B)] \in \mathcal{C}$. *End of proof of $\Rightarrow$.* ☐

We have "transitivity of closedness":

FACT 28.12. *Let X be a topological space, $B \subseteq X$, $A \subseteq B$. Assume: (A is closed in B) and (B is closed in X). Then A is closed in X.*

Proof. Since A is closed in B, by the restriction result for closed sets (*i.e.*, by (1) of Corollary 16.17), choose a closed subset C of X such that $A = B \cap C$. Since B and C are closed in X, it follows that $B \cap C$ is closed in X. So, since $A = B \cap C$, we see that A is closed in X. ☐

A similar argument can be used to prove "transitivity of openness".

Recall, by Theorem 25.15, that any nonempty, closed, bounded subset of $\mathbb{R}$ has a max and a min:

THEOREM 28.13. *Let $A \subseteq \mathbb{R}$. Assume that A is nonempty, closed in $\mathbb{R}$, and bounded in $\mathbb{R}$. Then $\exists w, x \in A$ s.t. $w \leqslant A \leqslant x$.*

Proof. This follows from Theorem 25.15. ☐

We call the next result "absoluteness of global continuity". It follows from Theorem 21.4, but we will prove it "from scratch", as an exercise in dealing with restriction and extension results for relative topologies.

REMARK 28.14. *Let X and Y be topological spaces, let $X_0 \subseteq X$ and let $Y_0 \subseteq Y$. Let $f : X_0 \dashrightarrow Y_0$. Then:*

$$[\ f \text{ is } (X_0, Y_0)\text{-continuous}\ ] \quad\Leftrightarrow\quad [\ f \text{ is } (X, Y)\text{-continuous}\ ].$$

Proof. Proof of $\Leftarrow$*:* Unassigned HW. *End of proof of* $\Leftarrow$.

Proof of $\Rightarrow$*:* Assume: f is (X_0, Y_0)-continuous. We wish to show: f is (X, Y)-continuous. That is, we wish to show that, $\forall p \in \mathrm{dom}[f]$, f is (X, Y)-continuous at p. Let $p \in \mathrm{dom}[f]$ be given. We wish to show that f is (X, Y)-continuous at p. We wish to show: $\forall V \in \mathcal{N}_Y(f(p))$, $\exists U \in \mathcal{N}_X(p)$ s.t. $f_*(U) \subseteq V$. Let $V \in \mathcal{N}_Y(f(p))$ be given. We wish to show: $\exists U \in \mathcal{N}_X(p)$ s.t. $f_*(U) \subseteq V$.

Let $V_0 := V \cap Y_0$. By (1) of Corollary 16.19, $V_0 \in \mathcal{N}_{Y_0}(f(p))$. Because f is (X_0, Y_0)-continuous, we know that f is (X_0, Y_0)-continuous at p. So, since $V_0 \in \mathcal{N}_{Y_0}(f(p))$, choose $U_0 \in \mathcal{N}_{X_0}(p)$ s.t. $f_*(U_0) \subseteq V_0$. By (2) of Corollary 16.19, choose $U \in \mathcal{N}_X(p)$ s.t. $U_0 = U \cap X_0$. We wish to show: $f_*(U) \subseteq V$. We wish to show: $\forall q \in \mathrm{dom}[f]$,

$$[\ q \in U\ ] \quad \Rightarrow \quad [\ f(q) \in V\ ].$$

Let $q \in \mathrm{dom}[f]$ be given. We wish to show:

$$[\ q \in U\ ] \quad \Rightarrow \quad [\ f(q) \in V\ ].$$

Assume: $q \in U$. We wish to show: $f(q) \in V$.

Since $f : X_0 \dashrightarrow Y_0$, it follows that $\mathrm{dom}[f] \subseteq X_0$ and $\mathrm{im}[f] \subseteq Y_0$. We have $q \in \mathrm{dom}[f] \subseteq X_0$. So, since $q \in U$, we get $q \in U \cap X_0$. So, since $U_0 = U \cap X_0$, we get $q \in U_0$. So, since $q \in \mathrm{dom}[f]$, we get $f(q) \in f_*(U_0)$. Then $f(q) \in f_*(U_0) \subseteq V_0 = V \cap Y_0 \subseteq V$, as desired. *End of proof of* $\Rightarrow$. $\square$

FACT 28.15. *Let* $x \in \mathbb{R}$*,* $U \in \mathcal{N}_{\mathbb{R}}(x)$*,* $b > x$*. Then* $U \cap (x, b) \neq \varnothing$*.*

Proof. By Theorem 17.15, $\mathcal{B}_{\mathbb{R}}(x)$ is a neighborhood base at x in $\mathbb{R}$. So, since $U \in \mathcal{N}_{\mathbb{R}}(x)$, choose $A \in \mathcal{B}_{\mathbb{R}}(x)$ s.t. $A \subseteq U$. Since $A \in \mathcal{B}_{\mathbb{R}}(x)$, choose $r > 0$ s.t. $A = B_{\mathbb{R}}(x, r)$. Let $s := b - x$. Let $\delta := \min\{r, s\}$. Let $t := x + (\delta/2)$. We wish to show: $t \in U \cap (x, b)$.

We have $d_{\mathbb{R}}(t, x) = \delta/2 < \delta$, so $t \in B_{\mathbb{R}}(x, \delta)$. Since $\delta \leqslant r$, we get $B_{\mathbb{R}}(x, \delta) \subseteq B_{\mathbb{R}}(x, r)$. Then $t \in B_{\mathbb{R}}(x, \delta) \subseteq B_{\mathbb{R}}(x, r) = A \subseteq U$. We wish to show: $t \in (x, b)$. We wish to show $x < t < b$.

Since $\delta > 0$, $\delta/2 > 0$, so $x + (\delta/2) > x$. Then $x < x + (\delta/2) = t$, so $x < t$. We have $\delta/2 < \delta \leqslant s = b - x$, so $x + (\delta/2) < b$. Then $t = x + (\delta/2) < b$, so $t < b$. Then $x < t < b$, as desired. $\square$

We started the proof of the next result, and indicated, using a picture, how to finish it. We'll go through the entire proof next time.

THEOREM 28.16. *Let $a, b, y \in \mathbb{R}$. Assume $a < b$. Let $f : [a, b] \to \mathbb{R}$ be continuous. Assume $f(a) \leqslant y \leqslant f(b)$. Then $\exists x \in [a, b]$ s.t. $f(x) = y$.*

Proof. Next class. □

29. Class 3 on 23 January 2018, Tu of Week 2

Assigned HW#2-1.

FACT 29.1. *Let Y and Z be topological spaces. Let $b \in Y$ and $c \in Z$. Let $W \in \mathcal{N}_{Y \times Z}((b, c))$. Then $\exists U \in \mathcal{N}_Y(b)$, $\exists V \in \mathcal{N}_Z(c)$ s.t. $U \times V \subseteq W$.*

Proof. Since $W \in \mathcal{N}_{Y \times Z}((b, c))$, choose an open set W_0 in $Y \times Z$ s.t. $(b, c) \in W_0 \subseteq W$. Let $\mathcal{S} := \{U \times V \mid (U \text{ open in } Y) \& (V \text{ open in } Z)\}$. Then, as W_0 is open in $Y \times Z$, we get $W_0 \in \langle \mathcal{S} \rangle_\cup$. Then: $\forall z \in W_0$, $\exists A \in \mathcal{S}$ s.t. $z \in A \subseteq W_0$. So, since $(b, c) \in W_0$, choose $A \in \mathcal{S}$ s.t. $(b, c) \in A \subseteq W_0$. Since $A \in \mathcal{S}$, choose U open in Y and V open in Z s.t. $A = U \times V$. As $(b, c) \in A = U \times V$, we get: $b \in U$ and $c \in V$. Since $b \in U$ and since U is open in Y, we get $U \in \mathcal{N}_Y(b)$. Since $c \in V$ and since V is open in Z, we get $V \in \mathcal{N}_Z(c)$. We wish to show: $U \times V \subseteq W$.

We have $U \times V = A \subseteq W_0 \subseteq W$, as desired. □

DEFINITION 29.2. *Let f and g be functions. Then (f, g) is the function defined by: $(f, g)(x) = (f(x), g(x))$.*

Assigned HW#2-2.

Recall that we proved last time:

FACT 29.3. *Let $x \in \mathbb{R}$, $U \in \mathcal{N}_{\mathbb{R}}(x)$, $b > x$. Then $U \cap (x, b) \neq \varnothing$.*

Proof. This is Fact 28.15. □

THEOREM 29.4. *Let $a, b, y \in \mathbb{R}$. Assume $a < b$. Let $f : [a, b] \to \mathbb{R}$ be continuous. Assume $f(a) \leqslant y \leqslant f(b)$. Then $\exists x \in [a, b]$ s.t. $f(x) = y$.*

Proof. Let $S := f^*((-\infty, y])$.

Since $f(a) \leqslant y$, we get $f(a) \in (-\infty, y]$. Then $a \in f^*((-\infty, y]) = S$. Then S is nonempty. Since $(-\infty, y]$ is closed in $\mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ is continuous, it follows that S is closed in $[a, b]$. So, since $[a, b]$ is closed in $\mathbb{R}$, we see, from Fact 28.12, that S is closed in $\mathbb{R}$.

We have $S = f^*((-\infty, y]) \subseteq \mathrm{dom}[f] = [a, b]$. So, since $[a, b]$ is bounded in $\mathbb{R}$, we see that S is bounded in $\mathbb{R}$. So, since S is nonempty and closed in $\mathbb{R}$, by Theorem 28.13, choose $x \in S$ s.t. $S \leqslant x$. Then

$x \in S \subseteq [a,b]$. We wish to show: $f(x) = y$. Since $x \in S = f^*((-\infty, y])$, we get $f(x) \in (-\infty, y]$, so $f(x) \leqslant y$. We wish to show: $f(x) \geqslant y$. Assume: $f(x) < y$. We aim for a contradiction.

Since $f(x) < y \leqslant f(b)$, we get $f(x) < f(b)$, so $f(x) \neq f(b)$, so $x \neq b$. So, since $x \in [a,b]$, we get $x \in [a,b)$. Then $x < b$. Let $\varepsilon := y - [f(x)]$. Since $f(x) < y$, we get $\varepsilon > 0$. Also $[f(x)] + \varepsilon = y$. Let $V := B_{\mathbb{R}}(f(x), \varepsilon)$. Then V is open in $\mathbb{R}$ and $f(x) \in V$, so $V \in \mathcal{N}_{\mathbb{R}}(f(x))$. Also, $V = ([f(x)] - \varepsilon, [f(x)] + \varepsilon) = ([f(x)] - \varepsilon, y)$. Then $V < y$.

By assumption, f is $([a,b], \mathbb{R})$-continuous. So, by Remark 28.14, we see that f is $(\mathbb{R}, \mathbb{R})$-continuous. So, since $x \in S \subseteq [a,b] = \mathrm{dom}[f]$, we see that f is $(\mathbb{R}, \mathbb{R})$-continuous at x. So, since $V \in \mathcal{N}_{\mathbb{R}}(f(x))$, choose $U \in \mathcal{N}_{\mathbb{R}}(x)$ s.t. $f_*(U) \subseteq V$. By Fact 29.3, $U \cap (x,b) \neq \varnothing$, so choose $t \in U \cap (x,b)$. Then $t \in U$ and $t \in (x,b)$. Since $x \in [a,b)$, we get $a \leqslant x$. Then $(x,b) \subseteq (a,b)$. Then $t \in (x,b) \subseteq (a,b) \subseteq [a,b] = \mathrm{dom}[f]$. So, since $t \in U$, we see that $f(t) \in f_*(U)$. Then $f(t) \in f_*(U) \subseteq V < y$. Then $f(t) \in (-\infty, y) \subseteq (-\infty, y]$. Then $t \in f^*((-\infty, y]) = S \leqslant x$, and it follows that $t \leqslant x$. Then $t \notin (x,b)$. Contradiction. $\square$

Assigned HW#2-3.

THEOREM 29.5. *Let $a, b, y \in \mathbb{R}$. Assume $a < b$. Let $f : [a,b] \to \mathbb{R}$ be continuous. Assume $f(a) \geqslant y \geqslant f(b)$. Then $\exists x \in [a,b]$ s.t. $f(x) = y$.*

Proof. Unassigned homework. $\square$

Assigned HW#2-4.

Let $X_0 := [1,2]$ and let $f := \chi_{X_0}^{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ be the characteristic function of X_0 in $\mathbb{R}$. Since f is neither continuous at 1 nor at 2. we see that f is not continuous on X_0. However, $f|X_0$ is constant, so $f|X_0 : X_0 \to \mathbb{R}$ is continuous.

DEFINITION 29.6. *$\forall \alpha, \beta \in \mathbb{R}$,* $\quad [\alpha|\beta] := [\min\{\alpha, \beta\}, \max\{\alpha, \beta\}]$.

For example, $[2|1] = [1,2]$.

The following is called the **Intermediate Value Theorem**:

THEOREM 29.7. *Let $\alpha, \beta \in \mathbb{R}$ and let $f : \mathbb{R} \dashrightarrow \mathbb{R}$. Assume that f is continuous on $[\alpha|\beta]$. Then $[f(\alpha)|f(\beta)] \subseteq f_*([\alpha|\beta])$.*

Proof. Let $a := \min\{\alpha, \beta\}$ and let $b := \max\{\alpha, \beta\}$. Then $a \leqslant b$ and $[\alpha|\beta] = [a,b]$ and $[f(\alpha)|f(\beta)] = [f(a)|f(b)]$. We know: f is continuous on $[a,b]$. We wish to show: $[f(a)|f(b)] \subseteq f_*([a,b])$.

Since f is continuous on $[a,b]$, it follows that $[a,b] \subseteq \mathrm{dom}[f]$. Let $g := f|[a,b]$. Then we have: $g(a) = f(a)$ and $g(b) = f(b)$ and $f_*([a,b]) = g_*([a,b])$. By HW#2-4, $g : [a,b] \to \mathbb{R}$ is continuous. We wish to show: $[g(a)|g(b)] \subseteq g_*([a,b])$. We wish to show: $\forall y \in [g(a)|g(b)]$, $y \in g_*([a,b])$. Let $y \in [g(a)|g(b)]$ be given. We wish to show: $y \in g_*([a,b])$. Since $\mathrm{dom}[g] = [a,b]$, it follows that $g_*([a,b]) = \mathrm{im}[g]$. We wish to show: $y \in \mathrm{im}[g]$. We wish to show: $\exists x \in [a,b]$ s.t. $g(x) = y$.

At least one of the following must be true:

(1) $g(a) \leqslant g(b)$ or
(2) $g(a) \geqslant g(b)$.

Case 1: We have $y \in [g(a)|g(b)] = [g(a), g(b)]$, so $g(a) \leqslant y \leqslant g(b)$. Then, by Theorem 29.4 (with f replaced by g), we see that $\exists x \in [a,b]$ s.t. $g(x) = y$, as desired. *End of Case 1.*

Case 2: We have $y \in [g(a)|g(b)] = [g(b), g(a)]$, so $g(a) \geqslant y \geqslant g(b)$. Then, by Theorem 29.5 (with f replaced by g), we see that $\exists x \in [a,b]$ s.t. $g(x) = y$, as desired. *End of Case 2.* □

Recall the following:

DEFINITION 29.8. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$.

(1) f *is* **semiincreasing** *means:*
$$\forall a, b \in dom[f], \quad [\,(a \leqslant b) \Rightarrow (f(a) \leqslant f(b))\,].$$
(2) f *is* **semidecreasing** *means:*
$$\forall a, b \in dom[f], \quad [\,(a \leqslant b) \Rightarrow (f(a) \geqslant f(b))\,].$$
(3) f *is* **strictly increasing** *means:*
$$\forall a, b \in dom[f], \quad [\,(a < b) \Rightarrow (f(a) < f(b))\,].$$
(4) f *is* **strictly decreasing** *means:*
$$\forall a, b \in dom[f], \quad [\,(a < b) \Rightarrow (f(a) > f(b))\,].$$
(5) f *is* **semimonotone** *means:*
(f is semiincreasing) or (f is semidecreasing).
(6) f *is* **strictly monotone** *means:*
(f is strictly increasing) or (f is strictly decreasing).

Assigned HW#2-5.

FACT 29.9. *Let* $a, b, c \in \mathbb{R}$. *Then both of the following are true:*

(1) $a \in [b|c]$ *or* $b \in [a|c]$ *or* $c \in [a|b]$.
(2) $[a|c] \subseteq [a|b] \cup [b|c]$.

Proof. At least one of the following must be true:

(α) $a \leqslant b \leqslant c$,
(β) $a \leqslant c \leqslant b$,
(γ) $b \leqslant a \leqslant c$,
(δ) $b \leqslant c \leqslant a$,
(ε) $c \leqslant a \leqslant b$,
(η) $c \leqslant b \leqslant a$.

Unassigned homework: Prove both (1) and (2) in cases $(\alpha), \ldots, (\delta)$.

Case (ε)*:* We have $a \in [c, b] = [b|c]$, so (1) holds. Also,

$$[a|c] \;=\; [c, a] \;\subseteq\; [c, b] \;=\; [b|c] \;\subseteq\; [a|b] \cup [b|c],$$

so (2) holds. *End of Case* (ε).

Case (η)*:* We have $b \in [c, a] = [a|c]$, so (1) holds. Also,

$$[a|c] \;=\; [c, a] \;=\; [c, b] \cup [b, a] \;=\; [b|c] \cup [a|b] \;=\; [a|b] \cup [b|c],$$

so (2) holds. *End of Case* (η). $\square$

DEFINITION 29.10. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $k \in \mathbb{N}$. Then f is k-**monotone** means: $\forall S \subseteq dom[f]$,*

$$[\ \#S \leqslant k\ ] \quad \Rightarrow \quad [\ f|S \textit{ is strictly monotone}\ ].$$

REMARK 29.11. *Let $\phi : \mathbb{R} \dashrightarrow \mathbb{R}$. Assume $\#(dom[\phi]) \leqslant 1$. Then ϕ is strictly increasing.*

Proof. Since $\#(\text{dom}[\phi]) \leqslant 1$, it follows that $\forall a, b \in \text{dom}[f], \quad [\, a = b\,]$. Then $\forall a, b \in \text{dom}[f], \quad [\,(a < b) \Rightarrow (f(a) < f(b))\,]$. $\square$

We note that Remark 29.11 remains true if we replace "strictly increasing" by any one of the following

- strictly decreasing,
- semiincreasing,
- semidecreasing,
- strictly monotone,
- semimonotone.

COROLLARY 29.12. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$. Then f is 1-monotone.*

Proof. We wish to show: $\forall S \subseteq \text{dom}[f]$,

$$[\ \#S \leqslant 1\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone}\ ].$$

Let $S \subseteq \mathrm{dom}[f]$ be given. We wish to show

$$[\ \#S \leqslant 1\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

Assume: $\#S \leqslant 1$. We wish to show: $f|S$ is strictly monotone.

Since $\#(\mathrm{dom}[f|S]) = \#S \leqslant 1$, by Remark 29.11, we see that $f|S$ is strictly increasing. Then $f|S$ is strictly monotone. □

REMARK 29.13. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$. *Then:*

$$[\ f \textit{ is } 2\textit{-monotone}\ ] \quad \Leftrightarrow \quad [\ f \textit{ is 1-1}\ ].$$

Proof. Proof of $\Rightarrow$*:* Assume: f is 2-monotone. We wish to show: f is 1-1. We wish to show: $\forall \alpha, \beta \in \mathrm{dom}[f]$, $[(f(\alpha) = f(\beta)) \Rightarrow (\alpha = \beta)]$. Let $\alpha, \beta \in \mathrm{dom}[f]$ be given. We wish to show: $[(f(\alpha) = f(\beta)) \Rightarrow (\alpha = \beta)]$. Assume: $f(\alpha) = f(\beta)$. We wish to show $\alpha = \beta$. Asume that $\alpha \neq \beta$. We aim for a contradiction.

Let $a := \min\{\alpha, \beta\}$ and $b := \max\{\alpha, \beta\}$. Then $a < b$. Since $f(\alpha) = f(\beta)$, we get: $f(a) = f(b)$. Let $g := f|\{a, b\}$. Then $g(a) = g(b)$. Since f is 2-monotone, it follows that g is strictly monotone. Then: (g is strictly increasing) or (g is strictly decreasing). So, since $a < b$, we see that: ($g(a) < g(b)$) or ($g(a) > g(b)$). Then $g(a) \neq g(b)$. Contradiction. *End of proof of* $\Rightarrow$.

Proof of $\Leftarrow$*:* Assume: f is 1-1. We wish to show: f is 2-monotone. We wish to show: $\forall S \subseteq \mathrm{dom}[f]$,

$$[\ \#S \leqslant 2\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

By Corollary 29.12, f is 1-monotone. We wish to show: $\forall S \subseteq \mathrm{dom}[f]$,

$$[\ \#S = 2\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

Let $S \subseteq \mathrm{dom}[f]$ be given. We wish to show:

$$[\ \#S = 2\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

Assume $\#S = 2$. We wish to show: $f|S$ is strictly monotone.

Choose $a, b \in S$ s.t. $a < b$. Then $S = \{a, b\}$. Since $a < b$, we have $a \neq b$. So, since f is 1-1, we have $f(a) \neq f(b)$. Then one of the following is true:

(1) $f(a) < f(b)$ or
(2) $f(a) > f(b)$.

Case 1: We wish to show that $f|S$ is strictly increasing.

Since $a < b$ and $f(a) < f(b)$, we see that $f|\{a,b\}$ is strictly increasing. So, since $S = \{a,b\}$, we see that $f|S$ is strictly increasing, as desired. *End of Case 1.*

Case 2: We wish to show that $f|S$ is strictly decreasing.

Since $a < b$ and $f(a) > f(b)$, we see that $f|\{a,b\}$ is strictly decreasing. So, since $S = \{a,b\}$, we see that $f|S$ is strictly decreasing, as desired. *End of Case 2. End of proof of* $\Leftarrow$. $\square$

REMARK 29.14. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$ *be strictly monotone. Then:*

(1) $\forall k \in \mathbb{N}$, f *is* k*-monotone* *and*
(2) f *is 1-1.*

Sketch of proof: Any restriction of strictly increasing function is strictly increasing. Any restriction of strictly decreasing function is strictly decreasing. Therefore, any restriction of strictly monotone function is strictly monotone. Then (1) holds.

By (1), f is 2-monotone. Then, by Remark 29.13, f is 1-1. $\square$

Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$. **Our first goal:**

if f is 3-monotone, then f is strictly monotone.

Then, by (1) of Remark 29.14, it will follow, for all $k \in [3..\infty)$, that k-monotone is equivalent to strictly monotone. Upshot: (1-monotone is always true), (2-monotone is equivalent to 1-1) and (all the other k-monotones are equivalent to strictly monotone). **Our second goal:**

if (dom$[f]$ is an interval)
and (f is continuous)
and (f is 2-monotone),
then f is strictly monotone.

We showed the graph of a function $f : \mathbb{R} \dashrightarrow \mathbb{R}$ such that

(1) dom$[f]$ is NOT an interval,
(2) f is continuous,
(3) f is 1-1, hence 2-monotone and
(4) f is NOT 3-monotone, hence NOT strictly monotone

REMARK 29.15. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$*. Assume that* f *is* 3*-monotone. Then* f *is* 4*-monotone.*

Proof. By assumption, f is 3-monotone, so, to show that f is 4-monotone, we need only show: $\forall S \subseteq \mathrm{dom}[f]$,

$$[\ \#S = 4\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

Let $S \subseteq \mathrm{dom}[f]$ be given. We wish to show:

$$[\ \#S = 4\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

Assume: $\#S = 4$. We wish to show: $f|S$ is strictly monotone.

Choose $a, b, c, d \in S$ s.t. $a < b < c < d$. Then $S = \{a, b, c, d\}$. Because f is 3-monotone, it follows that f is 2-monotone, so, by $\Rightarrow$ of Remark 29.13, f is 1-1. So, since $b \neq c$, it follows that $f(b) \neq f(c)$. Then one of the following must be true:

(1) $f(b) < f(c)$ or
(2) $f(b) > f(c)$.

Case (1): We wish to show that $f|S$ is strictly increasing.

Since $b < c$ and $f(b) < f(c)$, it follows that $f|\{a, b, c\}$ is not strictly decreasing. However, f is 3-monotone, so $f|\{a, b, c\}$ is strictly monotone. Then $f|\{a, b, c\}$ is strictly increasing. Then $f(a) < f(b) < f(c)$.

Since $b < c$ and $f(b) < f(c)$, it follows that $f|\{b, c, d\}$ is not strictly decreasing. However, f is 3-monotone, so $f|\{b, c, d\}$ is strictly monotone. Then $f|\{b, c, d\}$ is strictly increasing. Then $f(b) < f(c) < f(d)$.

Then $f(a) < f(b) < f(c) < f(d)$. Then $f|\{a, b, c, d\}$ is strictly increasing. So, since $S = \{a, b, c, d\}$, we conclude that $f|S$ is strictly increasing, as desired. *End of Case (1).*

Case (2): We wish to show that $f|S$ is strictly decreasing.

Since $b < c$ and $f(b) > f(c)$, it follows that $f|\{a, b, c\}$ is not strictly increasing. However f is 3-monotone, so $f|\{a, b, c\}$ is strictly monotone. Then $f|\{a, b, c\}$ is strictly decreasing. Then $f(a) > f(b) > f(c)$.

Since $b < c$ and $f(b) > f(c)$, it follows that $f|\{b, c, d\}$ is not strictly increasing. However f is 3-monotone, so $f|\{b, c, d\}$ is strictly monotone. Then $f|\{b, c, d\}$ is strictly decreasing. Then $f(b) > f(c) > f(d)$.

Then $f(a) > f(b) > f(c) > f(d)$. Then $f|\{a, b, c, d\}$ is strictly decreasing. So, since $S = \{a, b, c, d\}$, we conclude that $f|S$ is strictly decreasing, as desired. *End of Case (2).* □

The next result achieves our first goal.

THEOREM 29.16. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$. Assume that f is 3-monotone. Then f is strictly monotone.*

Proof. By Remark 29.15, f is 4-monotone. Let $X := \mathrm{dom}[f]$. Because f is 4-monotone, we know:

$$[\ \#X \leqslant 4\ ] \quad \Rightarrow \quad [\ f \text{ is strictly monotone }].$$

It therefore suffices to show

$$[\ \#X > 4\ ] \quad \Rightarrow \quad [\ f \text{ is strictly monotone }].$$

Assume: $\#X > 4$. We wish to show: f is strictly monotone.

Since $\#X > 1$, choose $a, b \in X$ s.t. $a < b$. Since f is 3-monotone, we see that f is 2-monotone, so, by $\Rightarrow$ of Remark 29.13, f is 1-1. So, as $a \neq b$, we get: $f(a) \neq f(b)$. Then one of the following must be true:

(1) $f(a) < f(b)$ or
(2) $f(a) > f(b)$.

Case (1): We wish to show: f is strictly increasing. As $X = \mathrm{dom}[f]$, we wish to show: $\forall c, d \in X$,

$$[\ c < d\ ] \quad \Rightarrow \quad [\ f(c) < f(d)\ ].$$

Let $c, d \in X$ be given. We wish to show:

$$[\ c < d\ ] \quad \Rightarrow \quad [\ f(c) < f(d)\ ].$$

Assume: $c < d$. We wish to show: $f(c) < f(d)$.

Since $a < b$ and $f(a) < f(b)$, it follows that $f|\{a, b, c, d\}$ is not strictly decreasing. However, f is 4-monotone, so $f|\{a, b, c, d\}$ is strictly monotone. Then $f|\{a, b, c, d\}$ is strictly increasing. Then, since $c < d$, we get $f(c) < f(d)$, as desired. *End of Case (1).*

Case (2): We wish to show: f is strictly decreasing. As $X = \mathrm{dom}[f]$, we wish to show: $\forall c, d \in X$,

$$[\ c < d\ ] \quad \Rightarrow \quad [\ f(c) > f(d)\ ].$$

Let $c, d \in X$ be given. We wish to show:

$$[\ c < d\ ] \quad \Rightarrow \quad [\ f(c) > f(d)\ ].$$

Assume: $c < d$. We wish to show: $f(c) > f(d)$.

Since $a < b$ and $f(a) > f(b)$, it follows that $f|\{a, b, c, d\}$ is not strictly increasing. However, f is 4-monotone, so $f|\{a, b, c, d\}$ is strictly monotone. Then $f|\{a, b, c, d\}$ is strictly decreasing. Then, since $c < d$, we get $f(c) > f(d)$, as desired. *End of Case (2).* $\square$

This completes our first goal. Recall **our second goal:** $\forall f : \mathbb{R} \to \mathbb{R}$,

$$\begin{array}{ll} \text{if} & \quad (\text{dom}[f] \text{ is an interval}) \\ & \text{and } (f \text{ is continuous}) \\ & \text{and } (f \text{ is 2-monotone}), \\ \text{then} & \quad f \text{ is strictly monotone.} \end{array}$$

LEMMA 29.17. *Let I be an interval and let $f : I \to \mathbb{R}$ be continuous and 1-1. For all $x \in I$, let $\overline{x} := f(x)$. Let $p, q, r \in I$. Assume: $q \notin [p|r]$. Then $\overline{q} \notin [\overline{p}|\overline{r}]$.*

Proof. Assume: $\overline{q} \in [\overline{p}|\overline{r}]$. We aim for a contradiction.

Since I is an interval and $p, r \in I$, it follows that $[p|r] \subseteq I$. By assumption $f : I \to \mathbb{R}$ is continuous. Then f is continuous on $[p|r]$. Then, by the Intermediate Value Theorem (Theorem 29.7), we get: $[f(p)|f(r)] \subseteq f_*([p|r])$ Then $\overline{q} \in [\overline{p}|\overline{r}] = [f(p)|f(r)] \subseteq f_*([p|r])$, so choose $x \in [p|r]$ s.t. $\overline{q} = f(x)$. Since $f(q) = \overline{q} = f(x)$ and since f is 1-1, it follows that $q = x$. Therefore, $q = x \in [p|r]$. However, by assumption, we have: $q \notin [p|r]$. Contradiction. □

We indicated how to use Lemma 29.17 to finish the second goal. In the next class, we'll write out the details.

30. Class 4 on 25 January 2018, Th of Week 2

Recall: $\forall \alpha, \beta \in \mathbb{R}$, we define

$$[\alpha|\beta] \quad := \quad [\ \min\{\alpha, \beta\}\ ,\ \max\{\alpha, \beta\}\ ].$$

So, for example, $[3|1] = [1, 3] = [1|3]$.

Recall: $\forall s, t, u \in \mathbb{R}$, we have both

$$\begin{array}{l} (\ (\ s \in [t|u]\ ) \text{ or } (\ t \in [s|u]\ ) \text{ or } (\ u \in [s|t]\ )\ ) \\ \text{and} \qquad (\ [s|u] \subseteq [s|t] \cup [t|u]\ ). \end{array}$$

Recall: Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $k \in \mathbb{N}$. Then f is ***k*-monotone** means: $\forall S \subseteq \text{dom}[f]$, $(\ [\ \#S \leqslant k\ ] \Rightarrow [\ f|S \text{ is strictly monotone}\ ]\ )$.

By Remark 29.13, 2-monotone is equivalent to injective.

By Theorem 29.16, 3-monotone is equivalent to strictly monotonic.

Goal: Show, for any $f : \mathbb{R} \dashrightarrow \mathbb{R}$, if

- dom$[f]$ is an interval,
- f is continuous and
- f is 2-monotone,

then f is strictly monotone. By Remark 29.13, this goal follows from $\Leftarrow$ of Theorem 30.1 below.

In Lemma 29.17, we proved: Let I be an interval and let $f : I \to \mathbb{R}$ be continuous and 1-1. For all $x \in I$, let $\overline{x} := f(x)$. Let $p, q, r \in I$. Assume: $q \notin [p|r]$. Then $\overline{q} \notin [\overline{p}|\overline{r}]$.

THEOREM 30.1. *Let I be an interval and let $f : I \to \mathbb{R}$ be continuous. Then:* $\quad$ [*f is strictly monotone*] $\quad\Leftrightarrow\quad$ [*f is 1-1*].

Proof. Proof of $\Rightarrow$: By Remark 29.13, [f is 2-monotone] $\Rightarrow$ [f is 1-1]. Then: [f is strictly monotone] $\Rightarrow$ [f is 2-monotone] $\Rightarrow$ [f is 1-1], as desired. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume: f is 1-1. We wish to prove: f is strictly montone. By Theorem 29.16, it suffices to show: f is 3-monotone. By Remark 29.13, f is 2-monotone. That is, $\forall S \subseteq \mathrm{dom}[f]$,

$$[\ \#S \leqslant 2\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

It thererfore suffices to show: $\forall S \subseteq \mathrm{dom}[f]$,

$$[\ \#S = 3\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

Let $S \subseteq \mathrm{dom}[f]$ be given. We wish to show:

$$[\ \#S = 3\ ] \quad \Rightarrow \quad [\ f|S \text{ is strictly monotone }].$$

Assume: $\#S = 3$. We wish to show: $f|S$ is strictly monotone.

Choose $a, b, c \in S$ s.t. $a < b < c$. Then $S = \{a, b, c\}$. For all $x \in I$, let $\overline{x} := f(x)$. Since $a < b < c$, we have $a \notin [b|c]$. Then, by Lemma 29.17, $\overline{a} \notin [\overline{b}|\overline{c}]$. Since $a < b < c$, we have $c \notin [a|b]$. Then, by Lemma 29.17, $\overline{c} \notin [\overline{a}|\overline{b}]$. Since $\overline{a} \notin [\overline{b}|\overline{c}]$ and since $\overline{c} \notin [\overline{a}|\overline{b}]$, by (1) of Fact 29.9, we see that $\overline{b} \in [\overline{a}|\overline{c}]$. Thus: $[\ (\ \overline{a} \leqslant \overline{b} \leqslant \overline{c}\ ) \quad \text{or} \quad (\ \overline{a} \geqslant \overline{b} \geqslant \overline{c}\ )\ ]$.

Since $a < b < c$, we see that $a \neq b \neq c$. So, since f is 1-1, we get $\overline{a} \neq \overline{b} \neq \overline{c}$. Then: $[\ (\ \overline{a} < \overline{b} < \overline{c}\ ) \quad \text{or} \quad (\ \overline{a} > \overline{b} > \overline{c}\ )\ ]$. So, since $S = \{a, b, c\}$, we conclude:

$$(\ f|S \text{ is strictly increasing }) \quad \text{or} \quad (\ f|S \text{ is strictly decreasing }).$$

Then $f|S$ is strictly monotone, as desired. *End of proof of $\Leftarrow$.* $\square$

Let Z be a topological space and let $S \subseteq Z$. We reviewed the definition of $\mathrm{Int}_Z S$ and discussed a pictorial example when $Z = \mathbb{R}^2$. We noted that $\mathrm{Int}_{\mathbb{R}}[3, 7] = (3, 7)$ and that $\mathrm{Int}_{\mathbb{R}}[0, \infty) = (0, \infty)$. We have the following quantified equivalence for interior points:

LEMMA 30.2. *Let Z be a topological space, let $S \subseteq Z$ and let $q \in Z$. Then:* $\quad [\ q \in \mathrm{Int}_Z S\ ] \quad \Leftrightarrow \quad [\ \exists V \in \mathcal{N}_Z(q)$ *s.t.* $V \subseteq S\ ]$.

Proof. Proof of $\Rightarrow$: Assume that $q \in \mathrm{Int}_Z S$. We wish to show that $\exists V \in \mathcal{N}_Z(q)$ s.t. $V \subseteq S$.

We have $q \in \mathrm{Int}_Z S = \bigcup \{\ V \text{ open in } Z \mid V \subseteq S\ \}$. Choose an open subset V of Z s.t. $V \subseteq S$ and s.t. $q \in V$. We wish to show: $V \in \mathcal{N}_Z(q)$.

Since V is open in Z and since $q \in V$, we conclude, by Remark 16.4, that $V \in \mathcal{N}_Z(q)$, as desired. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume: $\exists V \in \mathcal{N}_Z(q)$ s.t. $V \subseteq S$. Want: $q \in \mathrm{Int}_Z S$. We wish to show: $q \in \bigcup \{\ U \text{ open in } Z \mid U \subseteq S\ \}$. That is, we wish to show: $\exists U$ open in Z s.t. $U \subseteq S$ and s.t. $q \in U$.

Choose $V \in \mathcal{N}_Z(q)$ s.t. $V \subseteq S$. Since $V \in \mathcal{N}_Z(q)$, choose U open in Z s.t. $q \in U \subseteq V$. It remains to show: $U \subseteq S$.

We have: $U \subseteq V \subseteq S$, as desired. *End of proof of $\Leftarrow$.* $\square$

The next result is called one-dimensional **Invariance of Domain**:

THEOREM 30.3. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ be 1-1 and continuous, and let $U \subseteq dom[f]$. Assume that U is open in $\mathbb{R}$. Then $f_*(U)$ is open in $\mathbb{R}$.*

Proof. Let $S := f_*(U)$. We wish to show: S is open in $\mathbb{R}$. It suffices to show: $S = \mathrm{Int}_{\mathbb{R}} S$. Since $\mathrm{Int}_{\mathbb{R}} S \subseteq S$, we want: $S \subseteq \mathrm{Int}_{\mathbb{R}} S$. That is, we want: $\forall q \in S$, $q \in \mathrm{Int}_{\mathbb{R}} S$. Let $q \in S$ be given. Want: $q \in \mathrm{Int}_{\mathbb{R}} S$. Then, by Lemma 30.2, we wish to prove: $\exists V \in \mathcal{N}_{\mathbb{R}}(q)$ s.t. $V \subseteq S$.

Since $q \in S = f_*(U)$, choose $p \in U$ s.t. $q = f(p)$. For all $x \in \mathbb{R}$, let $\overline{x} := f(x)$. Then $q = \overline{p}$. By assumption, U is open in $\mathbb{R}$. So, since $p \in U$, we see, by Remark 16.4, that $U \in \mathcal{N}_{\mathbb{R}}(p)$. By Theorem 17.15, $\mathcal{B}_{\mathbb{R}}(p)$ is a neighborhood base at p in $\mathbb{R}$. So, since $U \in \mathcal{N}_{\mathbb{R}}(p)$, choose $A \in \mathcal{B}_{\mathbb{R}}(p)$ s.t. $A \subseteq U$. By definition of $\mathcal{B}_{\mathbb{R}}(p)$, choose $\delta > 0$ s.t. $A = B_{\mathbb{R}}(p, \delta)$. Let $a := p - (\delta/2)$, $b := p + (\delta/2)$ and $I := [a, b]$. Then I is an interval and

$$I \subseteq (\, p - \delta \,,\, p + \delta \,) = B(p, \delta) = A \subseteq U \subseteq \mathrm{dom}[f].$$

Since f is continuous, by Remark 28.14, $f|I$ is continuous. Also, since f is 1-1, it follows that $f|I$ is 1-1. Then, by Theorem 30.1, $f|I$ is strictly monotone. Therefore, one of the following must be true:

(1) $f|I$ is strictly increasing $\qquad$ or
(2) $f|I$ is strictly decreasing.

Case 1: We have $a, p, b \in I$ and $a < p < b$. So, since $f|I$ is strictly increasing, $\overline{a} < \overline{p} < \overline{b}$. Since $a < b$, we have $[a|b] = [a, b]$. Since $\overline{a} < \overline{b}$, we have $[\overline{a}|\overline{b}] = [\overline{a}, \overline{b}]$. Let $V := (\overline{a}, \overline{b}) \subseteq \mathbb{R}$. Then V is open in $\mathbb{R}$. Also, $q = \overline{p} \in V$. Then, by Remark 16.4, $V \in \mathcal{N}_{\mathbb{R}}(q)$. Want: $V \subseteq S$.

Since $f|I$ is continuous and $I = [a, b]$, by the Intermediate Value Theorem (Theorem 29.7), we get $[f(a)|f(b)] \subseteq f_*([a|b])$. Then we have $[\overline{a}, \overline{b}] = [\overline{a}|\overline{b}] = [f(a)|f(b)] \subseteq f_*([a|b]) = f_*([a, b])$. It follows that $V \subseteq [\overline{a}, \overline{b}] \subseteq f_*([a, b]) = f_*(I) \subseteq f_*(U) = S$. *End of Case 1.*

Case 2: We have $a, p, b \in I$ and $a < p < b$. So, since $f|I$ is strictly decreasing, $\overline{a} > \overline{p} > \overline{b}$. Since $a < b$, we have $[a|b] = [a, b]$. Since $\overline{a} > \overline{b}$, we have $[\overline{a}|\overline{b}] = [\overline{b}, \overline{a}]$. Let $V := (\overline{b}, \overline{a}) \subseteq \mathbb{R}$. Then V is open in $\mathbb{R}$. Also, $q = \overline{p} \in V$. Then, by Remark 16.4, $V \in \mathcal{N}_{\mathbb{R}}(q)$. Want: $V \subseteq S$.

Since $f|I$ is continuous and $I = [a, b]$, by the Intermediate Value Theorem (Theorem 29.7), we get $[f(a)|f(b)] \subseteq f_*([a|b])$. Then we have $[\overline{b}, \overline{a}] = [\overline{a}|\overline{b}] = [f(a)|f(b)] \subseteq f_*([a|b]) = f_*([a, b])$. It follows that $V \subseteq [\overline{b}, \overline{a}] \subseteq f_*([a, b]) = f_*(I) \subseteq f_*(U) = S$. *End of Case 2.* $\square$

Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Let $U := \mathbb{R}$. Then $f_*(U) = [0, \infty)$. This shows that the 1-1 hypothesis in Theorem 30.3 is needed.

We gave an example of an open interval $U \subseteq \mathbb{R}$ and a 1-1 function $f : U \to \mathbb{R}$ such that $f_*(U)$ is not open. This shows that the continuity hypothesis in Theorem 30.3 is needed.

Theorem 30.3 generalizes to all dimensions:

THEOREM 30.4. *Let $k \in \mathbb{N}$, and let $f : \mathbb{R}^k \dashrightarrow \mathbb{R}^k$ be 1-1 and continuous, and let $U \subseteq dom[f]$. Assume that U is open in $\mathbb{R}$. Then $f_*(U)$ is open in $\mathbb{R}$.*

Proof. Omitted. For $k \in [2..\infty)$, this is beyond the scope of our course. It is typically proved in a first-year graduate course on topology. $\square$

REMARK 30.5. *Let Z be a topological space, let $Z_0 \subseteq Z$ and let $U \subseteq Z_0$. Then:* $\quad$ *[U open in Z_0]* $\quad\Leftrightarrow\quad$ *[U open in Z_0].*

Proof. Unassigned HW. $\square$

THEOREM 30.6. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ be 1-1 and continuous. We define $X := dom[f]$ and $Y := im[f]$. Assume that X is open in $\mathbb{R}$. Then*

(1) *Y is open in $\mathbb{R}$* $\quad$ *and*
(2) *f is a homeomorphism from X onto Y.*

Proof. Proof of (1): By Theorem 30.4 (with U replaced by X), we see that $f_*(X)$ is open in $\mathbb{R}$. So, since $f_*(X) = f_*(\mathrm{dom}[f]) = \mathrm{im}[f] = Y$, we get: Y is open in $\mathbb{R}$, as desired. *End of proof of (1).*

Proof of (2): Since $f : X \hookrightarrow\!> Y$ is 1-1 and continuous, we need only show $f : X \to Y$ is open, *i.e.*, that $\forall$open U in X, $f_*(U)$ is open in Y. Let an open U in X be given. We wish to show: $f_*(U)$ is open in Y.

By $\Leftarrow$ of Remark 30.5 (with X replaced by $\mathbb{R}$ and Z_0 by X), we see that U is open in $\mathbb{R}$. Then, by Invariance of Domain (Theorem 30.3), we conclude that $f_*(U)$ is open in $\mathbb{R}$. Then, by $\Rightarrow$ of Remark 30.5 (with X replaced by $\mathbb{R}$, Z_0 by Y, and U by $f_*(U)$), it follows that $f_*(U)$ is open in Y, as desired. *End of proof of (2).* $\square$

THEOREM 30.7. *Let $f : \mathbb{R} \to \mathbb{R}$ be 1-1 and continuous. We define $X := dom[f]$ and $Y := im[f]$. Then $f_*(\mathrm{Int}_{\mathbb{R}}X) \subseteq \mathrm{Int}_{\mathbb{R}}Y$.*

Proof. Let $\mathcal{S} := \{V \subseteq Y \,|\, V \text{ is open in } \mathbb{R}\}$. Then $\mathrm{Int}_{\mathbb{R}}Y = \bigcup \mathcal{S}$. Let $U := \mathrm{Int}_{\mathbb{R}}X$. Let $V := f_*(U)$. Then $V = f_*(\mathrm{Int}_{\mathbb{R}}X)$. We wish to show: $V \subseteq \mathrm{Int}_{\mathbb{R}}Y$. That is, we wish to show: $V \subseteq \bigcup \mathcal{S}$. It suffices to show $V \in \mathcal{S}$. That is, we wish to show: ($V \subseteq Y$) and (V is open in $\mathbb{R}$). As $V = f_*(\mathrm{Int}_{\mathbb{R}}X) \subseteq \mathrm{im}[f] = Y$, we need only show: V is open in $\mathbb{R}$.

Since $\mathrm{Int}_{\mathbb{R}}X$ is open in $\mathbb{R}$ and $\mathrm{Int}_{\mathbb{R}}X \subseteq X = \mathrm{dom}[f]$, by Invariance of Domain (Theorem 30.3), we conclude that $f_*(\mathrm{Int}_{\mathbb{R}}X)$ is open in $\mathbb{R}$. That is, we conclude that V is open in $\mathbb{R}$, as desired. $\square$

Recall: Let Y and Z be topological spaces, let $\phi : Y \dashrightarrow Z$ and let $u \in Y$. Then

$$\begin{aligned} \operatorname*{LIMS}_u \phi \quad &:= \quad \{v \in Z \,|\, \phi \to v \text{ near } u\} \qquad \text{and} \\ \lim_u \phi \quad &:= \quad \mathrm{ELT}(\operatorname*{LIMS}_u \phi). \end{aligned}$$

When clarity is needed we sometimes indicate Y and Z by writing

$$(Y, Z)\text{-}\operatorname*{LIMS}_u \phi \qquad \text{and} \qquad (Y, Z)\text{-}\lim_u \phi,$$

and by writing $\phi \to v$ in Z near u in Y.

Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ be defined by $f(x) = 1/(x^2)$. I drew the graph of f. I asked the value of $\lim\limits_0 f$. Correct answer: It depends on how you think about f. Since $f : \mathbb{R} \dashrightarrow \mathbb{R}$, we might compute

$$(\mathbb{R}, \mathbb{R})\text{-}\lim_0 f \;=\; \text{☹}.$$

On the other hand, $f : \mathbb{R}^* \dashrightarrow \mathbb{R}^*$, and so we might compute

$$(\mathbb{R}^*, \mathbb{R}^*)\text{-}\lim_0 f \;=\; \infty.$$

The moral is that, if Y and Z are ambiguous, then we really need to indicate them for $\lim\limits_u \phi$ to have a clear meaning. On the other hand, if Y and Z are clear, then the notation $(Y,Z)\text{-}\lim\limits_u \phi$ is unnecessarily cluttered and can be hard on the reader. Some judgment is required.

Recall: Let Y and Z be topological spaces, let $\phi : Y \dashrightarrow Z$, let $u \in Y$ and let $v \in Z$. Then

$$(\ \lim_u \phi = v\ )\ \Rightarrow\ (\ v \in \operatorname*{LIMS}_u \phi\ )\ \Leftrightarrow\ (\ \phi \to v \text{ near } u\ ).$$

The $\Rightarrow$ cannot be replaced by $\Leftrightarrow$, as we explain in the next paragraph.

In this paragraph, LIMS will always mean $(\mathbb{R},\mathbb{R})$- LIMS. Let $Y := \mathbb{R}$, let $Z := \mathbb{R}$ and let $S := [1,2] \cup \{3\}$. Let $u := 3$ and let $v := 4$. Let $\phi := C_Z^4 : S \to \mathbb{R}$ be the constant function that is equal to 4 on S and undefined elsewhere. Then $\phi : \mathbb{R} \dashrightarrow \mathbb{R}$. By Proposition 24.7, $\operatorname*{LIMS}\limits_u \phi = \operatorname*{LIMS}\limits_3 \phi = \mathbb{R}$, so we have both

$$(\ v \in \operatorname*{LIMS}_u \phi\ ) \qquad \text{and} \qquad (\ \phi \to v \text{ near } u\ ).$$

On the other hand, $\lim\limits_u \phi = \mathrm{ELT}(\operatorname*{LIMS}\limits_u \phi) = \mathrm{ELT}\ \mathbb{R} = ☹$.

Recall: Let Y be a topological space and let $S \subseteq Y$. Then

$$\begin{aligned}
\mathrm{Int}_Y S &= \bigcup\{U \subseteq Y \mid (\,U \text{ is open in } Y\,) \,\&\, (\,U \subseteq S\,)\},\\
\mathrm{Cl}_Y S &= \bigcap\{C \subseteq Y \mid (\,C \text{ is closed in } Y\,) \,\&\, (\,C \supseteq S\,)\},\\
\mathrm{Isol}_Y S &= \{p \in S \mid \exists U \in \mathcal{N}_Y(p) \text{ s.t. } U \cap S = \{p\}\},\\
\partial_Y S &= [\mathrm{Cl}_Y S] \backslash [\mathrm{Int}_Y S] \qquad \text{and}\\
\mathrm{LP}_Y S &= [\mathrm{Cl}_Y S] \backslash [\mathrm{Isol}_Y S].
\end{aligned}$$

We worked out all five of these sets in an example where S was a subset of $\mathbb{R}^2$. We chose the subset S to be neither open nor closed, and to have exactly two isolated points.

Let $Y := \mathbb{R}$ and let $S := [1,2) \cup \{3,4\}$. We compute

$$\begin{aligned}
\mathrm{Int}_Y S &= (1,2),\\
\mathrm{Cl}_Y S &= [1,2] \cup \{3,4\},\\
\mathrm{Isol}_Y S &= \{3,4\},\\
\partial_Y S &= \{1,2,3,4\} \qquad \text{and}\\
\mathrm{LP}_Y S &= [1,2].
\end{aligned}$$

Recall: Let Y and Z be topological spaces and let $\phi : Y \dashrightarrow Z$. Then we define $\mathrm{LPD}_Y\phi := \mathrm{LP}_Y(\mathrm{dom}[\phi])$. Here, the notation "LPD" stand for "limit point of the domain".

Recall: Let Y and Z be topological spaces and let $\phi : Y \dashrightarrow Z$. Let $u \in Y\backslash[\mathrm{LPD}_Y\phi]$. Then both

(1) $\underset{u}{\mathrm{LIMS}}\,\phi = Z$ and
(2) $(\#Z \neq 1) \Rightarrow (\lim_u \phi = ☹)$.

It therefore makes little sense to try to compute a limit of a function at a point that is NOT a limit point of the domain. By contrast, in the next paragraph, we show that the situation improves greatly when a limit is computed at a point that IS a limit point of the domain.

Recall: Let Y and Z be topological spaces and let $\phi : Y \dashrightarrow Z$. Let $u \in \mathrm{LPD}_Y\phi$. Assume that Z is Hausdorff. Then both

(1) $\forall v, w \in \underset{u}{\mathrm{LIMS}}\,\phi,\quad v = w$ and
(2) $\#(\underset{u}{\mathrm{LIMS}}\,\phi) \leqslant 1$.

From this, we get:

COROLLARY 30.8. *Let Y and Z be topological spaces. Assume that Z is Hausdorff. Let $\phi : Y \dashrightarrow Z$. Let $u \in \mathrm{LPD}_Y\phi$. Then*

$$[\ \lim_u \phi = v\ ] \quad \Leftrightarrow \quad [\ \phi \to v \textit{ near } u\ ].$$

Recall: Let W be a vector space, let $f : \mathbb{R} \dashrightarrow W$ and let $p \in \mathrm{dom}[f]$. Then $SS_f^p : \mathbb{R} \dashrightarrow W$ is defined by

$$(SS_f^p)(h) \quad = \quad \frac{[f(p+h)] - [f(p)]}{h}.$$

We noted that

$$\mathrm{dom}[SS_f^p] \quad = \quad [\ (\,\mathrm{dom}[f]\,) - p\ ] \setminus \{\,0\,\},$$

and, therefore, that

$$\mathrm{LPD}_{\mathbb{R}}(SS_f^p) \quad = \quad (\,\mathrm{LPD}_{\mathbb{R}}f\,) - p,$$

and, therefore, that

$$[\ 0 \in \mathrm{LPD}_{\mathbb{R}}(SS_f^p)\ ] \quad \Leftrightarrow \quad [\ p \in \mathrm{LPD}_{\mathbb{R}}f\ ].$$

It therefore makes little sense to compute $\lim_0 SS_f^p$ unless $p \in \mathrm{LPD}_{\mathbb{R}}f$. In fact, if $p \notin \mathrm{LPD}_{\mathbb{R}}f$, then $\lim_0 SS_f^p = \mathrm{ELT}\ \mathbb{R} = ☹$.

Recall: Let W be a normed vector space and $f : \mathbb{R} \dashrightarrow W$. We define $f' : \mathbb{R} \dashrightarrow W$ by $f'(p) = \lim\limits_{0} SS_f^p$. By $\lim\limits_{0}$, we mean $(\mathbb{R}, W)$-$\lim\limits_{0}$.

Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ be defined by $f(x) = \sqrt{1 - x^2}$. We drew the graph of f, which is a closed semicircle in the plane. (It is the upper half of the circle of radius one about the origin.) We observed that

$$(\mathbb{R}^*, \mathbb{R}^*)\text{-}\lim_{0} SS_f^1 \quad = \quad -\infty.$$

However,

$$f'(1) \quad = \quad (\mathbb{R}, \mathbb{R})\text{-}\lim_{0} SS_f^1 \quad = \quad ☹.$$

Let $g := f \,|\, [-1/2, 1/2]$. We observed that $g'(1/2) \neq$ ☹.

Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Let $S := \{1, 1/2, 1/3, 1/4, \ldots\} \cup \{0\}$. Let $g := f|S$. The graph of f is a parabola, and the graph of g is a set of countably many points on that parabola, accumulating at the origin. Note that $\mathrm{dom}[g'] = \{0\}$ and that $g'(0) = 0$.

31. Class 5 on 30 January 2018, Tu of Week 3

Assigned HW#3-1.

Let X be a topological space and let $S \subseteq X$. Then

$$\begin{aligned} \mathrm{Isol}_X S &= \{p \in S \,|\, \exists V \text{ open in } X \text{ s.t. } V \cap S = \{p\}\} \\ &= \{p \in S \,|\, \{p\} \text{ is open in } S\}, \qquad \text{and} \\ \mathrm{LP}_X S &= [\mathrm{Cl}_X S] \setminus [\mathrm{Isol}_X S]. \end{aligned}$$

For example,

$$\begin{aligned} \mathrm{Isol}_{\mathbb{R}}([1,2) \cup \{3\}) &= \{3\} \qquad \text{and} \\ \mathrm{LP}_{\mathbb{R}}([1,2) \cup \{3\}) &= [1,2]. \end{aligned}$$

Let X be a topological space. Then $\forall B \subseteq X$, $\forall A \subseteq B$,

$$\begin{aligned} \mathrm{LP}_X A &= \{p \in X \,|\, \forall U \in \mathcal{N}_X^\times(p),\, U \cap A \neq \varnothing\} \\ &\subseteq \{p \in X \,|\, \forall U \in \mathcal{N}_X^\times(p),\, U \cap B \neq \varnothing\} \;=\; \mathrm{LP}_X B. \end{aligned}$$

That is, LP_X is set-theoretically monotonic.

REMARK 31.1. *Let X be a topological space and let $S \subseteq X$. Assume, for all $p \in S$, that $\{p\}$ is not open in X. Then* $\mathrm{Int}_X S \subseteq \mathrm{LP}_X S$.

Proof. Let $U := \mathrm{Int}_X S$. We wish to show: $U \subseteq \mathrm{LP}_X S$.

By definition of Int, we see that U is open in X and that $U \subseteq S$. Recall that $\mathrm{Isol}_X U = \{p \in U \,|\, \{p\} \text{ is open in } U\}$. Since U is open in X, it follows that every open subset of U is open in X. Consequently, $\{p \in U \,|\, \{p\} \text{ is open in } U\} \subseteq \{p \in U \,|\, \{p\} \text{ is open in } X\}$. As $U \subseteq S$, we see that $\{p \in U \,|\, \{p\} \text{ is open in } X\} \subseteq \{p \in S \,|\, \{p\} \text{ is open in } X\}$. By assumption, $\{p \in S \,|\, \{p\} \text{ is open in } X\} = \varnothing$. Then

$$\begin{aligned} \mathrm{Isol}_X U &= \{p \in U \,|\, \{p\} \text{ is open in } U\} \\ &\subseteq \{p \in U \,|\, \{p\} \text{ is open in } X\} \\ &\subseteq \{p \in S \,|\, \{p\} \text{ is open in } X\} \;=\; \varnothing, \end{aligned}$$

So $\mathrm{Isol}_X U = \varnothing$. Therefore $LP_X U = [\mathrm{Cl}_X U]\backslash[\mathrm{Isol}_X U] = \mathrm{Cl}_X U$. By definition of Cl, we have $U \subseteq \mathrm{Cl}_X U$. Also, since $U \subseteq S$, we get $\mathrm{LP}_X U \subseteq \mathrm{LP}_X S$. Then $U \subseteq \mathrm{Cl}_X U = \mathrm{LP}_X U \subseteq \mathrm{LP}_X S$, as desired. $\square$

In many topological spaces (*e.g.*, $\mathbb{R}$), no singleton set is open. So, by the preceding remark, we see, *e.g.*, that $\forall S \subseteq \mathbb{R}$, $\mathrm{Int}_{\mathbb{R}} S \subseteq \mathrm{LP}_{\mathbb{R}} S$. That is, in the real number system, every interior point of a set is a limit point of that set. The same result holds for $\mathbb{R}^2$, $\mathbb{R}^3$, etc.

On the other hand, there are those topological spaces that do have an open singleton set. For example, let $X := [1, 2) \cup \{3\}$, with the relative topology, inherited from $\mathbb{R}$. Let $S := \{3\}$. Then S is a clopen subset of X, so $\mathrm{Cl}_X S = S = \mathrm{Int}_X S$. Also, $\mathrm{Isol}_X S = S$. Then

$$\mathrm{LP}_X S \;=\; [\mathrm{Cl}_X S] \setminus [\mathrm{Isol}_X S] \;=\; S \setminus S \;=\; \varnothing.$$

Then $\mathrm{Int}_X S = S \nsubseteq \varnothing = \mathrm{LP}_S X$.

Recall: Let Y and Z be topological spaces, let $\phi : Y \dashrightarrow Z$ and let $p \in Y$. Then

(1) $\forall q \in Z, \Big[\quad \big(\phi \to q \text{ near } p \big) \quad \text{means} \quad \big(\forall V \in \mathcal{N}(q), \exists U \in \mathcal{N}^\times(p) \text{ s.t. } f_*(U) \subseteq V \big) \quad \Big]$

(1') $\forall q \in Z, \Big[\quad \big(\phi \to q \text{ near } p \big) \quad \Leftrightarrow \quad \big(\forall V \in \mathcal{N}(q), \exists U \in \mathcal{N}^\times(p) \text{ s.t.}, \forall x \in \mathrm{dom}[f], \; [\; x \in U \;] \;\Rightarrow\; [\; f(x) \in V \;] \;\big) \quad \Big].$

(2) $\underset{p}{\mathrm{LIMS}}\, \phi \quad := \quad \{q \in Z \,|\, \phi \to q \text{ near } p\}.$

(3) $\lim\limits_{p} \phi \quad := \quad \mathrm{ELT}\;(\; \underset{p}{\mathrm{LIMS}}\, \phi \;).$

(4) $\forall q \in Z$, $\Big[\quad \big(\lim\limits_p \phi = q\big) \quad \Rightarrow \quad \big(\phi \to q \text{ near } p\big) \quad \Big]$.

Recall: Let Y and Z be topological spaces. Assume that Z is Hausdorff. Let $\phi : Y \dashrightarrow Z$. Let $p \in \mathrm{LPD}_Y\phi$ and let $q \in Z$. Then:

$$(\quad \lim_p \phi = q \quad) \qquad \Leftrightarrow \qquad (\quad \phi \to q \text{ near } p \quad).$$

Recall: Every norm $|\bullet|$ has an associated metric $d_{|\bullet|}$. So every normed vector space has a standard metric. Every metric d has an associated topology $\mathcal{T}_d$. So every metric space has a standard topology.

Any metrizable topological space is Hausdorff. Therefore any metric space is Hausdorff, and, consequently, any normed vector space is Hausdorff. For this reason, for us, the Hausdorff condition is considered "tame"; any topological space of interest in this course will be Hausdorff.

Recall: Let W be a vector space, let $f : \mathbb{R} \dashrightarrow W$ and let $p \in \mathrm{dom}[f]$. Then $SS_p^f : \mathbb{R} \dashrightarrow W$ is defined by

$$(SS_f^p)(h) \quad = \quad \frac{[f(p+h)] \; - \; [f(p)]}{h}.$$

Since ☹ is used for any undefined quantity, it follows that, for any vector space W, for any $f : \mathbb{R} \dashrightarrow W$, for any p, if $p \notin \mathrm{dom}[f]$, then $SS_f^p = ☹$, and so, for all h, $(SS_f^p)(h) = ☹$.

Recall: Let W be a normed vector space and let $f : \mathbb{R} \dashrightarrow W$. Then $f' : \mathbb{R} \dashrightarrow W$ is defined by $f'(p) = \lim\limits_0 SS_f^p$. By lim, we mean $(\mathbb{R}, W)-$lim. When W is unclear we might write f'_W instead of f', but this is almost never necessary because of "absoluteness of differentation", which we describe in the next paragraph.

Let W be a normed vector space, let V be a subspace of W and let $f : \mathbb{R} \dashrightarrow V$. Then: $f'_V = f'_W$.

Note: Let W be a normed vector space W, and let $f : \mathbb{R} \dashrightarrow W$. As we pointed out above, $\forall p \in \mathbb{R}\backslash(\mathrm{dom}[f])$, we have $f'(p) = ☹$, *i.e.*, we have $p \in \mathbb{R}\backslash(\mathrm{dom}[f'])$. That is, $\mathbb{R}\backslash(\mathrm{dom}[f]) \subseteq \mathbb{R}\backslash(\mathrm{dom}[f'])$. Equivalently, $\mathrm{dom}[f'] \subseteq \mathrm{dom}[f]$.

DEFINITION 31.2. *Let Y and Z be topological spaces, $f : Y \dashrightarrow Z$. Then $dct[f] := \{p \in Y \,|\, f$ is continuous at $p\}$.*

In Definition 31.2, "dct" stands for **domain of continuity**.

In Definition 31.2, by "continuous" we mean "(Y, Z)-continuous". When Y and Z are unclear, we can write $\mathrm{dct}_{Y,Z}[f]$. However, this is almost never necessary, because of "absoluteness of continuity", which we describe in the next paragraph.

Let Y and Z be topological spaces. Let $Y_0 \subseteq Y$ and let $Z_0 \subseteq Z$. Let $f : Y_0 \dashrightarrow Z_0$. Then: $\mathrm{dct}_{Y_0,Z_0}[f] = \mathrm{dct}_{Y,Z}[f]$.

Let Y and Z be topological spaces and let $f : Y \dashrightarrow Z$. By Definition 21.3, we see $\forall p \in \mathrm{dct}[f]$, we have $p \in \mathrm{dom}[f]$. That is, $\mathrm{dct}[f] \subseteq \mathrm{dom}[f]$.

Let W be a normed vector space and let $f : R \dashrightarrow W$. Recall that $\mathrm{dom}[f'] \subseteq \mathrm{dom}[f]$. By Theorem 26.17, we see: $\forall p \in \mathrm{dom}[f']$, we have $p \in \mathrm{dct}[f]$. That is, $\mathrm{dom}[f'] \subseteq \mathrm{dct}[f]$. Then

$$\mathrm{dom}[f'] \quad \subseteq \quad \mathrm{dct}[f] \quad \subseteq \quad \mathrm{dom}[f].$$

That is, $\mathrm{dct}[f]$ is set-theoretically between $\mathrm{dom}[f']$ and $\mathrm{dom}[f]$.

Recall: Let W be a vector space, $f : \mathbb{R} \dashrightarrow W$, $p \in \mathrm{dom}[f]$. Then

(1) $\mathrm{dom}[SS_f^p] = [(\mathrm{dom}[f]) - p]_0^\times$,
(2) $\mathrm{LPD}_\mathbb{R}[SS_f^p] = (\mathrm{LPD}_\mathbb{R} f) - p$ and
(3) $[\, 0 \in \mathrm{LPD}_\mathbb{R}(SS_f^p) \,] \;\Leftrightarrow\; [\, p \in \mathrm{LPD}_\mathbb{R} f \,]$.

REMARK 31.3. *Let W be a normed vector space, let $f : \mathbb{R} \dashrightarrow W$ and let $v \in W$. Then:*

(1) *$\forall p \in \mathrm{dom}[f]$, ([$f'(p) = v$] $\Rightarrow$ [$SS_f^p \to v$ near 0]) and*
(2) *$\forall p \in \mathrm{LPD}_\mathbb{R} f$, ([$f'(p) = v$] $\Leftrightarrow$ [$SS_f^p \to v$ near 0]).*

Proof. Unassigned HW. *Hint:* Keep in mind that, for all $p \in \mathrm{dom}[f]$,

$$[\; f'(p) = v \;] \quad \Leftrightarrow \quad [\; \lim_0 SS_f^p = v \;].$$

Also, for all $p \in \mathrm{dom}[f]$, we have: $[\, 0 \in \mathrm{LPD}_\mathbb{R}(SS_f^p) \,] \;\Leftrightarrow\; [\, p \in \mathrm{LPD}_\mathbb{R} f \,]$. Now apply Remark 25.18 for (1), and Remark 25.19 for (2). $\square$

Example: Define $\phi : \mathbb{R} \to \mathbb{R}$ by $\phi(x) = x^2$. Let

$$S \quad := \quad \{\, 1\,,\, 1/2\,,\, 1/3\,,\, 1/4\,,\, \ldots \} \;\cup\; \{\, 0 \,\}.$$

Note that $\mathrm{LP}_\mathbb{R} S = \{0\}$. Let $W := \mathbb{R}$. Let $f := \phi | S : \mathbb{R} \dashrightarrow W$. The graph of ϕ is a parabola passing through $(0, 0)$, and the graph G of f is a countable collecction of points on that parabola. Note that $(0, 0) \in G$ and $\mathrm{LP}_{\mathbb{R}^2} G = \{(0, 0)\}$. For all $x \in S_0^\times$, we have:

- $x \notin \mathrm{LP}_\mathbb{R} S = \mathrm{LPD}_\mathbb{R} f$,
- $0 \notin \mathrm{LPD}_\mathbb{R}(SS_f^x)$,
- $\mathrm{LIMS}_0(SS_f^x) = \mathbb{R}$,
- $f'(x) = ☹$ and

- $SS_f^x \to 73$ near 0.

The use of 73 was arbitrary. In fact, as $\underset{0}{\mathrm{LIMS}}(SS_f^x) = \mathbb{R}$, it follows, for all $y \in \mathbb{R}$, that $SS_f^x \to y$ near 0. Note: While $SS_f^x \to 73$ near 0, it does NOT follow that $f'(x) = 73$; in fact, we have $f'(x) = ☹$. We leave it as unassigned HW to show: $f'(0) = 0$. Then $\mathrm{dom}[f'] = \{0\}$.

LEMMA 31.4. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$ and $p \in dom[f']$. Then $SS_f^p \to f'(p)$ near 0.*

Proof. Let $v := f'(p)$. By (1) of Remark 31.3, we have $SS_f^p \to v$ near 0. That is, $SS_f^p \to f'(p)$ near 0, as desired. $\square$

The next result says that any point of differentiability is a point of continuity and is NOT an isolated point in domain of the function. We need to assume that the target vector space is nonzero, which is a tame assumption: Who would want to study a constant function whose only value is zero?

REMARK 31.5. *Let W be a nonzero normed vector space and let $f : \mathbb{R} \dashrightarrow W$. Then $dom[f'] \subseteq (dct[f]) \backslash (\mathrm{Isol}_{\mathbb{R}}(dom[f]))$.*

Proof. We want to show: $\forall p \in \mathrm{dom}[f]$, $p \in (\mathrm{dct}[f]) \backslash (\mathrm{Isol}_{\mathbb{R}}(\mathrm{dom}[f]))$. Let $p \in \mathrm{dom}[f']$ be given. Want: $p \in (\mathrm{dct}[f]) \backslash (\mathrm{Isol}_{\mathbb{R}}(\mathrm{dom}[f]))$.

By Theorem 26.17, $p \in \mathrm{dct}[f]$. Want: $p \notin \mathrm{Isol}_{\mathbb{R}}(\mathrm{dom}[f])$. Assume $p \in \mathrm{Isol}_{\mathbb{R}}(\mathrm{dom}[f])$. We aim for a contradiction.

As $p \notin [\mathrm{Cl}_{\mathbb{R}}(\mathrm{dom}[f])] \backslash [\mathrm{Isol}_{\mathbb{R}}(\mathrm{dom}[f])] = \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[f]) = \mathrm{LPD}_{\mathbb{R}} f$, we get: $0 \notin \mathrm{LPD}_{\mathbb{R}}(SS_f^p)$. Then, by Proposition 24.7, $\underset{0}{\mathrm{LIMS}}(SS_f^p) = W$. Since W is nonzero, we get $\#W = \infty$, and so $\mathrm{ELT}\, W = ☹$. Then $f'(p) = \lim\limits_0 (SS_f^p) = \mathrm{ELT}(\underset{0}{\mathrm{LIMS}}(SS_f^p)) = \mathrm{ELT}\, W = ☹$. However, since $p \in \mathrm{dom}[f']$, it follows that $f'(p) \neq ☹$. Contradiction. $\square$

Assigned HW#3-2.

DEFINITION 31.6. *Let $k \in \mathbb{N}$. Let $f_1, \ldots, f_k$ be functions. Then $(f_1, \ldots, f_k)$ is the function defined by*

$$(\, f_1 \,,\, \ldots \,,\, f_k \,)\,(x) \quad = \quad (\, f_1(x) \,,\, \ldots \,,\, f_k(x) \,).$$

Note: Let $k \in \mathbb{N}$. Let $f_1, \ldots, f_k$ be functions. Then

$$\mathrm{dom}[\,(\, f_1 \,,\, \ldots \,,\, f_k \,)\,] \quad = \quad (\,\mathrm{dom}[f_1]\,) \;\cap\; \cdots \;\cap\; (\,\mathrm{dom}[f_k]\,).$$

Note: Let $k \in \mathbb{N}$. Let X be a set and let $Y_1, \ldots, Y_k$ be sets. Let $f_1 : X \dashrightarrow Y_1, \ldots, f_k : X \dashrightarrow Y_k$. Then

$$(\, f_1\,,\, \ldots\,,\, f_k\,) \quad : \quad X \quad \dashrightarrow \quad Y_1 \,\times\, \cdots \,\times\, Y_k.$$

Note: Let $k \in \mathbb{N}$. Let X be a topological space. Let $Y_1, \ldots, Y_k$ be topological spaces. Let $f_1 : X \dashrightarrow Y_1, \ldots, f_k : X \dashrightarrow Y_k$. Then $\mathrm{dct}[(f_1, \ldots, f_k)] = (\mathrm{dct}[f_1]) \cap \cdots \cap (\mathrm{dct}[f_k])$.

Note: Let $k \in \mathbb{N}$. Let X be a topological space. Let $S_1, \ldots, S_k \subseteq X$. Then $\mathrm{LP}_X(S_1 \cap \cdots \cap S_k) \subseteq (\mathrm{LP}_X S_1) \cap \cdots \cap (\mathrm{LP}_X S_k)$. Equality may not hold: $\mathrm{LP}_{\mathbb{R}}\,(\,(-\infty, 0] \,\cap\, [0, \infty)\,) \;=\; \mathrm{LP}_{\mathbb{R}}\{0\} \;=\; \varnothing$, while

$$(\,\mathrm{LP}_{\mathbb{R}}(-\infty, 0]\,) \;\cap\; (\,\mathrm{LP}_{\mathbb{R}}[0, \infty)\,) = (\,(-\infty, 0]\,) \;\cap\; (\,[0, \infty)\,) = \{0\}.$$

Note: Let $k \in \mathbb{N}$. Let X be a topological space. Let $Y_1, \ldots, Y_k$ be sets. Let $f_1 : X \dashrightarrow Y_1, \ldots, f_k : X \dashrightarrow Y_k$. Then

$$\mathrm{LPD}_X[\,(f_1, \ldots, f_k)\,] \quad \subseteq \quad (\,\mathrm{LPD}_X[f_1]\,) \;\cap\; \cdots \;\cap\; (\,\mathrm{LPD}_X[f_k]\,).$$

Assigned HW#3-3 and HW#3-4.

HW#3-4 illustrates why an analyst would be drawn into topology. For any normed vector space W, if two functions $\mathbb{R} \dashrightarrow W$ agree on an open set U in $\mathbb{R}$, then their derivatives also agree on U.

If we drop the requirement that U be open in $\mathbb{R}$, then this statement becomes false: Let $f, g : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = |x|$ and $g(x) = x$. Let $U := [0, \infty)$. Then $f = g$ on U. Since $f'(0) = ☹ \neq 1 = g'(0)$, we conclude: NOT ($f' = g'$ on U).

DEFINITION 31.7. *The functions* $\cos, \sin : \mathbb{R} \dashrightarrow \mathbb{R}$ *are defined by*

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots\,, \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\,. \end{aligned}$$

Following material that is typically covered in MATH 3283, the two power series in Definition 31.7 both have infinite radii of convergence. It follows that $\mathrm{dom}[\cos] = \mathbb{R}$, that $\mathrm{dom}[\sin] = \mathbb{R}$. Moreover, it follows that the term-by-term derivatives of those two power series yield power series that also have infinte radius of convergence. Also, it follows that those two term-by-term derivatives are power series for $\cos' x$ and $\sin' x$. This shows: $\cos' = -\sin$ and $\sin' = -\cos$.

Let $Z := \cos^*(\{0\})$. Since cos is differentiable on $\mathbb{R}$, $\cos : \mathbb{R} \to \mathbb{R}$ is continuous. So, since $\{0\}$ is closed in $\mathbb{R}$, it follows that Z is closed in $\mathbb{R}$. Let $S := Z \cap [0, \infty)$. Then S, being the intersection of two closed subsets of $\mathbb{R}$, is closed in $\mathbb{R}$. We have $\cos 0 = 1$. We have $\cos 2 < 1 - [(2^2)/(2!)] + [(2^4)/(4!)] < 0$. Then, by the Intermediate Value Theorem (Theorem 29.7), $\exists x \in (0, 2)$ s.t. $\cos x = 0$. Then $Z \cap (0, 2) \neq \varnothing$. So, since $Z \cap (0, 2) \subseteq Z \cap [0, \infty) = Y$, we see that $Y \neq \varnothing$. Finally, since $Y \subseteq [0, \infty) \geqslant 0$, we see that Y is bounded below. By Corollary 25.14, Y has a minimum. Note that $Y = \{x \in [0, \infty) \,|\, \cos x = 0\}$.

DEFINITION 31.8. *We define* $\pi := 2 \cdot (\min\{x \in [0, \infty) \,|\, \cos x = 0\})$.

Then $\pi/2 = \min\{x \in [0, \infty) \,|\, \cos x = 0\}$. In particular, $\cos(\pi/2) = 0$ and $0 \notin \cos_*([0, \pi/2))$. Since cos is continuous on $\mathbb{R}$, since $\cos(0) = 1$ and since $0 \notin \cos_*([0, \pi/2))$, it follows, from the Intermediate Value Theorem (Theorem 29.7), that $\cos_*([0, \pi/2)) > 0$.

We will prove the next three lemmas after we prove the Mean Value Theorem, but we will use them immediately. (To avoid circular reasoning, our proof of the Mean Value Theorem cannot use any of the next three lemmas!)

LEMMA 31.9. *We have* $\sin^2 + \cos^2 = C_{\mathbb{R}}^1$.

Proof. Deferred until after the proof of the Mean Value Theorem. □

Recall: $C_{\mathbb{R}}^1 : \mathbb{R} \to \{1\}$ is the constant function defined by $C_{\mathbb{R}}^1(x) = 1$. Then Lemma 31.9 is equivalent to

$$\forall x \in \mathbb{R}, \qquad (\sin^2 x) + (\cos^2 x) = 1.$$

LEMMA 31.10. *We have* $\sin(\pi/2) = 1$.

Proof. Deferred until after the proof of the Mean Value Theorem. □

LEMMA 31.11. *Let* $\alpha, \beta \in \mathbb{R}$. *Then both*

(1) $\sin(\alpha + \beta) = [\sin\alpha][\cos\beta] + [\cos\alpha][\sin\beta]$ *and*
(2) $\cos(\alpha + \beta) = [\cos\alpha][\cos\beta] - [\sin\alpha][\sin\beta]$.

Proof. Deferred until after the proof of the Mean Value Theorem. □

We have

$$\begin{aligned} \cos\pi &= \cos((\pi/2) + (\pi/2)) \\ &= [\cos(\pi/2)][\cos(\pi/2)] - [\sin(\pi/2)][\sin(\pi/2)] \\ &= [0][0] - [1][1] = -1. \end{aligned}$$

Also, we have

$$\begin{aligned}
\sin\pi &= \sin(\,(\pi/2) + (\pi/2)\,)\\
&= [\sin(\pi/2)][\cos(\pi/2)] + [\cos(\pi/2)][\sin(\pi/2)]\\
&= [1][0] - [0][1] = 0.
\end{aligned}$$

LEMMA 31.12. *Let $x \in \mathbb{R}$. Then all of the following are true:*

(1) $\cos(-x) = \cos x$,
(2) $\sin(-x) = -(\sin x)$,
(3) $\cos x = \sin(\ (\pi/2) - x\ )$,
(4) $\cos(\pi + x) = -(\cos x)$ *and*
(5) $\sin(\pi + x) = -(\sin x)$.

Proof. Proof of (1): This follows because, in Definition 31.7, the power series for cos involves only even powers of x. *End of proof of (1).*

Proof of (2): This follows because, in Definition 31.7, the power series for sin involves only odd powers of x. *End of proof of (2).*

Proof of (3): We have

$$\begin{aligned}
\sin(\ (\pi/2) - x\ ) &= [\sin(\pi/2)][\cos(-x)] + [\cos(\pi/2)][\sin x]\\
&= [1][\cos x] + [0][\sin x] = \cos x.
\end{aligned}$$

End of proof of (3).

Proof of (4): We have

$$\begin{aligned}
\cos(\ \pi + x\ ) &= [\cos\pi][\cos x] - [\sin\pi][\sin x]\\
&= [-1][\cos x] - [0][\sin x] = -(\cos x).
\end{aligned}$$

End of proof of (4).

Proof of (5): We have

$$\begin{aligned}
\sin(\ \pi + x\ ) &= [\sin\pi][\cos x] + [\cos\pi][\sin x]\\
&= [0][\cos x] + [-1][\sin x] = -(\sin x).
\end{aligned}$$

End of proof of (5). □

For all $x \in \mathbb{R}$, we have

$$\begin{aligned}
\sin(2\pi + x) &= \sin(\pi + \pi + x) = -[\sin(\pi + x)]\\
&= -[-(\sin x)] = \sin x \qquad \text{and}\\
\cos(2\pi + x) &= \cos(\pi + \pi + x) = -[\cos(\pi + x)]\\
&= -[-(\cos x)] = \cos x.
\end{aligned}$$

By the Principle of Mathematical Induction, we conclude:

(1) $\forall j \in \mathbb{N}, \quad \cos(2j\pi) = 1,$
(2) $\forall j \in \mathbb{N}, \quad \sin(2j\pi) = 0,$
(3) $\forall j \in \mathbb{N}, \quad \cos((2j-1)\pi) = -1$ and
(4) $\forall j \in \mathbb{N}, \quad \sin((2j-1)\pi) = 0,$

Assigned HW#3-5.

Define $f : \mathbb{R} \dashrightarrow \mathbb{R}$ by $f(x) = [x^2][\sin(x^{-3})]$. Let $\beta := \mathrm{adj}_0^0 f$. According to HW#3-5, we have $\beta'(0) = 1$. In particular, $0 \in \mathrm{dom}[\beta']$. Since $\beta = f$ on $\mathbb{R}_0^\times$, by HW#3-4, we see that $\beta' = f'$ on $\mathbb{R}_0^\times$. Then, for all $x \in \mathbb{R}_0^\times$, we have

$$\beta'(x) = f'(x) == [2x][\sin(x^{-3})] + [3x^{-2}][\cos(x^{-3})].$$

In particular, $\mathbb{R}_0^\times \subseteq \mathrm{dom}[\beta']$. So, as $0 \in \mathrm{dom}[\beta']$, we get $\mathrm{dom}[\beta'] = \mathbb{R}$. Define $p \in \mathbb{R}^\mathbb{N}$ by $p_j = (2j\pi)^{-1/3}$. For all $j \in \mathbb{N}$, we have $0 < p_j < 1$ and $p_j^{-3} = 2j\pi$ and $\sin(p_j^{-3}) = 0$ and $\cos(p_j^{-3}) = 1$, and so

$$\beta'(p_j) = [2p_j][0] - [3p_j^{-2}][1] = -3p_j^{-2} < -3.$$

Define $q \in \mathbb{R}^\mathbb{N}$ by $q_j = ((2j-1)\pi)^{-1/3}$. For all $j \in \mathbb{N}$, we have $0 < q_j < 1$ and $q_j^{-3} = (2j-1)\pi$ and $\sin(q_j^{-3}) = 0$ and $\cos(q_j^{-3}) = -1$, and so

$$\beta'(q_j) = [2q_j][0] - [3q_j^{-2}][-1] = 3q_j^{-2} > 3.$$

Since ($p_\bullet \to 0$ in $\mathbb{R}$) and ($q_\bullet \to 0$ in $\mathbb{R}$) and

$$(\, \beta'(p_j) < -3 \,) \quad \text{and} \quad (\, \beta'(q_j) > 3 \,),$$

we get $\lim\limits_0 \beta' = ☹$. Thus, the function β' is NOT continuous at 0, even though β' is defined at every real number, and, in particular, at 0.

32. Class 6 on 1 February 2018, Th of Week 3

LEMMA 32.1. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$. Let $p \in dom[f']$. Assume $f'(p) > 0$. Then $\exists U \in \mathcal{N}_\mathbb{R}(0)$ s.t. $(SS_f^p)_*(U) > 0$*

Proof. By Lemma 31.4, $SS_f^p \to f'(p)$ near 0. Let $\phi := SS_f^p$ and let $q := f'(p)$. Then $\phi \to q$ near 0. Also, $0 \notin \mathrm{dom}[SS_f^p] = \mathrm{dom}[\phi]$.

We have $q = f'(p) > 0$. By HW#3-2 (with p replaced by 0 and U by $U_\times$), choose $U_\times \in \mathcal{N}_\mathbb{R}^\times(0)$ s.t. $\phi_*(U_\times) > 0$. Let $U := U_\times \cup \{0\}$. Then $U \in \mathcal{N}_\mathbb{R}(0)$. We wish to show: $(SS_f^p)_*(U) > 0$.

Since $0 \notin \mathrm{dom}[\phi]$, we conclude that $U \cap (\mathrm{dom}[\phi]) = U_\times \cap (\mathrm{dom}[\phi])$. Then $\phi_*(U) = \phi_*(U_\times)$. Then $(SS_f^p)_*(U) = \phi_*(U) = \phi_*(U_\times) > 0$. $\square$

Geometrically, Lemma 32.1 says that, for any $f : \mathbb{R} \dashrightarrow \mathbb{R}$, for any $p \in \mathrm{dom}[f']$, if the tangent line to f at p runs uphill, then all the nearby secant lines to f at p will also run uphill. In the next lemma, we will show that it does NOT follow that, on some neighborhood of p, f is semiincreasing. (However, we will use β in place of f.)

LEMMA 32.2. *There is a function $\beta : \mathbb{R} \to \mathbb{R}$ such that*

(1) $\beta'(0) = 1$ *and*
(2) $\forall U \in \mathcal{N}_{\mathbb{R}}(0)$, $\beta|U$ *is NOT semiincreasing.*

Proof. Let $f : \mathbb{R}_0^\times \to \mathbb{R}$ be defined by $f(x) = x + [x^2][\sin(x^{-3})]$. Let $\beta := \mathrm{adj}_0^0 f$. Then $\beta : \mathbb{R} \to \mathbb{R}$. Define $\alpha, \gamma : \mathbb{R} \to \mathbb{R}$ by $\alpha(x) = x - x^2$ and $\gamma(x) = x + x^2$. Then $\alpha \leqslant \beta \leqslant \gamma$ on $\mathbb{R}$ and $\alpha'(0) = 1 = \gamma'(0)$. So, by HW#3-3, we see that $\beta'(0) = 1$, which shows that (1) is true. It remains to show that (2) is true. Let $U \in \mathcal{N}_{\mathbb{R}}(0)$ be given. We wish to show: $\beta|U$ is NOT semiincreasing. Assume: $\beta|U$ is semiincreasing. We aim for a contradiction.

As $\mathcal{B}_{\mathbb{R}}(0)$ is a neighborhood base at 0 in $\mathbb{R}$ and as $U \in \mathcal{N}_{\mathbb{R}}(0)$, choose $r > 0$ s.t. $B_{\mathbb{R}}(0, r) \subseteq U$. We have $0 \in \mathrm{dom}[\beta']$. Since $\beta = f$ on $\mathbb{R}_0^\times$, by HW#3-4, we conclude that $\beta' = f'$ on $\mathbb{R}_0^\times$. Then, for all $x \in \mathbb{R}_0^\times$,

$$\beta'(x) \;=\; f'(x) \;=\; 1 \;+\; [2x]\,[\sin(x^{-3})] \;+\; [3x^{-2}]\,[\cos(x^{-3})].$$

In particular, $\mathbb{R}_0^\times \subseteq \mathrm{dom}[\beta']$. So, as $0 \in \mathrm{dom}[\beta']$, we get $\mathrm{dom}[\beta'] = \mathbb{R}$. By the Archimedean Principle, choose $j \in \mathbb{N}$ s.t. $j > r^{-3}/(2\pi)$. Let $p = (2j\pi)^{-1/3}$. Then $p < r$. Also, $0 < p \leqslant (2\pi)^{-1/3} < 1$. Since $p < 1$, we get $p^{-2} > 1$, and so $-3p^{-2} < -3$. Also, $p^{-3} = 2j\pi$, and so $\sin(p^{-3}) = 0$ and $\cos(p^{-3}) = 1$. Then

$$\beta'(p) \;=\; 1 \;+\; [2p]\,[0] \;-\; [3p^{-2}]\,[1] \;=\; 1 \;-\; 3p^{-2} \;<\; 1 - 3 \;<\; 0.$$

Then $\beta'(p) < 0$. Let $y := \beta'(p)$. Then $y < 0$. Let $\varepsilon := -y$. Then $\varepsilon > 0$. We have $\lim\limits_0 (SS_\beta^p) = \beta'(p) = y$, and so $SS_\beta^p \to y$ near 0. Choose $\delta > 0$ s.t., $\forall h \in \mathrm{dom}[SS_\beta^p]$,

$$[\;|\,h \,-\, 0\,| \,<\, \delta\;] \quad\Rightarrow\quad [\;|\,[\,(SS_\beta^p)(h)\,] \,-\, y\,| \,<\, \varepsilon\;].$$

We have $-r < 0 < p < r$. Then $p \in (-r, r) = B_{\mathbb{R}}(0, r)$, so, by the Recentering Down Lemma (Lemma 14.2), choose $\eta > 0$ s.t. $B_{\mathbb{R}}(p, \eta) \subseteq B_{\mathbb{R}}(0, r)$. Let $h := [1/2][\min\{\eta, \delta\}]$. Then $h > 0$ and $|h| < \eta$ and $|h| < \delta$. Since $h > 0$, it follows that $h \in \mathbb{R}_0^\times$. Since $\mathrm{dom}[\beta] = \mathbb{R}$, we get $\mathrm{dom}[SS_\beta^p] = [\mathbb{R} - p]_0^\times = \mathbb{R}_0^\times$. Then $h \in \mathbb{R}_\times^0 = \mathrm{dom}[SS_\beta^p]$ and $|h| < \delta$.

Then, by the choice of δ, we get $|[(SS^p_\beta)(h)] - y| < \varepsilon$. Equivalently, we have $y - \varepsilon < (SS^p_\beta)(h) < y + \varepsilon$. Then $(SS^p_\beta)(h) < y + \varepsilon = y + (-y) = 0$.

As $|(p+h) - p| = |h| < \eta$, we get $p + h \in B_{\mathbb{R}}(p, \eta)$. Then

$$p\,,\,p+h \ \in\ B_{\mathbb{R}}(p,\eta)\ \subseteq\ B_{\mathbb{R}}(0,r)\ \subseteq\ U.$$

So, as $\beta|U$ is semiincreasing and as $p < p + h$, we get $\beta(p) \leqslant \beta(p+h)$. Then $[\beta(p+h)] - [\beta(p)] \geqslant 0$. Then, as $h > 0$, we get $(SS^p_\beta)(h) \geqslant 0$. Then $0 \leqslant (SS^p_\beta)(h) < 0$, so $0 < 0$. Contradiction. ☐

We next reivew the proof of the Product Rule and Chain Rule. After that, we will give a proof of the Quotient Rule.

DEFINITION 32.3. *For any object a, for any set S, by $a\,{}^*\!\!\in S$, we mean: [($a =$ ☹) or ($a \in S$)].*

For all $a, a_1, b, b_1 \,{}^*\!\!\in \mathbb{R}$, we have

$$\begin{aligned} a_1 b_1 - ab \quad &= \quad (a_1 - a)\,b \ +\ a\,(b_1 - b)\ + \\ &\qquad\qquad (a_1 - a)\,(b_1 - b). \end{aligned}$$

Setting $\triangle(ab) = a_1 b_1 - ab$, $\triangle a := a_1 - a$ and $\triangle b := b_1 - b$, we have:

$$\begin{aligned} \triangle(ab) \quad &= \quad (\triangle a)\,b \ +\ a\,(\triangle b)\ + \\ &\qquad\qquad (\triangle a)\,(\triangle b). \end{aligned}$$

Note the similarity to the product rule, but with the (perhaps) unexpected term $(\triangle a)\,(\triangle b)$. The formulas above are versions of the **precalculus product rule**, as are the three formulas in the next remark:

REMARK 32.4. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. Then*

(1) $\forall h \in \mathbb{R}$,

$$\begin{aligned} (fg)^T_p(h) \quad &= \quad [f^T_p(h)]\,[g(p)] \ +\ [f(p)]\,[g^T_p(h)] \ + \\ &\qquad\qquad [f^T_p(h)]\,[g^T_p(h)]. \end{aligned}$$

(2) $\forall h \in \mathbb{R}^\times_0$,

$$\begin{aligned} (SS^p_{fg})(h) \quad &= \quad [(SS^p_f)(h)]\,[g(p)] \ +\ [f(p)]\,[(SS^p_g)(h)] \ + \\ &\qquad\qquad [(SS^p_f)(h)]\,[(SS^p_g)(h)]\,[h]. \end{aligned}$$

(3) $SS^p_{fg} = [SS^p_f][C^{g(p)}_{\mathbb{R}}] + [C^{f(p)}_{\mathbb{R}}][SS^p_g] + [SS^p_f][SS^p_g][\mathrm{id}_{\mathbb{R}}]$.

Proof. Proof of (1): Let $h \in \mathbb{R}$ be given. We wish to prove

$$\begin{array}{rcl}(fg)_p^T(h) & = & [f_p^T(h)]\,[g(p)] \;+\; [f(p)]\,[g_p^T(h)] \;+ \\ & & [f_p^T(h)]\,[g_p^T(h)].\end{array}$$

By Definition 25.21,

$$\begin{array}{rcl}(fg)_p^T(h) & = & [f(p+h)]\,[g(p+h)] \;-\; [f(p)]\,[g(p)] \\ f_p^T(h) & = & [f(p+h)] \;-\; [f(p)] \\ g_p^T(h) & = & [g(p+h)] \;-\; [g(p)].\end{array}$$

Let $a := f(p)$, $b := g(p)$, $a_1 := f(p+h)$, $b_1 := g(p+h)$. Then

$$\begin{array}{rcl}(fg)_p^T(h) & = & a_1\,b_1 \;-\; a\,b \\ f_p^T(h) & = & a_1 \;-a \\ g_p^T(h) & = & b_1 \;-b.\end{array}$$

So, since

$$\begin{array}{rcl}a_1b_1 - ab & = & (a_1-a)\,b \;+\; a\,(b_1-b) \;+ \\ & & (a_1-a)\,(b_1-b),\end{array}$$

it follows that

$$\begin{array}{rcl}(fg)_p^T(h) & = & [f_p^T(h)]\,[g(p)] \;+\; [f(p)]\,[g_p^T(h)] \;+ \\ & & [f_p^T(h)]\,[g_p^T(h)],\end{array}$$

as desired. *End of proof of (1).*

Proof of (2): Unassigned HW. *Hint:* Take the formula in (1) and divide by h. Note that in the third term on the right hand side, we end up dividing by h twice and multiplying by h once. *End of proof of (2).*

Proof of (3): Let $L \;:=\; SS_{fg}^p$ and let

$$R \;:=\; [SS_f^p][C_{\mathbb{R}}^{g(p)}] + [C_{\mathbb{R}}^{f(p)}][SS_g^p] + [SS_f^p][SS_g^p][\mathrm{id}_{\mathbb{R}}].$$

We wish to show that $L = R$. Since $\mathbb{R}_0^\times$ is a common superdomain for L and R, it suffices to show, for all $h \in \mathbb{R}_0^\times$, that $L(h) = R(h)$. Let $h \in \mathbb{R}_0^\times$ be given. We wish to show: $L(h) = R(h)$.

By (2), we have

$$\begin{array}{rcl}(SS_{fg}^p)(h) & = & [(SS_f^p)(h)]\,[g(p)] \;+\; [f(p)]\,[(SS_g^p)(h)] \;+ \\ & & [(SS_f^p)(h)]\,[(SS_g^p)(h)]\,[h].\end{array}$$

Then

$$L(h) \;=\; (SS^p_{fg})(h) \;=\; [(SS^p_f)(h)]\,[g(p)] \;+\; [f(p)]\,[(SS^p_g)(h)] \;+\; [(SS^p_f)(h)]\,[(SS^p_g)(h)]\,[h] \;=\; R(h),$$

as desired. *End of proof of (3).* ☐

We next reprove the **Product Rule**.

THEOREM 32.5. *Let* $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ *and let* $p \in \mathrm{LPD}_{\mathbb{R}}(fg)$. *Then we have:* $(fg)'(p) =^* [f'(p)]\,[g(p)] + [f(p)]\,[g'(p)]$.

Proof. Let $L := (fg)'(p)$ and let $R := [f'(p)][g(p)] + [f(p)][g'(p)]$. Want: $L =^* R$. That is, want: $[R \neq ☹] \Rightarrow [L = R]$. Assume: $R \neq ☹$. Want: $L = R$. That is, want: $(fg)'(p) = R$. As $p \in \mathrm{LPD}_{\mathbb{R}}(fg)$, by $\Leftarrow$ of (2) of Remark 31.3, it suffices to show: $SS^p_{fg} \to R$ near 0.

Since $R \neq ☹$, it follows that $f'(p) \neq ☹ \neq g'(p)$. Then $p \in \mathrm{dom}[(f', g')]$. Then, by Lemma 31.4, we see both that $SS^p_f \to f'(p)$ near 0 and that $SS^p_g \to g'(p)$ near 0. By (3) of Remark 32.4, we conclude that $SS^p_{fg} = [SS^p_f][C^{g(p)}_{\mathbb{R}}] + [C^{f(p)}_{\mathbb{R}}][SS^p_g] + [SS^p_f][SS^p_g][\mathrm{id}_{\mathbb{R}}]$. We have

$$\begin{aligned} SS^p_f &\to f'(p) \text{ near } 0, \\ C^{g(p)}_{\mathbb{R}} &\to g(p) \text{ near } 0, \\ C^{f(p)}_{\mathbb{R}} &\to f(p) \text{ near } 0, \\ SS^p_g &\to g'(p) \text{ near } 0 \qquad \text{and} \\ \mathrm{id}_{\mathbb{R}} &\to 0 \text{ near } 0. \end{aligned}$$

Then $SS^p_{fg} \to [f'(p)][g(p)] + [f(p)][g'(p)] + [f'(p)][g'(p)][0]$ near 0. So, since $R = [f'(p)][g(p)] + [f(p)][g'(p)] + [f'(p)][g'(p)][0]$, we get $SS^p_{fg} \to R$ near 0, as desired. ☐

Our next focus is on the Chain Rule.

Recall: Let W be a normed vector space, let $g : \mathbb{R} \dashrightarrow W$ and let $p \in \mathrm{dom}[g']$. Then $STS^p_g := \mathrm{adj}^{g'(p)}_0(SS^p_g)$. That is, STS^p_g is obtained from SS^p_g by "removing the discontinuity at 0". In particular, STS^p_g is continuous at 0. Since $STS^p_g = \mathrm{adj}^{g'(p)}_0(SS^p_g)$. by definition of adj, we see, for all $h \in \mathbb{R}$, that

$$(STS^p_g)(h) \quad = \quad \begin{cases} (SS^p_g)(h), & \text{if } h \neq 0 \\ g'(p), & \text{if } h = 0. \end{cases}$$

Note: For any normed vector space W, for any $g : \mathbb{R} \dashrightarrow W$, for any object p, if $p \notin \mathrm{dom}[g']$, then $STS_g^p = ☹$.

In the next proof, keep in mind both of the following:

- $\forall a \,{}^*\!\!\in \mathbb{R}, \forall b \in \mathbb{R}, \qquad b + a - b = a$.
- $\forall a, c \,{}^*\!\!\in \mathbb{R}, \forall b \in \mathbb{R}_0^\times, \qquad a/c = (a/b)(b/c)$.

The next lemmas is the **Precalculus Chain Rule**:

LEMMA 32.6. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$, $p \in dom[f]$, $h \in \mathbb{R}_0^\times$, $q := f(p)$, $k := f_p^T(h)$. Assume $k \neq 0$. Then $(SS_{g\circ f}^p)(h) = [(SS_g^q)(k)][(SS_f^p)(h)]$.*

Proof. Since $p \in \mathrm{dom}[f]$, we have $f(p) \in \mathrm{im}[f] \subseteq \mathbb{R}$. We have $q = f(p)$ and $k = f_p^T(h) = [f(p+h)] - [f(p)]$. Then $q + k = f(p+h)$. We have

$$(SS_f^p)(h) \quad = \quad \frac{[f(p+h)] - [f(p)]}{h} \quad = \quad \frac{k}{h}.$$

Then

$$\begin{aligned}
(SS_{g\circ f}^p)(h) &= \frac{[(g\circ f)(p+h)] - [(g\circ f)(p)]}{h} \\
&= \frac{[g(f(p+h))] - [g(f(p))]}{h} \\
&= \frac{[g(q+k)] - [g(q)]}{h} \\
&= \left[\frac{[g(q+k)] - [g(q)]}{k}\right]\left[\frac{k}{h}\right] \\
&= [(SS_g^q)(k)][(SS_f^p)(h)],
\end{aligned}$$

as desired. □

The next lemma is the **Semicalculus Chain Rule**:

LEMMA 32.7. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$, $p \in dom[f]$, $q := f(p)$. Assume that g is differentiable at q. Then $SS_{g\circ f}^p = [(STS_g^q) \circ (f_p^T)] \cdot [SS_f^p]$.*

Proof. Let $L := SS_{g\circ f}^p$ and let $R := [(STS_g^q) \circ (f_p^T)] \cdot [SS_f^p]$. We wish to show: $L = R$. Since $\mathbb{R}_0^\times$ is a common superdomain for L and R, it suffices to show: $\forall h \in \mathbb{R}_0^\times$, $L(h) = R(h)$. Let $h \in \mathbb{R}_0^\times$ be given. We wish to show $L(h) = R(h)$.

Let $k := f_p^T(h)$. One of the following must be true:

(1) $k \neq 0$ or
(2) $k = 0$.

Case 1: By Lemma 32.6, $(SS^p_{g\circ f})(h) = [(SS^q_g)(k)][(SS^p_f)(h)]$. Since $k \neq 0$, we get $(STS^q_g)(k) = (SS^q_g)(k)$. Then

$$\begin{aligned} L(h) &= (SS^p_{g\circ f})(h) \;=\; [(SS^q_g)(k)]\,[(SS^p_f)(h)] \\ &= [(STS^q_g)(k)]\,[(SS^p_f)(h)] \\ &= [(STS^q_g)(f^T_p(h))]\,[(SS^p_f)(h)] \\ &= [((STS^q_g)\circ(f^T_p))(h))]\,[(SS^p_f)(h)] \\ &= (\;[(STS^q_g)\circ(f^T_p)]\cdot[SS^p_f]\;)\,(h) \;=\; R(h), \end{aligned}$$

as desired. *End of Case 1.*

Case 2: As $p \in \mathrm{dom}[f]$, we get $f(p) \in \mathrm{im}[f] \subseteq \mathbb{R}$. As g is differentiable at q, we get $q \in \mathrm{dom}[g']$. Then $f(p) = q \in \mathrm{dom}[g'] \subseteq \mathrm{dom}[g]$. Then $g(f(p)) \in \mathrm{im}[g] \subseteq \mathbb{R}$. We have

$$[f(p+h)] \;-\; [f(p)] \quad=\quad f^T_p(h) \quad=\quad k \quad=\quad 0.$$

Then $f(p+h) = f(p) \in \mathbb{R}$. Then $g(f(p+h)) = g(f(p)) \in \mathbb{R}$. Then

$$[g(f(p+h))] \;-\; [g(f(p))] \quad=\quad 0.$$

We calculate

$$\begin{aligned} L(h) &= (SS^p_{g\circ f})(h) \;=\; \frac{[(g\circ f)(p+h)] \;-\; [(g\circ f)(p)]}{h} \\ &= \frac{[g(f(p+h))] \;-\; [g(f(p))]}{h} \;=\; \frac{0}{h} \;=\; 0. \end{aligned}$$

By definition of STS^q_g, we have $(STS^q_g)(0) = g'(q)$. Since $q \in \mathrm{dom}[g']$, we get $g'(q) \in \mathrm{im}[g'] \subseteq \mathbb{R}$. Then, as $h \in \mathbb{R}^\times_0$, we get $[g'(q)][0/h] = 0$. Since $k = 0$, we have $(STS^q_g)(k) = (STS^q_g)(0)$. Then

$$\begin{aligned} R(h) &= [(STS^q_g)(k)]\,[(SS^p_f)(h)] \\ &= [(STS^q_g)(0)]\left[\frac{[f(p+h)] \;-\; [f(p)]}{h}\right] \\ &= [g'(q)]\left[\frac{0}{h}\right] \;=\; 0. \end{aligned}$$

Then $L(h) = 0 = R(h)$, as desired. *End of Case 2.* $\square$

The next result is called the **Chain Rule**.

THEOREM 32.8. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathrm{LPD}_{\mathbb{R}}(g\circ f)$. Then we have:* $(g\circ f)'(p) \;=^*\; [g'(f(p))]\,[f'(p)]$.

Proof. Let $L := (g \circ f)'(p)$ and $R := [g'(f(p))][f'(p)]$. Want: $L =^* R$. That is, want: $[R \neq ☹] \Rightarrow [L = R]$. Assume: $R \neq ☹$. Want: $L = R$. That is, want: $(g \circ f)'(p) = R$. As $p \in \mathrm{LPD}_{\mathbb{R}}(g \circ f)$, by $\Leftarrow$ of (2) of Remark 31.3, it suffices to show: $SS^p_{g \circ f} \to R$ near 0. Let $q := f(p)$. Then $R = [g'(q)] \cdot [f'(p)]$. Want: $SS^p_{g \circ f} \to [g'(q)] \cdot [f'(p)]$ near 0.

Since $[g'(q)] \cdot [f'(p)] = R \neq ☹$, we see that $g'(q) \neq ☹ \neq f'(p)$. That is, we have: both ($q \in \mathrm{dom}[g']$) and ($p \in \mathrm{dom}[f']$). That is, we have: both (g is differentiable at q) and (f is differentiable at p). By the Semicalculus Chain Rule (Lemma 32.7), $SS^p_{g \circ f} = [(STS^q_g) \circ (f^T_p)] \cdot [SS^p_f]$. It therefore suffices to show both of the following

$$(STS^q_g) \circ (f^T_p) \to g'(q) \text{ near } 0 \qquad \text{and}$$
$$SS^p_f \to f'(p) \text{ near } 0.$$

By Lemma 31.4, we conclude that $SS^p_f \to f'(p)$ near 0. It remains to show: $(STS^q_g) \circ (f^T_p) \to g'(q)$ near 0.

Since f is differentiable at p, by Theorem 26.17, f is continuous at p. Then, by $\Rightarrow$ of Fact 26.16, f^T_p is continuous at 0. That is, $f^T_p \to f^T_P(0)$ near 0. So, since $f^T_p(0) = 0$, we see that $f^T_p \to 0$ near 0. So, since STS^q_g is continuous at 0, by Fall Semester HW#60, we see that $(STS^q_g) \circ (f^T_p) \to (STS^q_g)(0)$ near 0. So, since $(STS^q_g)(0) = g'(q)$, we get $(STS^q_g) \circ (f^T_p) \to g'(q)$ near 0, as desired. □

Our next focus is on the Quotient Rule.

We first need the derivative of reciprocation:

LEMMA 32.9. *Define $r : \mathbb{R} \to \mathbb{R}$ by $r(x) = 1/x$. Then $r' = -r^2$.*

Proof. We have both

$$(\,\mathrm{dom}[r'] \subseteq \mathrm{dom}[r] = \mathbb{R}^\times_0\,) \quad \text{and} \quad (\,\mathrm{dom}[-r^2] = \mathrm{dom}[r] = \mathbb{R}^\times_0\,).$$

Then $\mathbb{R}^\times_0$ is a common superdomain for both r' and $-r^2$. It therefore suffices to show: $\forall p \in \mathbb{R}^\times_0$, $r'(p) = (-r^2)(p)$. Let $p \in \mathbb{R}^\times_0$ be given. We wish to show: $r'(p) = (-r^2)(p)$. We have $(-r^2)(p) = -[r(p)]^2 = -p^{-2}$. We wish to show: $r'(p) = -p^{-2}$.

We have $\mathrm{LPD}_{\mathbb{R}} r = \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[r]) = \mathrm{LP}_{\mathbb{R}}(\mathbb{R}^\times_0) = \mathbb{R}$. Then

$$p \quad \in \quad \mathbb{R}^\times_0 \quad \subseteq \quad \mathbb{R} \quad = \quad \mathrm{LPD}_{\mathbb{R}} r,$$

so, by $\Leftarrow$ of (2) of Remark 31.3, we want: $SS^p_r \to -p^{-2}$ near 0.

Let the function $\tau : \mathbb{R}_0^\times \to \mathbb{R}$ be defined by $\tau(h) = p + h$. Then $\tau(0) = ☹$, but $\tau \to p$ near 0. For all $h \in \mathbb{R}_0^\times$, we have

$$\begin{aligned}(SS_r^p)(h) &= \frac{[r(p+h)] - [r(p)]}{h} = \frac{1}{h}\left[\frac{1}{p+h} - \frac{1}{p}\right] \\ &= \frac{1}{h}\left[\frac{p-(p+h)}{(p+h)p}\right] = \frac{1}{h}\left[\frac{-h}{(p+h)p}\right] \\ &= -\frac{1}{[p+h]\cdot[p]} = -\frac{1}{[\tau(h)]\cdot[C_{\mathbb{R}}^p(h)]} \\ &= \left(-\frac{1}{\tau \cdot C_{\mathbb{R}}^p}\right)(h).\end{aligned}$$

So, as $\mathbb{R}_0^\times$ is a common superdomain for SS_f^p and $-\dfrac{1}{\tau \cdot C_{\mathbb{R}}^p}$, we get $SS_f^p = -\dfrac{1}{\tau \cdot C_{\mathbb{R}}^p}$. So, as $-p^{-2} = -\dfrac{1}{p^2}$, we want: $-\dfrac{1}{\tau \cdot C_{\mathbb{R}}^p} \to -\dfrac{1}{p^2}$ near 0.

Since ($\tau \to p$ near 0) and ($C_{\mathbb{R}}^p \to p$ near 0) and ($p \in \mathbb{R}_0^\times$), it follows that $-\dfrac{1}{\tau \cdot C_{\mathbb{R}}^p} \to -\dfrac{1}{p^2}$ near 0, as desired. □

We will prove the Quotient Rule in the next class.

33. Class 7 on 6 February 2018, Tu of Week 4

Assigned HW 4-1.

Recall: Let f be a function. Then, for any set A,

$$\begin{aligned} f_*(A) &:= \{f(x) \,|\, x \in A \cap (\mathrm{dom}[f])\} \qquad \text{and} \\ f^*(A) &:= \{x \in \mathrm{dom}[f] \,|\, f(x) \in A\}. \end{aligned}$$

Also, for any set $\mathcal{S}$ of sets,

$$\begin{aligned} f_{**}(\mathcal{S}) &:= \{f_*(A) \,|\, A \in \mathcal{S}\} \qquad \text{and} \\ f^{**}(\mathcal{S}) &:= \{f^*(A) \,|\, A \in \mathcal{S}\}. \end{aligned}$$

FACT 33.1. *Let f be a function. Let $\mathcal{S}$ be a set of sets. Then*

(1) $f_*(\bigcup \mathcal{S}) = \bigcup [f_{**}(\mathcal{S})]$,
(2) $f^*(\bigcup \mathcal{S}) = \bigcup [f^{**}(\mathcal{S})]$ *and*
(3) $f^*(\bigcap \mathcal{S}) = \bigcap [f^{**}(\mathcal{S})]$.

Proof. Omitted. □

FACT 33.2. *Let f be a function. Let A and B be sets. Then*

$$f^*(A\backslash B) \;=\; [f^*(A)]\backslash[f^*(B)].$$

Proof. Omitted. ∎

Let f be a 1-1 function. Then f_* is the same as $(f^{-1})^*$ and f_{**} is the same as $(f^{-1})^{**}$. Consequently, for any set $\mathcal{S}$ of sets we have:

$$f_*(\bigcap \mathcal{S}) \;=\; \bigcap[f_{**}(\mathcal{S})].$$

Also, for any sets A and B, we have

$$f_*(A\backslash B) \;=\; [f_*(A)]\backslash[f_*(B)].$$

If, in the last paragraph, we drop the assumption that f is 1-1, then neither formula is correct, and we gave examples to show why.

THEOREM 33.3. *Let X, Y be topological spaces, let $f : X \hookrightarrow\!\!> Y$ be a homeomorphism and let $S \subseteq X$. Then*

(1) $f_*(\mathrm{Int}_X S) \;=\; \mathrm{Int}_Y(f_*(S))$,
(2) $f_*(\mathrm{Cl}_X S) \;=\; \mathrm{Cl}_Y(f_*(S))$,
(3) $f_*(\mathrm{Isol}_X S) \;=\; \mathrm{Isol}_Y(f_*(S))$ *and*
(4) $f_*(\mathrm{LP}_X S) \;=\; \mathrm{LP}_Y(f_*(S))$.

Proof. Proof of (1): Let

$$\begin{aligned} \mathcal{A} &:= \{U \subseteq S \,|\, U \text{ is open in } X\} \qquad \text{and} \\ \mathcal{B} &:= \{V \subseteq f_*(S) \,|\, V \text{ is open in } Y\}. \end{aligned}$$

Then $\bigcup \mathcal{A} = \mathrm{Int}_X S$ and $\bigcup \mathcal{B} = \mathrm{Int}_Y(f_*(S)$.

We have: $\forall U \in \mathcal{A}$, $f_*(U) \in \mathcal{B}$. Then $f_{**}(\mathcal{A}) \subseteq \mathcal{B}$. We have: $\forall V \in \mathcal{B}$, $f^*(V) \in \mathcal{A}$. Then $f^{**}(\mathcal{B}) \subseteq \mathcal{A}$. Then $f_{**}(f^{**}(\mathcal{B})) \subseteq f_{**}(\mathcal{A})$. So, since $f_{**}(f^{**}(\mathcal{B})) = \mathcal{B}$, we get $\mathcal{B} \subseteq f_{**}(\mathcal{A})$. So, since $f_{**}(\mathcal{A}) \subseteq \mathcal{B}$, we get $f_{**}(\mathcal{A}) = \mathcal{B}$. Then $\bigcup[f_{**}(\mathcal{A})] = \bigcup \mathcal{B}$. So, since $\bigcup[f_{**}(\mathcal{A})] = f_*(\bigcup \mathcal{A})$, we get $f_*(\bigcup \mathcal{A}) = \bigcup \mathcal{B}$. So, since $\bigcup \mathcal{A} = \mathrm{Int}_X S$ and $\bigcup \mathcal{B} = \mathrm{Int}_Y(f_*(S)$, we get $f_*(\mathrm{Int}_X S) = \mathrm{Int}_Y(f_*(S))$, as desired. *End of proof of (1).*

Proof of (2): Unassigned HW. *End of proof of (1).*

Proof of (3): Unassigned HW. *End of proof of (3).*

Proof of (4): Unassigned HW. *End of proof of (4).* ∎

DEFINITION 33.4. *For all $S \in \mathbb{R}$, $-S := \{-x \,|\, x \in S\}$.*

Assigned HW#4-2.

Define $r : \mathbb{R} \dashrightarrow \mathbb{R}$ by $r(x) = 1/x$. Recall (Lemma 32.9): $r' = -r^2$.
We explained:

(1) $\forall a, b \mathrel{{}^*{\in}} \mathbb{R}$, both $a \cdot \left(\dfrac{1}{b}\right) = \dfrac{a}{b} = \dfrac{ba}{b^2}$ and $a + (-b) = a - b$,

and (2) $\forall a, b, c \mathrel{{}^*{\in}} \mathbb{R}$, both $\dfrac{a}{c} - \dfrac{b}{c} = \dfrac{a-b}{c}$ and $a \cdot \left(-\dfrac{1}{b}\right) \cdot c = -\dfrac{ac}{b}$.

THEOREM 33.5. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$. Let $p \in \mathrm{LPD}_{\mathbb{R}}(f/g)$. Then*

$$\left(\frac{f}{g}\right)'(p) \quad =^* \quad \frac{[g(p)]\,[f'(p)] \,-\, [f(p)]\,[g'(p)]}{[g(p)]^2}.$$

Proof. Define $r : \mathbb{R} \dashrightarrow \mathbb{R}$ by $r(x) = 1/x$. By Lemma 32.9, $r' = -r^2$. Let $\gamma := r \circ g$. Then $f/g = (f) \cdot (1/g) = (f) \cdot (r \circ g) = f\gamma$.

Since $p \in \mathrm{LPD}_{\mathbb{R}}(f/g) = \mathrm{LPD}_{\mathbb{R}}(f\gamma)$, by the Product Rule (Theorem 32.5), we have $(f\gamma)'(p) =^* [f'(p)][\gamma(p)] + [f(p)][\gamma'(p)]$.

We have $\gamma(p) = r(g(p)) = \dfrac{1}{g(p)}$.

Since $\mathrm{dom}[f\gamma] \subseteq \mathrm{dom}[\gamma]$, we see that $\mathrm{LPD}_{\mathbb{R}}(f\gamma) \subseteq \mathrm{LPD}_{\mathbb{R}}\gamma$. Since $p \in \mathrm{LPD}_{\mathbb{R}}(f\gamma) \subseteq \mathrm{LPD}_{\mathbb{R}}\gamma \subseteq \mathrm{LPD}_{\mathbb{R}}(r \circ g)$, by the Chain Rule (Theorem 32.8), we have $(r \circ g)'(p) =^* [r'(g(p))][g'(p)]$. Then

$$\begin{aligned} \gamma'(p) \quad &= \quad (r \circ g)'(p) \;=^*\; [r'(g(p))]\,[g'(p)] \\ &= \quad [(-r^2)(g(p))]\,[g'(p)] \;=\; \left[-\frac{1}{[g(p)]^2}\right][g'(p)] \end{aligned}$$

Then

$$\begin{aligned} \left(\frac{f}{g}\right)'(p) \quad &= \quad (f\gamma)'(p) \;=^*\; [f'(p)][\gamma(p)] + [f(p)][\gamma'(p)] \\ &=^* \quad [f'(p)]\left[\frac{1}{g(p)}\right] + [f(p)]\left[-\frac{1}{[g(p)]^2}\right][g'(p)] \\ &= \quad \frac{[g(p)][f'(p)]}{[g(p)]^2} - \frac{[f(p)][g'(p)]}{[g(p)]^2} \\ &= \quad \frac{[g(p)][f'(p)] - [f(p)][g'(p)]}{[g(p)]^2}, \end{aligned}$$

as desired. □

This completes our basic results about differentiation, including

- linearity of differentiation,
- the Product Rule
- the Chain Rule
- the Quotient Rule.

We now turn our attention to a new topic: Optimization.

DEFINITION 33.6. *For any f, by f is a* **functional***, we mean:*

$$(\ f \text{ is a function}\ ) \qquad \text{and} \qquad (\ im[f] \subseteq \mathbb{R}\ ).$$

DEFINITION 33.7. *Let f be a functional and let $p \in \mathbb{R}$. By f has a* **maximum** *at p, we mean:*

$$(\ p \in dom[f]\ ) \qquad \text{and} \qquad (\ f(p) \geqslant im[f]\ ).$$

By f has a **minimum** *at p, we mean:*

$$(\ p \in dom[f]\ ) \qquad \text{and} \qquad (\ f(p) \leqslant im[f]\ ).$$

By f has an **extremum** *at p, we mean:*

$$(\ f \text{ has a maximum at } p\ ) \qquad \text{or} \qquad (\ f \text{ has a minimum at } p\ ).$$

DEFINITION 33.8. *Let X be a top. space, $f : X \dashrightarrow \mathbb{R}$, $p \in X$. By f has a* **local maximum** *at p in X, we mean: $\exists V \in \mathcal{N}_X(p)$ s.t.*

$$(\ V \subseteq dom[f]\ ) \qquad \text{and} \qquad (\ f(p) \geqslant f_*(V)\ ).$$

By f has a **local minimum** *at p in X, we mean: $\exists V \in \mathcal{N}_X(p)$ s.t.*

$$(\ V \subseteq dom[f]\ ) \qquad \text{and} \qquad (\ f(p) \leqslant f_*(V)\ ).$$

By f has an **local extremum** *at p in X, we mean:*

$$(\ f \text{ has a local maximum at } p\ ) \text{ or } (\ f \text{ has a local minimum at } p\ ).$$

We sometimes omit "in X" if the choice of X is clear.

Recall: Let X be a toplogical space and let $p \in X$. By (2) of Remark 16.3, any superset of a neighborhood of p is again a neighborhood of p.

REMARK 33.9. *Let X be a topological space, $f : X \dashrightarrow \mathbb{R}$, $p \in X$. Assume that f has a local extremum at p in X. Then*

(1) $dom[f] \in \mathcal{N}_X(p)$,
(2) $p \in \mathrm{Int}_X(dom[f])$ *and*
(3) $(\ \forall z \in X, \{z\}$ *is not open in* $X\ ) \Rightarrow (\ p \in \mathrm{LPD}_X f\ )$.

Proof. Proof of (1): Since f has a local extremum at p in X, it follows that: $\exists V \in \mathcal{N}_X(p)$ such that $V \subseteq \mathrm{dom}[f]$. Then, by (2) of Remark 16.3, we conclude that $\mathrm{dom}[f] \in \mathcal{N}_X(p)$, as desired. *End of proof of (1).*

Proof of (2): Let $S := \mathrm{dom}[f]$. We wish to show: $p \in \mathrm{Int}_X S$.

By (1) of Remark 33.9, $S \in \mathcal{N}_X(p)$. Then by $\Leftarrow$ of HW#3-1, $p \in \mathrm{Int}_X S$, as desired. *End of proof of (2).*

Proof of (3): Assume: $\forall z \in X$, $\{z\}$ is not open in X. We wish to show: $p \in \mathrm{LPD}_X f$. Let $S := \mathrm{dom}[f]$. Then $\mathrm{LPD}_X f = \mathrm{LP}_X S$. We wish to show: $p \in \mathrm{LP}_X S$.

By (2) of Remark 33.9, we have $p \in \mathrm{Int}_X S$. By Remark 31.1, we have $\mathrm{Int}_X S \subseteq \mathrm{LP}_X S$. Then $p \in \mathrm{Int}_X S \subseteq \mathrm{LP}_X S$. *End of proof of (3).* $\square$

We drew a graph of a function $\mathbb{R} \dashrightarrow \mathbb{R}$ and identified and classified several extrema and local extrema. We observed that a local extremum need not be an extremum. We drew a closed line segment in the plane that was neither horizontal nor vertical. This line segment is the graph of a function f whose domain is a closed interval I. We observed that f has extrema at each endpoint I, but that f has no local extrema in $\mathbb{R}$. Note: the endpoints of I are local extrema of f in $\mathrm{dom}[f]$.

REMARK 33.10. *Let X be a topological space, let $f : X \dashrightarrow \mathbb{R}$ and let $p \in \mathrm{Int}_X(\mathit{dom}[f])$. Assume: f has a maximum at p. Then: f has a local maximum at p.*

Proof. Let $V := \mathrm{dom}[f]$. Then $p \in \mathrm{Int}_X(\mathrm{dom}[f]) = \mathrm{Int}_X V$. So, by $\Rightarrow$ of HW#3-1, we have $V \in \mathcal{N}_X(p)$. By Definition 33.7, $f(p) \geqslant \mathrm{im}[f]$. Then $f(p) \geqslant \mathrm{im}[f] = f_*(\mathrm{dom}[f]) = f_*(V)$. Then, by Definition 33.8, f has a local maximum at p. $\square$

REMARK 33.11. *Let X be a topological space, let $f : X \dashrightarrow \mathbb{R}$ and let $p \in \mathrm{Int}_X(\mathit{dom}[f])$. Assume: f has a minimum at p. Then: f has a local minimum at p.*

Proof. Let $V := \mathrm{dom}[f]$. Then $p \in \mathrm{Int}_X(\mathrm{dom}[f]) = \mathrm{Int}_X V$. So, by $\Rightarrow$ of HW#3-1, we have $V \in \mathcal{N}_X(p)$. By Definition 33.7, $f(p) \leqslant \mathrm{im}[f]$. Then $f(p) \leqslant \mathrm{im}[f] = f_*(\mathrm{dom}[f]) = f_*(V)$. Then, by Definition 33.8, f has a local minimum at p. $\square$

REMARK 33.12. *Let X be a topological space, let $f : X \dashrightarrow \mathbb{R}$ and let $p \in \mathrm{Int}_X(\mathit{dom}[f])$. Assume: f has an extremum at p. Then: f has a local extremum at p.*

Proof. Unassigned HW. $\square$

THEOREM 33.13. *Let $a, b \in \mathbb{R}$. Assume $a < b$. Let $f : [a, b] \to \mathbb{R}$ be continuous. Assume that $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that f has an extremum at c.*

Proof. By the Extreme Value Theorem (Theorem 25.16), we know that f has a maximum and a minimum, so choose $p \in [a,b]$ such that f has a minimum at p and choose $q \in [a,b]$ such that f has a maximum at q. Then at least one of the following must hold:

(1) $p \notin \{a,b\}$ or
(2) $q \notin \{a,b\}$ or
(3) $\{p,q\} \subseteq \{a,b\}$.

Case (1): We have $p \in [a,b]\backslash\{a,b\} = (a,b)$. Let $c := p$. We wish to show f has an extremum at c.

Since f has a minimum at p and $c = p$, it follows that f has a minimum at c. Then f has an extremum at c. *End of Case (1).*

Case (2): We have $q \in [a,b]\backslash\{a,b\} = (a,b)$. Let $c := q$. We wish to show f has an extremum at c.

Since f has a maximum at q and $c = p$, it follows that f has a maximum at c. Then f has an extremum at c. *End of Case (2).*

Case (3): Let $c := (a+b)/2$. Then $c \in (a,b)$. We wish to show: f has an extremum at c.

Since $\{p,q\} \subseteq \{a,b\}$, we get $f_*(\{p,q\}) \subseteq f_*(\{a,b\})$. Let $z := f(a)$. By assumption, $f(a) = f(b)$. Then $z = f(b)$. Then

$$f_*(\{a,b\}) \quad = \quad \{f(a), f(b)\} \quad = \quad \{z\}.$$

Then $f_*(\{p,q\}) \subseteq f_*(\{a,b\}) = \{z\}$. Then $f(p) = z = f(q)$. By choice of p and q, we have $f(p) \leqslant \mathrm{im}[f] \leqslant f(q)$. Then $z \leqslant \mathrm{im}[f] \leqslant z$. Then $\mathrm{im}[f] = \{z\}$. Since $f(c) \in \mathrm{im}[f] = \{z\}$, we get $f(c) = z$. Then $\mathrm{im}[f] \leqslant z = f(c)$. Then f has a maximum at c. Then f has an extremum at c, as desired. *End of Case (3).* □

REMARK 33.14. *Let $S \in \mathcal{N}_{\mathbb{R}}(0)$. Then $\exists h > 0$ s.t. $h \in S$.*

Proof. Since $\mathcal{B}_{\mathbb{R}}(0)$ is a neighborhood base at 0 in $\mathbb{R}$, choose $A \in \mathcal{B}_{\mathbb{R}}(0)$ such that $A \subseteq S$. Choose $\delta > 0$ such that $A = B_{\mathbb{R}}(0,\delta)$. Let $h := \delta/2$. Then $h > 0$, and we wish to show: $h \in S$.

We have $h = \delta/2 \in (0,\delta) \subseteq (-\delta,\delta) = B_{\mathbb{R}}(0,\delta) = A \subseteq S$. □

LEMMA 33.15. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in dom[f']$. Assume that f has a local maximum at p. Then $f'(p) \leqslant 0$.*

Proof. Assume $f'(p) > 0$. We aim for a contradiction.

By Lemma 32.1, choose $U \in \mathcal{N}_{\mathbb{R}}(0)$ s.t. $(SS_f^p)_*(U) > 0$.

By assumption, f has a local maximum at p. Choose $V \in \mathcal{N}_{\mathbb{R}}(p)$ s.t.

$$(\ V \subseteq \mathrm{dom}[f]\ ) \qquad \text{and} \qquad (\ f(p) \geqslant f_*(V)\ ).$$

Since $V \in \mathcal{N}_{\mathbb{R}}(p)$, we get $V - p \in \mathcal{N}_{\mathbb{R}}(0)$. Then, as $U \in \mathcal{N}_{\mathbb{R}}(0)$, we get $U \cap (V - p) \in \mathcal{N}_{\mathbb{R}}(0)$. Let $S := U \cap (V - p)$. Then $S \in \mathcal{N}_{\mathbb{R}}(0)$, so, by Remark 33.14, choose $h > 0$ such that $h \in S$. We then have $h \in S = U \cap (V - p) \subseteq V - p$. Then $p + h \in V \subseteq \mathrm{dom}[f]$. Then $f(p + h) \in f_*(V) \leqslant f(p)$. Let $\phi := SS_f^p$. We have

$$\phi(h) \quad = \quad (SS_f^p)(h) \quad = \quad \frac{[f(p+h)] - [f(p)]}{h}.$$

So, since $p + h \in \mathrm{dom}[f]$, since $p \in \mathrm{dom}[f'] \subseteq \mathrm{dom}[f]$ and since $h \neq 0$, we get $h \in \mathrm{dom}[\phi]$. Also, since $f(p+h) \leqslant f(p)$ and since $h > 0$, we get

$$\phi(h) \quad = \quad \frac{[f(p+h)] - [f(p)]}{h} \quad \leqslant \quad 0.$$

As $h \in \mathrm{dom}[\phi]$ and as $h \in S = U \cap (V - p) \subseteq U$, we get $\phi(h) \in \phi_*(U)$. Then $0 \geqslant \phi(h) \in \phi_*(U) = (SS_f^p)_*(U) > 0$, so $0 > 0$. Contradiction. $\square$

LEMMA 33.16. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in$ dom$[f']$. Assume that f has a local maximum at p. Then $f'(p) = 0$.*

Proof. By Lemma 33.15, $f'(p) \leqslant 0$. We wish to show: $f'(p) \geqslant 0$.

Define $\lambda : \mathbb{R} \to \mathbb{R}$ by $\lambda(x) = -x$. By HW#4-1, $\lambda' = C_{\mathbb{R}}^{-1}$. Let $f_0 := f \circ \lambda$ and let $p_0 := \lambda(p)$. By HW#4-3, we see that f_0 has a local maximum at p_0. Since $\mathrm{dom}[f_0] = -(\mathrm{dom}[f])$, by HW#4-2, we get $\mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[f_0]) = -(\mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[f]))$. That is $\mathrm{LPD}_{\mathbb{R}} f_0 = -(\mathrm{LPD}_{\mathbb{R}} f)$. By (3) of Remark 33.9, $p \in \mathrm{LPD}_{\mathbb{R}} f$. Then

$$p_0 \;=\; -p \;\in\; -(\mathrm{LPD}_{\mathbb{R}} f) \;=\; \mathrm{LPD}_{\mathbb{R}} f_0 \;=\; \mathrm{LPD}_{\mathbb{R}}(f \circ \lambda).$$

By the Chain Rule (Theorem 32.8), $(f \circ \lambda)'(p_0) =^* [f'(\lambda(p_0))][\lambda'(p_0)]$. So, since $f \circ \lambda = f_0$, since $\lambda(p_0) = -p_0 = p$ and since $\lambda'(p_0) = -1$, we get $f_0'(p_0) =^* -[f'(p)]$. Since $p \in \mathrm{dom}[f']$, we get $-[f'(p)] \neq ☹$. Then $f_0'(p_0) = -[f'(p)] \neq ☹$. Then $p_0 \in \mathrm{dom}[f_0']$ and $f'(p) = -[f_0'(p_0)]$.

Since f_0 has a local maximum at p_0 and since $p_0 \in \mathrm{dom}[f_0']$, it follows, from Lemma 33.15, that $f_0'(p_0) \leqslant 0$. Then $f'(p) = -[f_0'(p_0)] \geqslant 0$. $\square$

Assigned HW#4-3, HW#4-4 and HW#4-5.

The following is called **Fermat's Theorem**:

THEOREM 33.17. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in$ dom$[f']$. Assume that f has a local extremum at p. Then $f'(p) = 0$.*

Proof. This is HW#4-5. □

34. Class 8 on 8 February 2018, Th of Week 4

Let X and Y be topological spaces and let $f : X \dashrightarrow Y$. Let $S \subseteq X$. By f is **continuous** on S, we mean $S \subseteq \mathrm{dct}[f]$, *i.e.*, ($\forall p \in S$, f is continuous at p). Note that if S is continuous on S, then $S \subseteq \mathrm{dom}[f]$.

Let W be a normed vector space and let $f : \mathbb{R} \dashrightarrow W$. Let $S \subseteq \mathbb{R}$. By f is **differentiable** on S, we mean $S \subseteq \mathrm{dom}[f']$, *i.e.*, ($\forall p \in S$, f is differentiable at p). Note that if S is differentiable on S, then $S \subseteq \mathrm{dom}[f]$.

DEFINITION 34.1. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$ and $S \subseteq \mathbb{R}$. Then f is* **c/d** *on S means:*

$$(\ f \textit{ is continuous on } S\ ) \quad \textit{and} \quad (\ f \textit{ is differentiable on } \mathrm{Int}_{\mathbb{R}} S\ ).$$

In Definition 34.1, the condition "f is continuous on S" is equivalent to: $S \subseteq \mathrm{dct}[f]$. As $\mathrm{dct}[f] \subseteq \mathrm{dom}[f]$, it implies that $S \subseteq \mathrm{dom}[f]$.

In Definition 34.1, the condition "f is differentiable on $\mathrm{Int}_{\mathbb{R}} S$" is equivalent to: $\mathrm{Int}_{\mathbb{R}} S \subseteq \mathrm{dom}[f']$.

DEFINITION 34.2. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$. Then f is* **c/d** *means: f is c/d on dom$[f]$.*

We graphed the function $f : \mathbb{R} \dashrightarrow \mathbb{R}$ defined by $f(x) = \sqrt{1 - x^2}$. (The graph is the upper half of the unit circle about the origin in the plane.) We noted that f is differentiable neither at -1 nor at 1. We explained why f is c/d.

Let W be a normed vector space. Then, for any c/d $f, g : \mathbb{R} \dashrightarrow W$, it is unassigned HW to show that $f + g$ is c/d. For any $a \in \mathbb{R}$, for any $f : \mathbb{R} \dashrightarrow W$, it is unassigned HW to show that af is c/d. For any c/d $f : \mathbb{R} \dashrightarrow \mathbb{R}$, for any c/d $g : \mathbb{R} \dashrightarrow W$, it is unassigned HW to show that $g \circ f$ is c/d.

For any $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$, it is unassigned HW to show that fg and f/g are both c/d.

The following is called **Rolle's Theorem**.

THEOREM 34.3. *Let $a, b \in \mathbb{R}$. Assume $a < b$. Let $f : [a, b] \to \mathbb{R}$. Assume f is c/d. Assume $f(a) = f(b)$. Then $\exists c \in (a, b)$ s.t. $f'(c) = 0$.*

Proof. By Theorem 33.13, choose $c \in (a, b)$ such that f has an extremum at c. We wish to show: $f'(c) = 0$.

We have $c \in (a,b) = \mathrm{Int}_{\mathbb{R}}([a,b]) = \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f])$. Then, by Remark 33.12, f has a local extremum at c. Since f is c/d, we conclude that $\mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f]) \subseteq \mathrm{dom}[f']$. Then $c \in \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f]) \subseteq \mathrm{dom}[f']$. Then, by Fermat's Theorem (Theorem 33.17), $f'(c) = 0$, as desired. $\square$

DEFINITION 34.4. *Let W be a vector space and let $f : \mathbb{R} \dashrightarrow W$. The function $DQ_f : \mathbb{R} \times \mathbb{R} \dashrightarrow W$ is defined by*

$$DQ_f(p,q) \qquad = \qquad \frac{[\,f(q)\,] \;-\; [\,f(p)\,]}{q-p}.$$

Here, "DQ" stands for "difference quotient".

Let W be a vector space and let $f : \mathbb{R} \dashrightarrow W$. The function DQ_f is symmetric, *i.e.*, $\forall p,q \in \mathbb{R}$, $(DQ_f)(p,q) = (DQ_f)(q,p)$. Also, DQ_f is undefined on the diagonal, *i.e.*, $\forall p \in \mathbb{R}$, $(DQ_f)(p,p) = ☹$. Also, we have: $\forall p,h \in \mathbb{R}$, $(SS_f^p)(h) = (DQ_f)(p,p+h)$.

The next two results are precalculus results; they do not use limits or differentiation. They both relate natural properties of a function (injectivity, constancy, increasing, decreasing) to an understanding of its secant slopes.

LEMMA 34.5. *Let W be a vector space, $f : \mathbb{R} \dashrightarrow W$, $D \subseteq dom[f]$, $S := (DQ_f)(D \times D)$. Then all of the following hold:*

(1) $(\,0_W \notin S\,) \;\Leftrightarrow\; (\,f|D$ *is 1-1*$\,)$ *and*
(2) $(\,S \subseteq \{0_W\}\,) \;\Leftrightarrow\; (\,f|D$ *is constant*$\,)$.

Proof. Unassigned HW. $\square$

LEMMA 34.6. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$, $D \subseteq dom[f]$, $S := (DQ_f)_*(D \times D)$. Then all of the following hold:*

(1) $(\,S > 0\,) \;\Leftrightarrow\; (\,f|D$ *is strictly increasing*$\,)$,
(2) $(\,S \geqslant 0\,) \;\Leftrightarrow\; (\,f|D$ *is semiincreasing*$\,)$,
(3) $(\,S < 0\,) \;\Leftrightarrow\; (\,f|D$ *is strictly decreasing*$\,)$ *and*
(4) $(\,S \leqslant 0\,) \;\Leftrightarrow\; (\,f|D$ *is semidecreasing*$\,)$,

Proof. Unassigned HW. $\square$

In the next proof, we need linearity of differentiation, *i.e.*, we need both of the following two results: Let W be a normed vector space. First, for all $f,g : \mathbb{R} \dashrightarrow W$, for all $p \in \mathrm{LPD}_{\mathbb{R}}(f+g)$,

$$(f+g)'(p) \quad =^* \quad [f'(p)] \;+\; [g'(p)].$$

Second, for all $a \in \mathbb{R}$, for all $f : \mathbb{R} \dashrightarrow W$, for all $p \in \mathrm{LPD}_{\mathbb{R}}(af)$,

$$(af)'(p) \quad =^* \quad a \cdot [f'(p)].$$

The first result is Theorem 26.19. The second one was never quite state explcitly, but follows from the Product Rule (Theorem 32.5, with f replaced by $C_{\mathbb{R}}^a$ and g by f).

The following is called **Mean Value Theorem**.

THEOREM 34.7. *Let $a, b \in \mathbb{R}$. Assume $a < b$. Let $f : [a, b] \to \mathbb{R}$. Assume f is c/d. Then $\exists c \in (a, b)$ s.t. $f'(c) = (DQ_f)(a, b)$.*

Proof. Let $m := (DQ_f)(a, b)$. Define $\lambda : \mathbb{R} \to \mathbb{R}$ by $\lambda(x) = mx$. By HW#4-1, $\lambda' = C_{\mathbb{R}}^m$. Since λ is differentiable, λ is c/d. We have

$$m \quad = \quad (DQ_f)(a, b) \quad = \quad \frac{[f(b)] - [f(a)]}{b - a},$$

so, multiplying by $b - a$, we get $(b - a)m = [f(b)] - [f(a)]$. Then $mb - ma = [f(b)] - [f(a)]$, and so $[f(a)] - ma = [f(b)] - mb$. Let $\phi := f - \lambda$. Note that $\mathrm{dom}[\phi] = [a, b] \cap \mathbb{R} = [a, b]$. Also, as f and λ are both c/d, we conclude that ϕ is c/d. Also,

$$\begin{aligned} \phi(a) &= [f(a)] - [\lambda(a)] = [f(a)] - ma \\ &= [f(b)] - mb = [f(b)] - [\lambda(b)] = \phi(b). \end{aligned}$$

Then, by Rolle's Theorem (Theorem 34.3, with f replaced by ϕ), choose $c \in (a, b)$ s.t. $\phi'(c) = 0$. We wish to show: $f'(c) = (DQ_f)(a, b)$.

We have $c \in (a, b) \subseteq [a, b] = \mathrm{LP}_{\mathbb{R}}([a, b])$. Then

$$c \quad \in \quad \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[\phi]) \quad = \quad \mathrm{LPD}_{\mathbb{R}}\phi \quad = \quad \mathrm{LPD}_{\mathbb{R}}(f - \lambda).$$

Then $(f - \lambda)'(c) \;=^*\; [f'(c] - [\lambda'(c)]$. So, because $f - \lambda = \phi$, we conclude that $\phi'(c) \;=^*\; [f'(c] - [\lambda'(c)]$. Since $\phi'(c) = 0$ and since $\lambda'(c) = C_{\mathbb{R}}^m(c) = m$, we get $0 \;=^*\; [f'(c] - m$. As f is c/d, we have: $\mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f]) \subseteq \mathrm{dom}[f']$. Then

$$c \;\in\; (a, b) \;=\; \mathrm{Int}_{\mathbb{R}}([a, b]) \;=\; \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f]) \;\subseteq\; \mathrm{dom}[f'].$$

Then $f'(c) \in \mathrm{im}[f'] \subseteq \mathbb{R}$. Then $0 \;=^*\; [f'(c] - m \neq ☹$. It follows that $0 = [f'(c)] - m$. Then $f'(c) = m = (DQ_f)(a, b)$, as desired. $\square$

Our buzz phrase for the Mean Value Theorem (Theorem 34.7) is: "Any secant slope is a tangent slope somewhere in between." In Corollary 34.10 below, we'll see how useful this is. Basically, by studying the derivative, we get information about tangent slopes. This calculus information then translates into information about secant slopes, which then, by Lemma 34.6, gives basic precalculus information about the function. For example, we can now use calculus to figure out on which intervals a function is increasing and decreasing, and then use that to maximize or minimize the function. Since optimization is useful even to a person who may not care about calculus, another buzz phrase for the Mean Value Theorem might be, "Calculus is useful."

FACT 34.8. *Let W be a normed vector space and let $f : \mathbb{R} \dashrightarrow W$. Let $D \subseteq dom[f]$. Then all of the following hold:*

(1) (*f is continuous on D*) $\Rightarrow$ (*$f|D$ is continuous*),
(2) $\forall x \in \mathrm{Int}_{\mathbb{R}} D, \qquad (f|D)'(x) = f'(x)$,
(3) (*f is c/d*) $\Rightarrow$ (*$f|D$ is c/d*).

Proof. Proof of (1): True by Theorem 22.3. *End of proof of (1).*

Proof of (2): Let $g := f|D$, $U := \mathrm{Int}_{\mathbb{R}} D$. We want: $g' = f'$ on U.

By definition of restriction, $g = f$ on D. So, since $U \subseteq D$, we get: $g = f$ on U. So, since U is an open subset of $\mathbb{R}$, by HW#3-4, we conclude that $g' = f'$ on U, as desired. *End of proof of (2).*

Proof of (3): Assume that f is c/d. We wish to show: $f|D$ is c/d.

Since f is c/d, it follows that f is continuous. So, since $D \subseteq \mathrm{dom}[f]$, we see that f is continuous on D. Then, by (1) of Fact 34.8, we see that $f|D$ is continuous. It remains to show: $f|D$ is differentiable on $\mathrm{Int}_{\mathbb{R}} D$. We wish to show: $\forall x \in \mathrm{Int}_{\mathbb{R}} D$, $(f|D)'(x) \neq ☹$. Let $x \in \mathrm{Int}_{\mathbb{R}} D$ be given. We wish to show: $(f|D)'(x) \neq ☹$.

Since $D \subseteq \mathrm{dom}[f]$, $\mathrm{Int}_{\mathbb{R}} D \subseteq \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f])$. As f is c/d, we see that f is differentiable on $\mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f])$. So, as $x \in \mathrm{Int}_{\mathbb{R}} D \subseteq \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f])$, we see that f is differentiable at x. That is, we have $f'(x) \neq ☹$. By (2) of Fact 34.8, $(f|D)'(x) = f'(x)$. Then $(f|D)'(x) = f'(x) \neq ☹$, as desired. *End of proof of (3).* □

The next result is the **Mean Value Inclusion**. It says: "Along an interval, any secant slope is a tangent slope from the interior."

THEOREM 34.9. *Let $f : \mathbb{R} \to \mathbb{R}$ and let I be an interval. Assume that f is c/d on I. Then $(DQ_f)_*(I \times I) \subseteq f'_*(\mathrm{Int}_{\mathbb{R}} I)$.*

Proof. Let $\phi := DQ_f$. We wish to show: $\phi_*(I \times I) \subseteq f'_*(\mathrm{Int}_{\mathbb{R}} I)$. We wish to show: $\forall v \in \mathrm{dom}[\phi]$,

$$[\ v \in I \times I\ ] \quad \Rightarrow \quad [\ \phi(v) \in f'_*(\mathrm{Int}_{\mathbb{R}} I)\ ].$$

Let $v \in \mathrm{dom}[\phi]$ be given. We wish to show:

$$[\ v \in I \times I\ ] \quad \Rightarrow \quad [\ \phi(v) \in f'_*(\mathrm{Int}_{\mathbb{R}} I)\ ].$$

Assume $v \in I \times I$. We wish to show: $\phi(v) \in f'_*(\mathrm{Int}_{\mathbb{R}} I)$.

Since $v \in I \times I$, choose $\alpha, \beta \in I$ s.t. $v = (\alpha, \beta)$. Since ϕ is symmetric, we have $\phi(\alpha, \beta) = \phi(\beta, \alpha)$. We have $(\alpha, \beta) = v \in \mathrm{dom}[\phi] = \mathrm{dom}[DQ_f]$, and it follows, from the definition of DQ_f, that $\alpha \neq \beta$.

Let $a := \min\{\alpha, \beta\}$ and let $b := \max\{\alpha, \beta\}$. Then $a < b$. Also, either $(a, b) = (\alpha, \beta)$ or $(a, b) = (\beta, \alpha)$. Then $a, b \in I$ and $\phi(\alpha, \beta) = \phi(a, b)$.

Since $a, b \in I$ and since I is an interval, it follows that $[a, b] \subseteq I$. Let $D := [a, b]$. Then $D \subseteq I$, so $\mathrm{Int}_{\mathbb{R}} D \subseteq \mathrm{Int}_{\mathbb{R}} I$. Since f is c/d on I, we see that $I \subseteq \mathrm{dct}[f]$. Then $D \subseteq I \subseteq \mathrm{dct}[f] \subseteq \mathrm{dom}[f]$. Let $g := f|D$. Then, by (3) of Fact 34.8, g is c/d. By the Mean Value Theorem (Theorem 34.7, with f replaced by g), choose $c \in (a, b)$ such that $g'(c) = (DQ_g)(a, b)$. By (2) of Fact 34.8, we have: $\forall x \in (a, b)$, $g'(x) = f'(x)$. Then $g'(c) = f'(c)$.

Since $f(a) = g(a)$ and $f(b) = g(b)$, we get $(DQ_f)(a, b) = (DQ_g)(a, b)$. Then $\phi(a, b) = (DQ_f)(a, b) = (DQ_g)(a, b) = g'(c)$. We have

$$c \ \in \ (a, b) \ = \ \mathrm{Int}_{\mathbb{R}}(\,[a, b]\,) \ = \ \mathrm{Int}_{\mathbb{R}} D \ \subseteq \ \mathrm{Int}_{\mathbb{R}} I.$$

As f is c/d on I, we get $\mathrm{Int}_{\mathbb{R}} I \subseteq \mathrm{dom}[f']$. Then $c \in \mathrm{Int}_{\mathbb{R}} I \subseteq \mathrm{dom}[f']$. Then $f'(c) \in f'_*(\mathrm{Int}_{\mathbb{R}} I)$. As $v = (\alpha, \beta)$, we get $\phi(v) = \phi(\alpha, \beta)$. Then $\phi(v) = \phi(\alpha, \beta) = \phi(a, b) = g'(c) = f'(c) \in f'_*(\mathrm{Int}_{\mathbb{R}} I)$, as desired. □

COROLLARY 34.10. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$ *and let* I *be an interval. Assume that* f *is c/d on* I. *Let* $T := f'_*(\mathrm{Int}_{\mathbb{R}} I)$. *Then all of the following hold:*

(1) $(\ 0 \notin T\ ) \quad \Rightarrow \quad (\ f|I$ *is 1-1* $)$,
(2) $(\ T \subseteq \{0\}\ ) \quad \Leftrightarrow \quad (\ f|I$ *is constant* $)$,
(3) $(\ T > 0\ ) \quad \Rightarrow \quad (\ f|I$ *is strictly increasing* $)$,
(4) $(\ T \geqslant 0\ ) \quad \Leftrightarrow \quad (\ f|I$ *is semiincreasing* $)$,
(5) $(\ T < 0\ ) \quad \Rightarrow \quad (\ f|I$ *is strictly decreasing* $)$ *and*
(6) $(\ T \leqslant 0\ ) \quad \Leftrightarrow \quad (\ f|I$ *is semidecreasing* $)$.

Proof. Let $S := (DQ_f)_*(I \times I)$.

Proof of (1): Assume: $0 \notin T$. We wish to show: $f|I$ is 1-1.

By the Mean Value Inclusion (Theorem 34.9), we get $S \subseteq T$. So, since $0 \notin T$, we get $0 \notin S$. Then, by (1) of Lemma 34.5, $f|I$ is 1-1, as desired. *End of proof of (1).*

Proof of (2): Proof of $\Rightarrow$*:* Assume: $T \subseteq \{0\}$. Want: $f|I$ is constant.

By the Mean Value Inclusion (Theorem 34.9), we get $S \subseteq T$. Since $S \subseteq T \subseteq \{0\}$, by (1) of Lemma 34.5, we conclude that $f|I$ is constant, as desired. *End of proof of* $\Rightarrow$*.*

Proof of $\Leftarrow$*:* Unassigned HW. *End of proof of* $\Leftarrow$*. End of proof of (2).*

Proof of (3): Assume: $T > 0$. Want: $f|I$ is strictly increasing.

By the Mean Value Inclusion (Theorem 34.9), we get $S \subseteq T$. Since $S \subseteq T > 0$, by (1) of Lemma 34.5, we conclude that $f|I$ is strictly increasing, as desired. *End of proof of (3).*

Proof of (4): Proof of$\Rightarrow$*:* Assume $T \geqslant 0$. Want: $f|I$ is semiincreasing.

By the Mean Value Inclusion (Theorem 34.9), we get $S \subseteq T$. Since $S \subseteq T \geqslant 0$, by (1) of Lemma 34.5, we conclude that $f|I$ is semiincreasing, as desired. *End of proof of* $\Rightarrow$*.*

Proof of $\Leftarrow$*:* Unassigned HW. *Hint:* Similar to the proof of $\Leftarrow$ of (6) of Corollary 34.10; see below. *End of proof of* $\Leftarrow$*. End of proof of (4).*

Proof of (5): Assume: $T < 0$. Want: $f|I$ is strictly decreasing.

By the Mean Value Inclusion (Theorem 34.9), we get $S \subseteq T$. Since $S \subseteq T < 0$, by (1) of Lemma 34.5, we conclude that $f|I$ is strictly decreasing, as desired. *End of proof of (5).*

Proof of (6): Proof of$\Rightarrow$*:* Assume $T \leqslant 0$. Want: $f|I$ is semidecreasing.

By the Mean Value Inclusion (Theorem 34.9), we get $S \subseteq T$. Since $S \subseteq T \leqslant 0$, by (1) of Lemma 34.5, we conclude that $f|I$ is semidecreasing, as desired. *End of proof of* $\Rightarrow$*.*

Proof of $\Leftarrow$*:* Assume: $f|I$ is semidecreasing. We wish to show: $T \leqslant 0$. We wish to show: $\forall y \in T$, $y \leqslant 0$. Let $y \in T$ be given,. We wish to show: $y \leqslant 0$. Assume $y > 0$. We aim for a contradiction.

Let $I_0 := \mathrm{Int}_{\mathbb{R}} I$. Then I_0 is open in $\mathbb{R}$, $I_0 \subseteq I$ and $T = f'_*(I_0)$. Since $y \in T = f'_*(I_0)$, choose $p \in I_0 \cap (\mathrm{dom}[f'])$ such that $f'(p) = y$. Note that $p \in I_0$. Since $f'(p) = y > 0$, by Lemma 32.1, choose $U \in \mathcal{N}_{\mathbb{R}}(0)$ such that $(SS_f^p)_*(U) > 0$. As I_0 is open in $\mathbb{R}$, we see that $I_0 - p$ is open in $\mathbb{R}$. Since $p \in I_0$, we get $0 \in I_0 - p$. So, since $I_0 - p$ is open in $\mathbb{R}$, by Remark 16.4, we see that $I_0 - p \in \mathcal{N}_{\mathbb{R}}(0)$. So, since $U \in \mathcal{N}_{\mathbb{R}}(0)$, we get: $(I_0 - p) \cap U \in \mathcal{N}_{\mathbb{R}}(0)$. So, since $\{0\} \notin \mathcal{N}_{\mathbb{R}}(0)$ and since $\varnothing \notin \mathcal{N}_{\mathbb{R}}(0)$, we get: $(I_0 - p) \cap U \nsubseteq \{0\}$. Choose $h \in (I_0 - p) \cap U$ such that $h \neq 0$. Since $h \in I_0 - p$, we get $p + h \in I_0$. So, since $p \in I_0$, we have $(p, p + h) \in I_0 \times I_0$. Since f is c/d on I, we get $I \subseteq \mathrm{dom}[f]$. Then $I_0 \subseteq I \subseteq \mathrm{dom}[f]$, so $h \in I_0 - p \subseteq (\mathrm{dom}[f]) - p$. So, since $h \neq 0$, we see that $h \in [(\mathrm{dom}[f]) - p]_0^\times$. So, since $\mathrm{dom}[SS_f^p] = [(\mathrm{dom}[f]) - p]_0^\times$, we get $h \in \mathrm{dom}[SS_f^p]$. So, since $h \in (I_0 - p) \cap U \subseteq U$, we conclude that $(SS_f^p)(h) \in (SS_f^p)_*(U)$. Then, by the choice of U, $(SS_f^p)(h) > 0$.

Since $h \in \mathrm{dom}[SS_f^p]$, we conclude that $(SS_f^p)(h) \neq ☹$. It follows that $(DQ_f)(p, p+h) = (SS_f^p)(h) \neq ☹$, so $(p, p+h) \in \mathrm{dom}[DQ_f]$. So, since $(p, p+h) \in I_0 \times I_0 \subseteq I \times I$, we get $(DQ_f)(p, p+h) \in (DQ_f)_*(I \times I)$. As $f|I$ is semidecreasing, by (4) of Lemma 34.6, we get $(DQ_f)_*(I \times I) \leqslant 0$. Then $0 < (SS_f^p)(h) = (DQ_f)(p, p+h) \in (DQ_f)_*(I \times I) \leqslant 0$, so $0 < 0$. Contradiction. *End of proof of $\Leftarrow$. End of proof of (6).* $\square$

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^3$ and let $I := \mathbb{R}$. Then I is an interval and f is c/d on I. Also $f|I = f$ and f is 1-1 and f is strictly increasing. For all $x \in \mathbb{R}$, we have $f'(x) = 3x^2$. Let $T := f'_*(\mathrm{Int}_{\mathbb{R}} I)$. Then $0 = f'(0) \in T$. So, since $f|I$ is 1-1, we see that the converse of (1) of Corollary 34.10 fails. Since $0 \in T$, we have: NOT$(T > 0)$. So, as $f|I$ is strictly increasing, the converse of (3) of Corollary 34.10 fails.

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = -x^3$ and let $I := \mathbb{R}$. Then I is an interval and f is c/d on I. Also $f|I = f$ and f is strictly decreasing. For all $x \in \mathbb{R}$, we have $f'(x) = -3x^2$. Let $T := f'_*(\mathrm{Int}_{\mathbb{R}} I)$. Then $0 = f'(0) \in T$, and so we have: NOT$(T < 0)$. So, as $f|I$ is strictly decreasing, the converse of (5) of Corollary 34.10 fails.

The buzz phrase for the Mean Value Theorem asserts, "Any secant slope is a tangent slope somewhere in between." As the preceding two paragraphs show, it is not hard for a function to "level out" for an instant. This creates a horizontal tangent line, but does not necessarily create horizontal secant lines. So remember that it is NOT true that any tangent slope is a secant slope.

COROLLARY 34.11. *Let $f : \mathbb{R} \to \mathbb{R}$. Assume that $f' = C_{\mathbb{R}}^0$. Then f is constant, i.e., there exists $a \in \mathbb{R}$ such that $f = C_{\mathbb{R}}^a$.*

Proof. Let $I := \mathbb{R}$. Let $T := f'_*(\mathrm{Int}_{\mathbb{R}} I)$. Then $T = f'_*(\mathbb{R}) = \{0\}$. Then by $\Rightarrow$ of (2) of Corollary 34.10, we see that f is constant, as desired. $\square$

Recall: For all $x \in \mathbb{R}$, we have

$$\begin{aligned}
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots , \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots .
\end{aligned}$$

Note that $\cos 0 = 1$ and that $\sin 0 = 0$. Also, note that, for all $x \in \mathbb{R}$,

$$\begin{aligned} \cos(-x) &= \cos x \\ \sin(-x) &= -(\sin x). \end{aligned}$$

FACT 34.12. *We have* $\cos' = -\sin$ *and* $\sin' = \cos$.

Proof. Should be proved in MATH 3283. □

FACT 34.13. *We have* $\cos'' = -\cos$ *and* $\sin'' = -\sin$.

Proof. We have

$$\begin{aligned} \cos'' &= (-\sin)' = -(\sin') = -\cos \qquad \text{and} \\ \sin'' &= \cos' = -\sin, \end{aligned}$$

as desired. □

THEOREM 34.14. *Let* $f : \mathbb{R} \to \mathbb{R}$. *Assume* $f'' = -f$. *Assume* $f(0) = f'(0) = 0$. *Then* $f = C_{\mathbb{R}}^0$.

Proof. Want: $\forall x \in \mathbb{R}$, $f(x) = 0$. Let $x \in \mathbb{R}$ be given. Want: $f(x) = 0$.

Let $E := f^2 + (f')^2$. Then

$$E' = 2ff' + 2f'f'' = 2ff' + (2f') \cdot (-f) = C_{\mathbb{R}}^0,$$

so, by Corollary 34.11, choose $a \in \mathbb{R}$ such that $E = C_{\mathbb{R}}^a$. We have

$$a = C_{\mathbb{R}}^a(0) = E(0) = [f(0)]^2 + [f'(0)]^2 = 0^2 + 0^2 = 0.$$

Then $E = C_{\mathbb{R}}^a = C_{\mathbb{R}}^0$. Let $y := f(x)$ and $z := f'(x)$. Want: $y = 0$.

Since $y^2 \geqslant 0$ and $z^2 \geqslant 0$, we have $y^2 + z^2 \geqslant y^2 \geqslant 0$. So, because $y^2 + z^2 = [f(x)]^2 + [f'(x)]^2 = E(x) = C_{\mathbb{R}}^0(x) = 0$, we get $0 \geqslant y^2 \geqslant 0$. Then $y^2 = 0$. Then $y = 0$, as desired. □

Using Theorem 34.14 and Corollary 34.11, we will prove a variety of algebraic facts about sin and cos, *e.g.*, $\forall x \in \mathbb{R}$, $\sin(2\pi + x) = \sin x$. That is, we can use calculus ideas to prove results in trigonometry.

35. Class 9 on 13 February 2018, Tu of Week 5

Recall: $\forall x \in \mathbb{R}$,

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots, \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots. \end{aligned}$$

Note that $\cos(0) = 1$ and that $\sin(0) = 0$. Note: $\forall x \in \mathbb{R}$,

$$\begin{aligned} \cos(-x) &= \cos x, \\ \sin(-x) &= -(\sin x). \end{aligned}$$

Recall: $\sin' = \cos$ and $\cos' = -\sin$.

THEOREM 35.1. *We have:* $\sin^2 + \cos^2 = C^1_{\mathbb{R}}$.

Proof. Let $f := \sin^2 + \cos^2$. We want: $f = C^1_{\mathbb{R}}$.

Claim: $f' = C^0_{\mathbb{R}}$. *Proof of Claim:* Want: $\forall x \in \mathbb{R}$, $f'(x) = C^0_{\mathbb{R}}(x)$. Let $x \in \mathbb{R}$ be given. We wish to show: $f'(x) = C^0_{\mathbb{R}}(x)$.

We have

$$x \in \mathbb{R} = \mathrm{LP}_{\mathbb{R}}\mathbb{R} = \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[\sin^2 + \cos^2]) = \mathrm{LPD}_{\mathbb{R}}(\sin^2 + \cos^2).$$

Then $(\sin^2 + \cos^2)'(x) =^* [(\sin^2)'(x)] + [(\cos^2)'(x)]$.

Define $\sigma : \mathbb{R} \to \mathbb{R}$ by $\sigma(y) = y^2$. Then: $\forall y \in \mathbb{R}$, $\sigma'(y) = 2y$. Also, we have $\sin^2 = \sigma \circ \sin$ and $\cos^2 = \sigma \circ \cos$. We have

$$x \in \mathbb{R} = \mathrm{LP}_{\mathbb{R}}\mathbb{R} = \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[\sigma \circ \sin]) = \mathrm{LPD}_{\mathbb{R}}(\sigma \circ \sin).$$

Then $(\sigma \circ \sin)'(x) =^* [\sigma'(\sin(x))][\sin'(x)]$. We have

$$x \in \mathbb{R} = \mathrm{LP}_{\mathbb{R}}\mathbb{R} = \mathrm{LPD}_{\mathbb{R}}(\cos^2) = \mathrm{LPD}_{\mathbb{R}}(\sigma \circ \cos).$$

Then $(\sigma \circ \cos)'(x) =^* [\sigma'(\cos(x))][\cos'(x)]$. Then

$$\begin{aligned} f'(x) &= (\sin^2 + \cos^2)'(x) \\ &=^* [(\sin^2)'(x)] + [(\cos^2)'(x)] \\ &= [(\sigma \circ \sin)'(x)] + [(\sigma \circ \cos)'(x)] \\ &=^* [\sigma'(\sin(x))][\sin'(x)] + [\sigma'(\cos(x))][\cos'(x)] \\ &= [2 \cdot (\sin(x))][\cos(x)] + [2 \cdot (\cos(x))][-\sin(x)] \\ &= 0 \quad \neq \quad ☹. \end{aligned}$$

Then $f'(x) = 0$. Then $f'(x) = 0 = C^0_{\mathbb{R}}(x)$. *End of proof of Claim.*

By the Claim and Corollary 34.11, choose $a \in \mathbb{R}$ s.t. $f = C^a_{\mathbb{R}}$. Because $a = C^a_{\mathbb{R}}(0) = f(0) = 0^2 + 1^2 = 1$, we get $f = C^a_{\mathbb{R}} = C^1_{\mathbb{R}}$, as desired. $\square$

THEOREM 35.2. *Let* $f : \mathbb{R} \to \mathbb{R}$. *Assume that* $f'' = -f$. *Let* $a := f(0)$ *and let* $b := f'(0)$. *Then* $f = a \cdot \cos + b \cdot \sin$.

Proof. Let $\phi := f - a \cdot \cos - b \cdot \sin$. We wish to show $\phi = C^0_{\mathbb{R}}$.

Claim 1: $\phi' = f' + a \cdot \sin - b \cdot \cos$. *Proof of Claim 1:* Unassigned HW. *End of proof of Claim 1.*

Claim 2: $\phi'' = f'' + a \cdot \cos + b \cdot \sin$. *Proof of Claim 2:* Unassigned HW. *End of proof of Claim 2.*

Since $f'' = -f$, by Claim 2, we have $\phi'' = -f + a \cdot \cos + b \cdot \sin$. So, since $-\phi = -f + a \cdot \cos + b \cdot \sin$, we get $\phi'' = -\phi$.

Since $\phi = f - a \cdot \cos - b \cdot \sin$ and since $f(0) = a$, we calculate $\phi(0) = a - a \cdot 1 - b \cdot 0 = 0$. By Claim 1, $\phi' := f' + a \cdot \sin - b \cdot \cos$. So, since $f'(0) = b$, we calculate $\phi'(0) = b + a \cdot 0 - b \cdot 1 = 0$.

Then, by Theorem 34.14 (with f replaced by ϕ), we get $\phi = C_{\mathbb{R}}^0$. □

THEOREM 35.3. *For all $w, x \in \mathbb{R}$, we have*

$$\sin(w + x) \quad = \quad (\sin w) \cdot (\cos x) \; + \; (\cos w) \cdot (\sin x).$$

Proof. Let $w \in \mathbb{R}$ be given. We wish to show: $\forall x \in \mathbb{R}$,

$$\sin(w + x) \quad = \quad (\sin w) \cdot (\cos x) \; + \; (\cos w) \cdot (\sin x).$$

Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \sin(w + x)$.

Claim 1: $\forall x \in \mathbb{R}$, $f'(x) = \cos(w + x)$. *Proof of Claim 1:* Unassigned HW. *End of proof of Claim 1.*

Claim 2: $\forall x \in \mathbb{R}$, $f''(x) = (-\sin)(w + x)$. *Proof of Claim 2:* Unassigned HW. *End of proof of Claim 2.*

Claim 3: $f'' = -f$. *Proof of Claim 3:* We wish to show: $\forall x \in \mathbb{R}$, $f''(x) = (-f)(x)$. Let $x \in \mathbb{R}$ be given. Want: $f''(x) = (-f)(x)$.

By Claim 2, we have $f''(x) = (-\sin)(w + x)$. By definition of f, we have $f(x) = \sin(w + x)$. Then

$$f''(x) = (-\sin)(w + x) = -[\sin(w + x)] = -[f(x)] = (-f)(x),$$

as desired. *End of proof of Claim 3.*

Let $a := f(0)$ and let $b := f'(0)$. By Claim 3 and Theorem 35.2, $f = a \cdot \cos + b \cdot \sin$. By definition of f, we have $f(0) = \sin(w + 0)$. Then $a = f(0) = \sin(w + 0) = \sin w$. By Claim 1, we have $f'(0) = \cos(w + 0)$. Then $b = f'(0) = \cos(w + 0) = \cos w$. Then: $\forall x \in \mathbb{R}$,

$$\begin{aligned} \sin(w + x) \quad &= \quad f(x) \quad = \quad (a \cdot \cos + b \cdot \sin)(x) \\ &= \quad a \cdot (\cos x) + b \cdot (\sin x) \\ &= \quad (\sin w) \cdot (\cos x) + (\cos w) \cdot (\sin x), \end{aligned}$$

as desired. □

Assigned HW#5-1 and HW#5-2.

Recall: $\min\{x \geqslant 0 \,|\, \cos x = 0\} \neq ☹$. Recall that, from the defintion of ϕ, we have: $\pi/2 = \min\{x \geqslant 0 \,|\, \cos x = 0\}$. Then $\min\{x \geqslant 0 \,|\, \cos x = 0\} \geqslant \pi/2$. Note that $\cos(\pi/2) = 0$. Also, we have the following:

REMARK 35.4. *For all* $u \in (0, \pi/2)$, *we have* $\cos u > 0$.

Proof. Let $u \in (0, \pi/2)$ be given. We wish to show: $\cos u > 0$. Assume $\cos u \leqslant 0$. We aim for a contradiction.

By the Intermediate Value Theorem (Theorem 29.7), we conclude that $[\cos 0 | \cos u] \subseteq \cos_*([0, u])$. Since $\cos u \leqslant 0 < 1 = \cos 0$, we get $[\cos 0 | \cos u] = [\cos u, \cos 0]$. We have $\cos u \leqslant 0 \leqslant \cos 0$, and so $0 \in [\cos u, \cos 0]$. Then $0 \in [\cos u, \cos 0] = [\cos 0 | \cos u] \subseteq \cos_*([0, u])$. Choose $t \in [0, u]$ such that $0 = \cos t$.

Then $t \in \{x \geqslant 0 \,|\, \cos x = 0\} \geqslant \pi/2$, and so we get $\pi/2 \leqslant t$. Then $\pi/2 \leqslant t \in [0, u] \leqslant u \in (0, \pi/2) < \pi/2$, so $\pi/2 < \pi/2$. Contradiction. $\square$

REMARK 35.5. *We have* $\sin(\pi/2) = 1$.

Proof. Let $c := \cos$ and let $s := \sin$. We wish to show: $s(\pi/2) = 1$.

By Theorem 35.1, we have $\sin^2 + \cos^2 = C^1_{\mathbb{R}}$, *i.e.*, $s^2 + c^2 = C^1_{\mathbb{R}}$. Then $[s(\pi/2)]^2 + [c(\pi/2)]^2 = C^1_{\mathbb{R}}(\pi/2)$. So, as $c(\pi/2) = \cos(\pi/2) = 0$ and $C^1_{\mathbb{R}}(\pi/2) = 1$, we get $[s(\pi/2)]^2 + 0^2 = 1$. Then $[s(\pi/2)]^2 = 1$, so $s(\pi/2) \in \{1, -1\}$. It suffices to show: $s(\pi/2) > 0$.

Let $I := [0, \pi/2]$. Then $\mathrm{Int}_{\mathbb{R}} I = (0, \pi/2)$. By Remark 35.4, we know that: $\forall u \in (0, \pi/2)$, $\cos u > 0$. Then $\cos_*(\,(0, \pi/2)\,) > 0$. That is, $c_*(\,(0, \pi/2)\,) > 0$. Then $s'_*(\mathrm{Int}_{\mathbb{R}} I) = c_*(\,(0, \pi/2)\,) > 0$. Then, by (5) of Corollary 34.10, $s|I$ is strictly increasing. We have $0, \pi/2 \in I$ and $0 < \pi/2$. Then $(s|I)(0) < (s|I)(\pi/2)$. Then

$$s(\pi/2) \;=\; (s|I)(\pi/2) \;>\; (s|I)(0) \;=\; s(0) \;=\; \sin 0 \;=\; 0,$$

as desired. $\square$

REMARK 35.6. *We have* $\sin \pi = 0$ *and* $\cos \pi = -1$.

Proof. By Remark 35.5, $\sin(\pi/2) = 1$. Recall that $\cos(\pi/2) = 0$.

We have $\sin \pi = \sin(2 \cdot [\pi/2])$. By HW#5-2,

$$\sin(\,2 \cdot [\pi/2]\,) \quad = \quad 2 \cdot [\,\sin(\,\pi/2\,)\,] \cdot [\,\cos(\,\pi/2\,)\,].$$

Then

$$\begin{aligned} \sin \pi &= \sin(\,2 \cdot [\pi/2]\,) \\ &= 2 \cdot [\,\sin(\,\pi/2\,)\,] \cdot [\,\cos(\,\pi/2\,)\,] \\ &= 2 \cdot [1] \cdot [0] \;=\; 0. \end{aligned}$$

It remains to show: $\cos\pi = -1$.

We have $\cos\pi = \cos(2\cdot[\pi/2])$. By HW#5-2,

$$\cos\,(\,2\cdot[\pi/2]\,) \quad = \quad [\,\cos^2(\,\pi/2\,)\,] - [\,\sin^2(\,\pi/2\,)\,].$$

Then

$$\begin{aligned}\cos\pi &= \cos(\,2\cdot[\pi/2]\,)\\ &= [\,\cos^2(\,\pi/2\,)\,] - [\,\sin^2(\,\pi/2\,)\,]\\ &= [\,0^2\,] - [\,1^2\,] \;=\; -1,\end{aligned}$$

as desired. $\square$

REMARK 35.7. *For all $x \in \mathbb{R}$, we have $\sin((\pi/2) - x) = \cos x$.*

Proof. By Remark 35.5, $\sin(\pi/2) = 1$. Recall that $\cos(\pi/2) = 0$. Let $x \in \mathbb{R}$ be given. We wish to show: $\sin((\pi/2) - x) = \cos x$.

Recall that $\cos(-x) = \cos x$. By Theorem 35.3, we conclude that: $\sin((\pi/2) - x) = [\sin(\pi/2)]\cdot[\cos(-x)] + [\cos(\pi/2)]\cdot[\sin(-x)]$. Then

$$\begin{aligned}\sin((\pi/2) - x) &= [\sin(\pi/2)]\cdot[\cos(-x)] + [\cos(\pi/2)]\cdot[\sin(-x)]\\ &= [\,1\,]\cdot[\,\cos x\,] + [\,0\,]\cdot[\,\sin\,(\,-x\,)\,] \;=\; \cos x,\end{aligned}$$

as desired. $\square$

REMARK 35.8. *For all $x \in \mathbb{R}$, we have $\sin(\pi + x) = -\sin x$ and $\cos(\pi + x) = -\cos x$.*

Proof. By Remark 35.6, we have $\sin\pi = 0$ and $\cos\pi = -1$. By Theorem 35.3, we have: $\sin(\pi + x) = [\sin\pi]\cdot[\cos x] + [\cos\pi]\cdot[\sin x]$. By HW#5-1, we have: $\cos(\pi + x) = [\cos\pi]\cdot[\cos x] - [\sin\pi]\cdot[\sin x]$. Then

$$\begin{aligned}\sin(\pi + x) &= [\sin\pi]\cdot[\cos x] + [\cos\pi]\cdot[\sin x]\\ &= [\,0\,]\cdot[\,\cos x\,] + [\,-1\,]\cdot[\,\sin x\,] \;=\; -\sin x\end{aligned}$$

and

$$\begin{aligned}\cos(\pi + x) &= [\cos\pi]\cdot[\cos x] - [\sin\pi]\cdot[\sin x]\\ &= [\,-1\,]\cdot[\,\cos x\,] - [\,0\,]\cdot[\,\sin x\,] \;=\; -\cos x,\end{aligned}$$

as desired. $\square$

DEFINITION 35.9. *The function $exp : \mathbb{R} \dashrightarrow \mathbb{R}$ is defined by*

$$exp\,x \quad = \quad 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots.$$

Following material that is typically covered in MATH 3283, the power series in Definition 31.7 has an infinite radius of convergence. Then $\mathrm{dom}[\exp] = \mathbb{R}$. Moreover, the term-by-term derivative of this power series yield a power series that also has an infinte radius of convergence. Also, that term-by-term derivative is a power series for $\exp' x$. This shows: $\exp' = \exp$.

In Theorem 35.14, we will prove Let $f : \mathbb{R} \to \mathbb{R}$. Assume $f' = f$. Assume $f(0) = 0$. Then $f = C_{\mathbb{R}}^0$. We'll call this our "f-result", and will prove it in Theorem 35.14 below.

Assigned HW#5-3 and HW#5-4 and HW#5-5.

NOTE: You are allowed to use the "f-result" above in your solutions to both HW#5-3 and HW#5-4.

Using HW#5-5, to prove our "f-result", it is enough to prove: Let $g : \mathbb{R} \to [0, \infty)$. Assume $g' = 2g$. Assume $g(0) = 0$. Then $g = C_{\mathbb{R}}^0$. We'll call this our "g-result", and will prove it in Lemma 35.13 below.

LEMMA 35.10. *Let $g : \mathbb{R} \to [0, \infty)$. Assume $g' = 2g$. Then g is semiincreasing.*

Proof. Since $g : \mathbb{R} \to [0, \infty)$, we get $\mathrm{im}[g] \geqslant 0$. Then $\mathrm{im}[2g] \geqslant 0$. We have $g'_*(\mathbb{R}) = \mathrm{im}[g'] = \mathrm{im}[2g] \geqslant 0$. Then, as $\mathbb{R}$ is an interval and $g|\mathbb{R} = g$, by (4) of Corollary 34.10, we see that g is semiincreasing. $\square$

LEMMA 35.11. *Let $g : \mathbb{R} \to [0, \infty)$. Assume $g' = 2g$. Assume $g(0) = 0$. Then $g = 0$ on $(-\infty, 0]$.*

Proof. We wish to show: $\forall x \in (-\infty, 0]$, $g(x) = 0$. Let $x \in (-\infty, 0]$ be given. Want: $g(x) = 0$. We have $g(x) \in \mathrm{im}[g] \geqslant 0$. Want: $g(x) \leqslant 0$.

We have $x \in (-\infty, 0] \leqslant 0$. By Lemma 35.10, g is semiincreasing. So, since $x \leqslant 0$, we get $g(x) \leqslant g(0)$. Then $g(x) \leqslant g(0) = 0$, as desired. $\square$

LEMMA 35.12. *Let $g : \mathbb{R} \to [0, \infty)$ and let $a \in \mathbb{R}$. Assume $g' = 2g$. Assume $g(a) = 0$. Then $g = 0$ on $[a, a + (1/3)]$.*

Proof. Let $b := a + (1/3)$, $K := [a, b]$. Want: $g = 0$ on K. As $g : \mathbb{R} \to [0, \infty)$, we get $g_*(K) \geqslant 0$ It suffices to show: $g_*(K) \leqslant 0$.

By Lemma 35.10, g is semiincreasing. Then $g_*([a, b]) \leqslant g(b)$. Let $M := g(b)$. Then $g_*(K) = g_*([a, b]) \leqslant g(b) = M$. Want: $M \leqslant 0$.

By the Mean Value Theorem (Theorem 34.7), choose $c \in (a, b)$ such that $g'(c) = (DQ_g)(a, b)$. By assumption, $g(a) = 0$. It follows that

$[g(b)] - [g(a)] = g(b) = M$. Also, $b - a = 1/3$. Therefore

$$g'(c) \quad = \quad (DQ_g)(a,b) \quad = \quad \frac{[g(b)] - [g(a)]}{b - a} \quad = \quad \frac{M}{1/3} \quad = \quad 3M.$$

We have $c \in (a,b) \subseteq [a,b] = K$. Then $g(c) \in g_*(K)$. So, because $g_*(K) \leqslant M$, we see that $g(c) \leqslant M$. By assumption, $g' = 2g$. Then $3M = g'(c) = (2g)(c) = 2 \cdot [g(c)] \leqslant 2M$. Then $M = 3M - 2M \leqslant 0$. $\square$

We can now prove our g-result:

LEMMA 35.13. *Let $g : \mathbb{R} \to [0,\infty)$. Assume $g' = 2g$. Assume $g(0) = 0$. Then $g = C_{\mathbb{R}}^0$.*

Proof. We want: $g = 0$ on $\mathbb{R}$. For all $j \in \mathbb{N}$, let $I_j := (-\infty,(j-1)/3]$. Then $\mathbb{R} = I_1 \cap I_2 \cap I_3 \cap \cdots$, so it suffices to show: $\forall j \in \mathbb{N}$, $g = 0$ on I_j. For all $j \in \mathbb{N}$, let $P_j := [\, g = 0 \text{ on } I_j \,]$. Want: $\forall j \in \mathbb{N}$, P_j.

We have $I_1 = (-\infty,(1-1)/3] = (-\infty,0]$. So, by Lemma 35.11, we see that $g = 0$ on I_1. That is, P_1 is true. So, by the Principle of Mathematical Induction, it suffices to show: $\forall j \in \mathbb{N}$, $(\, P_j \Rightarrow P_{j+1}\,)$. Let $j \in \mathbb{N}$ be given. We want to show: $P_j \Rightarrow P_{j+1}$. Assume: P_j. We want to show: P_{j+1}. We know: $g = 0$ on I_j. We want: $g = 0$ on I_{j+1}.

Let $a := (j-1)/3$, $b := j/3$. Then $I_j = (-\infty,a]$, $I_{j+1} = (-\infty,b]$. Then $I_j \cup [a,b] = I_{j+1}$. Since $a \in (-\infty,a] = I_j$ and since $g = 0$ on I_j, we conclude that $g(a) = 0$. So, since $a + (1/3) = b$, it follows, from Lemma 35.12, that $g = 0$ on $[a,b]$. Since $g = 0$ both on I_j and on $[a,b]$, and since $I_j \cup [a,b] = I_{j+1}$, we get: $g = 0$ on I_{j+1}. $\square$

We can now prove our f-result:

THEOREM 35.14. *Let $f : \mathbb{R} \to \mathbb{R}$. Assume both that $f' = f$ and that $f(0) = 0$. Then $f = C_{\mathbb{R}}^0$.*

Proof. We wish to show: $\forall x \in \mathbb{R}$, $f(x) = C_{\mathbb{R}}^0(x)$. Let $x \in \mathbb{R}$ be given. We wish to show: $f(x) = C_{\mathbb{R}}^0(x)$. We wish to show: $f(x) = 0$.

Let $g := f^2$. Then $g : \mathbb{R} \to [0,\infty)$. By HW#5-5, we get: $g' = 2g$. Also, $g(0) = [f(0)]^2 = 0^2 = 0$. Then, by Lemma 35.13, $g = C_{\mathbb{R}}^0$. Since $[f(x)]^2 = g(x) = (C_{\mathbb{R}}^0)(0) = 0$, we get $f(x) = 0$, as desired. $\square$

LEMMA 35.15. *For all $x \in \mathbb{R}$, we have: $\exp x > 0$.*

Proof. Since $\exp : \mathbb{R} \to \mathbb{R}$ is differentiable, $\exp : \mathbb{R} \to \mathbb{R}$ is continuous. We have $\mathrm{dom}[\exp] = \mathbb{R} = (-\infty,\infty)$, so $\mathrm{dom}[\exp]$ is an interval. So, by the Intermediate Value Theorem (Theorem 29.7), as $\exp 0 = 1 > 0$, it suffices to show: $\forall x \in \mathbb{R}$, $\exp x \neq 0$. Given $x \in \mathbb{R}$. Want: $\exp x \neq 0$.

By HW#5-4, $[\exp x] \cdot [\exp(-x)] = \exp 0$. So, since $\exp 0 = 1 \neq 0$, we conclude that $\exp x \neq 0$, as desired. $\square$

LEMMA 35.16. *The function* $\exp : \mathbb{R} \to \mathbb{R}$ *is strictly increasing.*

Proof. As $\exp' = \exp$, by Lemma 35.15, we know: $\forall x \in \mathbb{R}$, $\exp'(x) > 0$, Then, by (3) of Corollary 34.10, $\exp : \mathbb{R} \to \mathbb{R}$ is strictly increasing. $\square$

LEMMA 35.17. *We have* $[1, \infty) \subseteq \mathrm{im}[\exp]$.

Proof. We wish to show: $\forall a \in [1, \infty)$, $a \in \mathrm{im}[\exp]$. Let $a \in [1, \infty)$ be given. We wish to show: $a \in \mathrm{im}[\exp]$.

As $a > 0$, we get $1 + a + \dfrac{a^2}{2!} + \cdots \geqslant a$, *i.e.*, that $\exp a \geqslant a$. Also, $a \geqslant 1 = \exp 0$. Then $\exp 0 \leqslant a \leqslant \exp a$, *i.e.*, $a \in [\exp 0, \exp a]$. By the Intermediate Value Theorem (Theorem 29.7), we conclude that $[\exp 0|\exp a] \subseteq \exp_*([0|a])$.

Then $a \in [\exp 0, \exp a] = [\exp 0|\exp a] \subseteq \exp_*([0|a]) \subseteq \mathrm{im}[\exp]$. $\square$

LEMMA 35.18. *We have* $\exp : \mathbb{R} \hookrightarrow\!\!> (0, \infty)$.

Proof. By Lemma 35.16, $\exp : \mathbb{R} \to \mathbb{R}$ is strictly increasing, and is therefore 1-1. It remains to show: $\mathrm{im}[\exp] = (0, \infty)$. By Lemma 35.15, $\mathrm{im}[\exp] \subseteq (0, \infty)$. We therefore wish to show: $(0, \infty) \subseteq \mathrm{im}[\exp]$. By Lemma 35.17, we have $[1, \infty) \subseteq \mathrm{im}[\exp]$. It therefore suffices to show $(0, 1) \subseteq \mathrm{im}[\exp]$. We want to show: $\forall b \in (0, 1)$, $b \in \mathrm{im}[\exp]$. Let $b \in (0, 1)$ be given. We wish to show: $b \in \mathrm{im}[\exp]$.

Let $a := 1/b$. Since $b \in (0, 1)$, we get $a \in (1, \infty)$. Also, $1/a = b$. Since $a \in (1, \infty) \subseteq [1, \infty) \subseteq \mathrm{im}[\exp]$, choose $t \in \mathbb{R}$ s.t. $a = \exp t$. By HW#5-4, $[\exp t] \cdot [\exp(-t)] = \exp 0$. Then

$$\exp(-t) \quad = \quad \frac{\exp 0}{\exp t} \quad = \quad \frac{1}{a} \quad = \quad b.$$

Then $b = \exp(-t) \in \mathrm{im}[\exp]$, as desired. $\square$

We define $0^0 = 1$. Then, $\forall x \in \mathbb{R}$, $x^0 = 1$. Define

$$\ln \ := \ \exp^{-1} \ : \ (0, \infty) \ \hookrightarrow\!\!> \ \mathbb{R}.$$

For all $p > 0$, we define $0^p = 0$. For all $x > 0$, for all $p \in \mathbb{R}$, we define $x^p = \exp(p \cdot [\ln x])$. Finally, we define

$$\tan \ := \ \sin / \cos \qquad \text{and} \qquad \cot \ := \ \cos / \sin .$$

Recall Definition 22.14:

DEFINITION 35.19. *Let Y, Z be topological spaces, $f : Y \dashrightarrow Z$. Then f is* **open** *or* (Y, Z)**-open** *means:*

$$\forall \textit{open } U \textit{ in } Y,\ f_*(U) \textit{ is open in } Z.$$

DEFINITION 35.20. *Let Y, Z be topological spaces, $f : Y \dashrightarrow Z$. Then f is* **closed** *or* (Y, Z)**-closed** *means:*

$$\forall \textit{closed } C \textit{ in } Y,\ f_*(C) \textit{ is closed in } Z.$$

LEMMA 35.21. *Let Y and Z be topological spaces, $f : Y \hookrightarrow\!\!> Z$. Then:* $\qquad$ *(f is open) $\Leftrightarrow$ (f is closed).*

Proof. Proof of $\Rightarrow$: Unassigned HW. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume: f is closed. We wish to show: f is open. We wish to show: $\forall$ open U in Y, $f_*(U)$ is open in Z. Let an open U in Y be given. We wish to show: $f_*(U)$ is open in Z.

Since $U \subseteq Y$, we get $Y \backslash (Y \backslash U) = U$. Since U is open in Y, we see that $Y \backslash U$ is closed in Y. Let $C := Y \backslash U$. Then C is closed in Y. Since C is closed in Y and since f is closed, we get: $f_*(C)$ is closed in Z. Then $Z \backslash [f_*(C)]$ is open in Z. We want: $f_*(U) = Z \backslash [f_*(C)]$.

We have $f_*(U) = f_*(Y \backslash C)$. Because f is 1-1, we conclude that $f_*(Y \backslash C) = [f_*(Y)] \backslash [f_*(C)]$. Since $f : Y \to\!\!> Z$, it follows that $f_*(Y) = Z$. We have $U = Y \backslash (Y \backslash U) = Y \backslash C$. Then

$$f_*(U) \quad = \quad f_*(Y \backslash C) \quad = \quad [f_*(Y)] \backslash [f_*(C)] \quad = \quad Z \backslash [f_*(C)],$$

as desired. *End of proof of $\Leftarrow$.* ☐

Recall: Let X be a topological space X and let $A \subseteq X$. The **sequential closure** of A in X is

$$\mathrm{sCl}_X A \quad := \quad \{ z \in X \,|\, \exists s \in A^{\mathbb{N}} \text{ s.t. } s_\bullet \to z \text{ in } X \}.$$

Then $\mathrm{sCl}_X A \subseteq \mathrm{Cl}_X A$. Also, if X is metrizable, then $\mathrm{sCl}_X A = \mathrm{Cl}_X A$. Thus, in the situations we care about, closure = sequential closure.

DEFINITION 35.22. *Let X be a topological space and let $A \subseteq X$. Then A is* **sequentially closed** *in X means:* $\mathrm{sCl}_X A = A$.

Let X be a topological space and let $A \subseteq X$. Then:

$$\begin{aligned} (\, A \text{ is closed in } X\,) \quad &\Leftrightarrow \quad (\, \mathrm{Cl}_X A \subseteq A\,) \\ &\Rightarrow \quad (\, \mathrm{sCl}_X A \subseteq A\,) \\ &\Leftrightarrow \quad (\, A \text{ is sequentially closed in } X\,). \end{aligned}$$

Moreover, if X is metrizable, then

$$\begin{aligned}(\,A\text{ is closed in }X\,)\quad&\Leftrightarrow\quad(\,\mathrm{Cl}_X A\subseteq A\,)\\&\Leftrightarrow\quad(\,\mathrm{sCl}_X A\subseteq A\,)\\&\Leftrightarrow\quad(\,A\text{ is sequentially closed in }X\,).\end{aligned}$$

Thus, in the situations we care about, closed = sequentially closed.

LEMMA 35.23. *Let K be a sequentially compact topological space and let $C \subseteq K$. Assume that C is sequentially closed in K. Then C is sequentially compact.*

Proof. We wish to show: $\forall s \in C^{\mathbb{N}}$, $s_\bullet$ is subconvergent in C. Let $s \in C^{\mathbb{N}}$ be given. We wish to show: $s_\bullet$ is subconvergent in C.

Since $s_\bullet \in C^{\mathbb{N}} \subseteq K^{\mathbb{N}}$ and since K is sequentially compact, we see that $s_\bullet$ is subconvergent in K. Choose a subsequence t of $s_\bullet$ such that $t_\bullet$ is convergent in K. Choose $z \in K$ such that $t_\bullet \to z$ in K. Since $s_\bullet \in C^{\mathbb{N}}$ and since $t_\bullet$ is a subsequence of $s_\bullet$, we conclude that $t_\bullet \in C^{\mathbb{N}}$. So, since $t_\bullet \to z$ in K, we get $z \in \mathrm{sCl}_K C$. Since C is sequentially closed in K, we have $\mathrm{sCl}_K C = C$. Then $z \in \mathrm{sCl}_K C = C$.

By $\Leftarrow$ of Theorem 19.9, since $t_\bullet \to z$ in K, we see that $t_\bullet \to z$ in C. Then $t_\bullet$ is convergent in C. So, since $t_\bullet$ is a subsequence of $s_\bullet$, we see that $s_\bullet$ is subconvergent in C, as desired. $\square$

36. Class 10 on 15 February 2018, Th of Week 5

Announced Midterm 1 on Thursday 22 February 2018, Th of Week 6. The midterm will, as usual be during the last hour of class, and we will, as usual, review and have a break before the midterm.

Recall (Lemma 35.21): Let Y and Z be topological spaces, and let $f : Y \hookrightarrow\!\!> Z$. Then [(f is open) $\Leftrightarrow$ (f is closed)].

Recall: Let X be a metrizable topological space and let $C \subseteq X$. Then [(C is closed in X) $\Leftrightarrow$ (C is sequentially closed in X)].

Recall (Lemma 35.23): Let K be a sequentially compact topological space and let C be a sequentially closed subset of K. Then C is sequentially compact.

Recall (Theorem 25.6): Let K be a sequentially compact topological space, let Y be a topological space and let $\phi : K \to Y$ be continuous. Then $\mathrm{im}[\phi]$ is sequentially compact.

LEMMA 36.1. *let X and Y be topological spaces, let K be a sequentially compact subset of X and let $f : X \to Y$ be continuous. Then $f_*(K)$ is sequentially compact.*

Proof. Let $\phi := f|K$. Then, by HW#2-4, $\phi : K \to Y$ is continuous. By Theorem 25.6, $\mathrm{im}[\phi]$ is sequentially compact. So, as $\mathrm{im}[\phi] = f_*(K)$, we conclude that $f_*(K)$ is sequentially compact, as desired. $\square$

FACT 36.2. *Let X be a Hausdorff topological space. Let K be a sequentially compact subset of X. Then K is sequentially closed in X.*

Proof. We wish to show: $\mathrm{sCl}_X K = K$. Since $\mathrm{sCl}_X K \supseteq K$, it suffices to show: $\mathrm{sCl}_X K \subseteq K$. We wish to show: $\forall z \in \mathrm{sCl}_X K$, $z \in K$. Let $z \in \mathrm{sCl}_X K$ be given. We wish to show: $z \in K$.

Since $z \in \mathrm{sCl}_X K$, choose $s \in K^{\mathbb{N}}$ s.t. $s_\bullet \to z$ in X. Since K is sequentially compact, $s_\bullet$ is subconvergent in K. Choose a subsequence t of $s_\bullet$ s.t. $t_\bullet$ is convergent in K. Choose $y \in K$ s.t. $t_\bullet \to y$ in K. By Theorem 19.9, $t_\bullet \to y$ in X. Since $s_\bullet \to z$ in X and since $t_\bullet$ is a subsequence of $s_\bullet$, by Theorem 23.14, we see that $t_\bullet \to z$ in X. Since

$$(\ s_\bullet \to y \text{ in } X\ ) \quad \text{and} \quad (\ s_\bullet \to z \text{ in } X\ ),$$

and since X is Hausdorff, by Corollary 24.9, we get $y = z$. Then $z = y \in K$, as desired. $\square$

COROLLARY 36.3. *Let X be a metrizable topological space and let K be a sequentially compact subset of X. Then K is closed in X.*

Proof. By Fact 36.2, we see that K is sequentially closed in X. So, since X is metrizable, K is closed in X, as desired. $\square$

Recall (Theorem 25.1): Let X be a metric space. Let K be a sequentially compact subset of X. Then K is closed and bounded in X. Using this, we have another proof of Corollary 36.3:

Proof. Choose $d \in \mathcal{M}(X)$ such that $\mathcal{T}_d$ is the topology on X. By Theorem 25.1, K is closed in $(X, \mathcal{T}_d)$. That is, K is closed in X. $\square$

Recall one-dimensional Invariance of Domain (Theorem 30.3): Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ be 1-1 and continuous, and let $U \subseteq \mathrm{dom}[f]$. Assume that U is open in $\mathbb{R}$. Then $f_*(U)$ is open in $\mathbb{R}$.

Recall: Let $X := [1, 2) \cup [3, 4)$ and let $Y := [5, 7)$. Then there exists $f : X \hookrightarrow\!\!> Y$ such that f is continuous, but not (X, Y)-open. This is unfortunate, but Theorem 30.3 has the following corollary, which says that if a function $\mathbb{R} \dashrightarrow \mathbb{R}$ is both injective and continuous, then it is "open on the interior of its domain". Often that interior is most of the domain, and so "most" of the function is a homeomorphism.

COROLLARY 36.4. *Let $X \subseteq \mathbb{R}$, $f : X \hookrightarrow \mathbb{R}$. Let $X_0 := \mathrm{Int}_{\mathbb{R}} X$, let $Y_0 := f_*(X_0)$ and let $f_0 := f|X_0$. Then*

(1) *Y_0 is open in $\mathbb{R}$ and*
(2) *$f_0 : X_0 \to Y_0$ is a homeomorphism.*

Proof. Proof of (1): Since X_0 is open in $\mathbb{R}$ and $X_0 \subseteq X = \mathrm{dom}[f]$, we see, by Theorem 30.3, that $f_*(X_0)$ is open in $\mathbb{R}$. That is, Y_0 is open in $\mathbb{R}$, as desired. *End of proof of (1).*

Proof of (2): Want: $\forall$open subset U of X_0, $f_*(U)$ is open in Y_0. Let an open subset U of X_0 be given. Want: $f_*(U)$ is open in Y_0.

Since U is open in X_0 and X_0 is open in $\mathbb{R}$, it follows that U is open in $\mathbb{R}$. So, since $U \subseteq X_0 \subseteq X = \mathrm{dom}[f]$, it follows, from Theorem 30.3, that $f_*(U)$ is open in $\mathbb{R}$. Then $[f_*(U)] \cap Y_0$ is open in Y_0. Since $f_*(U) \subseteq f_*(X_0) = Y_0$, we get $[f_*(U)] \cap Y_0 = f_*(U)$. Then $f_*(U)$ is open in Y_0, as desired. *End of proof of (2).* $\square$

The next result is called the **Topological Inverse Function Theorem for open subsets of** $\mathbb{R}$ or the **TIFT for open in** $\mathbb{R}$:

THEOREM 36.5. *Let $X, Y \subseteq \mathbb{R}$ and let $f : X \hookrightarrow\!> Y$ be continuous. Assume X is open in $\mathbb{R}$. Then*

(A) *Y is open in $\mathbb{R}$ and*
(B) *$f : X \to Y$ is a homeomorphism.*

Proof. Let $X_0 := \mathrm{Int}_{\mathbb{R}} X$, $Y_0 := f_*(X_0)$, $f_0 := f|X_0$. By Corollary 36.4,

(1) Y_0 is open in $\mathbb{R}$ and
(2) $f_0 : X_0 \to Y_0$ is a homeomorphism.

Since X is open in $\mathbb{R}$, we have $\mathrm{Int}_{\mathbb{R}} X = X$. That is $X_0 = X$. Then $Y_0 = f_*(X_0) = f_*(X) = \mathrm{im}[f] = Y$. Then $f_0 = f|X_0 = f|X = f$. Then (A) follows from (1), and (B) follows from (2). $\square$

The next result is called the **Topological Inverse Function Theorem for intervals** or the **TIFT for intervals**:

THEOREM 36.6. *Let $X, Y \subseteq \mathbb{R}$ and let $f : X \hookrightarrow\!> Y$ be continuous. Assume X is an interval. Then*

(A) *Y is an interval and*

(B) *$f : X \to Y$ is a homeomorphism.*

Proof. Omitted. Note that, by Theorem 30.1, f is stricly monotone. So by looking at f or $-f$, we may reduce to the case where f is strictly increasing. We will try to double back to this later, and add a proof. It's within the scope of this course, but is complicated, because there are many different kinds of intervals (open, closed, compact, *etc.*). $\square$

Finally, we have the **Topological Inverse Function Theorem for compcata** or the **TIFT for compacta**:

THEOREM 36.7. *Let X and Y be topological spaces. Assume that X is sequentially compact, and that Y is metrizable. Let $f : X \hookrightarrow\!> Y$ be continuous. Then*

(A) *Y is sequentially compact and*
(B) *$f : X \to Y$ is a homeomorphism.*

Proof. By Theorem 25.6, $\mathrm{im}[f]$ is sequentially compact. So, since $\mathrm{im}[f] = Y$, we see that Y is sequentially compact, proving (A). It remains to prove (B). Since $f : X \hookrightarrow\!> Y$ is continuous, we need only show that $f : X \hookrightarrow\!> Y$ is open. Then, by Lemma 35.21, we wish to show that $f : X \to Y$ is closed. That is, we wish to show: For any closed subset C of X, $f_*(C)$ is closed in Y. Let a closed subset C of X be given. We wish to show: $f_*(C)$ is closed in Y.

Since C is closed in X, it follows that C is sequentially closed in X. Then, since X is sequentially compact, by Lemma 35.23, we see that C is sequentially compact. Then, by Lemma 36.1, $f_*(C)$ is sequentially compact as well. Then, by Corollary 36.3, $f_*(C)$ is closed in Y. $\square$

We now turn from topological inverse function theory to differential inverse function theory. Our next main result is Theorem 36.11, which is a kind of "anti-IFT". However, we first need two preliminary remarks.

REMARK 36.8. *Let X and Y be topological spaces, $\phi : X \dashrightarrow Y$, $w \in X$, $z \in Y$. Assume: $\forall h \in X$, $\phi(h) \overset{*}{=} z$. Then $\phi \to z$ near w.*

Proof. Want: $\forall V \in \mathcal{N}_Y(z)$, $\exists U \in \mathcal{N}_X^\times(w)$ such that $\phi_*(U) \subseteq V$. Let $V \in \mathcal{N}_Y(z)$ be given. Want: $\exists U \in \mathcal{N}_X^\times(w)$ such that $\phi_*(U) \subseteq V$.

By Remark 16.4, we conclude that $X \in \mathcal{N}_X(w)$. Let $U := X_w^\times$. Then $U \in \mathcal{N}_X^\times(w)$, and we wish to show: $\phi_*(U) \subseteq V$. Want: $\forall h \in \mathrm{dom}[\phi]$,

$$[\ h \in U\ ] \quad \Rightarrow \quad [\ \phi(h) \in V\ ].$$

Let $h \in \mathrm{dom}[\phi]$ be given. We wish to show:

$$[\ h \in U\ ] \quad \Rightarrow \quad [\ \phi(h) \in V\ ].$$

Assume $h \in U$. Want: $\phi(h) \in V$.

By assumption, $\phi(h) \stackrel{*}{=} z$. Since $h \in \mathrm{dom}[\phi]$, we have $\phi(h) \neq ☹$. Then $\phi(h) = z$. As $V \in \mathcal{N}_Y(z)$, we get $z \in V$. Then $\phi(h) = z \in V$. $\square$

REMARK 36.9. *Let $X \subseteq \mathbb{R}$ and let $p \in \mathbb{R}$. Then $(\mathrm{id}_X)'(p) \stackrel{*}{=} 1$.*

Proof. Let $i := \mathrm{id}_X$. Want: $i'(p) \stackrel{*}{=} 1$. That is, we wish to show:

$$[\ i'(p) \neq ☹\ ] \quad \Rightarrow \quad [\ i'(p) = 1\ ].$$

Assume: $i'(p) \neq ☹$. We wish to show: $i'(p) = 1$.

For all $x \in \mathbb{R}\backslash(\mathrm{LPD}_{\mathbb{R}} i)$, we have $f'(x) = ☹$. So, as $p \in \mathbb{R}$ and $i'(p) \neq ☹$, we get $p \in \mathrm{LPD}_{\mathbb{R}} i$. It suffices to show: $SS_i^p \to 1$ near 0.

For all $h \in \mathbb{R}$, we have

$$(SS_i^p)(h) \quad = \quad \frac{[i(p+h)] - [i(p)]}{h} \ \stackrel{*}{=}\ \frac{[p+h] - [p]}{h} \ \stackrel{*}{=}\ 1.$$

Then, by Remark 36.8, we have $SS_i^p \to 1$ near 0, as desired. $\square$

LEMMA 36.10. *Let $X, Y \subseteq \mathbb{R}$, $f : X \hookrightarrow\!> Y$, $p \in dom[f']$. Then*

(1) $p \in \mathrm{LP}_{\mathbb{R}} X$ *and*
(2) $[\ f'(p) \neq 0\ ] \quad \Rightarrow \quad [\ f(p) \in \mathrm{LP}_{\mathbb{R}} Y\ ]$.

Proof. By Remark 31.5, we have $\mathrm{dom}[f'] \subseteq X\backslash(\mathrm{Isol}_{\mathbb{R}} X)$. Then we have $p \in \mathrm{dom}[f'] \subseteq X\backslash(\mathrm{Isol}_{\mathbb{R}} X) \subseteq \mathrm{LP}_{\mathbb{R}} X$, proving (1). It remains to prove that (2) holds. Assume that $f'(p) \neq 0$. We wish to show that $f(p) \in \mathrm{LP}_{\mathbb{R}} Y$. Let $q := f(p)$. We wish to show that $q \in \mathrm{LP}_{\mathbb{R}} Y$.

Since $p \in \mathrm{dom}[f'] \subseteq \mathrm{dom}[f]$, it follows that $f(p) \in \mathrm{im}[f]$. Since $\mathrm{LP}_{\mathbb{R}} Y = [\mathrm{Cl}_{\mathbb{R}} Y]\backslash[\mathrm{Isol}_{\mathbb{R}} Y]$ and since $q = f(p) \in \mathrm{im}[f] = Y \subseteq \mathrm{Cl}_{\mathbb{R}} Y$, we want: $q \notin \mathrm{Isol}_{\mathbb{R}} Y$. Assume: $q \in \mathrm{Isol}_{\mathbb{R}} Y$. We aim for a contradiction.

Since $q \in \mathrm{Isol}_{\mathbb{R}} Y$, choose $W \in \mathcal{N}_{\mathbb{R}}(q)$ such that $W \cap Y = \{q\}$. Since $p \in \mathrm{dom}[f'] \subseteq \mathrm{dct}[f]$, we see that f is continuous at p. So, since $W \in \mathcal{N}_{\mathbb{R}}(q) = \mathcal{N}_{\mathbb{R}}(f(p))$, by definition of continuity, choose $V \in \mathcal{N}_{\mathbb{R}}(p)$ such that $f_*(V) \subseteq W$.

Let $m := f'(p)$ and let $S := \mathbb{R}\backslash\{0\}$. Then $SS_f^p \to m$ near 0, and S is an open subset of $\mathbb{R}$. We have $p \in \mathrm{dom}[f']$ and $m = f'(p) \neq 0$, so $m \in S$. By Remark 16.4 any open set is a neighborhood of each of its points, so, since S is open in $\mathbb{R}$, we get $S \in \mathcal{N}_{\mathbb{R}}(m)$. So, since $SS_f^p \to m$ near 0, choose $U \in \mathcal{N}_{\mathbb{R}}^\times(0)$ such that $(SS_f^p)_*(U) \subseteq S$.

Let $U' := U + p$. Since $U \in \mathcal{N}_{\mathbb{R}}^{\times}(0)$, we get $U' \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$. Let $V' := U' \cap V$. Since $U \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$ and $V \in \mathcal{N}_{\mathbb{R}}(p)$, we get $V' \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$.

By (1) of Lemma 36.10, we have $p \in \mathrm{LP}_{\mathbb{R}}X$. So, since $V' \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$, by (2) of Fact 24.3, we conclude that $V' \cap X \neq \varnothing$. Choose $z \in V' \cap X$. We have $z \in V' \cap X = U' \cap V \cap X$. Since $z \in X = \mathrm{dom}[f]$, we get $f(z) \in \mathrm{im}[f]$. Since $z \in \mathrm{dom}[f]$ and $z \in V$, we get $f(z) \in f_*(V)$. Then $f(z) \in f_*(V) \subseteq W$. Also, $f(z) \in \mathrm{im}[f] = Y$. Then $f(z) \in W \cap Y = \{q\}$. Then $f(z) = q = f(p)$. Since $z \in U' = U + p$, we get $z - p \in U$. Let $h := z - p$. Then $h \in U$. Since $z \in U'$ and $U' \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$, we see that $z \neq p$. Then $h = z - p \neq 0$. We have $p \in \mathrm{dom}[f'] \subseteq \mathrm{dom}[f]$ and $f(p+h) = f(z) = q = f(p)$, and so $[f(p+h)] - [f(p)] = 0$. Therefore, because $h \neq 0$, we conclude $(SS_f^p)(h) = 0$. Then $h \in \mathrm{dom}[SS_f^p]$. Therefore, because $h \in U$, it follows that $(SS_f^p)(h) \in (SS_f^p)_*(U)$.

Then $0 = (SS_f^p)(h) \in (SS_f^p)_*(U) \subseteq S = \mathbb{R}\backslash\{0\}$. Contradiction. $\square$

THEOREM 36.11. *Let $X, Y \subseteq \mathbb{R}$, $f : X \hookrightarrow\!> Y$, $p \in dom[f']$. Let $q := f(p)$, $g := f^{-1}$. Assume $f'(p) = 0$. Then $g'(q) = ☹$.*

Proof. By (1) of Lemma 36.10, we have $p \in \mathrm{LP}_{\mathbb{R}}X$. Therefore, we have $p \in \mathrm{LP}_{\mathbb{R}}X = \mathrm{LPD}_{\mathbb{R}}(g \circ f)$. So, by the Chain Rule (Theorem 32.8), we get $(g \circ f)'(p) =^* [g'(f(p))] \cdot [f'(p)]$. So, since $f(p) = q$ and since $f'(p) = 0$, we have $(g \circ f)'(p) =^* [g'(q)] \cdot [0]$. By Remark 36.9, $(\mathrm{id}_X)'(p) \;^*\!= 1$. Then $1 =^* (\mathrm{id}_X)'(p) = (g \circ f)'(p) =^* [g'(q)] \cdot [0]$, so $1 =^* [g'(q)] \cdot [0]$. Assume $g'(q) \neq ☹$. We aim for a contradiction.

Since $g'(q) \neq ☹$, we get $g'(q) \in \mathbb{R}$, and so $[g'(q)] \cdot [0] = 0$. Then $1 =^* 0 \neq ☹$, and so $1 = 0$. Contradiction. $\square$

THEOREM 36.12. *Let $X, Y \subseteq \mathbb{R}$, $f : X \hookrightarrow\!> Y$, $p \in dom[f']$. Let $q := f(p)$ and let $g := f^{-1}$. Assume that g is continuous at q. Then $g'(q) = 1/[f'(p)]$.*

Proof. Next class. $\square$

We noted that, in Theorem 36.12, if we drop the hypothesis that g is continuous at q, then the result becomes false, as follows: Define $X := [1, 2) \cup [3, 4)$, define $Y := [5, 7)$, define $f : X \to Y$ by

$$f(x) = \begin{cases} x + 4, & \text{if } x \in [1, 2) \\ x + 3, & \text{if } x \in [3, 4), \end{cases}$$

and then let $p := 3$, $q := f(p)$, $g := f^{-1}$. Then $f'(p) = 1$, $q = 6$ and g is NOT continuous at q. Then $g'(q) = ☹ \neq 1/1 = 1/[f'(1)]$.

37. Class 11 on 20 February 2018, Tu of Week 6

Midterm 1 on Thursday 22 February.

Recall the Topological Inverse Function Theorem for intervals (Theorem 36.6): Let $X, Y \subseteq \mathbb{R}$ and let $f : X \hookrightarrow\!\!> Y$ be continuous. Assume that X is an interval. Then

(A) Y is an interval and
(B) $f : X \to Y$ is a homeomorphism.

We did not prove this result, as the proof involves a lot of bookkeeping. However, our next goal is to prove enough of this theorem that students should be able to see how to do the rest if they want to.

NOTE TO SELF: In fact, we could prove: Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ be strictly monotone. Assume that $\mathrm{dom}[f]$ is an interval. Then f^{-1} is continuous. (Interestingly, we don't need to assume that f is continuous. *Proof:* If f^{-1} had a discontinuity, it would have to be a jump discontinuity, but then $\mathrm{im}[f^{-1}]$ would not be an interval, and, as $\mathrm{dom}[f] = \mathrm{im}[f^{-1}]$, this contradicts the assumption that $\mathrm{dom}[f]$ is an interval. QED)

NOTE: What we are calling topological inverse function theorems are, more commonly, called **open mapping theorems**, just in case you might wish to look for some of these kinds of theorems online.

FACT 37.1. *Let $S \subseteq \mathbb{R}$. Then:*

$$[\ S \textit{ is an interval}\ ] \Leftrightarrow [\ (\ \forall a, b \in S,\ [a|b] \subseteq S\ )\ \&\ (\ S \neq \varnothing\ )\ ].$$

Proof. Omitted □

Assigned HW#6-1.

LEMMA 37.2. *Let Y, Z be topological spaces. Let $g, h : Y \dashrightarrow Z$. Let $q \in Y$ and let $Y_0 \in \mathcal{N}_Y(q)$. Assume that h is continuous at q. Assume that $g = h$ on Y_0. Then g is continuous at q.*

Proof. We wish to show: $\forall W \in \mathcal{N}_Z(g(q))$, $\exists U \in \mathcal{N}_Y(q)$ s.t. $g_*(U) \subseteq W$. Let $W \in \mathcal{N}_Z(g(q))$ be given. We want: $\exists U \in \mathcal{N}_Y(q)$ s.t. $g_*(U) \subseteq W$.

Since $q \in Y_0$ and since $g = h$ on Y_0, we conclude: $g(q) = h(q)$. Then $W \in \mathcal{N}_Z(g(q)) = \mathcal{N}_Z(h(q))$. So, since h is continuous at q, choose $V \in \mathcal{N}_Y(q)$ s.t. $h_*(V) \subseteq W$. Since $V, Y_0 \in \mathcal{N}_Y(q)$, it follows that $V \cap Y_0 \in \mathcal{N}_Y(q)$. Let $U := V \cap Y_0$. Then $U \in \mathcal{N}_Y(q)$ and we wish to show: $g_*(U) \subseteq W$. We wish to show: $\forall x \in \mathrm{dom}[g]$,

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ g(x) \in W\ ].$$

Let $x \in \mathrm{dom}[g]$ be given. We wish to show:

$$[\ x \in U\ ] \quad \Rightarrow \quad [\ g(x) \in W\ ].$$

Assume: $x \in U$. We wish to show: $g(x) \in W$.

We have $x \in U = V \cap Y_0$, so both $x \in U$ and $x \in Y_0$. Since $x \in Y_0$ and since $g = h$ on Y_0, we get $g(x) = h(x)$. Since $x \in \mathrm{dom}[g]$, we get $g(x) \neq ☹$. Since $h(x) = g(x) \neq ☹$, we get $x \in \mathrm{dom}[h]$. So, since $x \in V$, we get $h(x) \in h_*(V)$. Then $g(x) = h(x) \in h_*(V) \subseteq W$, as desired. $\square$

Assigned HW#6-2, HW#6-3 and HW#6-4.

Unassigned HW: Prove a variant of HW#6-4, in which "strictly increasing" is replaced by "stricly decreasing".

Recall (Theorem 30.1): Let $f : \mathbb{R} \to \mathbb{R}$. Assume $\mathrm{dom}[f]$ is an interval. Then: $(\ [\ f \text{ is strictly monotone }\ ] \quad \Leftrightarrow \quad [\ f \text{ is 1-1 }\ ]\ )$.

LEMMA 37.3. *Let $X, Y \subseteq \mathbb{R}$, $a \in \mathbb{R}$. Let $f : X \hookrightarrow\!> Y$ be continuous. Assume: $X = [a, \infty)$. Then: f^{-1} is continuous at $f(a)$.*

Proof. If f is strictly increasing, then, by HW#6-4, f^{-1} is continuous at $f(a)$. If f is strictly decreasing, then, by the Unassigned HW above, f^{-1} is continuous at $f(a)$. Then it suffices to show that f is either strictly increasing or strictly decreasing, *i.e.*, that f is strictly monotone.

By Theorem 30.1, f is strictly monotone. $\square$

THEOREM 37.4. *Let $X, Y \subseteq \mathbb{R}$ and let $f : X \hookrightarrow\!> Y$ be continuous. Assume: $\exists a \in \mathbb{R}$ s.t. $X = [a, \infty)$. Then: $f^{-1} : Y \to X$ is continuous.*

Proof. Choose $a \in \mathbb{R}$ s.t. $X = [a, \infty)$. By Lemma 37.3, f^{-1} is continuous at $f(a)$. We wish to show: f^{-1} is continuous on $Y \setminus \{f(a)\}$.

Let $X_0 := \mathrm{Int}_{\mathbb{R}} X$. By HW#6-2, f^{-1} is continuous on $f_*(X_0)$. We wish to show: $f_*(X_0) = Y \setminus \{f(a)\}$.

We have $X_0 = \mathrm{Int}_{\mathbb{R}} X = \mathrm{Int}_{\mathbb{R}}[a, \infty) = (a, \infty)$. Also, as

$$a \ \in \ [a, \infty) \ = \ X \ = \ \mathrm{dom}[f],$$

we conclude that $f_*(\{a\}) = \{f(a)\}$. Since f is 1-1, it follows that $f_*(\,X \setminus \{a\}\,) = [f_*(X)] \setminus [f_*(\{a\})]$. We therefore calculate:

$$\begin{aligned} f_*(X_0) \ &= \ f_*(\,(a, \infty)\,) \ = \ f_*(\,[a, \infty) \setminus \{a\}\,) \ = \ f_*(\,X \setminus \{a\}\,) \\ &= \ [f_*(X)] \setminus [f_*(\{a\})] \ = \ Y \setminus \{f(a)\}, \end{aligned}$$

as desired. $\square$

Assigned HW#6-5.

LEMMA 37.5. *Let $X, Y \subseteq \mathbb{R}$, $a \in \mathbb{R}$. Let $f : X \hookrightarrow\!> Y$ be continuous. Assume: $\exists b \in (a, \infty)$ s.t. $X = [a, b)$. Then: f^{-1} is continuous at $f(a)$.*

Proof. Unassigned HW. *Hint:* Similar to the proof of Lemma 37.3, but with the following outline. First, prove a variant of HW#6-5, in which "strictly increasing" is replaced by "stricly decreasing". Then combine HW#6-5 with this new variant, and use Theorem 30.1. $\square$

THEOREM 37.6. *Let $X, Y \subseteq \mathbb{R}$ and let $f : X \hookrightarrow\!> Y$ be continuous. Assume: $\exists a, b \in \mathbb{R}$ s.t. $a < b$ and $X = [a, b)$. Then $f^{-1} : Y \to X$ is continuous.*

Proof. Choose $a, b \in \mathbb{R}$ s.t. $a < b$ and $X = [a, b)$. By Lemma 37.5, f^{-1} is continuous at $f(a)$. We want: f^{-1} is continuous on $Y \setminus \{f(a)\}$.

Let $X_0 := \mathrm{Int}_{\mathbb{R}} X$. By HW#6-2, f^{-1} is continuous on $f_*(X_0)$. We wish to show: $f_*(X_0) = Y \setminus \{f(a)\}$.

We have $X_0 = \mathrm{Int}_{\mathbb{R}} X = \mathrm{Int}_{\mathbb{R}}[a, b) = (a, b)$. Also, as

$$a \quad \in \quad [a, b) \quad = \quad X \quad = \quad \mathrm{dom}[f],$$

we conclude that $f_*(\{a\}) = \{f(a)\}$. Since f is 1-1, it follows that $f_*(\,X \setminus \{a\}\,) = [f_*(X)] \setminus [f_*(\{a\})]$. We therefore calculate:

$$\begin{aligned} f_*(X_0) \quad &= \quad f_*(\,(a, b)\,) \quad = \quad f_*(\,[a, b) \setminus \{a\}\,) \quad = \quad f_*(\,X \setminus \{a\}\,) \\ &= \quad [f_*(X)] \setminus [f_*(\{a\})] \quad = \quad Y \setminus \{f(a)\}, \end{aligned}$$

as desired. $\square$

We can now indicate, in more detail, how to prove Theorem 36.6:

THEOREM 37.7. *Let $X, Y \subseteq \mathbb{R}$ and let $f : X \hookrightarrow\!> Y$ be continuous. Assume X is an interval. Then*

(A) *Y is an interval* *and*
(B) *$f : X \to Y$ is a homeomorphism.*

Proof. Since $X = \mathrm{dom}[f]$, $f_*(X) = \mathrm{im}[f] = Y$. So, by HW#6-1, Y is an interval. Want: $f : X \to Y$ is a homeomorphism. Since f is, by assumption, continuous, we need only show: f^{-1} is continuous.

Since X is an interval, at least one of the following must be true:

(1) X is compact or
(2) X is open in $\mathbb{R}$ or
(3) $\exists a \in \mathbb{R}$ s.t. $X = [a, \infty)$ or
(4) $\exists a, b \in \mathbb{R}$ s.t. $a < b$ and $X = [a, b)$ or
(5) $\exists b \in \mathbb{R}$ s.t. $X = (-\infty, b]$ or
(6) $\exists a, b \in \mathbb{R}$ s.t. $a < b$ and $X = (a, b]$.

Case (1): By Theorem 36.7, f^{-1} is continuous. *End of Case (1).*
Case (2): By Theorem 36.5, f^{-1} is continuous. *End of Case (2).*
Case (3): By Theorem 37.4, f^{-1} is continuous. *End of Case (3).*
Case (4): By Theorem 37.6, f^{-1} is continuous. *End of Case (4).*
Case (5): Unassigned HW. *End of Case (5).*
Case (6): Unassigned HW. *End of Case (6).* $\square$

Recall Theorem 36.11: Let $X, Y \subseteq \mathbb{R}$, $f : X \hookrightarrow\!> Y$, $p \in \mathrm{dom}[f']$. Let $q := f(p)$, $g := f^{-1}$. Assume: $f'(p) = 0$. Then: $g'(q) = ☹$.

Recall Lemma 36.10: Let $X, Y \subseteq \mathbb{R}$, $f : X \hookrightarrow\!> Y$, $p \in \mathrm{dom}[f']$. Then: $(\ p \in \mathrm{LP}_{\mathbb{R}} X\ )$ and $(\ [\ f'(p) \neq 0\ ] \quad \Rightarrow \quad [\ f(p) \in \mathrm{LP}_{\mathbb{R}} Y\ ]\ )$.

We can now prove:

THEOREM 37.8. *Let $X, Y \subseteq \mathbb{R}$, $f : X \hookrightarrow\!> Y$, $p \in \mathrm{dom}[f']$. Let $q := f(p)$ and let $g := f^{-1}$. Assume that g is continuous at q. Then $g'(q) = 1/[f'(p)]$.*

Proof. By Theorem 36.11, if $f'(p) = 0$, then $g'(q) = ☹$. Let $m := f'(p)$. We wish to show: $g'(q) = 1/m$. We know:

$$[\ m = 0\ ] \quad \Rightarrow \quad [\ g'(q) = ☹ = 1/0 = 1/m\ ].$$

It therefore suffices to show:

$$[\ m \neq 0\ ] \quad \Rightarrow \quad [\ g'(q) = 1/m\ ].$$

Assume $m \neq 0$. We wish to show: $g'(q) = 1/m$.

Since $m = f'(p)$, by Lemma 25.18, we get: $SS_f^p \to m$ near 0. By (2) of Lemma 36.10, we conclude that $f(p) \in \mathrm{LP}_{\mathbb{R}} Y$. It follows that $q = f(p) \in \mathrm{LP}_{\mathbb{R}} Y = \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[g]) = \mathrm{LPD}_{\mathbb{R}} g$. So, by Remark 25.19, it suffices to show: $SS_g^q \to 1/m$ near 0.

Let $\phi := SS_f^p$ and $\psi := SS_g^q$. Then $\phi \to m$ near 0, and we want: $\psi \to 1/m$ near 0. We wish to show: $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t., $\forall k \in \mathrm{dom}[\psi]$,

$$[\ 0 < |k| < \delta\ ] \quad \Rightarrow \quad [\ |\,[\psi(k)] - [1/m]\,| < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \delta > 0$ s.t., $\forall k \in \mathrm{dom}[\psi]$,

$$[\ 0 < |k| < \delta\ ] \quad \Rightarrow \quad [\ |\,[\psi(k)] - [1/m]\,| < \varepsilon\ ].$$

Since $p \in \mathrm{dom}[f']$ and $f' : \mathbb{R} \dashrightarrow \mathbb{R}$, it follows that $f'(p) \in \mathbb{R}$. That is, $m \in \mathbb{R}$. So, as $m \neq 0$, we get $1/m \neq ☹$. Since $\phi \to m$ near 0, by Lemma 20.5, we get: $1/\phi \to^* 1/m$ near 0. So, since $1/m \neq ☹$, we get: $1/\phi \to 1/m$ near 0. So choose $\gamma > 0$ s.t., $\forall h \in \mathrm{dom}[1/\phi]$,

$$[\ 0 < |h| < \gamma\ ] \quad \Rightarrow \quad [\ |\,[(1/\phi)(h)] - [1/m]\,| < \varepsilon\ ].$$

Since g is continuous at q, choose $\delta > 0$ s.t., $\forall z \in \mathrm{dom}[q]$,

$$[\ |\,z - q\,| < \delta\ ] \quad \Rightarrow \quad [\ |\,[g(z)] - [g(q)]\,| < \gamma\ ].$$

We wish to show: $\forall k \in \mathrm{dom}[\psi]$,

$$[\ 0 < |k| < \delta\ ] \quad \Rightarrow \quad [\ |\,[\psi(k)] - [1/m]\,| < \varepsilon\ ].$$

Let $k \in \mathrm{dom}[\psi]$ be given. We wish to show:

$$[\ 0 < |k| < \delta\ ] \quad \Rightarrow \quad [\ |\,[\psi(k)] - [1/m]\,| < \varepsilon\ ].$$

Assume $0 < |k| < \delta$ We wish to show: $|\,[\psi(k)] - [1/m]\,| < \varepsilon$.

By assumption, g is continuous at q. Then $q \in \mathrm{dct}[g] \subseteq \mathrm{dom}[g]$. Then $g(q) \in \mathrm{im}[g] = X \subseteq \mathbb{R}$. We have

$$\psi(k) \quad = \quad (SS_g^q)(k) \quad = \quad \frac{[\,g(q+k)\,] - [\,g(q)\,]}{k}.$$

So, since $k \in \mathrm{dom}[\psi]$, $\psi(k) \neq ☹$. Then $g(q+k) \neq ☹$, so $q+k \in \mathrm{dom}[g]$. Then $g(q+k) \in \mathrm{im}[g] = X \subseteq \mathbb{R}$. Let $h := [g(q+k)] - [g(q)]$. Then $\psi(k) = h/k$. Since $g(q+k) \in \mathbb{R}$ and $g(q) \in \mathbb{R}$, we get: $h \in \mathbb{R}$.

By assumption, $f : X \hookrightarrow\!> Y$ and $g = f^{-1}$. Then $g : Y \hookrightarrow\!> X$. We have $0 < |k|$, so $k \neq 0$, so $q + k \neq q$. So, as g is 1-1, we see that $g(q+k) \neq g(q)$. Then $h = [g(q+k)] - [g(q)] \neq 0$. Then $1/(k/h) = h/k$.

By assumption, $q = f(p)$. So, since $g = f^{-1}$, we get $f(q) = p$. Since $g(q) \in \mathbb{R}$, we get $[g(q)] - [g(q)] = 0$. Then, adding $p = g(q)$ to $h = [g(q+k)] - [g(q)]$ gives $p + h = g(q+k)$. So, since $g = f^{-1}$, we get $f(p+h) = q + k$. Since $q \in \mathbb{R}$, we get $q - q = 0$. Subtracting $f(p) = q$ from $f(p+h) = q + k$ gives $[f(p+h)] - [f(p)] = k$. Then

$$\phi(h) \quad = \quad (SS_f^p)(h) \quad = \quad \frac{[\,f(p+h)\,] - [\,f(p)\,]}{h} \quad = \quad \frac{k}{h}.$$

Then $(1/\phi)(h) = 1/(k/h) = h/k = \psi(k)$. So, since $\psi(k) \neq ☹$, we get $(1/\phi)(h) \neq ☹$. It follows that $h \in \mathrm{dom}[1/\phi]$. Let $z := q + k$. Then $z \in \mathrm{dom}[g]$ and $|z - q| = |k| < \delta$, so, by choice of δ, we get

$|\,[g(z)] - [g(q)]\,| < \gamma$. So, since $[g(z)]-[g(q] = [g(q+k)]-[g(q)] = h$, we conclude that $|h| < \gamma$. So, since $h \neq 0$, we get $0 < |h| < \gamma$. So, since $h \in \mathrm{dom}[1/\phi]$, by choice of γ, we get $|\,[(1/\phi)(h)] - [1/m]\,| < \varepsilon$. So, since $(1/\phi)(h) = \psi(k)$, we get $|\,[\psi(k)] - [1/m]\,| < \varepsilon$, as desired. ☐

Theorem 37.8 will be called the **Differential Inverse Function Theorem** or **DIFT**. In Theorem 37.8, the hypothesis that g is continuous at q will be called the **unfortunate DIFT hyptothesis**. It means that, to use the DIFT on f, one has to verify an analytic condition on the inverse of f, something that may be very hard to do.

Fortunately, we have three *Topological* Inverse Function Theorems (Theorem 36.7 and Theorem 36.5 and Theorem 36.6). In Theorem 38.8, we will combine them with our DIFT to get a version of the DIFT that avoids any unfortunate hypotheses.

38. Class 12 on 27 February 2018, Tu of Week 7

BEGIN remarks about the exam

REMARK 38.1. *Let X be a topological space. Then*

$$\begin{aligned} \mathrm{Isol}_X X &= \{w \in X \,|\, \{w\} \textit{ is open in } X\} \qquad \textit{and} \\ \mathrm{LP}_X X &= X \backslash [\mathrm{Isol}_X X]. \end{aligned}$$

Proof. Omitted. ☐

REMARK 38.2. *Let X and Y be topological spaces. Let $w \in X$ and $z \in Y$. Assume that $w \in \mathrm{LP}_X X$. Then $\lim\limits_w C_X^z = z$.*

Proof. Omitted. ☐

In Remark 38.2, the hypothesis that $w \in \mathrm{LP}_X X$ is necessary: Let $X := [1,2] \cup \{3\}$, $Y := \mathbb{R}$, $w := 3$, $z := 4$. Since $3 \notin \mathrm{LP}_X X$, by Proposition 24.7, we see that $\mathop{\mathrm{LIMS}}\limits_3 C_X^4 = Y$. Then

$$\mathop{\mathrm{LIMS}}_w C_X^z \quad = \quad \mathop{\mathrm{LIMS}}_3 C_X^4 \quad = \quad Y \quad = \quad \mathbb{R}.$$

Since $\#\mathbb{R} = \infty$, we get $\mathrm{ELT}(\mathbb{R}) = ☹$ Then

$$\lim_w C_X^z \;=\; \mathrm{ELT}(\mathop{\mathrm{LIMS}}_w C_X^z) \;=\; \mathrm{ELT}(\mathbb{R}) \;=\; ☹ \;\neq\; 4 \;=\; z.$$

The topological space $\mathbb{R}$ has no "open points", *i.e.*, $\mathrm{LP}_\mathbb{R}\mathbb{R} = \mathbb{R}$. So, for $X = \mathbb{R}$, the problem described in the last paragraph goes away:

COROLLARY 38.3. *Let Y be a topological space. Let $w \in \mathbb{R}$ and $z \in Y$. Then $\lim_{w} C_{\mathbb{R}}^{z} = z$.*

Proof. Let $X := \mathbb{R}$. Then $w \in \mathbb{R} = \mathrm{LP}_{\mathbb{R}}\mathbb{R}$. So, by Remark 38.2, $\lim_{w} C_X^z = z$. That is, $\lim_{w} C_{\mathbb{R}}^{z} = z$, as desired. □

Recall: $\forall$sets X, Y and Z, $\forall f : X \dashrightarrow Y$, $\forall g : X \dashrightarrow Z$, the function $(f, g) : X \dashrightarrow Y \times Z$ is defined by $(f, g)(x) = (f(x), g(x))$.

I have two warnings about the formula $(f, g)'(x) = (f'(x), g'(x))$:

First, define $f : (-\infty, 0] \to \mathbb{R}$ and $g : [0, \infty) \to \mathbb{R}$ by

$$f(x) \;=\; 3x \qquad \text{and} \qquad g(x) \;=\; 4x.$$

Then $\mathrm{dom}[(f, g)] = \{0\}$, so $(f, g)'(0) = ☹$. Also $(f'(0), g'(0)) = (3, 4)$. Then $(f, g)'(0) \neq (f'(0), g'(0))$.

Second, define $f : \mathbb{R} \to \mathbb{R}$ and $g : [0, \infty) \to \mathbb{R}$ by

$$f(x) \;=\; |x| \qquad \text{and} \qquad g(x) \;=\; 4x.$$

Then $\mathrm{dom}[(f, g)] = [0, \infty)$. Also, $\forall x \in [0, \infty)$,

$$(f, g)(x) \quad = \quad (|x|, 4x) \quad = \quad (x, 4x).$$

Then $(f, g)'(0) = (1, 4)$. Also $f'(0) = ☹$, so $(f'(0), g'(0)) = ☹$. Then $(f, g)'(0) \neq (f'(0), g'(0))$.

THEOREM 38.4. *Let both V and W be normed vector spaces. Let $f : \mathbb{R} \dashrightarrow V$ and $g : \mathbb{R} \dashrightarrow W$. Let $p \in \mathrm{LPD}_{\mathbb{R}}(f, g)$. Then*

$$(f, g)'(p) \quad =^* \quad (\, f'(p) \, , \, g'(p) \,).$$

Proof. Omitted. □

END remarks about the exam

BEGIN Via homework, you prove the second derivative test

DEFINITION 38.5. *Let S be a set, $f : S \dashrightarrow \mathbb{R}$, $p \in S$. Then*

(1) *f has a* **unique maximum** *at p means: $f_*(S_p^\times) < f(p)$ and*
(2) *f has a* **unique minimum** *at p means: $f_*(S_p^\times) > f(p)$.*

Assigned HW#7-1.

DEFINITION 38.6. *Let X be a topological space and let f be a function. Assume $dom[f] \subseteq X$. Then $\mathrm{IntD}_X f := \mathrm{Int}_X(dom[f])$.*

Assigned HW#7-2, HW#7-3 and HW#7-4.

FACT 38.7. *There exists a continuous $f : \mathbb{R} \to \mathbb{R}$ such that*

(1) $f'(0) = 0$,
(2) $f''(0) = -2$,
(3) *NOT (f has a local extremum at* 0 *),*
(4) $0 \notin \mathrm{IntD}_{\mathbb{R}}(f')$ *and*
(5) $0 \in \mathrm{Int}_{\mathbb{R}}(dct[f])$.

Proof. Choose a continuous $g : [0, \infty) \to [0, 1]$ s.t. $\mathrm{dom}[g'] = \varnothing$. (One says that g is "nowhere differentiable" to indicate that g' is the empty function. We will not take the time to prove that such a continuous function g exists, so this is only a sketch of a proof. See me if you want more detail.) Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) \quad = \quad \begin{cases} -x^2, & \text{if } x < 0, \\ x^2 + x^2 \cdot [g(x)], & \text{if } x \geqslant 0. \end{cases}$$

Then, for all $x \in (-\infty, 0)$, $f'(x) = -2x$. Then f is differentiable on $(-\infty, 0)$. Also, f is continuous on $(0, \infty)$. Also, for all $x \in (0, \infty)$, f is NOT differentiable at x. Also,

(A) $f < 0$ on $(-\infty, 0)$,
(B) $f(0) = 0$ and
(C) $f > 0$ on $(0, \infty)$.

Since f is differentiable on $(-\infty, 0]$, it follows that f is continuous on $(-\infty, 0]$. Recall tht f is continuous on $(0, \infty)$. Then f is continuous on $(-\infty, 0] \cup (0, \infty)$. So, since $(-\infty, 0] \cup (0, \infty) = \mathbb{R} = \mathrm{dom}[f]$, we see that f is continuous.

For all $x \in \mathbb{R}$, we have $-x^2 \leqslant f(x) \leqslant 2x^2$. So, by HW#3-3, $f'(0) = 0$, proving (1). Then $\mathrm{dom}[f'] = (-\infty, 0]$ and, for all $x \in (-\infty, 0]$, we have $f'(x) = -2x$. Then $f''(0) = -2$, proving (2).

By (A), (B) and (C), we see that (3) holds. Moreover, because we have $\mathrm{IntD}_{\mathbb{R}}(f') = \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f']) = \mathrm{Int}_{\mathbb{R}}((-\infty, 0]) = (-\infty, 0)$, we see that (4) holds. Since f is continuous, $\mathrm{dct}[f] = \mathrm{dom}[f]$. Then $\mathrm{Int}_{\mathbb{R}}(\mathrm{dct}[f]) = \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[f]) = \mathrm{Int}_{\mathbb{R}}\mathbb{R} = \mathbb{R}$, so (5) holds. $\square$

Assigned HW#7-5.

Unassigned HW: Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$. Asume both that $0 \in \mathrm{IntD}_{\mathbb{R}} f$ and that $f(0) = 0$. Define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = x^3$. Assume that $f/g \to 1$ near 0. Show: NOT(f has a local extremum at 0).

By HW#7-5, we can use asymptotics to find local extrema. Later, we'll show how to combine asymptotics with L'Hospital's Rule to get another proof of the Second Derivative Test. We'll also show that that approach leads to a third derivative test, a fourth derivative test, *etc.*

END Via homework, you prove the second derivative test

We now pick where we left off at the end of Class 12.

Recall: Let $X, Y \subseteq \mathbb{R}$, $f : X \hookrightarrow\!> Y$, $p \in \mathrm{dom}[f']$. Let $q := f(p)$, $g := f^{-1}$. Assume that g is continuous at q. Then $g'(q) = 1/[f'(p)]$. (This is Theorem 37.8.)

Recall: Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$, $S \subseteq \mathbb{R}$. Then f is **c/d** on S means

(1) f is continuous on S and
(2) f is differentiable on S.

Recall: Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$. Then f is **c/d** means: f is c/d on $\mathrm{dom}[f]$. (That is, f is continuous on $\mathrm{dom}[f]$ and differentiable on $\mathrm{IntD}_{\mathbb{R}} f$.)

The next result will be called **DIFT redux**. (Recall that DIFT stands for "Differential Inverse Function Theorem".)

THEOREM 38.8. *Let $X, Y \subseteq \mathbb{R}$. Let $f : X \hookrightarrow Y$ be c/d. Assume that at least one of the following holds:*

(1) X is compact or
(2) X is open in $\mathbb{R}$ or
(3) X is an interval.

Let $g := f^{-1}$, $X_0 := \mathrm{Int}_{\mathbb{R}} X$, $Y_0 := \mathrm{Int}_{\mathbb{R}} Y$. Let $q \in Y_0$. Then both

(1) $g(q) \in X_0$ and
(2) $g'(q) = 1/[f'(g(q))]$.

Proof. We have $q \in Y_0 \subseteq Y = \mathrm{dom}[g]$. Then $g(q) \in g_*(Y_0)$.

Since $Y_0 = \mathrm{Int}_{\mathbb{R}} Y$, we conclude that Y_0 is open in $\mathbb{R}$. By our three Topological Inverse Function Theorems (Theorem 36.7 and Theorem 36.5 and Theorem 36.6), we see that $f : X \to Y$ is a homeomorphism, so $g : Y \to X$ is continuous. So, since $g : Y \to X$ is also 1-1, by Theorem 30.7, $g_*(Y_0) \subseteq X_0$. Then $g(q) \in g_*(Y_0) \subseteq X_0$, proving (1). It remains to prove (2). Let $p := g(q)$. We want: $g'(q) = 1/[f'(p)]$.

Since $q \in \mathrm{dom}[g]$ and since g is continuous, it follows that g is continuous at q. Then, by Theorem 37.8, $g'(q) = 1/[f'(p)]$, as desired. $\square$

DEFINITION 38.9. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$.
Then f *is a* **(c/d)-diffeomorphism** *means:*

(1) f *is 1-1,*
(2) f *is c/d* *and*
(3) f^{-1} *is c/d.*

Let $X, Y \subseteq \mathbb{R}$. Let $f : X \hookrightarrow\!> Y$ be a (c/d)-diffeomorphism. Let $g := f^{-1}$, $X_0 := \mathrm{Int}_{\mathbb{R}} X$, $Y_0 := \mathrm{Int}_{\mathbb{R}} Y$. Then all of the following hold:

(1) $f_*(X_0) = Y_0$,
(2) $g_*(Y_0) = X_0$,
(3) $\forall p \in X_0, \quad 1 = (g \circ f)'(p) = [g'(f(p))][f'(p)]$,
(4) $\forall q \in Y_0, \quad 1 = (f \circ g)'(q) = [f'(g(q))][g'(q)]$,
(3) $\forall p \in X_0, \quad f'(p) = 1/[g'(f(p))]$ and
(3) $\forall q \in Y_0, \quad g'(q) = 1/[f'(g(q))]$.

THEOREM 38.10. *Let* X *be an interval. Let* $f : X \to \mathbb{R}$ *be c/d. Assume:* $\forall p \in \mathrm{Int}_{\mathbb{R}} X$, $f'(p) \neq 0$. *Then* f *is a (c/d)-diffeomorphism.*

Proof. By (1) of Corollary 34.10, f is 1-1. Let $g := f^{-1}$. We wish to show that g is c/d.

Let $Y := \mathrm{im}[f]$. Then $Y = \mathrm{dom}[g]$. Let $X_0 := \mathrm{Int}_{\mathbb{R}} X$, $Y_0 := \mathrm{Int}_{\mathbb{R}} Y$. By Theorem 36.6, $f : X \to Y$ is a homeomorphism. Then $g : Y \to X$ is continuous. It remains to show: g is differentiable on Y_0. We wish to show: $\forall q \in Y_0$, g is differentiable at q. Let $q \in Y_0$ be given. We wish to show: g is differentiable at q.

Since $g : Y \to X$ is 1-1 and continuous, by Theorem 30.7, we get $g_*(Y_0) \subseteq X_0$. Since $q \in Y_0 \subseteq Y = \mathrm{dom}[g]$, we get $g(q) \in g_*(Y_0)$. Let $p := g(q)$. Since f is c/d, we get $X_0 \subseteq \mathrm{dom}[f']$. Then

$$p \quad = \quad g(q) \quad \in \quad g_*(Y_0) \quad \subseteq \quad X_0 \quad \subseteq \quad \mathrm{dom}[f'].$$

Since $q \in \mathrm{dom}[g]$ and g is continuous, we see that g is continuous at q. Then, by Theorem 37.8, $g'(q) = 1/[f'(p)]$. We have $p \in X_0 = \mathrm{Int}_{\mathbb{R}} X$, so, by assumption, we get $f'(p) \neq 0$. So, since $p \in \mathrm{dom}[f']$, we conclude that $f'(p) \in \mathbb{R}_0^\times$. Then $1/[f'(p)] \neq ☹$. Since $g'(q) = 1/[f'(p)] \neq ☹$, we conclude that $q \in \mathrm{dom}[g']$, as desired. ☐

39. Class 13 on 1 March 2018, Th of Week 7

Recall: $\forall k \in \mathbb{N}$, $[1..k]$ denotes the set $[1, k] \cap \mathbb{Z} = \{1, \ldots, k\}$ of all integers j such that $1 \leqslant j \leqslant k$.

Recall: Let S be a set and let $k \in \mathbb{N}$. Then $S^k := S^{[1..k]}$. That is, S^k is the set of all functions $[1..k] \to S$. For any $z \in S^k$, we have $z : [1..k] \to S$. For any $z \in S^k$, for any $j \in [1..k]$, we typically denote $z(j)$ by z_j and call z_j the "jth component of z".

Recall: Let S be a set and let $k \in \mathbb{N}$. For all $j \in [1..k]$, let $x_j \in S$. Then $(x_1, \ldots, x_k) \in S^k$ is defined by $(x_1, \ldots, x_k)_j = x_j$. For example,

$$(3, 8, 7, -5) \quad = \quad \left(\begin{array}{c} 1 \mapsto 3 \\ 2 \mapsto 8 \\ 3 \mapsto 7 \\ 4 \mapsto -5 \end{array} \right) \quad \in \quad \mathbb{R}^{[1..4]} \quad = \quad \mathbb{R}^4.$$

DEFINITION 39.1. *Let S be a set. Then, for all z, by z is an* **arrow** *in S, we mean: $z \in S^2$.*

Let S be a set and let $z \in S^2$. Then $z_1, z_2 \in S$. We call z_1 the **footpoint** of z, and we call z_2 the **endpoint** of z. We picture z as an arrow that goes from z_1 to z_2.

DEFINITION 39.2. *$\forall$vector space W, $\forall z \in W^2$,* $\quad \operatorname{vec} z := z_2 - z_1$.

That is, for any arrow z in a vector space, the **vector** of z is

(the endpoint of z) minus (the footpoint of z),

and is denoted by $\operatorname{vec} z$.

DEFINITION 39.3. *Let W be a vector space and let $y, z \in W^2$. Then $y|||z$ in W means:* $\operatorname{vec} y = \operatorname{vec} z$.

We indicate $y|||z$ by saying "y is a **translate** of z" or "z is a translate of y" or "y and z are translates (of each other)".

For example, let $z := (\,(4,6)\,,\,(8,5)\,) \in (\mathbb{R}^2)^2$. Then z is an arrow in $\mathbb{R}^2$. The footpoint of z is $z_1 = (4,6)$ and the endpoint of z is $z_2 = (8,5)$. We sometimes say that z is "footed" at $(4,6)$ and that z "ends" at $(8,5)$. We picture z as an arrow in a coordiate plane that runs from the point $(4,6)$ to the point $(8,5)$. The vector of z is $\operatorname{vec} z = (8,5) - (4,6) = (4,-1)$. Let $v := \operatorname{vec} z$, so $v = (4,-1)$. In our picture, the arrow z should be labeled "z" and not "v". However, the vector of an arrow is often thought of as its most important data, and so people sometimes do label an arrow with its vector. Note that z runs $v_1 = 4$ units. Note that z rises $v_2 = -1$ units (*i.e.*, falls 1 unit). Note that $v_1 = (z_2 - z_1)_1 = (z_2)_1 - (z_1)_1 = (8,5)_1 - (4,6)_1 = 8 - 4 = 4$. Note that $v_2 = (z_2 - z_1)_2 = (z_2)_2 - (z_1)_2 = (8,5)_2 - (4,6)_2 = 5 - 6 = -1$.

We searched for an arrow $y \in \mathbb{R}^2$ footed at the origin $(0,0)$ s.t. $y|||z$. We found that there is only one solution: $y = (\,(0,0)\,,\,(4,-1)\,)$.

DEFINITION 39.4. *Let W be a vector space and let $f : \mathbb{R} \dashrightarrow W$. Then $SA_f : \mathbb{R}^2 \dashrightarrow W^2$ and $SV_f : \mathbb{R}^2 \dashrightarrow W$ are defined by*

$$\begin{aligned}(SA_f)(s,t) &= (\,f(s)\,,\,f(t)\,) \qquad \textit{and}\\ (SV_f)(s,t) &= [\,f(t)\,] - [\,f(s)\,].\end{aligned}$$

Note that $(SV_f)(s,t)$ is the vector of the arrow $(SA_f)(s,t)$. The notations SA and SV stand for "secant arrow" and "secant vector".

We visualized a path of a particle $f : \mathbb{R} \dashrightarrow \mathbb{R}^2$ in a coordinate plane. We marked one point as $f(s)$ and another as $f(t)$. We drew the secant arrow between them. We labeled it as $(SA_f)(s,t)$. We commented that it might sometimes be labeled as $(SV_f)(s,t)$, or, equivalently, as $[f(t)] - [f(s)]$. We noted that $(SV_f)(s,t) = (f_s^T)(t-s)$ and that $(DQ_f)(s,t) = [(SV_f)(s,t)]/[t-s]$. Thus, we can connect secant vectors with double translates and difference quotients.

DEFINITION 39.5. *Let W be a normed vector space, $f : \mathbb{R} \dashrightarrow W$. Then $TA_f : \mathbb{R} \dashrightarrow W^2$ and $TV_f : \mathbb{R} \dashrightarrow W$ are defined by*

$$\begin{aligned}(TA_f)(t) &= (\,f(t)\,,\,[f(t)] + [f'(t)]\,) \qquad \textit{and}\\ (TV_f)(t) &= f'(t).\end{aligned}$$

Note that $(TV_f)(t)$ is the vector of the arrow $(TA_f)(t)$. The notations TA and TV stand for "tangent arrow" and "tangent vector".

We visualized a path of a particle $f : \mathbb{R} \dashrightarrow \mathbb{R}^2$ in a coordinate plane. We marked a point as $f(t)$. We drew the velocity arrow at that point. We labeled it as $(TA_f)(t)$. We commented that it might sometimes be labeled as $(TV_f)(s,t)$, or, equivalently, as $f'(t)$. We noted that $TV_f = f'$. The notation TV_f is redundant; generally, f' is preferred.

We posed the Mean Value Question: Let $f : [0,1] \to \mathbb{R}^2$ be c/d. Does there necessarily exist $c \in (a,b)$ s.t. $f'(c) = (DQ_f)(0,1)$? Note that $f'(c) = (TV_f)(c)$ and that

$$(DQ_f)(0,1) \quad = \quad \frac{(SV_f)(0,1)}{1-0} \quad = \quad (SV_f)(0,1).$$

So, an equivalent question is: does there necessarily exist $c \in (a,b)$ s.t. $(TV_f)(c) = (SV_f)(0,1)$? Since $(TV_f)(c) = \mathrm{vec}((TA_f)(c))$ and since $(SV_f)(0,1) = \mathrm{vec}((SA_f)(0,1)$, we form yet another equivalent question:

does there necessarily exist $c \in (a,b)$ s.t. $(TA_f)(c)|||(SA_f)(0,1)$? That is, is there a tangent arrow that is a translate of the secant arrow?

We answered this question in the negative, showing the path of a partcle in the plane such that the only point where the two arrows could possibly be translates didn't work because one of the arrows ran in the opposite direction of the other.

So, if we want to generalize the Mean Value Theorem from $\mathbb{R} \dashrightarrow \mathbb{R}$ to $\mathbb{R} \dashrightarrow \mathbb{R}^2$, then we will need to discuss what it means for two arrows to be parallel, a weaker condation than being translates. We begin by defining parallel for vectors, and will define parallel for arrows later.

DEFINITION 39.6. *Let W be a vector space and let $u, v \in W$. Then $u||v$ in W means: $\{u, v\}$ is linearly dependent in W.*

We indicate $u||v$ by saying "u is a **parallel** to v" or "v is parallel to u" or "u and v are parallel (to each other)".

For example, $(3,4)||(6,8)$ in $\mathbb{R}^2$. Also, $(2,3)||(0,0)||(1,4)$ in $\mathbb{R}^2$, but we have: $\quad$ NOT $(\,(2,3)\,||\,(1,4)$ in $\mathbb{R}^2\,)$.

Let W be a vector space. Then, $\forall u, v \in \mathbb{R}$, we have

$$[\ u||v\ ] \qquad \text{iff} \qquad [\ (\ u \in \mathbb{R}v\ ) \text{ or } (\ v \in \mathbb{R}u\ )\ ].$$

Unassigned homework: $\forall u, v, v' \in W$, if $\mathbb{R}v = \mathbb{R}v'$, then

$$[\ u||v\ ] \qquad \text{iff} \qquad [\ u||v'\ ].$$

We extend the definition of parallel to arrows in a vector space W:

DEFINITION 39.7. *Let W be a vector space, and let $y, z \in W^2$. Then $y||z$ means: $(\operatorname{vec} y)\,||\,(\operatorname{vec} z)$.*

We now begin to look for simple tests that tell us whether two vectors in a vector space are parallel. We start with the **slope function** on vectors in $\mathbb{R}^2$.

DEFINITION 39.8. *Define $\operatorname{sl} : \mathbb{R}^2 \dashrightarrow \mathbb{R}$ by $\operatorname{sl} v = v_2/v_1$.*

Here, "sl" stands for "slope". Note that $\operatorname{sl}(0,1) = ☹$, and that $\operatorname{dom}[\operatorname{sl}] = \mathbb{R}_0^\times \times \mathbb{R}$. We extend the definition of slope to arrows in $\mathbb{R}^2$:

DEFINITION 39.9. *Define $\operatorname{asl} : (\mathbb{R}^2)^2 \dashrightarrow \mathbb{R}$ by $\operatorname{asl} z = \operatorname{sl}(\operatorname{vec} z)$.*

Here, "asl" stands for "arrow slope".

While Definition 39.7 and Definition 39.9 are natural, and are often used in conversation (or intuition), it typically turns out that, in formal proofs, arrows are often not of great importance. When we convert a conversation or an intuitive "picture of a proof" into the formal proof, the arrow is usually forgotten and only its vector is needed.

Recall: For all $k \in \mathbb{N}$, the zero vector in the vector space $\mathbb{R}^k$ is denoted by $0_k = 0_{\mathbb{R}^k} = (0, \ldots, 0) \in \mathbb{R}^k$.

The following is the slope test for parallel arrows in $\mathbb{R}^2$:

FACT 39.10. *Let $u, v \in \mathbb{R}^2$. Then*

$$[\ u||v\ ] \quad \Leftrightarrow \quad [\ (\ u = 0_2\ )\ or\ (\ u = 0_2\ )\ or\ (\ \mathrm{sl}\, u = \mathrm{sl}\, v\ )\ ].$$

Proof. Unassigned HW. ☐

This test is simple to implement, if you are given u and v, but, to use it in a proof often requires breaking the proof into three cases. A better test comes from determinant, which we now describe:

First, note that $(3, 4)||(6, 8)$ in $\mathbb{R}^2$. The numbers $3, 4, 6, 8$ can be put into an array

$$\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix},$$

and, if you know about determinants, then you'll know that the determinant of this array is $3 \cdot 8 - 6 \cdot 4 = 0$. The idea is that, in general, by checking a dterminant, we can tell whether two vectors are parallel. So we need to define matrices and to talk about the simplest kinds of determinants, and then we can develop our determinant test for parallel vectors.

DEFINITION 39.11. *Let S be a set and let $k, \ell \in \mathbb{N}$. Then $S^{k\times\ell} := S^{[1..k]\times[1..\ell]}$.*

That is, $S^{k\times\ell}$ is the set of all functions $[1..k] \times [1..\ell] \to S$. For any $M \in S^{k\times\ell}$, we have $M : [1..k] \times [1..\ell] \to S$. For any $M \in S^k$, for any $i \in [1..k]$, for any $j \in [1..\ell]$, we typically denote $M(i, j)$ by M_{ij} or $M_{i,j}$, and call M_{ij} the "i, j entry of z".

DEFINITION 39.12. *Let S be a set and let $k, \ell \in \mathbb{N}$. For all $i \in [1..k]$, for all $j \in [1..\ell]$, let $x_{ij} \in S$. Then*

$$\begin{bmatrix} x_{11} & \cdots & x_{1\ell} \\ \vdots & & \vdots \\ x_{k1} & \cdots & x_{k\ell} \end{bmatrix} \quad \in \quad S^{k\times\ell}$$

is defined by

$$\begin{bmatrix} x_{11} & \cdots & x_{1\ell} \\ \vdots & & \vdots \\ x_{k1} & \cdots & x_{k\ell} \end{bmatrix}_{ij} \quad = \quad x_{ij}.$$

For example,

$$\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix} \quad = \quad \begin{pmatrix} (1,1) \mapsto 3 \\ (1,2) \mapsto 4 \\ (2,1) \mapsto 6 \\ (2,2) \mapsto 8 \end{pmatrix} \in \mathbb{R}^{[1..2]\times[1..2]} \quad = \quad \mathbb{R}^{2\times 2}.$$

We can also form a "matrix of vectors", *e.g.*,

$$\begin{bmatrix} (3,4) \\ (6,8) \end{bmatrix} \quad = \quad \begin{pmatrix} (1,1) \mapsto (3,4) \\ (1,2) \mapsto (6,8) \end{pmatrix} \in (\mathbb{R}^2)^{2\times 1}.$$

It is not hard to set up a bijection $\mathbb{R}^{2\times 2} \hookrightarrow\!> (\mathbb{R}^2)^{2\times 1}$, so the two sets $\mathbb{R}^{2\times 2}$ and $(\mathbb{R}^2)^{2\times 1}$ are very closly related. Typically, a determinant function is set up with domain $\mathbb{R}^{2\times 2}$, but it turns out that, for what we want, it's more convenient to work in $(\mathbb{R}^2)^{2\times 1}$, as follows:

DEFINITION 39.13. *Define Det* : $(\mathbb{R}^2)^{2\times 1} \to \mathbb{R}$ *by*

$$Det \begin{bmatrix} (a,b) \\ (c,d) \end{bmatrix} \quad = \quad ad - bc.$$

So, for example,

$$\mathrm{Det} \begin{bmatrix} (3,4) \\ (6,8) \end{bmatrix} \quad = \quad 3 \cdot 4 - 6 \cdot 8 \quad = \quad 0.$$

The following is the determinant test for parallel arrows in $\mathbb{R}^2$:

FACT 39.14. *Let* $u, v \in \mathbb{R}^2$. *Then*

$$(\, u \,||\, v \,) \quad \Leftrightarrow \quad \left(Det \begin{bmatrix} u \\ v \end{bmatrix} = 0 \right).$$

Proof. Unassigned HW. ☐

FACT 39.15. *Let* $w \in \mathbb{R}^2$. *Then Det* $\begin{bmatrix} w \\ w \end{bmatrix} = 0$.

Proof. Det $\begin{bmatrix} w \\ w \end{bmatrix} = w_2 w_1 - w_1 w_2 = 0.$ ☐

REMARK 39.16. *Let $u, v, w \in \mathbb{R}^2$. Then*

$$Det \begin{bmatrix} u - v \\ w \end{bmatrix} \quad = \quad \left(Det \begin{bmatrix} u \\ w \end{bmatrix} \right) \; - \; \left(Det \begin{bmatrix} v \\ w \end{bmatrix} \right).$$

Proof. We have

$$\begin{aligned} \mathrm{Det} \begin{bmatrix} u - v \\ w \end{bmatrix} &= (u-v)_1 \cdot w_2 \; - \; (u-v)_2 \cdot w_1 \\ &= (u_1 - v_1) \cdot w_2 \; - \; (u_2 - v_2) \cdot w_1 \\ &= (u_1 w_2 - v_1 w_2) \; - \; (u_2 w_1 - v_2 w_1) \\ &= (u_1 w_2 - u_2 w_1) \; - \; (v_1 w_2 - v_2 w_1) \\ &= \left(\mathrm{Det} \begin{bmatrix} u \\ w \end{bmatrix} \right) \; - \; \left(\mathrm{Det} \begin{bmatrix} v \\ w \end{bmatrix} \right), \end{aligned}$$

as desired. $\square$

DEFINITION 39.17. *Let $w \in \mathbb{R}^2$. Then $Det \begin{bmatrix} \bullet \\ w \end{bmatrix} : \mathbb{R}^2 \to \mathbb{R}$ is defined by $\left(Det \begin{bmatrix} \bullet \\ w \end{bmatrix} \right)(v) = Det \begin{bmatrix} v \\ w \end{bmatrix}$.*

For exmaple, let $L := \mathrm{Det} \begin{bmatrix} \bullet \\ (4,7) \end{bmatrix}$. Then: $\forall x, y \in \mathbb{R}$,

$$L(x,y) \quad = \quad \mathrm{Det} \begin{bmatrix} (x,y) \\ (4,7) \end{bmatrix} \quad = \quad 7x - 4y.$$

Note that $L : \mathbb{R}^2 \to \mathbb{R}$ is continuous and linear. The choice of $(4,7)$ is irrelevant to that continuity and linearity:

FACT 39.18. *Let $w \in \mathbb{R}^2$. Then the function $Det \begin{bmatrix} \bullet \\ w \end{bmatrix} : \mathbb{R}^2 \to \mathbb{R}$ is both continuous and linear.*

Proof. Unassigned HW. $\square$

REMARK 39.19. *Let V and W be normed vector spaces. Assume $V \neq \{0_V\}$. Let $L : V \to W$ be continuous and linear. Let $f : \mathbb{R} \dashrightarrow V$. Let $t \in dom[f']$. Then $(L \circ f)'(t) = L(f'(t))$.*

The assumption that $V \neq \{0_V\}$ is, of course, very tame. However, if we drop it, then Remark 39.19 can fail: Say $V = \{0_V\}$, $W = \mathbb{R}$, $f = C_{\{2\}}^{0_V} : \mathbb{R} \dashrightarrow V$, $L = C_V^0 : V \to W$ and $t = 2$. Then: $f'(t) = 0_V$, and so $L(f'(t)) = 0$, but $L \circ f = C_{\{2\}}^0 : \mathbb{R} \dashrightarrow W$, and so $(L \circ f)'(t) = ☹$.

In the case $V = W = \mathbb{R}$ and, for simplicity, $\mathrm{dom}[f'] = \mathbb{R}$, we can prove Remark 39.19 from the Chain Rule (Theorem 32.8). For example, suppose: $\forall x \in \mathbb{R}$, $L(x) = 5x$. Then $L(f'(t)) = 5 \cdot [f'(t)]$. Differentiating L, we see: $\forall x \in \mathbb{R}$, $L'(x) = 5$. Then $L'(f(t)) = 5$. The Chain Rule gives $(L \circ f)'(t) = [L'(f(t))] \cdot [f'(t)]$. Then

$$\begin{aligned}(L \circ f)'(t) &= [L'(f(t))] \cdot [f'(t)] \\ &= 5 \cdot [f'(t)] = L(f'(t)),\end{aligned}$$

as desired. To prove Remark 39.19 in general:

Proof. By Remark 31.5, $\mathrm{dom}[f'] \subseteq \mathrm{LPD}_{\mathbb{R}} f$. As $t \in \mathrm{dom}[f'] \subseteq \mathrm{LPD}_{\mathbb{R}} f$, by (2) of Remark 31.3, it suffices to show: $SS^t_{L\circ f} \to L(f'(t))$ near 0.

For all $h \in \mathbb{R}$, we have

$$\begin{aligned}(SS^t_{L\circ f})(h) &= \frac{[(L \circ f)(t+h)] - [(L \circ f)(t)]}{h} \\ &= \frac{[L(f(t+h))] - [L(f(t))]}{h} \\ &= \frac{L\left([f(t+h)] - [f(t)]\right)}{h} \\ &= L\left(\frac{[f(t+h)] - [f(t)]}{h}\right) \\ &= L((SS^t_f)(h)) = (L \circ (SS^t_f))(h).\end{aligned}$$

Then $SS^t_{L\circ f} = L \circ (SS^t_f)$.

Since $t \in \mathrm{dom}[f']$, by Lemma 31.4, we get: $SS^t_f \to f'(t)$ near 0. So, since L is continuous at $f'(t)$, by HW#60 from Fall 2017, we conclude that $L \circ (SS^t_f) \to L(f'(t))$ near 0. So, since $SS^t_{L\circ f} = L \circ (SS^t_f)$, this yields $SS^t_{L\circ f} \to L(f'(t))$ near 0, as desired. □

COROLLARY 39.20. *Let V and W be normed vector spaces. Assume $V \neq \{0_V\}$. Let $S \subseteq \mathbb{R}$ and let $f : \mathbb{R} \dashrightarrow V$ be c/d on S. Let $L : V \to W$ be continuous and linear. Then both of the following hold:*

(1) *$L \circ f$ is c/d on S and*
(2) *$\forall t \in \mathrm{Int}_{\mathbb{R}} S$, $(L \circ f)'(t) = L(f'(t))$.*

Proof. Let $S_0 := \mathrm{Int}_{\mathbb{R}} S$. Since f is c/d on S, we get: $S \subseteq \mathrm{dct}[f]$ and $S_0 \subseteq \mathrm{dom}[f']$. Then, for all $t \in S_0$, we have $t \in \mathrm{dom}[f']$, so, by Remark 39.19, $(L \circ f)'(t) = L(f'(t))$. This proves (2). It remains to prove that (1) is true; that is, we wish to show:

(A) $L \circ f$ is continuous on S and

(B) $L \circ f$ is differentiable on S_0.

Proof of (A): We wish to show: $\forall t \in S$, $L \circ f$ is continuous at t. Let $t \in S$ be given. We wish to show: $L \circ f$ is continuous at t.

We have $t \in S \subseteq \mathrm{dct}[f]$. That is, f is continuous at t. Since $t \in \mathrm{dct}[f] \subseteq \mathrm{dom}[f]$ and since $f : \mathbb{R} \dashrightarrow V$, we get $f(t) \in V$. So, since $L : V \to W$ is continuous, we conclude that L is continuous at $f(t)$.

Since f is continuous at t and L is continuous at $f(t)$, we conclude, by Theorem 22.1, that $L \circ f$ is continuous at t. *End of proof of (A).*

Proof of (B): We wish to show: $\forall t \in S_0$, $L \circ f$ is differentiable at t. Let $t \in S_0$ be given. We wish to show: $L \circ f$ is differentiable at t.

Since $t \in S_0 \subseteq \mathrm{dom}[f']$ and $f' : \mathbb{R} \dashrightarrow V$, we get $f'(t) \in V$. Thus $f'(t) \in V = \mathrm{dom}[L]$, so $L(f'(t)) \neq ☹$, and so, by (2), $(L \circ f)'(t) \neq ☹$. Thus, $L \circ f$ is differentiable at t. *End of proof of (B).* $\square$

Next time we will prove the Cauchy Mean Value Theorem. After that, we will prove L'Hospital's Rule. Then we'll use that to prove the higher order derivative tests (second derivative test, third derivative test, fourth derivative test, *etc*).

40. Class 14 on 6 March 2018, Tu of Week 8

Recall: Let W be a vector space and let $u, v \in W$. Then $u \| v$ means: [$\{u, v\}$ is linearly dependent in W], *i.e.*, [($u \in \mathbb{R}v$) or ($v \in \mathbb{R}u$)].

Assigned HW#8-1.

Recall: For all $v \in \mathbb{R}^2$, $\mathrm{sl}\, v = v_2/v_1$.

Assigned HW#8-2.

Recall: For all $u, v \in \mathbb{R}^2$, $\mathrm{Det} \begin{bmatrix} u \\ v \end{bmatrix} = u_1 v_2 - v_2 u_1$.

Assigned HW#8-3.

Recall: For all $w \in \mathbb{R}^2$, $\mathrm{Det} \begin{bmatrix} \bullet \\ w \end{bmatrix} : \mathbb{R}^2 \to \mathbb{R}$ is continuous and linear. Also, recall both of the following:

(1) $\forall w \in \mathbb{R}^2$, $\mathrm{Det} \begin{bmatrix} w \\ w \end{bmatrix} = 0$.

(2) $\forall u, v, w \in \mathbb{R}^2$, $\mathrm{Det} \begin{bmatrix} u - v \\ w \end{bmatrix} = \left(\mathrm{Det} \begin{bmatrix} u \\ w \end{bmatrix} \right) - \left(\mathrm{Det} \begin{bmatrix} v \\ w \end{bmatrix} \right)$.

Finally, recall Corollary 39.20:

COROLLARY 40.1. *Let V and W be normed vector spaces. Assume $V \neq \{0_V\}$. Let $S \subseteq \mathbb{R}$ and let $f : \mathbb{R} \dashrightarrow V$ be c/d on S. Let $L : V \to W$ be continuous and linear. Then both of the following hold:*

(1) *$L \circ f$ is c/d on S* *and*
(2) *$\forall t \in \mathrm{Int}_{\mathbb{R}} S, \quad (L \circ f)'(t) = L(f'(t))$.*

The following is called the **Cauchy Mean Value Theorem**:

THEOREM 40.2. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}^2$ and let $a, b \in \mathbb{R}$. Assume $a < b$. Assume f is c/d on $[a, b]$. Then $\exists c \in (a, b)$ s.t. $f'(c) \parallel (SV_f)(a, b)$.*

Proof. Let $w := (SV_f)(a, b)$. We wish to show: $\exists c \in (a, b)$ s.t. $f'(c) \parallel w$.

Let $L := \mathrm{Det} \begin{bmatrix} \bullet \\ w \end{bmatrix}$. Let $\phi := L \circ f$. By Corollary 39.20,

(1) ϕ is c/d on $[a, b]$ and
(2) $\forall t \in (a, b), \quad \phi'(t) = L(f'(t))$.

By linearity of L, we have $L([f(b)] - [f(a)]) = [L(f(b))] - [L(f(a))]$. Also, by definition of SV_f, we have $SV_f(a, b) = [f(b)] - [f(a)]$. Then

$$\begin{aligned} [\phi(b)] - [\phi(a)] &= [(L \circ f)(b)] - [(L \circ f)(a)] \\ &= [L(f(b))] - [L(f(a))] \\ &= L([f(b)] - [f(a)]) \\ &= L((SV_f)(a, b)) \\ &= L(w) = \mathrm{Det} \begin{bmatrix} w \\ w \end{bmatrix} = 0. \end{aligned}$$

Then $\phi(a) = \phi(b)$. So, by (1) above combined with Rolle's Theorem (Theorem 34.3), choose $c \in (a, b)$ s.t. $\phi'(c) = 0$. We want: $f'(c) \parallel w$.

By (2) above, $L(f'(c)) = \phi'(c)$. Then

$$\mathrm{Det} \begin{bmatrix} f'(c) \\ w \end{bmatrix} = L(f'(c)) = \phi'(c) = 0,$$

so, by HW#8-3, $f'(c) \parallel w$, as desired. ∎

REMARK 40.3. *Let X, Y, Z be topological spaces. Let $\phi : X \dashrightarrow Y$ and $\psi : X \dashrightarrow Z$. Let $q \in X$, $s \in Y$ and $t \in Z$. Assume that $\phi \to s$ near q and that $\psi \to t$ near q. Then $(\phi, \psi) \to (s, t)$ near q.*

Proof. We want: $\forall W \in \mathcal{N}_{Y\times Z}(s,t)$, $\exists A \in \mathcal{N}_X^\times(q)$ s.t. $(\phi,\psi)_*(A) \subseteq W$. Given $W \in \mathcal{N}_{Y\times Z}(s,t)$. Want: $\exists A \in \mathcal{N}_X^\times(q)$ s.t. $(\phi,\psi)_*(A) \subseteq W$.

By Fact 29.1, choose $U \in \mathcal{N}_Y(s)$ and $V \in \mathcal{N}_Z(t)$ s.t. $U \times V \subseteq W$. Since $\phi \to s$ near q, choose $B \in \mathcal{N}_X^\times(q)$ s.t. $\phi_*(B) \subseteq U$. Since $\psi \to t$ near q, choose $C \in \mathcal{N}_X \times (q)$ s.t. $\psi_*(C) \subseteq V$. Then $B \cap C \ni \mathcal{N}_X^\times(q)$. Let $A := B \cap C$. We wish to show: $(\phi,\psi)_*(A) \subseteq W$.

Unassigned HW: Show $(\phi,\psi)_*(A) \subseteq [\phi_*(A)] \times [\psi_*(A)]$. Then we have $(\phi,\psi)_*(A) \subseteq [\phi_*(A)] \times [\psi_*(A)] \subseteq U \times V \subseteq W$, as desired. □

Assigned HW#8-4.

REMARK 40.4. *Let X, Y and Z be topological spaces. Assume that Y and Z are both Hausdorff. Let $\phi : X \dashrightarrow Y$ and let $\psi : X \dashrightarrow Z$. Let $q \in \mathrm{LPD}_X(\phi,\psi)$. Then $\lim\limits_q (\phi,\psi) \;\; =^* \;\; (\lim\limits_q \phi\,,\, \lim\limits_q \psi\,)$.*

Proof. Let $L := \lim\limits_q (\phi,\psi)$ and let $R := (\lim\limits_q \phi\,,\, \lim\limits_q \psi\,)$. We wish to show: $[R \neq ☹] \Rightarrow [L = R]$. Assume: $R \neq ☹$. We wish to show: $L = R$.

Let $s := \lim\limits_q \phi$ and $t := \lim\limits_q \psi$. Then $R = (s,t)$. So, since $R \neq ☹$, we get: $s \neq ☹ \neq t$. We want to show that $\lim\limits_q (\phi,\psi) = (s,t)$.

Since Y and Z are both Hausdorff, it follows that $Y \times Z$ is Hausdorff. So, since $q \in \mathrm{LPD}_X(\phi,\psi)$, by Remark 25.19, we see that it suffices to show that $(\phi,\psi) \to (s,t)$ near q.

Since $\lim\limits_q \phi = s$ and $\lim\limits_q \psi = t$, by Remark 25.18, we have $\phi \to s$ near q and $\psi \to t$ near q. Then, by Remark 40.3, $(\phi,\psi) \to (s,t)$ near q. □

Assigned HW#8-5.

LEMMA 40.5. *Let $x, y : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $S \subseteq \mathbb{R}$. Assume x and y are both c/d on S. Then both of the following hold:*

(1) *(x,y) is c/d on S* *and*
(2) *$\forall p \in \mathrm{Int}_{\mathbb{R}} S$, $\quad (x,y)'(p) \;=\; (\,x'(p)\,,\, y'(p)\,)$.*

Proof. Let $S_0 := \mathrm{Int}_{\mathbb{R}} S$. We wish to show:

(A) (x,y) is continuous on S,
(B) $\forall p \in S_0$, $\quad (x,y)'(p) \;=\; (\,x'(p)\,,\, y'(p)\,)$.
(C) (x,y) is differentiable on S_0.

Proof of (A): We wish to show: $\forall p \in S$, (x,y) is continuous at p. Let $p \in S$ be given. We wish to show: (x,y) is continuous at p.

Since x and y are both c/d on S, we know that x and y are both continuous on S. Then x and y are both continuous at p, and so, by HW#8-4, (x, y) is continuous at p. *End of proof of (A).*

Proof of (B): Let $p \in S_0$ be given. We want: $(x, y)'(p) = (x'(p), y'(p))$.

By HW#8-5, it suffices to show: $p \in \mathrm{LPD}_{\mathbb{R}}(x, y)$. By (A), we have $S \subseteq \mathrm{dct}[(x, y)]$. So, since $\mathrm{dct}[(x, y)] \subseteq \mathrm{dom}[(x, y)]$, we get $S \subseteq \mathrm{dom}[(x, y)]$. Then $\mathrm{LP}_{\mathbb{R}} S \subseteq \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[(x, y)])$. Then

$$p \in S_0 = \mathrm{Int}_{\mathbb{R}} S \subseteq \mathrm{LP}_{\mathbb{R}} S \subseteq \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[(x, y)]) = \mathrm{LPD}_{\mathbb{R}}(x, y),$$

as desired. *End of proof of (B).*

Proof of (C): We wish to show: $\forall p \in S_0$, (x, y) is differentiable at p. Let $p \in S$ be given. We wish to show: (x, y) is differentiable at p.

Since x and y are both c/d on S, we know that x and y are both differentiable on S_0. Then x and y are both differentiable at p. Then $x(p) \neq ☹ \neq y(p)$. Then $(x(p), y(p)) \neq ☹$. So, by (B), $(x, y)'(p) \neq ☹$. Then (x, y) is differentiable at p. *End of proof of (C).* $\square$

Recall (Definition 29.6): $\forall \alpha, \beta \in \mathbb{R}$, we defined:

$$[\alpha|\beta] \quad := \quad [\,\min\{\alpha, \beta\}\,,\, \max\{\alpha, \beta\}\,].$$

Similarly:

DEFINITION 40.6. *Let $\alpha, \beta \in \mathbb{R}$. Assume that $\alpha \neq \beta$. Then we define* $(\alpha|\beta) := (\,\min\{\alpha, \beta\}\,,\, \max\{\alpha, \beta\}\,)$.

The following is analogous to Definition 34.4.

DEFINITION 40.7. *Let $x, y : \mathbb{R} \dashrightarrow \mathbb{R}$. Then $DQ_{x,y} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is defined by*

$$DQ_{x,y}(p, q) \qquad = \qquad \frac{[\,y(q)\,] \;-\; [\,y(p)\,]}{[x(q)] - [x(p)]}.$$

Let $x, y : \mathbb{R} \dashrightarrow \mathbb{R}$. The function $DQ_{x,y}$ is symmetric, *i.e.*, we have: $\forall p, q \in \mathbb{R}$, $(DQ_{x,y})(p, q) = (DQ_{x,y})(q, p)$. Also, $DQ_{x,y}$ is undefined on the diagonal, *i.e.*, $\forall p \in \mathbb{R}$, $(DQ_{x,y})(p, p) = ☹$.

Let $x, y : \mathbb{R} \dashrightarrow \mathbb{R}$. Let $f := (x, y) : \mathbb{R} \dashrightarrow \mathbb{R}^2$. For all $p, q \in \mathbb{R}$,

$$(SA_f)(p, q) \;=\; (\,f(p)\,,\, f(q)\,) \;=\; (\,(x(p), y(p))\,,\, (x(q), y(q))\,),$$

and so

$$(SV_f)(p, q) \quad = \quad (\quad [x(q)] - [x(p)] \quad , \quad [y(q)] - [y(p)] \quad),$$

and so $(DQ_{x,y})(p, q) = \mathrm{sl}\,(\,(SV_f)(p, q)\,)$. (Recall: sl is “slope”.)

Let $i := \mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$. For any $f : \mathbb{R} \dashrightarrow \mathbb{R}$, we have $DQ_f = DQ_{i,f}$.

We will call the next result the **Cauchy Mean Value Corollary**.

THEOREM 40.8. *Let $x, y : \mathbb{R} \to \mathbb{R}$ and let $\alpha, \beta \in \mathbb{R}$. Assume that $\alpha \neq \beta$. Assume that x and y are both c/d on $[\alpha|\beta]$. Assume: $\forall s \in (\alpha|\beta)$, $x'(s) \neq 0$. Then there exists $c \in (\alpha|\beta)$ s.t.*

$$(y'/x')(c) \quad = \quad (DQ_{x,y})(\alpha, \beta).$$

Proof. By symmetry of $DQ_{x,y}$, we have $(DQ_{x,y})(\alpha, \beta) = (DQ_{x,y})(\beta, \alpha)$. Let $a := \min\{\alpha, \beta\}$, $b := \max\{\alpha, \beta\}$. Then $a < b$. Also,

- $(a, b) \ \in \ \{\ (\alpha, \beta)\ ,\ (\beta, \alpha)\ \}$,
- $(DQ_{x,y})(a, b) = (DQ_{x,y})(\alpha, \beta)$,
- x and y are both c/d on $[a, b]$ and
- $\forall s \in (a, b)$, $x'(s) = 0$.

We wish to show: $\exists c \in (a, b)$ s.t. $(y'/x')(c) = (DQ_{x,y})(a, b)$.

Let $f := (x, y) : \mathbb{R} \dashrightarrow \mathbb{R}^2$. By (1) of Lemma 40.5, f is c/d on $[a, b]$. By the Cauchy Mean Value Theorem (Theorem 40.2) choose $c \in (a, b)$ s.t. $f'(c) \parallel (SV_f)(a, b)$. We wish to show $(y'/x')(c) = (DQ_{x,y})(a, b)$.

As $f'(c) \parallel (SV_f)(a, b)$, by HW#8-2, at least one of the following holds:

(a) $f'(c) = 0_2$ or
(b) $(SV_f)(a, b) = 0_2$ or
(c) $\mathrm{sl}\,(\,f'(c)\,) \ = \ \mathrm{sl}\,(\,(SV_f)(a, b)\,)$.

By (2) of Lemma 40.5, $f'(c) = (x'(c), y'(c))$. By assumption, we know for all $s \in (\alpha|\beta)$, that $x'(s) \neq 0$. So, since $c \in (a, b)$, we get $x'(c) \neq 0$. Then $f'(c) \neq 0_2$, so (a) is false. By (1) of Corollary 34.10, we conclude that $x|[a, b]$ is 1-1. So, since $a \neq b$, we get $x(a) \neq x(b)$. Then $f(a) \ = \ (\,x(a)\,,\,y(a)\,) \ \neq \ (\,x(b)\,,\,y(b)\,) \ = \ f(b)$. Then

$$(SV_f)(a, b) \quad = \quad [f(b)] - [f(a)] \quad \neq \quad 0_2.$$

Then (b) is false. Since (a) and (b) are both false, (c) must be true. That is, we have: $\mathrm{sl}\,(\,f'(c)\,) \ = \ \mathrm{sl}\,(\,(SV_f)(a, b)\,a)$. Then

$$\begin{aligned}
(y'/x')(c) \quad &= \quad \frac{y'(c)}{x'(c)} \quad = \quad \mathrm{sl}\,(\,(x'(c)\,,\,y'(c)\,) \quad = \quad \mathrm{sl}\,(\,f'(c)\,) \\
&= \quad \mathrm{sl}\,(\,(SV_f)(a, b)\,) \quad = \quad \mathrm{sl}\,(\,[f(b)] \,-\, [f(a)]\,) \\
&= \quad \mathrm{sl}\,(\,(\,[x(b)] \,-\, [x(a)]\,,\,[y(b)] \,-\, [y(a)]\,)\,) \\
&= \quad \frac{[y(b)] \,-\, [y(a)]}{[x(b)] \,-\, [x(a)]} \quad = \quad (DQ_{x,y})(a, b),
\end{aligned}$$

as desired. $\square$

We next aim to use the Cauchy Mean Value Corollary (Theorem 40.8) to prove L'Hospital's Rule, and to use that to get the asymptotic results needed to prove the higher order derivative tests.

We require some preliminaries. First, we have:

41. Class 15 on 8 March 2018, Th of Week 8

Spring Break is next week (12-16 March).

Midterm 2 is two weeks from today (Thursday 22 March).

DEFINITION 41.1. *Let Y be a topological space, let Z be a set, let $f, g : Y \dashrightarrow Z$ and let $p \in Y$. Then f* **follows** *g near p means: $\forall V \in \mathcal{N}_Y^\times(p)$, $\exists U \in \mathcal{N}_Y^\times(p)$ s.t. $f_*(U) \subseteq g_*(V)$.*

LEMMA 41.2. *Let Y and Z be toplogical spaces. Let $f, g : Y \dashrightarrow Z$. Let $p \in Y$ and let $q \in Z$. Assume: $g \to q$ near p. Assume: f follows g near p. Then: $f \to q$ near p.*

Proof. We wish to show: $\forall W \in \mathcal{N}_Z(q)$, $\exists U \in \mathcal{N}_Y(p)$ s.t. $f_*(U) \subseteq W$. Let $W \in \mathcal{N}_Z(q)$ be given. We want: $\exists U \in \mathcal{N}_Y(p)$ s.t. $f_*(U) \subseteq W$.

As $g \to q$ near p, choose $V \in \mathcal{N}_Y^\times(p)$ s.t. $g_*(V) \subseteq W$. As f follows g near p, choose $U \in \mathcal{N}_Y^\times(p)$ s.t. $f_*(U) \subseteq g_*(V)$. Want: $f_*(U) \subseteq W$.

We have $f_*(U) \subseteq g_*(V) \subseteq W$, as desired. $\square$

LEMMA 41.3. *Let Y, Z be toplogical spaces, and $f, g : Y \dashrightarrow Z$, and $p \in \mathrm{LPD}_Y f$. Assume: (Z is Hausdorff) and (f follows g near p). Then $\lim\limits_p f \;\;=^*\;\; \lim\limits_p g$.*

Proof. Let $L := \lim\limits_p f$ and $R := \lim\limits_p g$. Want: $(R \neq ☹) \Rightarrow (L = R)$. Assume: $R \neq ☹$. Want: $L = R$. That is, we want to prove: $\lim\limits_p f = R$. Since $p \in \mathrm{LPD}_Y f$ and since Z is Hausdorff, by Remark 25.19, it suffices to show that $f \to R$ near p.

Since $\lim\limits_p g = R \neq ☹$, by Remark 25.18, we have $g \to R$ near p. So, since f follows g near p, it follows, from Lemma 41.2 (with q replaced by R), that $f \to R$ near p, as desired. $\square$

Recall:

- $\forall$set A, $\forall p$, $\quad A_p^\times := A \backslash \{p\}$,
- $\forall$set $\mathcal{S}$ of sets, $\forall p$, $\quad \mathcal{S}_p^{\times\times} := \{A_p^\times \mid A \in \mathcal{S}\}$,

- $\forall$topological space Y, $\forall p \in Y$, $\quad \mathcal{N}_Y^\times(p) := [\mathcal{N}_Y(p)]_p^{\times\times}$ and
- $\forall$metric space Y, $\forall p \in Y$, $\quad \mathcal{B}_Y^\times(p) := [\mathcal{B}_Y(p)]_p^{\times\times}$.

DEFINITION 41.4. *For any set A, for any p, $A_p^+ := A \cup \{p\}$.*

DEFINITION 41.5. *Let $\mathcal{S}$ be a set of sets. Then, for any p, we define $\mathcal{S}_p^{++} := \{A_p^+ \mid A \in \mathcal{S}\}$.*

FACT 41.6. *Let Y be a top. space, $p \in Y$. Then $[\mathcal{N}_Y^\times(p)]_p^{++} = \mathcal{N}_Y(p)$.*

Proof. Omitted. □

FACT 41.7. *Let Y be a metric space, $p \in Y$. Then $[\mathcal{B}_Y^\times(p)]_p^{++} = \mathcal{B}_Y(p)$.*

Proof. Omitted. □

LEMMA 41.8. *Let Y be a topological space, $p \in Y$, $\mathcal{B}$ be a neighborhood base at p in Y. Let Z be a set, $f, g : Y \dashrightarrow Z$. Assume that: $\forall V \in \mathcal{B}_p^{\times\times}$, $\exists U \in \mathcal{N}_Y^\times(p)$ s.t. $f_*(U) \subseteq g_*(V)$. Then f follows g near p.*

Proof. We want: $\forall W \in \mathcal{N}_Y^\times(p)$, $\exists U \in \mathcal{N}_Y^\times(p)$ s.t. $f_*(U) \subseteq g_*(W)$. Let $W \in \mathcal{N}_Y^\times(p)$ be given. We want: $\exists U \in \mathcal{N}_Y^\times(p)$ s.t. $f_*(U) \subseteq g_*(W)$.

Since $W_p^+ \in \mathcal{N}_Y(p)$ and since $\mathcal{B}$ is a neighborhood base at p in Y, choose $V \in \mathcal{B}$ s.t. $V \subseteq W_p^+$. Since $V \in \mathcal{B}$, it follows that $V_p^\times \in \mathcal{B}_p^{\times\times}$. Then, by assumption, choose $U \in \mathcal{N}_Y^\times(p)$ s.t. $f_*(U) \subseteq g_*(V_p^\times)$. Since $V \subseteq W_p^+$, it follows that $V_p^\times \subseteq (W_p^+)_p^\times$. So, since $(W_p^+)_p^\times = W_p^\times \subseteq W$, we get $V_p^\times \subseteq W$, and so $g_*(V_p^\times) \subseteq g_*(W)$.

Then $f_*(U) \subseteq g_*(V_p^\times) \subseteq g_*(W)$, as desired. □

REMARK 41.9. *Let Y be a metric space, and let $p \in Y$. Then*

(1) $\forall U, V \in \mathcal{B}_Y(p), \quad U \cap V \in \{U, V\}$,
(2) $\forall U, V \in \mathcal{B}_Y^\times(p), \quad U \cap V \in \{U, V\}$,
(3) $\mathcal{B}_Y(p) \subseteq \mathcal{N}_Y(p)$ *and*
(4) $\mathcal{B}_Y^\times(p) \subseteq \mathcal{N}_Y^\times(p)$.

Proof. Unassigned HW. □

REMARK 41.10. *Let $p \in \mathbb{R}$, $A \in \mathcal{B}_\mathbb{R}^\times(p)$ and $t \in A$. Then $(p|t) \subseteq A$.*

Proof. We have $A_p^+ \in \mathcal{B}_\mathbb{R}(p)$, so choose $\alpha > 0$ s.t. $A_p^+ = B_\mathbb{R}(p, \alpha)$. Then

$$\begin{aligned} t \in A &= (A_p^+)_p^\times = (B_\mathbb{R}(p,\alpha))_p^\times = (\,(p-\alpha, p+\alpha)\,)_p^\times \\ &= (p-\alpha, p+\alpha)\backslash\{p\} = (p-\alpha, p) \cup (p, p+\alpha). \end{aligned}$$

Then one of the following must hold:

(1) $t \in (p - \alpha, p)$ or
(2) $t \in (p, p + \alpha)$.

Case (1): Since $t < p$, we get $(p|t) = (t, p)$. Since $p - \alpha < t$, we get $(t, p) \subseteq (p - \alpha, p)$. Then

$$(p|t) \ = \ (t, p) \ \subseteq \ (p - \alpha, p) \ \subseteq \ (p - \alpha, p) \cup (p, p + \alpha) \ = \ A,$$

as desired. *End of Case (1).*

Case (2): Since $p < t$, we get $(p|t) = (p, t)$. Since $t < p + \alpha$, we get $(p, t) \subseteq (p, p + \alpha)$. Then

$$(p|t) \ = \ (p, t) \ \subseteq \ (p, p + \alpha) \ \subseteq \ (p - \alpha, p) \cup (p, p + \alpha) \ = \ A,$$

as desired. *End of Case (2).* $\square$

LEMMA 41.11. *Let $p \in \mathbb{R}$ and let $A \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$. Let Z be a set and let $f, g : \mathbb{R} \dashrightarrow Z$. Assume that $A \subseteq dom[f]$. Assume that: $\forall t \in A$, $\exists s \in (p|t)$ s.t. $g(s) = f(t)$. Then f follows g near p.*

Proof. Because $\mathcal{B}_{\mathbb{R}}(p)$ is a neighborhood base at p in $\mathbb{R}$ and because $[\mathcal{B}_{\mathbb{R}}(p)]_p^{\times} = \mathcal{B}_{\mathbb{R}}^{\times}(p)$, by Lemma 41.8 (with Y replaced by $\mathbb{R}$, and $\mathcal{B}$ by $\mathcal{B}_{\mathbb{R}}(p)$), we wish to show: $\forall V \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$, $\exists U \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$ s.t. $f_*(U) \subseteq g_*(V)$. Let $V \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$ be given. We want: $\exists U \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$ s.t. $f_*(U) \subseteq g_*(V)$.

Since $A, V \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$, by (2) of Remark 41.9, we get $A \cap V \in \{A, V\}$. So, since $\{A, V\} \subseteq \mathcal{B}_{\mathbb{R}}^{\times}(p)$, we see that $A \cap V \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$. By (4) of Remark 41.9, $\mathcal{B}_{\mathbb{R}}^{\times}(p) \subseteq \mathcal{N}_{\mathbb{R}}^{\times}(p)$. Then $A \cap V \in \mathcal{N}_{\mathbb{R}}^{\times}(p)$. Let $U := A \cap V$. We want: $f_*(U) \subseteq g_*(V)$. We wish to show: $\forall t \in \text{dom}[f]$,

$$[\ t \in U\ ] \quad \Rightarrow \quad [\ f(t) \in g_*(V)\ ].$$

Let $t \in \text{dom}[f]$ be given. We wish to show:

$$[\ t \in U\ ] \quad \Rightarrow \quad [\ f(t) \in g_*(V)\ ].$$

Assume: $t \in U$. We want: $f(t) \in g_*(V)$.

We have $t \in U = A \cap V \subseteq A$. So, by assumption, choose $s \in (p|t)$ s.t. $g(s) = f(t)$. As $t \in \text{dom}[f]$, $f(t) \neq ☹$. Then $g(s) = f(t) \neq ☹$, so $s \in \text{dom}[g]$. We have $U = A \cap V \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$. So, since $t \in U$, by Remark 41.10 (with A replaced by U), we get $(p|t) \subseteq U$. Then $s \in (p|t) \subseteq U = A \cap V \subseteq V$. So, since $s \in \text{dom}[g]$, we get $g(s) \in g_*(V)$. Then $f(t) = g(s) \in g_*(V)$, as desired. $\square$

LEMMA 41.12. *Let $p \in \mathbb{R}$, $x : \mathbb{R} \dashrightarrow \mathbb{R}$. Assume that x is continuous at p and that $x(p) = 0$. Let $y : \mathbb{R} \dashrightarrow \mathbb{R}$. Let $A \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$ and assume that $A \subseteq dom[y'/x']$. Then $A \subseteq dom[y/x]$.*

Proof. Want: $\forall t \in A$, $t \in \mathrm{dom}[y/x]$. Given $t \in A$. Want: $t \in \mathrm{dom}[y/x]$.

Since $A \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$, we get $p \notin A$. So, since $t \in A$, we get $t \neq p$. Since

$$(41.1) \qquad t \in A \;\subseteq\; \mathrm{dom}[y'/x'] \;\subseteq\; (\,\mathrm{dom}[x']\,) \cap (\,\mathrm{dom}[y']\,)$$

$$(41.2) \qquad \subseteq\; (\,\mathrm{dom}[x]\,) \cap (\,\mathrm{dom}[y]\,),$$

we only need to show that $x(t) \neq 0$.

By Remark 41.10, $(p|t) \subseteq A$. By assumption, x is continuous at p. Since $(p|t) \subseteq A \subseteq \mathrm{dom}[x'] \subseteq \mathrm{dct}[x]$, x is continuous on $(p|t)$. Since $t \in A \subseteq \mathrm{dom}[x'] \subseteq \mathrm{dct}[x]$, x is continuous at t. Then x is continuous on $[p|t]$. Since $(p|t) \subseteq A \subseteq \mathrm{dom}[x']$, x is differentiable on $(p|t)$. Then x is c/d on $[p|t]$. Let $I := [c|t]$. Then x is c/d on I.

As $\mathrm{Int}_{\mathbb{R}} I = (p|t) \subseteq A \subseteq \mathrm{dom}[y'/x']$, we know: $\forall s \in \mathrm{Int}_{\mathbb{R}} I$, $x'(s) \neq 0$. Then $x'_*(\mathrm{Int}_{\mathbb{R}} I) > 0$. So, by (1) of Corollary 34.10, $x|I$ is 1-1. So, since $t, p \in [p|t] = I$ and since $t \neq p$, it follows that $x(t) \neq x(p)$. By assumption, $x(p) = 0$. Then $x(t) \neq 0$, as desired. $\square$

DEFINITION 41.13. *Let Y be a topological space, $p \in Y$. Then $\mathcal{N}_Y^{\circ}(p) := [\mathcal{N}_Y(p)] \cup [\mathcal{N}_Y^{\times}(p)]$. For any $U \subseteq Y$, by U is a* **possibly punctured neighborhood** *of p in Y we mean: $U \in \mathcal{N}_Y^{\circ}(p)$.*

REMARK 41.14. *Let Y be a topological space, $p \in Y$, $U \subseteq Y$. Then:*

$$[\; U \in \mathcal{N}_Y^{\circ}(p) \;] \;\Leftrightarrow\; [\; U_p^{+} \in \mathcal{N}_Y(p) \;] \;\Leftrightarrow\; [\; U_p^{\times} \in \mathcal{N}_Y^{\times}(p) \;].$$

Proof. Unassigned HW. $\square$

REMARK 41.15. *Let Y be a metric space, $p \in Y$, $U \subseteq Y$. Then:*

$$[\; U \in \mathcal{N}_Y^{\circ}(p) \;] \;\Leftrightarrow\; [\; \exists A \in \mathcal{B}_Y^{\times}(p) \textit{ s.t. } A \subseteq U \;].$$

Proof. Unassigned HW. $\square$

REMARK 41.16. *Let $p \in \mathbb{R}$ and $A \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$. Then $p \in \mathrm{LP}_{\mathbb{R}} A$.*

Proof. Unassigned HW. $\square$

The following is **L'Hospital's Rule, continuous version**

THEOREM 41.17. *Let $x, y : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. Assume:*

(1) *$dom[y'/x'] \in \mathcal{N}_{\mathbb{R}}^{\circ}(p)$,*
(2) *x and y are continuous at p, and*
(3) *$x(p) = 0 = y(p)$.*

Then $\lim\limits_p (y/x) \;=^*\; \lim\limits_p (y'/x')$.

Proof. Since $\mathrm{dom}[y'/x'] \in \mathcal{N}_{\mathbb{R}}^{\circ}(p)$, by Remark 41.15, choose $A \in \mathcal{B}_{\mathbb{R}}^{\times}(p)$ s.t. $A \subseteq \mathrm{dom}[y'/x']$. Then, by Lemma 41.12, we have $A \subseteq \mathrm{dom}[y/x]$. Let $f := y/x$, $g := y'/x'$. Then $A \subseteq \mathrm{dom}[f]$. Want: $\lim\limits_p f \;=^*\; \lim\limits_p g$.

By Lemma 41.16, $p \in \mathrm{LP}_{\mathbb{R}}A$. Since $A \subseteq \mathrm{dom}[f]$, $\mathrm{LP}_{\mathbb{R}}A \subseteq \mathrm{LPD}_{\mathbb{R}}f$. Since $p \in \mathrm{LP}_{\mathbb{R}}A \subseteq \mathrm{LPD}_{\mathbb{R}}f$, and since $\mathbb{R}$ is Hausdorff, by Lemma 41.3, it suffices to show that f follows g near p. Then, by Lemma 41.11, we wish to show: $\forall t \in A$, $\exists s \in (p|t)$ s.t. $g(s) = f(t)$. Let $t \in A$ be given. We wish to show: $\exists s \in (p|t)$ s.t. $g(s) = f(t)$.

By Theorem 40.8 (with α replaced by p, β by t and c by s), choose $s \in (p|t)$ s.t. $(y'/x')(s) = (DQ_{x,y})(p,t)$. We want: $g(s) = f(t)$.

We have

$$\begin{aligned} g(s) &= (y'/x')(s) = (DQ_{x,y})(p,t) \\ &= \frac{[y(t)] - [y(p)]}{[x(t)] - [x(p)]} = \frac{[y(t)] - 0}{[x(t)] - 0} = \frac{y(t)}{x(t)} \\ &= (y/x)(t) = f(t), \end{aligned}$$

as desired. □

THEOREM 41.18. *Let Y, Z be topological spaces, $\phi, \psi : Y \dashrightarrow Z$. Let $p \in Y$. Let $A \in \mathcal{B}_Y^{\times}(p)$. Assume $\phi = \psi$ on A. Then $\lim\limits_p \phi = \lim\limits_p \psi$.*

Proof. It suffices to show: $\mathrm{LIMS}_p\, \phi = \mathrm{LIMS}_p\, \psi$. Want:

(1) $\mathrm{LIMS}_p\, \phi \subseteq \mathrm{LIMS}_p\, \psi$ and

(2) $\mathrm{LIMS}_p\, \phi \supseteq \mathrm{LIMS}_p\, \psi$.

Proof of (1): We wish to show: $\forall q \in \mathrm{LIMS}_p\, \phi$, $q \in \mathrm{LIMS}_p\, \psi$. Let $q \in \mathrm{LIMS}_p\, \phi$ be given. We wish to show: $q \in \mathrm{LIMS}_p\, \psi$. Since $q \in \mathrm{LIMS}_p\, \phi$, we get $\phi \to q$ near p. We wish to show: $\psi \to q$ near p. We wish to show: $\forall W \in \mathcal{N}_Z$, $\exists U \in \mathcal{N}_Y^{\times}(p)$ s.t. $\psi_*(U) \subseteq W$. Let $W \in \mathcal{N}_Z$ be given. We wish to show: $\exists U \in \mathcal{N}_Y^{\times}(p)$ s.t. $\psi_*(U) \subseteq W$.

Since $\phi \to q$ near p, choose $V \in \mathcal{N}_Y^{\times}(p)$ s.t. $\phi_*(V) \subseteq W$. Since $A \in \mathcal{N}_Y^{\times}(p)$ and $V \in \mathcal{N}_Y^{\times}(p)$, it follows that $A \cap V \in \mathcal{N}_Y^{\times}(p)$. Let $U := A \cap V$. We wish to show: $\psi_*(U) \subseteq W$.

Since $U = A \cap V \subseteq A$ and since $\phi = \psi$ on A, it follows that $\phi_*(U) = \psi_*(U)$. Since $U = A \cap V \subseteq V$, we get $\phi_*(U) \subseteq \phi_*(V)$. Then $\psi_*(U) = \phi_*(U) \subseteq \phi_*(V) \subseteq W$, as desired. *End of proof of (1).*

Proof of (2): Unassigned HW, similar to (1). *End of proof of (2).* □

The following is **L'Hospital's Rule, limit version**

THEOREM 41.19. *Let* $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ *and let* $p \in \mathbb{R}$. *Assume:*

(A) $dom[g'/f'] \in \mathcal{N}^{\circ}_{\mathbb{R}}(p)$ *and*
(B) $\lim\limits_{p} f = 0 = \lim\limits_{p} g$.

Then $\lim\limits_{p}(g/f) \;\;=^*\;\; \lim\limits_{p}(g'/f')$.

Proof. Let $x := \text{adj}^0_p f$, $y := \text{adj}^0_p g$. Then x and y are both continuous at p. Also, $x(p) = 0 = y(p)$. Let $U := \mathbb{R}^{\times}_p$. Then U is open in $\mathbb{R}$ and $U \in \mathcal{N}^{\times}_{\mathbb{R}}(p)$. As $x = f$ on U, by HW#3-4, we have: $x' = f'$ on U. As $y = g$ on U, by HW#3-4, we have: $y' = g'$ on U. Then $y/x = g/f$ on U. Also, $y'/x' = g'/f'$ on U. Then, by Theorem 41.18, we get $\lim\limits_{p}(y/x) = \lim\limits_{p}(g/f)$ and $\lim\limits_{p}(y'/x') = \lim\limits_{p}(g'/f')$. Let $S := \text{dom}[y'/x']$, $T := \text{dom}[g'/f']$. Since $y'/x' = g'/f'$ on U, it follows that $S \cap U = T \cap U$. By assumption, $T \in \mathcal{N}^{\circ}_{\mathbb{R}}(p)$. So, since $U \in \mathcal{N}^{\times}_{\mathbb{R}}(p)$, we get $T \cap U \in \mathcal{N}^{\times}_{\mathbb{R}}(p)$. So, since $T \cap U = S \cap U \subseteq S$, we get $S \in \mathcal{N}^{\circ}_{\mathbb{R}}(p)$. That is, $\text{dom}[y'/x'] \in \mathcal{N}^{\circ}_{\mathbb{R}}(p)$. Then, by Theorem 41.17, we have $\lim\limits_{p}(y/x) \;\;=^*\;\; \lim\limits_{p}(y'/x')$. So, since $\lim\limits_{p}(y/x) = \lim\limits_{p}(g/f)$ and $\lim\limits_{p}(y'/x') = \lim\limits_{p}(g'/f')$, we get $\lim\limits_{p}(g/f) \;\;=^*\;\; \lim\limits_{p}(g'/f')$, as desired. $\square$

After Spring Break, we will discuss the necessity of the hypotheses to Theorem 41.17 and Theorem 41.19. We will also show how to use L'Hospital's Rule to prove the "fourth derivative test".

Then we will begin analysis on functions of more than one variable.

42. Class 16 on 20 March 2018, Tu of Week 9

Midterm on Thursday 22 March.

Recall: Let S be a set and let W be a vector space. Then W^S is a vector space under the linear operations given by

$$(f+g)(x) \;=\; [f(x)] + [g(x)] \qquad \text{and}$$
$$(cf)(x) \;=\; c \cdot [f(x)].$$

DEFINITION 42.1. *Let* V *and* W *be vector spaces. Then*

$$L(V, W) \qquad := \qquad \{\, T \in W^V \mid T \text{ is linear} \,\}.$$

REMARK 42.2. *Let* V *and* W *be vector spaces. Then* $L(V, W)$ *is a vector subspace of* W^V.

Proof. Unassigned HW. $\square$

DEFINITION 42.3. *For any vector space X,*

$$\begin{aligned} \mathcal{T}_X &:= \{\textit{open subsets of } X\}, \qquad \textit{and} \\ \mathcal{T}_X^c &:= \{\textit{closed subsets of } X\}. \end{aligned}$$

For any metric space X, d_X denotes the metric on X. For any normed vector space V, $|\bullet|_V$ denotes the norm on V.

DEFINITION 42.4. *Let V be a normed vector space. Then*

$$\begin{aligned} S_V &:= \{u \in V \textit{ s.t. } |u|_V = 1\}, \\ B_V &:= \{x \in V \textit{ s.t. } |x|_V < 1\} \qquad \textit{and} \\ \overline{B}_V &:= \{x \in V \textit{ s.t. } |x|_V \leqslant 1\}. \end{aligned}$$

FACT 42.5. *Let V be a nonzero normed vector space and let $x \in V$. Then there exists $a \geqslant 0$ and $u \in S_V$ s.t. $x = au$.*

Proof. Let $a := |x|_V$. We wish to show: $\exists u \in S_V$ s.t. $x = au$.

One of the following must be true:

(1) $a = 0$ or

(2) $a \neq 0$.

Case (1): Since $V \neq \{0_V\}$, choose $y \in V \backslash \{0_V\}$. Let $b := |y|_V$. Since $y \neq 0_V$, we get $b \neq 0$. Then $|y/b|_V = |y|_V/b = b/b = 1$, so $y/b \in S_V$. Let $u := y/b$. We wish to show: $x = au$.

Since $|x|_V = a = 0$, we have $x = 0_V$. Then $x = 0_V = 0 \cdot u = au$, as desired. *End of Case (1).*

Case (2): Since $|x/a|_V = |x|_V/a = a/a = 1$, we get $x/a \in S_V$. Let $u := x/a$. We wish to show: $x = au$.

We have $x = a \cdot (x/a) = au$, as desired. *End of Case (2).* $\square$

For any metric space X, let $\mathcal{B}_X := \mathcal{B}_{d_X}$ be the set of balls in X.

Recall (Definition 23.1): Let X be a metric space and let $A \subseteq X$. Then A is **bounded** in X means: $\exists B \in \mathcal{B}_X$ s.t. $A \subseteq B$.

The next lemma will be called the **Recentering Up Lemma**.

LEMMA 42.6. *Let X be a metric space, let $B \in \mathcal{B}_X$ and let $p \in X$. Then $\exists C \in \mathcal{B}_X(p)$ such that $B \subseteq C$.*

Proof. Choose $q \in X$ and $r > 0$ s.t. $B = B_X(q,r)$. Let $s := d_X(p,q)$. Let $C := B_X(p, r+s)$. Then $C \in \mathcal{B}_X(p)$, and we wish to prove: $B \subseteq C$. We want: $\forall z \in B$, $z \in C$. Let $z \in B$ be given. We want: $z \in C$.

Since $z \in B = B_X(q,r)$, we get $d_X(z,q) < r$. By the triangle inequality, $d_X(z,p) \leqslant [d_X(z,q)] + [d_X(q,p)]$. So, since $d_X(z,q) < r$ and $d_X(q,p) = d_X(p,q) = s$, we get $d_X(z,p) < r+s$. Then $z \in B_X(p,r+s)$. That is $z \in C$, as desired. □

DEFINITION 42.7. *Let V be a normed vector space and let $A \subseteq V$. Then $|A|_V := \{|x|_V \text{ s.t. } x \in A\}$.*

REMARK 42.8. *Let V be a normed vector space and let $A \subseteq V$. Then:* $\quad$ *(A is bounded in V)* $\quad\Leftrightarrow\quad$ *($\sup |A|_V < \infty$)*

Proof. Proof of $\Rightarrow$: Assume A is bounded in V. Want: $\sup |A|_V < \infty$.

Choose $x \in V$ and $r > 0$ s.t. $A \subseteq B_V(x,r)$. By the Recentering Up Lemma (Lemma 42.6), choose $s > 0$ s.t. $B_V(x,r) \subseteq B_V(0_V,s)$. Then $A \subseteq B_V(x,r) \subseteq B_V(0_V,s) \subseteq \overline{B}_V(0_V,s)$, so $|A|_V \subseteq |\overline{B}(0_V,s)|_V$.

Then $|A|_V \subseteq |\overline{B}(0_V,s)|_V \leqslant s$, so $\sup |A|_V \leqslant s$. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume $\sup |A|_V < \infty$. Want: A is bounded in V.

Let $s := \sup |A|_V$. Then $|A|_V \leqslant s$, and so we have $A \subseteq \overline{B}_V(0_V,s)$. Let $B := B_V(0_V,s+1)$. Since $B \in \mathcal{B}_V$ and since

$$A \quad\subseteq\quad \overline{B}_V(0_V,s) \quad\subseteq\quad B_V(0_V,s+1) \quad=\quad B,$$

we conclude that A is bounded in V. *End of proof of $\Leftarrow$.* □

DEFINITION 42.9. *Let V and W be normed vector spaces and let $T \in L(V,W)$. Then*

$$\begin{aligned} E\ell\ell_T &:= T_*(S_V), \\ \widehat{T} &:= \sup |\, E\ell\ell_T \,|_W \qquad \textit{and} \\ \check{T} &:= \inf |\, E\ell\ell_T \,|_W. \end{aligned}$$

Let V and W be normed vector spaces and let $T \in L(V,W)$. For clarity, we sometimes write $E\ell\ell_T^{VW}$ for $E\ell\ell_T$. For clarity, we sometimes write $\widehat{T}_{VW}$ for $\widehat{T}$. For clarity, we sometimes write $\check{T}_{VW}$ for $\check{T}$.

DEFINITION 42.10. *Let V and W be normed vector spaces and let $T \in L(V,W)$. Then*

(1) T *is* **bounded** *means: $\widehat{T} < \infty$,*
(2) T *is* **bounded below** *means: $\check{T} > 0$,*
(3) $\forall K \geqslant 0$, $\quad T$ *is* K**-bounded** *means: $\widehat{T} \leqslant K$,* $\qquad$ *and*
(4) $\forall \varepsilon > 0$, $\quad T$ *is* ε**-bounded below** *means: $\check{T} \geqslant \varepsilon$.*

Assigned HW#9-1 and HW#9-2.

HW#9-1 is a quantified equivalence for K-bounded. Here's another:

REMARK 42.11. *Let V, W be normed vector spaces, $T \in L(V,W)$, $K \geqslant 0$. Then:* $[\ T \textit{ is } K\textit{-bounded}\ ] \Leftrightarrow [\ \forall u \in S_V,\ |Tu|_W \leqslant K\ ]$.

Proof. Unassigned HW. □

Assigned HW#9-3.

REMARK 42.12. *Let Y and Z be metric spaces. Let $f : Y \dashrightarrow Z$ be Lipschitz. Let $A \subseteq Y$ be bounded. Then $f_*(A)$ is bounded in Z.*

Proof. Choose $K > 0$ s.t. f is K-Lipschitz. Since A is bounded in Y, choose $x \in Y$ and $r > 0$ s.t. $A \subseteq B_Y(x,r)$. By HW#9-3, we have $f_*(A) \subseteq B_Z(f(x), Kr)$. Then $f_*(A)$ is bounded in Z, as desired. □

Assigned HW#9-4.

Recall: Let M and N be metric spaces and let $f : M \dashrightarrow N$. Then:

$$\begin{aligned}(\, f \text{ is Lipschitz}\,) \quad &\Rightarrow \quad (\, f \text{ is uniformaly continuous}\,)\\ &\Rightarrow \quad (\, f \text{ is continuous}\,).\end{aligned}$$

THEOREM 42.13. *Let V and W be normed vector spaces and let $T \in L(V,W)$. Then the following are all equivalent:*

(1) *T is continuous at 0_V.*
(2) *T is bounded.*
(3) *T is Lipschitz.*
(4) *T is uniformly continuous.*
(5) *T is continuous.*

Proof. It suffices to show:

$$(1) \quad \Rightarrow \quad (2) \quad \Rightarrow \quad (3) \quad \Rightarrow \quad (4) \quad \Rightarrow \quad (5) \quad \Rightarrow \quad (1).$$

By HW#9-4, (1)⇒(2). By HW#9-1, (2)⇒(3). Since Lipschitz implies uniformly continuous, we have (3)⇒(4). Since uniformly continuous implies continuous, we have (4)⇒(5). Want: (5)⇒(1).

Since $0_V \in V = \mathrm{dom}[T]$, we have (5)⇒(1), as desired. □

We now return to L'Hospital's Rule.

Recall Remark 41.15: Let Z be a metric space, $S \subseteq Z$, $p \in Z$. Then:

$$[\ S \in \mathcal{N}_Z^\circ(p)\ ] \quad \Leftrightarrow \quad [\ \exists \delta > 0 \text{ s.t. } B_Z^\times(p,\delta) \subseteq S\ ].$$

Recall our last version of L'Hospital's Rule:

THEOREM 42.14. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. Assume:*

(A) $\exists \delta > 0$ *s.t.* $B^{\times}_{\mathbb{R}}(p, \delta) \subseteq dom[g'/f']$ *and*

(B) $\lim\limits_{p} f = 0 = \lim\limits_{p} g$.

Then $\lim\limits_{p}(g/f) \; =^* \; \lim\limits_{p}(g'/f')$.

In the proof given, you should take $\lim\limits_{p}$ to mean $(\mathbb{R}^*, \mathbb{R}^*)$-$\lim\limits_{p}$. So, for exmaple, if $\lim\limits_{p}(g'/f') = \infty$, then $\lim\limits_{p}(g/f) = \infty$.

We next address the question of whether, in Theorem 42.14, we can replace "$=^*$" by "$=$", and the answer is no: Define $g_\times : \mathbb{R} \dashrightarrow \mathbb{R}$ by $g_\times(t) = t^2 \cdot [\sin(t^{-3})]$. Let $f := \mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ and let $g := \mathrm{adj}^0_0(g_\times) : \mathbb{R} \to \mathbb{R}$. We have studied g before, and we leave it as an unassigned exercise to show that

$$\lim_{0}(g/f) = 0 \qquad \text{and} \qquad \lim_{0}(g'/f') = \frown.$$

Thus the contingency on the equality in Theorem 42.14 is needed.

We next address whether, in Theorem 42.14, we can replace (A) by

(A') $p \in \mathrm{LPD}_{\mathbb{R}}(g'/f')$,

and the answer is no: Let $p := 0$. In class, we described piecewise linear functions f and g such that (A') and (B) hold, but

$$\lim_{0}(g/f) = 1 \qquad \text{and} \qquad \lim_{0}(g'/f') = 0.$$

Specifically: take any decreasing sequence $a \in \mathbb{R}^{\mathbb{N}}$ such that $a_\bullet \to 0$. Define g to be

$$\begin{gathered} 1 \text{ on } [a_1|a_2], \quad 1/2 \text{ on } [a_3|a_4], \quad 1/3 \text{ on } [a_5|a_6], \quad \textit{etc.} \\ \text{linear on } [a_2|a_3], \quad \text{linear on } [a_4|a_5], \quad \textit{etc.}. \end{gathered}$$

Then define f to be

$$\begin{gathered} 1 \text{ at } a_1, \quad 1/2 \text{ on } [a_2|a_3], \quad 1/3 \text{ on } [a_4|a_5], \quad \textit{etc.} \\ \text{linear on } [a_1|a_2], \quad \text{linear on } [a_3|a_4], \quad \textit{etc.} \end{gathered}$$

We graphed g/f, and observed that $\lim\limits_{0}(g/f) = 1$, although we did not write out a full symbolic proof. We noted that

$$\mathrm{dom}[g'/f'] \quad = \quad (a_1|a_2) \;\cup\; (a_3|a_4) \;\cup (a_5|a_6) \;\cup \cdots .$$

We noted also that ($\forall t \in \mathrm{dom}[g'/f']$, $(g'/f')(t) = 0$). From this, we have: $0 \in \mathrm{LPD}_{\mathbb{R}}(g'/f')$ and $\lim\limits_{0}(g'/f') = 0$.

We call Theorem 42.14 the "0/0 near p version of L'Hospital's Rule". We wish to state the "$\infty/(-\infty)$ near p^+ version of L'Hospital's Rule". First, we define right handed limits, denoted $\lim\limits_{p^+}$.

DEFINITION 42.15. *Let Y be a topological space, let $\phi : \mathbb{R}^* \dashrightarrow Y$ and let $p \in \mathbb{R}$. Then*

(1) *$\forall q \in Y$, by $\phi \to q$ near p^+, we mean:*
$$\forall V \in \mathcal{N}_Y(q),\ \exists \delta > 0\ s.t.,\ \forall x \in dom[\phi],$$
$$[\ p < x < p + \delta\ ] \quad \Rightarrow \quad [\ \phi(x) \in V\ ],$$

(2) $\operatorname*{LIMS}\limits_{p^+} \phi \ := \ \{ q \in Y \,|\, \phi \to q \text{ near } p^+ \}$ *and*

(3) $\lim\limits_{p^+} \phi \ := \ \mathrm{ELT}(\operatorname*{LIMS}\limits_{p^+} \phi)$.

The following is the $\infty/(-\infty)$ near p^+ version of L'Hospital's Rule:

THEOREM 42.16. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. Assume:*

(A) $\exists \delta > 0$ *s.t.* $(p, p + \delta) \subseteq dom[g'/f']$ *and*

(B) $\lim\limits_{p^+} f = -\infty$ *and* $\lim\limits_{p^+} g = \infty$.

Then $\lim\limits_{p^+}(g/f) \ =^* \ \lim\limits_{p^+}(g'/f')$.

Proof. Omitted. ☐

The following is the $(-\infty)/\infty$ near $-\infty$ version of L'Hospital's Rule:

THEOREM 42.17. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$. Assume:*

(A) $\exists N \in \mathbb{R}$ *s.t.* $(-\infty, N) \subseteq dom[g'/f']$ *and*

(B) $\lim\limits_{-\infty} f = \infty$ *and* $\lim\limits_{-\infty} g = -\infty$.

Then $\lim\limits_{-\infty}(g/f) \ =^* \ \lim\limits_{-\infty}(g'/f')$.

Proof. Omitted. ☐

The following is the 0/0 near p^+ version of L'Hospital's Rule:

THEOREM 42.18. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. Assume:*

(A) $\exists \delta > 0$ *s.t.* $(p, p + \delta) \subseteq dom[g'/f']$ *and*

(B) $\lim\limits_{p^+} f = 0$ *and* $\lim\limits_{p^+} g = 0$.

Then $\lim\limits_{p^+}(g/f) \ =^* \ \lim\limits_{p^+}(g'/f')$.

Proof. Omitted. ☐

Theorem 42.14 is the "0/0 near p version of L'Hospital's Rule". We can replace "0/0" by any of the following:

$$0/0, \quad \infty/\infty \quad (-\infty)/\infty \quad \infty/(-\infty) \quad (-\infty)/(-\infty).$$

We can replace "near p" by any of the following:

$$\text{near } -\infty, \qquad \text{near } p^-, \qquad \text{near } p, \qquad \text{near } p^+, \qquad \text{near } \infty.$$

Thus we obtain $5 \times 5 = 25$ different versions of L'Hospital's Rule, all of them true. In all of them, lim means $(\mathbb{R}^*, \mathbb{R}^*)$- lim.

We can't prove all 25 versions, for lack of time. The five 0/0 versions are not so hard to prove, and we'll indicate how below. The proofs of the other 20 are similar, in that they are based on the Cauchy Mean Value Corollary (Theorem 40.8). However, they are technically more complicated than the proofs of the 0/0 versions.

Currently (21 March 2018), all 25 proofs are unified into one in the Wikipedia article "L'Hôpital's rule", in the section "General proof".

FACT 42.19. *Let $h : \mathbb{R} \dashrightarrow \mathbb{R}$. Let $\eta : (0, \infty) \dashrightarrow \mathbb{R}$ be defined by $\eta(t) = h(1/t)$. Then both of the following are true:*

(1) $\displaystyle\lim_{\infty} h \;=\; \lim_{0^+} \eta$ *and*

(2) $\forall t \in (0, \infty)$, $\quad \eta'(t) = -\dfrac{h'(1/t)}{t^2}$.

Proof. Unassigned HW. □

Assigned HW#9-5.

Hint: Theorem 42.18 and Fact 42.19 are both useful in HW#9-5.

43. Class 17 on 27 March 2018, Tu of Week 10

Recall (Fact 42.5): Let V be a nonzero normed vector space, $x \in V$. Then there exists $a \geqslant 0$ and $u \in S_V$ s.t. $x = au$.

Recall: Let V and W be normed vector spaces, $T \in L(V, W)$. Then

(1) T is **bounded** means: $\widehat{T} < \infty$,
(2) T is **bounded below** means: $\check{T} > 0$,
(3) $\forall K \geqslant 0$, $\quad T$ is **K-bounded** means: $\widehat{T} \leqslant K$, and
(4) $\forall \varepsilon > 0$, $\quad T$ is **ε-bounded below** means: $\check{T} \geqslant \varepsilon$.

Assigned HW#10-1.

HW#10-1 is a quant. equiv. for ε-bounded below. Here's another:

REMARK 43.1. *Let V, W be normed vector spaces, $T \in L(V,W)$, and $\varepsilon > 0$. Then: $[\,T$ is ε-bounded below$\,] \Leftrightarrow [\,\forall u \in S_V,\ |Tu|_W \geqslant \varepsilon\,]$.*

Proof. Unassigned HW. $\square$

Let V be a vector space, let $|\bullet| \in \mathcal{N}(V)$ and let $K \geqslant 0$. We define $K|\bullet| : V \to [0,\infty)$ by $(K|\bullet|)(v) = K|v|$. Note:

$$[\quad K \ > \ 0 \quad] \qquad \Rightarrow \qquad [\quad K|\bullet| \ \in \ \mathcal{N}(V) \quad].$$

For any vector space V, for any $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$, for any $K > 0$, by $|\bullet| \leqslant K\|\bullet\|$, we mean: $\forall x \in V$, $|x| \leqslant K\|x\|$.

REMARK 43.2. *Let V be a vector space, let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$ and let $K > 0$. Let $I := \mathrm{id}_V : V \to V$. Let $V' := (V, |\bullet|)$ and $V'' := (V, \|\bullet\|)$. Then: $[\,|\bullet| \leqslant K\|\bullet\|\,] \Leftrightarrow [\,I : V'' \to V'$ is K-bounded$\,]$.*

Proof. Proof of $\Rightarrow$: Assume that $|\bullet| \leqslant K\|\bullet\|$. We wish to prove: $I : V'' \to V'$ is K-bounded. We want: $\forall x \in V''$, $|I(x)|_{V'} \leqslant K|x|_{V''}$. Let $x \in V''$ be given. We want to show: $|I(x)|_{V'} \leqslant K|x|_{V''}$.

Because $|\bullet| \leqslant K\|\bullet\|$, we get $|x| \leqslant K\|x\|$. Also, $Ix = \mathrm{id}_V(x) = x$. Then $|Ix|_{V'} = |x|_{V'} = |x| \leqslant K\|x\| = K|x|_{V''}$. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume that $I : V'' \to V'$ is K-bounded. We wish to prove: $|\bullet| \leqslant K\|\bullet\|$. We want to show: $\forall x \in V$, $|x| \leqslant K\|x\|$. Let $x \in V$ be given. We want to prove: $|x| \leqslant K\|x\|$.

As $I : V'' \to V'$ is K-bounded, $|Ix|_{V'} \leqslant K|x|_{V''}$. Also, $Ix = \mathrm{id}_V(x) = x$. Then $|x| = |x|_{V'} = |Ix|_{V'} \leqslant K|x|_{V''} = K\|x\|$. *End of proof of $\Leftarrow$.* $\square$

DEFINITION 43.3. *Let V be a vector space and let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Then, by $|\bullet| << \|\bullet\|$, we mean: $\exists K > 0$ s.t. $|\bullet| \leqslant K\|\bullet\|$.*

REMARK 43.4. *Let V be a vector space and let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Let $I := \mathrm{id}_V : V \to V$. Let $V' := (V, |\bullet|)$ and $V'' := (V, \|\bullet\|)$. Then: $[\,|\bullet| << \|\bullet\|\,] \Leftrightarrow [\,I : V'' \to V'$ is bounded$\,]$.*

Proof. Proof of $\Rightarrow$: Assume that $|\bullet| << \|\bullet\|$. We wish to show that: $I : V'' \to V'$ is bounded.

Since $|\bullet| << \|\bullet\|$, choose $K > 0$ s.t. $|\bullet| \leqslant K\|\bullet\|$. Then, by $\Rightarrow$ of Remark 43.2, $I : V'' \to V'$ is K-bounded. Then $I : V'' \to V'$ is bounded. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume: $I : V'' \to V'$ is bounded. Want: $|\bullet| << \|\bullet\|$.

Let $J := \widehat{I}_{V'',V'}$. Since $I : V'' \to V'$ is bounded, we get $J < \infty$. Then $J \in [0,\infty)$. Let $K := J + 1$. Then $K > 0$. Want: $|\bullet| \leqslant K\|\bullet\|$.

Since $\widehat{I}_{V'',V'} = J < K$, it follows that $I : V'' \to V'$ is K-bounded. Then, by $\Leftarrow$ of Remark 43.2, $|\bullet| \leqslant K\|\bullet\|$. *End of proof of* $\Leftarrow$. $\square$

Recall: Let $m \in \mathbb{N}$. Then:

(1) $\forall p \geqslant 1$, define $|\bullet|_{m,p} \in \mathcal{N}(\mathbb{R}^m)$ by $|x|_{m,p} = [|x_1|^p + \cdots + |x_m|^p]^{1/p}$,
&(2) define $|\bullet|_{\infty,p} \in \mathcal{N}(\mathbb{R}^m)$ by $|x|_{\infty,p} = \max\{|x_1|, \ldots, |x_m|\}$.

For all $p \in [1, \infty]$, for all $x \in \mathbb{R}^m$, in the notation $|x|_{m,p}$, we often omit the subscript m, and write $|x|_p$. The standard norm on $\mathbb{R}^m$ is $|\bullet|_2$.

We drew pictures in $\mathbb{R}^2$ of $S_{|\bullet|_\infty}$ and $S_{|\bullet|_2}$ and $S_{|\bullet|_1}$. We observed:

$$B_{|\bullet|_\infty} \ \supseteq \ B_{|\bullet|_2} \ \supseteq \ B_{|\bullet|_1} \ \supseteq \ [1/2] \cdot B_{|\bullet|_\infty} \ = \ B_{2|\bullet|_\infty}.$$

By the "compensation principle" (big norms have small balls and small norms have big balls), we concluded:

$$|\bullet|_\infty \ \leqslant \ |\bullet|_2 \ \leqslant \ |\bullet|_1 \ \leqslant \ 2|\bullet|_\infty.$$

More precisely: $|\bullet|_{2,\infty} \leqslant |\bullet|_{2,2} \leqslant |\bullet|_{2,1} \leqslant 2|\bullet|_{2,\infty}$.

Assigned HW#10-2.

From HW#10-2, we concluded: $\forall m \in \mathbb{N}$,

$$|\bullet|_{m,\infty} \quad << \quad |\bullet|_{m,2} \quad << \quad |\bullet|_{m,1} \quad << \quad |\bullet|_{m,\infty}.$$

DEFINITION 43.5. *Let V be a vector space and let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Then, by $|\bullet| \approx \|\bullet\|$, we mean: both $|\bullet| << \|\bullet\|$ and $\|\bullet\| << |\bullet|$.*

Let V be a vector space, $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Then, by $|\bullet|$ and $\|\bullet\|$ are **equivalent**, we mean: $|\bullet| \approx \|\bullet\|$.

From HW#10-2, we see: $\forall m \in \mathbb{N}$, $|\bullet|_{m,\infty} \approx |\bullet|_{m,2} \approx |\bullet|_{m,1}$. We will eventually show that, on any finite dimensional vector space, any two norms are equivalent.

REMARK 43.6. *Let V be a vector space, let Z be a set and let $p \in Z$. Let $\mu : V \dashrightarrow Z$. Then $dom[\mu(p + \bullet)] = (dom[\mu]) - p$.*

Proof. Unassigned HW. $\square$

LEMMA 43.7. *Let $\phi, \psi : \mathbb{R} \dashrightarrow \mathbb{R}$. Let $q \in \mathrm{LPD}_{\mathbb{R}}(\psi/\phi)$. Then $\lim_q(\psi/\phi) =^* [\lim_q \psi]/[\lim_q \phi]$.*

Proof. Unassigned HW. *Hint:* Use Theorem 20.6. $\square$

Assigned HW#10-3, HW#10-4 and HW#10-5.

We will call the next result **Easy L'Hospital's Rule**.

THEOREM 43.8. *Let* $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$, *and let* $p \in \mathrm{LPD}_{\mathbb{R}}(g/f)$. *Assume:* $f(p) = 0 = g(p)$. *Then:* $\lim\limits_p (g/f) =^* [g'(p)]/[f'(p)]$.

Proof. Let $R := [g'(p)]/[f'(p)]$. Want: $\lim\limits_p (g/f) =^* R$.

Let $\lambda := (g/f)(p + \bullet)$. Then, by Remark 43.6, we conclude that $\mathrm{dom}[\lambda] = (\mathrm{dom}[g/f]) - p$. It follows that $\mathrm{LPD}_{\mathbb{R}}\lambda = (\mathrm{LPD}_{\mathbb{R}}(g/f)) - p$. By assumption, $p \in \mathrm{LPD}_{\mathbb{R}}(g/f)$. Then $0 = p - p \in \mathrm{LPD}_{\mathbb{R}}\lambda$.

By HW#10-3, we have $\lim\limits_0 \lambda = \lim\limits_p (g/f)$. We want: $\lim\limits_0 \lambda =^* R$.

Let $\phi := SS_f^p$, $\psi := SS_g^p$. Then $f'(p) = \lim\limits_0 \phi$ and $g'(p) = \lim\limits_0 \psi$. Also $\phi(0) = ☹$ and $\psi(0) = ☹$. Then $(\psi/\phi)(0) = ☹$. By assumption, $f(p) = 0 = g(p)$. Then $[g(p)]/[f(p)] = ☹$. Then

$$\lambda(0) \quad = \quad (g/f)(p + 0) \quad = \quad (g/f)(p) \quad = \quad [g(p)]/[f(p)] \quad = \quad ☹.$$

Then $\lambda(0) = ☹ = (\psi/\phi)(0)$. For all $h \in \mathbb{R}_0^\times$, we have:

$$\begin{aligned} \lambda(h) \quad &= \quad (g/f)(p+h) \quad = \quad \frac{g(p+h)}{f(p+h)} \\ &= \quad \frac{[g(p+h)] - [g(p)]}{[f(p+h)] - [f(p)]} \\ &= \quad \frac{(\ [g(p+h)] - [g(p)]\ )\ /\ h}{(\ [f(p+h)] - [f(p)]\ )\ /\ h} \\ &= \quad \frac{(SS_g^p)(h)}{(SS_f^p)(h)} \quad = \quad \frac{\psi(h)}{\phi(h)} \quad = \quad (\psi/\phi)(h). \end{aligned}$$

So, since $\lambda(0) = (\psi/\phi)(0)$, we see that $\lambda = \psi/\phi$ on $\mathbb{R}$. Since $\mathbb{R}$ is a common superdomain of λ and ψ/ϕ, we conclude that $\lambda = \psi/\phi$. Then $0 \in \mathrm{LPD}_{\mathbb{R}}\lambda = \mathrm{LPD}_{\mathbb{R}}(\psi/\phi)$. Then, by Lemma 43.7, we have $\lim\limits_0(\psi/\phi) =^* [\lim\limits_0 \psi]/[\lim\limits_0 \phi]$. Then

$$\lim_0 \lambda \quad = \quad \lim_0 \frac{\psi}{\phi} \quad =^* \quad \frac{\lim\limits_0 \psi}{\lim\limits_0 \phi} \quad = \quad \frac{g'(p)}{f'(p)},$$

as desired. □

The tame hypothesis in Theorem 41.19 reads: $\mathrm{dom}[g'/f'] \in \mathcal{N}_{\mathbb{R}}^{\circ}(p)$. The tame hypothesis in Theorem 43.8, $p \in \mathrm{LPD}_{\mathbb{R}}(g/f)$, is weaker, and this means that there are times when Theorem 43.8 can be used even though Theorem 41.19 cannot.

However, care is required, because the conclusion of Theorem 43.8 is also weaker, in a substantial way. In particular, in Theorem 43.8, if $f'(p) = 0$, we get no information whatsoever. By contrast, in Theorem 41.19, the contingent equality has $\lim_p (g'/f')$ on the right hand side, and, even if the denominator f' is approaching zero, one can still hope to show that the limit $\lim_p (g'/f')$ exists, possibly by applying L'Hospital's Rule a second time, and studying $\lim_p (g''/f'')$.

DEFINITION 43.9. *Let X be a top. space, $f : X \dashrightarrow \mathbb{R}$, $p \in X$. By f has a* **local unique max** *at p in X, we mean: $\exists V \in \mathcal{N}_X(p)$ s.t.*

$$(V \subseteq dom[f]) \qquad and \qquad (f(p) > f_*(V_p^\times)).$$

By f has a **local unique min** *at p in X, we mean: $\exists V \in \mathcal{N}_X(p)$ s.t.*

$$(V \subseteq dom[f]) \qquad and \qquad (f(p) < f_*(V_p^\times)).$$

The next result is the **Asymptotic Opimization Th'm for umin**.

THEOREM 43.10. *Let Y be a topological space. Let $f, g : Y \dashrightarrow \mathbb{R}$. Let $p \in \mathrm{IntD}_Y(f,g)$. Assume:*

(1) *$f(p) = 0 = g(p)$,*
(2) *$f/g \to 1$ near p and*
(3) *f has a local unique minimum at p in Y.*

Then g has a local unique minimum at p in Y.

Proof. Let $D := \mathrm{dom}[(f,g)]$. Then

$$p \;\in\; \mathrm{IntD}_Y(f,g) \;=\; \mathrm{Int}_Y(\mathrm{dom}[(f,g)]) \;=\; \mathrm{Int}_Y D,$$

so, by HW#3-1, we conclude that $D \in \mathcal{N}_Y(p)$.

By (2), choose $U \in \mathcal{N}_Y^\times(p)$ s.t., $\forall t \in \mathrm{dom}[g/f]$,

$$[\ t \in U\ ] \quad \Rightarrow \quad [\ |\,[(g/f)(t)] - 1\,| < 1/2\ ].$$

Note that $U_p^+ \in \mathcal{N}_Y(p)$ and that $(U_p^+)_p^\times = U$.

By (3), choose $V \in \mathcal{N}_Y(p)$ such that

$$[\ V \subseteq \mathrm{dom}[f]\ ] \quad \text{and} \quad [\ f_*(V_p^\times) > f(p)\ ].$$

Want: $\exists W \in \mathcal{N}_Y(p)$ such that

$$[\ W \subseteq \mathrm{dom}[g]\ ] \quad \text{and} \quad [\ g_*(W_p^\times) > g(p)\ ].$$

Let $W := D \cap (U_p^+) \cap V$. Then $W \in \mathcal{N}_Y(p)$ and we want:

$$[\ W \subseteq \mathrm{dom}[g]\ ] \quad \text{and} \quad [\ g_*(W_p^\times) > g(p)\ ].$$

We have $W \subseteq D = \mathrm{dom}[(f,g)] \subseteq \mathrm{dom}[g]$. Want: $g_*(W_p^\times) > g(p)$. We want: $\forall y \in g_*(W_p^\times)$, $y > g(p)$. Given $y \in g_*(W_p^\times)$. Want: $y > g(p)$.

Since $y \in g_*(W_p^\times)$, choose $t \in W_p^\times$ such that $y = g(t)$. Because $W := D \cap (U_p^+) \cap V$ and $(U_p^+)_p^\times = U$, we get $W_p^\times = (D_p^\times) \cap U \cap (V_p^\times)$. Then $t \in D_p^\times$ and $t \in U$ and $t \in V_p^\times$. Since $t \in D_p^\times \subseteq D = \mathrm{dom}[(f,g)]$, it follows both that $t \in \mathrm{dom}[f]$ and that $t \in \mathrm{dom}[g]$.

By assumption, $f(p) = 0 = g(p)$. Since $t \in \mathrm{dom}[f]$ and $t \in V_p^\times$, we get $f(t) \in f_*(V_p^\times)$. Then $f(t) \in f_*(V_p^\times) > f(p) = 0$, so $f(t) > 0$.

Since $t \in \mathrm{dom}[f]$ and $t \in \mathrm{dom}[g]$ and $f(t) \neq 0$, we conclude that $t \in \mathrm{dom}[g/f]$. So, since $t \in U$, by the choice of U, it follows that $|\,[(g/f)(t)] - 1\,| < 1/2$. Then $1 - (1/2) < (g/f)(t) < 1 + (1/2)$. Then

$$\frac{g(t)}{f(t)} \quad = \quad (g/f)(t) \quad > \quad 1 - (1/2) \quad = \quad 1/2.$$

So, since $f(t) > 0$, we get $g(t) \geqslant [1/2] \cdot [f(t)]$. So, since $f(t) > 0$, we get $g(t) > 0$. Then $y = g(t) > 0 = g(p)$, as desired. $\square$

REMARK 43.11. *Let* $\gamma : \mathbb{R} \dashrightarrow \mathbb{R}$ *and let* $q \in \mathbb{R}$. *Then* $(\gamma - C_{\mathbb{R}}^q)' = \gamma'$.

Proof. Unassigned HW. $\square$

Recall Remark 41.16:

REMARK 43.12. *Let* $p \in \mathbb{R}$ *and let* $A \in \mathcal{B}_{\mathbb{R}}(p)$. *Then* $p \in \mathrm{LP}_{\mathbb{R}}(A_p^\times)$.

Proof. Unassigned HW. $\square$

The following is called the **Fourth Derivative Test for umin**.

THEOREM 43.13. *Let* $\gamma : \mathbb{R} \dashrightarrow \mathbb{R}$ *and let* $p \in \mathrm{IntD}_{\mathbb{R}}(\gamma''')$. *Assume*

(1) $0 = \gamma'(p) = \gamma''(p) = \gamma'''(p)$ *and*
(2) $0 < \gamma''''(p) \neq$ ☹.

Then γ *has a local unique minimum at* p.

Proof. Next class. $\square$

44. Class 18 on 29 March 2018, Th of Week 10

We recalled:

- the Asmptotic Optimization Theorem for umin (Theorem 43.10)
- Easy L'Hospital's Rule (Theorem 43.8)
- various HW problems (HW#10-3, HW#10-4, HW#10-5)
- various remarks (Remark 43.6, Remark 41.16, Remark 43.11)

Here is a restatement of L'Hospital's Rule for 0/0 near p (Theorem 41.19):

THEOREM 44.1. *Let $f, g : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. Assume:*

(A) $A \in \mathcal{B}_p$ *s.t.* $A_p^\times \subseteq dom[g'/f']$ *and*
(B) $\lim\limits_p f = 0 = \lim\limits_p g$.

Then $\lim\limits_p (g/f) \;\; =^* \;\; \lim\limits_p (g'/f')$.

The following is called the **Fourth Derivative Test for umin**.

THEOREM 44.2. *Let $\gamma : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathrm{IntD}_{\mathbb{R}}(\gamma''')$. Assume*

(1) $0 = \gamma'(p) = \gamma''(p) = \gamma'''(p)$ *and*
(2) $0 < \gamma''''(p) \neq ☹$.

Then γ has a local unique minimum at p.

Proof. Let $q := \gamma(p)$ and $g := \gamma - C_{\mathbb{R}}^q$. Then $g(p) = 0$. By Remark 43.11, $g' = \gamma'$. Then $g'' = \gamma''$, $g''' = \gamma'''$, $g'''' = \gamma''''$. Then $p \in \mathrm{IntD}_{\mathbb{R}}(g''')$. Also,

(A) $0 = g(p) = g'(p) = g''(p) = g'''(p)$ and
(B) $0 < g''''(p) \neq ☹$.

By HW#10-5, we wish to show: g has a local unique minimum at p.

Let $\sigma := [1/24][g''''(p)]$. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = \sigma(t-p)^4$. Since $\sigma > 0$, it follows that f has a local unique minimum at p, Because $\mathrm{dom}[g'''] \subseteq \mathrm{dom}[g] = \mathrm{dom}[(f,g)]$, we get $\mathrm{IntD}_{\mathbb{R}}(g''') \subseteq \mathrm{IntD}_{\mathbb{R}}(f,g)$. So, since $p \in \mathrm{IntD}_{\mathbb{R}}(g''')$, we conclude that $p \in \mathrm{IntD}_{\mathbb{R}}(f,g)$. So, since $f(p) = 0 = g(p)$ and since f has a local unique minimum at p, by the Asmptotic Optimization Theorem for umin (Theorem 43.10), it suffices to show: $g/f \to 1$ near p. It therefore suffices to show: $\lim_p (g/f) = 1$.

For all $t \in \mathbb{R}$, we have

$$\begin{array}{rcl} f(t) &=& \sigma(t-p)^4, \\ f'(t) &=& 4\sigma(t-p)^3, \\ f''(t) &=& 12\sigma(t-p)^2, \\ f'''(t) &=& 24\sigma(t-p) \qquad \text{and} \\ f''''(t) &=& 24\sigma. \end{array}$$

Then $f''''(p) = g''''(p) = 24\sigma \neq 0$, so $[g''''(p)]/[f''''(p)] = 1$. We have $\mathrm{dom}[f'] = \mathrm{dom}[f''] = \mathrm{dom}[f'''] = \mathbb{R}$. Also, we have

$$\begin{aligned} \{t \in \mathbb{R} \,|\, f'(t) = 0\} &= \{p\}, \\ \{t \in \mathbb{R} \,|\, f''(t) = 0\} &= \{p\} \qquad \text{and} \\ \{t \in \mathbb{R} \,|\, f'''(t) = 0\} &= \{p\} \end{aligned}$$

Then

$$\begin{aligned} \mathrm{dom}[g'/f'] &= (\mathrm{dom}[g'])_p^\times, \\ \mathrm{dom}[g''/f''] &= (\mathrm{dom}[g''])_p^\times \qquad \text{and} \\ \mathrm{dom}[g'''/f'''] &= (\mathrm{dom}[g'''])_p^\times. \end{aligned}$$

By assumption, $p \in \mathrm{IntD}_{\mathbb{R}}(g''')$. That is, $p \in \mathrm{Int}_{\mathbb{R}}(\mathrm{dom}[g'''])$. By HW#10-4, choose $A \in \mathcal{B}_{\mathbb{R}}(p)$ such that $A \subseteq \mathrm{dom}[g''']$. Then

$$A \quad \subseteq \quad \mathrm{dom}[g'''] \quad \subseteq \quad \mathrm{dom}[g''] \quad \subseteq \quad \mathrm{dom}[g'],$$

so

$$A_p^\times \quad \subseteq \quad (\mathrm{dom}[g'''])_p^\times \quad \subseteq \quad (\mathrm{dom}[g''])_p^\times \subseteq \quad (\mathrm{dom}[g'])_p^\times.$$

That is,

$$A_p^\times \quad \subseteq \quad \mathrm{dom}[g'''/f'''] \quad \subseteq \quad \mathrm{dom}[g''/f''] \subseteq \quad \mathrm{dom}[g'/f'].$$

Then, by three applications of Theorem 44.1, we have

$$\lim_0 (g/f) \quad =^* \quad \lim_0 (g'/f') \quad =^* \quad \lim_0 (g''/f'') \quad =^* \quad \lim_0 (g'''/f''').$$

It therefore suffices to show: $\lim\limits_0 (g'''/f''') = 1$.

By Remark 43.12, $p \in \mathrm{LP}_{\mathbb{R}}(A_p^\times)$. Since $A_p^\times \subseteq \mathrm{dom}[g'''/f''']$, we get $\mathrm{LP}_{\mathbb{R}}(A_p^\times) \subseteq \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[g'''/f'''])$. Then

$$p \;\in\; \mathrm{LP}_{\mathbb{R}}(A_p^\times) \;\subseteq\; \mathrm{LP}_{\mathbb{R}}(\mathrm{dom}[g'''/f''']) \;=\; \mathrm{LPD}_{\mathbb{R}}(g'''/f''').$$

Also, $f'''(p) = 0 = g'''(p)$. Then, by Easy L'Hospital's Rule (Theorem 43.8, with f replaced by f''' and g by g'''), we conclude that

$$\lim_p (g'''/f''') \quad =^* \quad [g''''(p)]/[f''''(p)].$$

Recall that $[g''''(p)]/[f''''(p)] = 1$. Then

$$\lim_p (g'''/f''') \quad =^* \quad [g''''(p)]/[f''''(p)] \quad = \quad 1 \quad \neq \quad ☹,$$

so $\lim\limits_p (g'''/f''') = 1$, as desired. □

Recall: Let V and W be vector spaces and let $T \in L(V,W)$. Then $\ker[T] := T^*(\{0_W\})$.

FACT 44.3. *Let V and W be vector spaces and let $T \in L(V,W)$. Then*

(1) $\ker[T]$ *is a vector subspace of* V,
(2) *im*$[T]$ *is a vector subspace of* W,
(3) (T *is 1-1*) $\Leftrightarrow$ ($\ker[T] = \{0_V\}$) *and*
(4) (T *is onto* W) $\Leftrightarrow$ (*im*$[T] = W$).

Proof. Proof of (1): Unassigned HW. *End of proof of (1).*

Proof of (2): Unassigned HW. *End of proof of (2).*

Proof of (3): Proof of $\Rightarrow$: Unassigned HW. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume $\ker[T] = \{0_V\}$. We wish to show that T is 1-1. That is, we wish to show: $\forall x, y \in V$,

$$[\ Tx = Ty\ ] \quad \Rightarrow \quad [\ x = y\ ].$$

Let $x, y \in V$ be given. We wish to show:

$$[\ Tx = Ty\ ] \quad \Rightarrow \quad [\ x = y\ ].$$

Assume $Tx = Ty$. We wish to show: $x = y$.

We have $T(x - y) = (Tx) - (Ty) = 0_V$. Then

$$x - y \quad \in \quad T^*(\{0_V\}) \quad = \quad \ker[T] \quad = \quad \{0_V\},$$

so $x - y = 0_V$, so $x = y$. *End of proof of $\Leftarrow$. End of proof of (3).*

Proof of (4): Unassigned HW. *End of proof of (4).* □

FACT 44.4. *Let V and W be vector spaces and let $T \in L(V,W)$. Assume that $T : V \to W$ is 1-1. Then $T^{-1} \in L(im[T], V)$.*

Proof. Let $Z := \mathrm{im}[T]$ and let $U := T^{-1}$. We wish to show: $U \in L(Z,V)$. Since $T : V \hookrightarrow\!\!> Z$, we get $U : Z \hookrightarrow\!\!> V$. It remains to show that $U : Z \to V$ is linear. That is, we wish to show:

(1) $\forall a \in \mathbb{R},\ \forall y \in Z, \quad U(ay) = a \cdot [Uy]$, and
(2) $\forall x, y \in Z, \quad U(x + y) = [Ux] + [Uy]$.

Proof of (1): Given $a \in \mathbb{R}$, $y \in Z$. We want to show: $U(ay) = a \cdot [Uy]$. Since T is 1-1, it suffices to show: $T(U(ay)) = T(a \cdot [Uy])$.

Since $U = T^{-1}$, we get $T(U(ay)) = ay$ and $T(Uy) = y$. By linearity of T, $T(a \cdot [Uy]) = a \cdot [T(Uy)]$. Then

$$T(U(ay)) \quad = \quad ay \quad = \quad a \cdot [T(Uy)] \quad = \quad T(a \cdot [Uy]),$$

as desired. *End of proof of (1).*

Proof of (2): Unassigned HW. *End of proof of (2).* □

Unassigned HW: Let V be a vector space, $|\bullet| \in \mathcal{N}(V)$, $K > 0$. Show:

(1) $K|\bullet| \in \mathcal{N}(V)$ and
(2) $B_{K|\bullet|} = [1/K] \cdot B_{|\bullet|}$.

Unassigned HW: Let V be a vector space, $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Show:

$$[\,|\bullet| \leqslant \|\bullet\|\,] \quad \Leftrightarrow \quad [\,B_{|\bullet|} \supseteq B_{\|\bullet\|}\,].$$

Recall (Definition 43.3 and Definition 43.5): Let V be a vector space and let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Then

(1) $|\bullet| << \|\bullet\|$ means: $(\,\exists K > 0$ s.t. $|\bullet| \leqslant K\|\bullet\|\,)$ and
(2) $|\bullet| \approx \|\bullet\|$ means: both $(\,|\bullet| << \|\bullet\|\,)$ and $(\,\|\bullet\| << |\bullet|\,)$.

Recall (HW#10-2): Let $m \in \mathbb{N}$. Then $|\bullet|_{m,1} \approx |\bullet|_{m,2} \approx |\bullet|_{m,\infty}$.

We will eventually show that any two norms on a finite dimensional vector space are equivalent.

LEMMA 44.5. *Let V be a vector space and let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Assume $|\bullet| << \|\bullet\|$. Then $\mathcal{T}_{|\bullet|} \subseteq \mathcal{T}_{\|\bullet\|}$.*

Proof. We wish to show: $\forall X \in \mathcal{T}_{|\bullet|}$, $X \in \mathcal{T}_{\|\bullet\|}$. Let $X \in \mathcal{T}_{|\bullet|}$ be given. We wish to show: $X \in \mathcal{T}_{\|\bullet\|}$. That is, we wish to show: $X \in \langle \mathcal{B}_{\|\bullet\|} \rangle_\cup$. By Quantified Equivalence for union closure, we wish to show: $\forall p \in X$, $\exists U \in \mathcal{B}_{\|\bullet\|}$ s.t. $p \in U \subseteq X$. Let $p \in X$ be given. We wish to show: $\exists U \in \mathcal{B}_{\|\bullet\|}$ s.t. $p \in U \subseteq X$.

We have $X \in \mathcal{T}_{|\bullet|} = \langle \mathcal{B}_{|\bullet|} \rangle_\cup$. Then $p \in X \in \langle \mathcal{B}_{|\bullet|} \rangle_\cup$, so choose $W \in \mathcal{B}_{|\bullet|}$ s.t. $p \in W \subseteq X$. By the Recentering Down Lemma (Lemma 14.2), choose $V \in \mathcal{B}_{|\bullet|}(p)$ s.t. $V \subseteq W$. Since $V \in \mathcal{B}_{|\bullet|}(p)$, choose $r > 0$ s.t. $V = B_{|\bullet|}(p, r)$. By assumption $|\bullet| << \|\bullet\|$, so choose $K > 0$ s.t. $|\bullet| \leqslant K\|\bullet\|$. Let $U := B_{\|\bullet\|}(p, r/K)$. Then $U \in \mathcal{B}_{\|\bullet\|}$, and we wish to show: $p \in U \subseteq X$.

Because $\|p - p\| = 0 < r/K$, we see that $p \in B_{\|\bullet\|}(p, r/K)$. Then $p \in B_{\|\bullet\|}(p, r/K) = U$, and it remains to show: $U \subseteq X$. We wish to show: $\forall q \in U$, $q \in X$. Let $q \in U$ be given. We wish to show $q \in X$.

We have $q \in U = B_{\|\bullet\|}(p, r/K)$, so $\|q - p\| < r/K$, so $K\|q - p\| < r$. Since $|\bullet| \leqslant K\|\bullet\|$, we get $|q-p| \leqslant K\|q-p\|$. Then $|q-p| \leqslant K\|q-p\| < r$, so $q \in B_{|\bullet|}(p, r)$. Then $q \in B_{|\bullet|}(p, r) = V \subseteq W \subseteq X$, as desired. □

COROLLARY 44.6. *Let V be a vector space and let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Assume $|\bullet| \approx \|\bullet\|$. Then $\mathcal{T}_{|\bullet|} = \mathcal{T}_{\|\bullet\|}$.*

Proof. Since $|\bullet| \approx \|\bullet\|$, we have both $|\bullet| << \|\bullet\|$ and $\|\bullet\| << |\bullet|$. Then, by Lemma 44.5, both $\mathcal{T}_{|\bullet|} \subseteq \mathcal{T}_{\|\bullet\|}$. and $\mathcal{T}_{\|\bullet\|} \subseteq \mathcal{T}_{|\bullet|}$. Then $\mathcal{T}_{|\bullet|} = \mathcal{T}_{\|\bullet\|}$. $\square$

COROLLARY 44.7. *Let $m \in \mathbb{N}$. Then $\mathcal{T}_{|\bullet|_{m,1}} = \mathcal{T}_{|\bullet|_{m,2}} = \mathcal{T}_{|\bullet|_{m,\infty}}$.*

Proof. By HW#10-2, $|\bullet|_{m,1} \approx |\bullet|_{m,2} \approx |\bullet|_{m,\infty}$. Then, by Corollary 44.6, we get $\mathcal{T}_{|\bullet|_{m,1}} = \mathcal{T}_{|\bullet|_{m,2}} = \mathcal{T}_{|\bullet|_{m,\infty}}$, as desired. $\square$

Recall: Let $m \in \mathbb{N}$. Then the *standard* topology on $\mathbb{R}^m$ is $\mathcal{T}_{|\bullet|_{m,2}}$. Corollary 44.7 asserts that the topologies coming from the two norms $|\bullet|_{m,1}$ and $|\bullet|_{m,\infty}$ are both equal to the standard topology.

Let V be a finite dimensional vector space. We will eventually show that any two norms on V are equivalent. So, by Corollary 44.6, we will be able to conclude that, even though V has many norms, they all give rise to the one single topology, called the *standard* topology on V. This only works in the world of *finite dimensional* vector spaces. In infinite dimensions, there are many norms *and* many topologies to track, making it a more complicated subject. Fortunately, undergraduate mathematics is focused on finite dimensional vector spaces.

Recall: Let V be a normed vector space. Give V its standard topology $\mathcal{T}_V$. Then $\mathcal{T}_V = \mathcal{T}_{|\bullet|_V} = \langle \mathcal{B}_{|\bullet|_V} \rangle_\cup$. That is, $\mathcal{T}_V$ is the union-closure of the set of $|\bullet|_V$-balls in V. By HW#37 from Fall semester, the map $|\bullet|_V : V \to \mathbb{R}$ is distance-semidecreasing, *i.e.* 1-Lipschitz. Then $|\bullet|_V : V \to \mathbb{R}$ is Lipschitz, hence unif. continuous, hence continuous.

REMARK 44.8. *Let $m \in \mathbb{N}$ and give $\mathbb{R}^m$ its standard topology. Then*

$$\begin{aligned} &|\bullet|_{m,1} : \mathbb{R}^m \to \mathbb{R} \qquad \textit{and} \\ &|\bullet|_{m,2} : \mathbb{R}^m \to \mathbb{R} \qquad \textit{and} \\ &|\bullet|_{m,\infty} : \mathbb{R}^m \to \mathbb{R} \end{aligned}$$

are all continuous.

Proof. Let $\mathcal{T}$ be the standard topology on $\mathbb{R}^m$. Then $\mathcal{T} = \mathcal{T}_{|\bullet|_{m,2}}$, so, by Corollary 44.7, we have $\mathcal{T} = \mathcal{T}_{|\bullet|_{m,1}}$ and $\mathcal{T} = \mathcal{T}_{|\bullet|_{m,\infty}}$.

Since $\mathcal{T} = \mathcal{T}_{|\bullet|_{m,1}}$, $|\bullet|_{m,1} : \mathbb{R}^m \to \mathbb{R}$ is continuous.

Since $\mathcal{T} = \mathcal{T}_{|\bullet|_{m,2}}$, $|\bullet|_{m,2} : \mathbb{R}^m \to \mathbb{R}$ is continuous.

It remains to show: $|\bullet|_{m,\infty} : \mathbb{R}^m \to \mathbb{R}$ is continuous.

Since $\mathcal{T} = \mathcal{T}_{|\bullet|_{m,\infty}}$, $|\bullet|_{m,\infty} : \mathbb{R}^m \to \mathbb{R}$ is continuous, as desired. $\square$

45. Class 19 on 3 April 2018, Tu of Week 11

Assigned HW#11-1.

DEFINITION 45.1. *Let V and W be vector spaces, let $f : V \dashrightarrow W$ and let $q, u \in V$. Then $SS_f^{q,u} : \mathbb{R} \dashrightarrow W$ is defined by*

$$(SS_f^{q,u})(h) \quad = \quad \frac{[f(q+hu)] - [f(q)]}{h}.$$

The fn SS_f^p is the "**secant slope function** of f at p in direction u".

DEFINITION 45.2. *Let V and W be normed vector spaces, let $f : V \dashrightarrow W$ and let $u \in V$. Then $\partial_u f : V \dashrightarrow W$ is defined by*

$$(\partial_u f)(q) \quad = \quad \lim_0 SS_f^{q,u}.$$

The fn $\partial_u f$ is the "**directional derivative** of f in direction u".

DEFINITION 45.3. *Let $m \in \mathbb{N}$, let W be a normed vector space, let $f : \mathbb{R}^m \dashrightarrow W$ and let $j \in [1..m]$. Let e be the standard ordered basis of $\mathbb{R}^m$. Then $\partial_j f := \partial_{e_j} f$.*

The function $\partial_j f$ is called the "jth **partial derivative** of f".

We went through three examples showing that: mere existence of directional derivatives at a point tells us very little, even about continuity of the function at that point.

First, let $V := \mathbb{R}^2$ and let $S := (R \times \{0\}) \cup (\{0\} \times \mathbb{R})$ be the union of the coordinate axes in V. Let $f := \chi_S^V : V \to \mathbb{R}$. Then f is not continuous at 0_2. However, f is "partial differentiable at 0_2", in the sense that: $\forall j \in [1..2]$, $0_2 \in \mathrm{dom}[\partial_j f]$.

Second, let $V := \mathbb{R}^2$ and let $S := \{(x, y) \in \mathbb{R}^2 \,|\, (y = x^2)\&(x > 0)\}$. Let $f := \chi_S^V : V \to \mathbb{R}$. Then f is not continuous at 0_2. However, the function f is "directional differentiable at 0_2", in the sense that: $\forall u \in V$, $0_2 \in \mathrm{dom}[\partial_u f]$.

Third, let $V := \mathbb{R}^2$. We will construct a fuction $f : \mathbb{R}^2 \to \mathbb{R}$ that is "infinitely directional differentiable at every point of $\mathbb{R}^2$", in the sense that: $\forall k \in \mathbb{N}$, $\forall u_1, \ldots, u_k \in V$, $0_2 \in \mathrm{dom}[\partial_{u_1} \cdots \partial_{u_k} f]$. The construction is contained in the following discussion.

Discussion about infinite directional differentiability:

Let $V := \mathbb{R}^2$. For all $k \in \mathbb{N}$, for all $s \in V^k$, for all $\phi : V \dashrightarrow \mathbb{R}$, let $\partial_s \phi := \partial_{s_1} \cdots \partial_{s_k} \phi$. Let $S := V \cup V^2 \cup V^3 \cup \cdots$. We will say that a function $\phi : V \to \mathbb{R}$ is **infinitely directionally differentiable** if, for all $s \in S$, we have $\mathrm{dom}[\partial_s \phi] = V$.

Choose an infinitely differentiable function $g : \mathbb{R} \to \mathbb{R}$ such that ($g = 0$ on $(-\infty, 1] \cup [3, \infty)$) and ($g(2) = 1$) and ($\mathrm{im}[g] \subseteq [0, 1]$). Define $f : V \to \mathbb{R}$ by $f(x, y) = \begin{cases} g(y/x^2), & \text{if } x \neq 0; \\ 0, & \text{if } x = 0. \end{cases}$

FACT 45.4. *The function $f : V \to \mathbb{R}$ is infinitely directionally differentiable, and is discontinuous at 0_2.*

Proof. We define

$$\begin{aligned} U &:= \{(x, y) \in V \,|\, x^2 < y < 3x^2\} \qquad \text{and} \\ \overline{u} &:= \{(x, y) \in V \,|\, x^2 \leqslant y \leqslant 3x^2\}. \end{aligned}$$

Then $0_V = (0, 0) \in \overline{u}$. Also, U is the interior in V of $\overline{u}$, so $V \backslash U$ is the closure in V of $V \backslash \overline{u}$. Let Φ be the set of all $\phi : V \to \mathbb{R}$ such that

- $\phi = 0$ on $V \backslash U$ and
- for all $s \in S$, $\partial_s \phi$ is continuous on $V \backslash \{0_V\}$.

Claim 1: $f \in \Phi$. Proof of Claim 1: By construction, $f = 0$ on $V \backslash U$. Let $s \in S$ be given. We wish to show that $\partial_s \phi$ is continuous on $V \backslash \{0_V\}$.

By the Multivariable Chain Rule (to be proved later), $\forall x, y \in \mathbb{R}$, we have: if $x \neq 0$, then $\partial_s f$ is continuous at (x, y). Let $Y := \{(0, y) | y \in \mathbb{R}\}$ denote the y-axis. Then $\partial_s f$ is continuous on $V \backslash Y$. It therefore suffices to show that $\partial_s \phi$ is continuous on $Y \backslash \{0_V\}$. Let $v \in Y \backslash \{0_V\}$ be given. We wish to show: $\partial_s \phi$ is continuous at v.

Let $V_0 := V \backslash \overline{u}$. We have $v \in Y \backslash \{0_V\} \subseteq V \backslash \overline{U} = V_0$. So, since V_0 is open in V, we see that V_0 is an open nbd in V of v. Since $V_0 \subseteq V \backslash U$ and since $f = 0$ on $V \backslash U$, we see that $f = 0$ on V_0. So, since V_0 is open in V, it follows that $\partial_s \phi = 0$ on V_0. So, since V_0 is an open nbd in V of v, it follows that $\partial_s \phi$ is continuous at v. *End of proof of Claim 1.*

Claim 2: Let $\phi \in \Phi$ and let $v \in V$. Then we have $(\partial_v \phi)(0_V) = 0$. *Proof of Claim 2:* Let $L := \mathbb{R}v \subseteq V$ denote the line through v in V, and give L the relative topology inherited from V. Choose a nbd L_0 of 0_V in L such that $L_0 \subseteq V \backslash U$. Since $\phi = 0$ on $V \backslash U$, we conclude that $\phi = 0$ on L_0. Then $(\partial_v \phi)(0_V) = 0$, as desired. *End of proof of Claim 2.*

Claim 3: Let $\phi \in \Phi$, let $v \in V$ and let $s \in S$. Then $\partial_s \partial_v \phi$ is continuous on $V \backslash \{0_V\}$. *Proof of Claim 3:* Because $s \in S = V \cup V^2 \cup V^3 \cup \cdots$, choose $k \in \mathbb{N}$ such that $s \in V^k$. Let $t := (s_1, \ldots, s_k, v)$. Then we have $t \in V^{k+1} \subseteq S$ and $\partial_t \phi = \partial_s \partial_v \phi$. Since $\phi \in \Phi$, we see that $\partial_t \phi$ is continuous on $V \backslash \{0_V\}$. Then $\partial_s \partial_v \phi$ is continuous on $V \backslash \{0_V\}$, as desired. *End of proof of Claim 3.*

Claim 4: Let $\phi \in \Phi$ and let $v \in V$. Then we have $\partial_v \phi \in \Phi$. *Proof of Claim 4:* By Claim 3, it suffices to show that $\partial_v \phi = 0$ on $V \backslash U$. Since $\phi \in \Phi$, we know that $\phi = 0$ on $V \backslash U$. So, since $V \backslash \overline{u} \subseteq V \backslash U$, it follows that $\phi = 0$ on $V \backslash \overline{u}$. So since $V \backslash \overline{u}$ is open in V, we see that $\partial_v \phi = 0$ on $V \backslash \overline{u}$. By Claim 2, $(\partial_v \phi)(0_V) = 0$. Let $U^* := \overline{u} \backslash (U \cup \{0_V\})$. It remains to show that $\partial_v \phi = 0$ on U^*. Let $x \in U^*$ be given. We wish to show that $(\partial_v \phi)(x) = 0$.

As $\phi \in \Phi$ and $v \in V \subseteq S$, we know that $\partial_v \phi$ is continuous on $V \backslash \{0_V\}$. We have $x \in U^* \subseteq V \backslash \{0_V\}$. Then $\partial_v \phi$ is continuous at x. Also, $x \in \overline{u} \backslash U \subseteq V \backslash U$, and $V \backslash U$ is the closure in V of $V \backslash \overline{u}$. Then x is an element of the closure in V of $V \backslash \overline{u}$, so, since $\partial_v \phi = 0$ on $V \backslash \overline{u}$ and since $\partial_v \phi$ is continuous at x, we get $(\partial_v \phi)(x) = 0$. *End of proof of Claim 4.*

Claim 5: Let $\phi \in \Phi$. Then we have: $\forall k \in \mathbb{N}$, $\forall s \in V^k$, $\partial_s \phi \in \Phi$. *Proof of Claim 5:* We argue by induction on k. By Claim 4, we have: $\forall s \in V$, $\partial_s \phi \in \Phi$, proving the base case. Let $k \in \mathbb{N}$ be given and make the induction assumption: $\forall s \in V^k$, $\partial_s \phi \in \Phi$. We wish to show: $\forall s \in V^{k+1}$, $\partial_s \phi \in \Phi$. Let $s \in V^{k+1}$ be given. We wish to show: $\partial_s \phi \in \Phi$.

Let $r := (s_2, \ldots, s_{k+1}) \in V^k$ and $\psi = \partial_r \phi$. Then $\partial_s \phi = \partial_{s_1} \partial_r \phi = \partial_{s_1} \psi$. By the induction assumption, $\partial_r \phi \in \Phi$. That is, we have $\psi \in \Phi$. Then, by Claim 4, $\partial_{s_1} \psi \in \Phi$. Then $\partial_s \phi = \partial_{s_1} \psi \in \Phi$. *End of proof of Claim 5.*

Let $C := \{(x, y) \in V \mid y = 2x^2\}$. By construction, know both that $f = 1$ on $C \backslash \{0_V\}$ and that $f(0_V) = 0$. Thus f is discontinuous at 0_V. It remains to show that f is infinitely directionally differentiable. Let $s \in S$ be given. We wish to show that $\mathrm{dom}[\partial_s \phi] = V$.

Since $s \in S = V \cup V^2 \cup V^3 \cup \cdots$, choose $k \in \mathbb{N}$ such that $s \in V^k$. By Claim 1, we have $f \in \Phi$. Then, by Claim 5, we have $\partial_s \phi \in \Phi$. Then, by definition of Φ, we see that $\mathrm{dom}[\partial_s \phi] = V$, as desired. $\square$

We say a function $\phi : V \to \mathbb{R}$ is **directionally differentiable** if, for all $v \in V$, $\mathrm{dom}[\partial_v \phi] = V$. Recall: $f : V \to \mathbb{R}$ is defined by

$$f(v) = \begin{cases} g(y/x^2), & \text{if } x \neq 0; \\ 0, & \text{if } x = 0. \end{cases}$$

Let $\sigma : V \to \mathbb{R}$ be an infinitely differentiable function satisfying $\sigma = 0$ on $V \backslash (B(0_2, 1))$ and $\sigma(0_2) = 1$ and $\mathrm{im}[\sigma] \subseteq [0, 1]$. For all $\varepsilon > 0$, define

$\sigma_\varepsilon : V \to \mathbb{R}$ by $\sigma_\varepsilon(x) = \varepsilon \cdot [\sigma(x/\varepsilon)]$; then $\sigma_\varepsilon = 0$ on $V \backslash (B(0_2, \varepsilon))$ and $\sigma_\varepsilon(0_2) = \varepsilon$ and $\mathrm{im}[\sigma_\varepsilon] \subseteq [0, \varepsilon]$. For all $\varepsilon > 0$, for all $p \in V$, we define $f_{p,\varepsilon} : V \to \mathbb{R}$ by $f_{p,\varepsilon}(x) = [\sigma_\varepsilon(x - p)][f(x - p)]$; then

- $f_{p,\varepsilon} = 0$ on $V \backslash (B(p, \varepsilon))$,
- $\mathrm{im}[f_{p,\varepsilon}] \subseteq [0, \varepsilon]$,
- $f_{p,\varepsilon}$ is infinitely directionally differentiable and
- $f_{p,\varepsilon}$ is discontinuous at p.

For any sequence $p_1, p_2, \ldots$ of distinct points in V, there is a sequence $\varepsilon_1, \varepsilon_2, \ldots$ of positive real numbers s.t. $f_{p_1,\varepsilon_1} + f_{p_2,\varepsilon_2} + \cdots$ is both directionally differentiable and discontinuous at each point of $\{p_1, p_2, \ldots\}$. In particular, there is a directionally differentiable $g : V \to \mathbb{R}$ such that g is discontinuous at each point of $\mathbb{Q}^2$.

A subset S of a topological space X is said to be **interior-free in** X if the interior in X of S is empty. A subset S of a topological space X is said to be **meager in** X if S is a subset of a countable union of closed interior-free subsets of X. A subset S of a topological space X is said to be **comeager in** X if $X \backslash S$ is meager in X. By the Baire Category Theorem, for any $n \in \mathbb{N}$, any comeager subset of $\mathbb{R}^n$ is to be dense in $\mathbb{R}^n$; in particular, a comeager subset of $\mathbb{R}^n$ is nonempty.

For any $\phi : V \to \mathbb{R}$, define $C_\phi := \{x \in V \,|\, \phi \text{ is continuous at } x\}$, and let $D_\phi := V \backslash C_\phi$ be the set of points of discontinuity of ϕ.

FACT 45.5. *Let $\phi : V \to \mathbb{R}$ and assume: $\forall i \in \{1, 2\}$, $dom[\partial_i \phi] = V$. Then C_ϕ is comeager in V.*

Proof. For all $x \in \mathbb{R}$, the maps $\phi(x, \bullet) : \mathbb{R} \to \mathbb{R}$ and $\phi(\bullet, x) : \mathbb{R} \to \mathbb{R}$ are both differentiable, hence continuous. Then, by *e.g.*, Namioka, Separate continuity and joint continuity, Pacific Journal of Mathematics, Volume 51, Number 2, 1974, we see that C_ϕ is comeager in V. □

FACT 45.6. *Let $\phi : V \to \mathbb{R}$. Assume: $\forall i, j \in \{1, 2\}$, $dom[\partial_i \partial_j \phi] = V$. Then C_ϕ contains a dense open subset of V.*

Proof. Since $D_\phi = V \backslash C_\phi$, we wish to show that D_ϕ is nowhere dense in V. Let $\overline{D}$ be the closure in V of D_ϕ. Let U be the interior in V of $\overline{D}$. We wish to show: $U = \varnothing$. Assume, for a contradiction, that $U \neq \varnothing$.

For all $j \in \{1, 2\}$, let $\psi_j := \partial_j \phi$. For all $i \in \{1, 2\}$, $\mathrm{dom}[\partial_i \psi_1] = V$, so, by Fact 45.5, we see that C_{ψ_1} is comeager in V. For all $i \in \{1, 2\}$, $\mathrm{dom}[\partial_i \psi_2] = V$, so, by Fact 45.5, we see that C_{ψ_2} is comeager in V. Let $C := C_{\psi_1} \cap C_{\psi_2}$. Then C is comeager in V, so, by the Baire

Category Theorem, C is dense in V. So, as U is a nonempty open subset of V, we conclude that $C \cap U \neq \varnothing$. Choose $x \in C \cap U$. Since $x \in C = C_{\psi_1} \cap C_{\psi_2}$, we see that ψ_1 and ψ_2 are both continuous at x. For all $j \in \{1, 2\}$, choose an open neighborhood U_j in U of x such that ψ_j is bounded on U_j. Then $U_1 \cap U_2$ is an open neighborhood in U of x, so choose open intervals $I, J \subseteq \mathbb{R}$ such that $x \in I \times J \subseteq U_1 \cap U_2$. Then ψ_1 and ψ_2 are bounded on $I \times J$. That is, $\partial_1 \phi$ and $\partial_2 \phi$ are bounded on $I \times J$. Then, by the Mean Value Theorem, ϕ is continuous on $I \times J$. Let $W := I \times J$. Then $W \subseteq C_\phi$, so $W \cap D_\phi = \varnothing$.

We have both $x \in I \times J = W$ and $W = I \times J \subseteq U_1 \cap U_2$. Also, $U_1 \subseteq U$ and $U_2 \subseteq U$. Also, by definition of U, we have $U \subseteq \overline{D}$. Then $x \in W \subseteq U_1 \cap U_2 \subseteq U \subseteq \overline{D}$. Since I and J are open intervals in $\mathbb{R}$ and since $W = I \times J$, it follows that W is an open subset of V. So, since $x \in W$, we see that W is an open neighborhood in V of x. So, since $x \in \overline{D}$ and since $\overline{D}$ is the closure in V of D_ϕ, it follows that $W \cap D_\phi \neq \varnothing$. However, we showed that $W \cap D_\phi = \varnothing$. Contradiction. □

End of discussion about infinite directional differentiability.

Assigned HW#11-2.

REMARK 45.7. *Let $m \in \mathbb{N}$, let $V := (\mathbb{R}^m, |\bullet|_{m,1})$ and let $S \subseteq V$. Assume S is closed and bounded in V. Then S is sequentially compact.*

Proof. We wish to show that $(S, \mathcal{T}_V|S)$ is sequentially compact.

Let $W := (\mathbb{R}^m, |\bullet|_{m,2})$. By HW#10-2, $|\bullet|_{m,2} \approx |\bullet|_{m,1}$. That is, $|\bullet|_W \approx |\bullet|_V$. Then by Corollary 44.6, $\mathcal{T}_W = \mathcal{T}_V$. It therefore suffices to show that $(S, \mathcal{T}_W|S)$ is sequentially compact.

By Corollary 24.14, W is proper. So, by Theorem 25.3, it suffices to show that S is closed and bounded in W. Since S is closed in V and since $\mathcal{T}_W = \mathcal{T}_V$, it follows that S is closed in W. It remains to show that S is bounded in W.

Since $|\bullet|_W \approx |\bullet|_V$, it follows that $|\bullet|_V << |\bullet|_W$. So, since S is bounded in V, we see, by HW#11-2, that S is bounded in W. □

COROLLARY 45.8. *Let $m \in \mathbb{N}$ and let $V := (\mathbb{R}^m, |\bullet|_{m,1})$. Then S_V is sequentially compact.*

Proof. Since $S_V \subseteq \overline{B}_V(0, 1) \subseteq B_V(0_v, 2)$, it follows that S_V is bounded in V. So, by Remark 45.7, it suffices to show that S_V is closed in V.

Let $f := |\bullet|_V : V \to \mathbb{R}$. By HW#37 from Fall Semester, the function $f : V \to \mathbb{R}$ is distance-semidecreasing. That is, $f : V \to \mathbb{R}$

is 1-Lipschitz. Then $f : V \to \mathbb{R}$ is Lipschitz, hence uniformly continuous, hence continuous. So, since $\{1\}$ is closed in $\mathbb{R}$, we conclude, from Theorem 28.11, that $f^*(\{1\})$ is closed in V. So, because

$$S_V \quad = \quad \{ \, x \in V \mid f(x) = 1 \, \} \quad = \quad f^*(\{1\}),$$

it follows that S_V is closed in V, as desired. □

Let V and W be normed vector spaces and let $T \in L(V, W)$. Recall:

$$\begin{aligned} E\ell\ell_T &= E\ell\ell_T^{VW} = T_*(S_V), \\ \widehat{T} &= \widehat{T}_{VW} = \sup |\, E\ell\ell_T \,|_W \qquad \text{and} \\ \check{T} &= \check{T}_{VW} = \inf |\, E\ell\ell_T \,|_W. \end{aligned}$$

Recall:

(1) T is **bounded** means: $\widehat{T} < \infty$,
(2) T is **bounded below** means: $\check{T} > 0$,
(3) $\forall K \geqslant 0$, T is K**-bounded** means: $\widehat{T} \leqslant K$, and
(4) $\forall \varepsilon > 0$, T is ε**-bounded below** means: $\check{T} \geqslant \varepsilon$.

Recall: For all $K \geqslant 0$, we have

$$\begin{aligned} &[\ T \text{ is } K\text{-bounded}\ ] \quad \Leftrightarrow \\ &[\ \forall u \in S_V,\ \ |Tu|_W \leqslant K\ ] \quad \Leftrightarrow \\ &[\ \forall x \in V,\ \ |Tx|_W \leqslant K \cdot |x|_V\ ]. \end{aligned}$$

Recall: For all $\varepsilon > 0$, we have

$$\begin{aligned} &[\ T \text{ is } \varepsilon\text{-bounded below}\ ] \quad \Leftrightarrow \\ &[\ \forall u \in S_V,\ \ |Tu|_W \geqslant \varepsilon\ ] \quad \Leftrightarrow \\ &[\ \forall x \in V,\ \ |Tx|_W \geqslant \varepsilon \cdot |x|_V\ ]. \end{aligned}$$

Assigned HW#11-3 and HW#11-4.

LEMMA 45.9. *Let V, W be normed vector spaces. Let $T : V \hookrightarrow\!\!> W$ be a vector space isomorphism. Assume that $T : V \to W$ is bounded. Then $T^{-1} : W \to V$ is bounded below.*

Proof. Since $T : V \to W$ is bounded, we have $\widehat{T}_{VW} < \infty$. Then $\widehat{T}_{VW} \in [0, \infty)$. Let $K := (\widehat{T}_{VW}) + 1$. Then $K \in (0, \infty)$ and T is K-bounded. Let $\varepsilon := 1/K$. Then $\varepsilon > 0$. We wish to show: $T^{-1} : W \to V$ is ε-bounded below. We wish to show: $\forall y \in W$, $|T^{-1}(y)|_V \geqslant \varepsilon \cdot |y|_W$. Let $y \in W$ be given. We wish to show: $|T^{-1}(y)|_V \geqslant \varepsilon \cdot |y|_W$. Let $x := T^{-1}(y)$. Then $Tx = y$, and we wish to show: $\varepsilon \cdot |y|_W \leqslant |x|_V$.

Since $T : V \to W$ is K-bounded, we get: $|Tx|_W \leqslant K \cdot |x|_V$. So, since $Tx = y$, we get $|y|_W \leqslant K \cdot |x|_V$. So, since $\varepsilon > 0$, we conclude that $\varepsilon \cdot |y|_W \leqslant \varepsilon \cdot K \cdot |x|_V$. So, since $\varepsilon \cdot K = 1$, we get $\varepsilon \cdot |y|_W \leqslant |x|_V$. $\square$

Assigned HW#11-5.

Recall: Let W, X be normed vector spaces. Let $T \in L(W, X)$. Then

$$[\ T : W \to X \text{ is bounded }\ ] \quad \Leftrightarrow \quad [\ T : W \to X \text{ is continuous }\ ].$$

It may seem, in light of this, that, for linear transformations between normed vector spaces, the word "bounded" is useless; it can always be replaced by "continuous". However, one may then ask, "How continuous?", and so it's really K-bounded that is useful, because it gives us a quantification of continuity.

Let W, X be normed vector spaces. Assume that W is finite dimensional. Let $T \in L(W, X)$. Then we will eventually show:

$$[\ T : W \to X \text{ is bounded below }\ ] \quad \Leftrightarrow \quad [\ T : W \to X \text{ is 1-1 }\ ].$$

It may seem, in light of this, that, for linear transformations between finite dimensional normed vector spaces, the phrase "bounded below" is useless; it can always be replaced by "injective". However, one may then ask, "How injective?", and so it's really ε-bounded below that is useful, because it gives us a quantification of injectivity.

Mostly we will focus on finite dimensional normed vector spaces. However, we went through one example, just to show how different things become in the infinite dimensional setting. For all $j \in \mathbb{N}$, let $e_j := \chi_{\{j\}}^{\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Then

$$\begin{array}{rcl} e_1 &=& (\,1\,,\,0\,,\,0\,,\,0\,,\,0\,,\,0\,,\,0\,,\,\ldots\,), \\ e_2 &=& (\,0\,,\,1\,,\,0\,,\,0\,,\,0\,,\,0\,,\,0\,,\,\ldots\,), \\ e_3 &=& (\,0\,,\,0\,,\,1\,,\,0\,,\,0\,,\,0\,,\,0\,,\,\ldots\,), \\ e_4 &=& (\,0\,,\,0\,,\,0\,,\,1\,,\,0\,,\,0\,,\,0\,,\,\ldots\,), \\ \vdots && \end{array}$$

Let $V := \langle \{e_1, e_2, e_3, e_4, \ldots\} \rangle_{\text{lin}}$ be the set of all *finite* linear combinations of the vectors $e_1, e_2, e_3, e_4, \ldots$. For all $x \in \mathbb{R}^{\mathbb{N}}$, we define $\operatorname{supp} x := \{j \in \mathbb{N} \,|\, x_j \neq 0\}$. Then, for all $x \in \mathbb{R}^{\mathbb{N}}$, we have:

$$[\ x \in V\ ] \quad \Rightarrow \quad [\ \#(\operatorname{supp} x) < \infty\ ].$$

So, for example

$$(0,0,0,2,-1,0,8,0,0,0,0,0,0,0,\ldots) \;=\; 2e_4 \;-\; e_5 \;+\; 8e_7 \;\in\; V,$$

while $(1,2,3,4,5,\ldots) \notin V$. Note that $\mathrm{supp}(2e_4 - e_5 + 8e_7) = \{4,5,7\}$ and that $\mathrm{supp}\,(1,2,3,4,5,\ldots) = \mathbb{N}$. Define $|\bullet| \in \mathcal{N}(V)$ by

$$|x| \quad = \quad \max\{\,|x_1|\,,\,|x_2|\,,\,|x_3|\,,\,|x_4|\,,\,|x_5|\,,\,\ldots\}.$$

(This norm is called the ℓ^∞ norm on V.) Define $T : V \to V$ by

$$Tx \quad = \quad (\,x_1\,,\,2x_2\,,\,3x_3\,,\,4x_4\,,\,5x_5\,,\,\ldots).$$

Then $T : V \hookrightarrow V$ is a vector space isomorphism, and we have: $\forall x \in V$,

$$T^{-1}x \quad = \quad (\,x_1\,,\,x_2/2\,,\,x_3/3\,,\,x_4/4\,,\,x_5/5\,,\,\ldots).$$

Note that $T^{-1} : V \to V$ is 1-1, and $(T^{-1})^{-1} = T$. Let $W := (V, |\bullet|)$. Then $T : W \to W$ is *NOT* bounded. So, by HW#11-5, $T^{-1} : W \to W$ is *NOT* bounded below (for, if it were, then $(T^{-1})^{-1}$ would be bounded, but $(T^{-1})^{-1} = T$ and we know that T is *NOT* bounded). So, in the setting of infinite dimensional normed vector spaces,

- not all linear transformations are bounded, and
- not all injective linear transformations are bounded below.

In the next class, we will show that, in the world of *finite* dimensional normed vector spaces, things are much better.

46. Class 20 on 5 April 2018, Th of Week 11

Recall (Corollary 45.8): Let $m \in \mathbb{N}$ and let $V := (\mathbb{R}^m, |\bullet|_{m,1})$. Then S_V is sequentially compact.

Recall (Corollary 25.12 and Theorem 25.1): Let $K \subseteq \mathbb{R}$. Assume that K is nonempty and sequentially compact. Then K has a maximum and a minimum.

Recall (HW#11-1): Let $S \subseteq \mathbb{R}$. Assume that S has a minimum. Then $\inf S = \min S \in S$.

THEOREM 46.1. *Let $m \in \mathbb{N}$ and let $V := (\mathbb{R}^m, |\bullet|_{m,1})$. Let W be a normed vector space and let $T \in L(V,W)$. Then*

(1) *$T : V \to W$ is bounded and*
(2) *(T is 1-1) $\Rightarrow$ ($T : V \to W$ is bounded below).*

Proof. Proof of (1): Let e be the standard ordered basis of $\mathbb{R}^m$. Let $K := \max\{|Te_1|_W, \ldots, |Te_m|_W\}$. We want: $T : V \to W$ is K-bounded. We wish to show that: $\forall x \in V$, $|Tx|_W \leqslant K|x|_V$. Let $x \in V$ be given. We wish to show that: $|Tx|_W \leqslant K|x|_V$.

We have $|x|_V = |x|_{m,1} = |x_1| + \cdots + |x_m|$. Since

$$x \quad = \quad (\, x_1\,,\,\ldots\,,\,x_m\,) \quad = \quad x_1 e_1 \,+\, \cdots \,+\, x_m e_m,$$

we get $Tx = x_1 \cdot [Te_1] + \cdots + x_m \cdot [Te_m]$. Then

$$\begin{aligned} |Tx|_W &= |\, x_1 \cdot [Te_1] \;+\; \cdots \;+\; x_m \cdot [Te_m]\,|_W \\ &\leqslant |x_1| \cdot |Te_1|_W \;+\; \cdots \;+\; |x_m| \cdot |Te_m|_W \\ &\leqslant |x_1| \cdot K \;+\; \cdots \;+\; |x_m| \cdot K \\ &= K \cdot [\, |x_1| \;+\; \cdots \;+\; |x_m|\,] \quad = \quad K|x|_V, \end{aligned}$$

as desired. *End of proof of (1).*

Proof of (2): Assume that T is 1-1. We wish to show: T is bounded below. That is, we wish to show that $\check{T} > 0$.

By Corollary 45.8, S_V is sequentially compact. By (1) of Theorem 46.1, $T : V \to W$ is bounded, so, by (2)⇒(5) of Theorem 42.13, $T : V \to W$ is continuous. So, since S_V is sequentially compact, by Lemma 36.1, we see that $T_*(S_V)$ is sequentially compact. Let $E := E\ell\ell_T$. Then $E = T_*(S_V)$, so E is sequentially compact. By HW#37 from Fall Semester, the function $|\bullet|_W : W \to \mathbb{R}$ is distance-semidecreasing, hence 1-Lipschitz, hence Lipschitz, hence uniformly continuous, hence continuous. So, since E is sequentially compact, by Lemma 36.1, we conclude that $|E|_W$ is sequentially compact. So, by Corollary 25.12 and Theorem 25.1, $|E|_W$ has a maximum and a minimum. Then, by HW#11-1, $\inf |E|_W = \min |E|_W \in |E|_W$. Then

$$\check{T} \;=\; \inf |E\ell\ell_T|_W \;=\; \inf |E|_W \;\in\; |E|_W \;=\; |T_*(S_V)|_W,$$

so choose $u \in S_V$ such that $\check{T} = |T(u)|_W$. Since $u \in S_V$, it follows that $|u|_V = 1$. Since $|u|_V = 1 \neq 0$, we get $u \neq 0_V$, so, as T is 1-1, we get $Tu \neq T(0_V)$. Since $Tu \neq T(0_V) = 0_W$, we get $|Tu|_W \neq 0$. So, since $|Tu|_W \geqslant 0$, $|Tu|_W > 0$. Then $\check{T} = |Tu|_W > 0$. *End of proof of (2).* □

Unassigned HW: Let X and Y and Z all be sets. Let $f : X \to Y$ and let $f : Y \to Z$. Assume

- $f : X \to Y$ is 1-1 and
- $g : Y \to Z$ is 1-1.

Then $g \circ f : X \to Z$ is 1-1.

Unassigned HW: Let X and Y and Z all be normed vector spaces. Let $f : X \to Y$ and let $f : Y \to Z$. Assume

- $f : X \to Y$ is bounded and
- $g : Y \to Z$ is bounded.

Then $g \circ f : X \to Z$ is bounded.

Unassigned HW: Let X and Y and Z all be normed vector spaces. Let $f : X \to Y$ and let $f : Y \to Z$. Assume

- $f : X \to Y$ is bounded below and
- $g : Y \to Z$ is bounded below.

Then $g \circ f : X \to Z$ is bounded below.

Recall (Definition 13.8) the definition of L_B.

THEOREM 46.2. *Let W be a finite dimensional normed vector space. Let X be a normed vector space and let $T \in L(W, X)$. Then*

(1) *$T : W \to X$ is bounded and*
(2) *(T is 1-1) $\Rightarrow$ ($T : W \to X$ is bounded below).*

Proof. Let $m := \dim W$. Let B be an ordered basis of W. Let $L := L_B$. Then $L : \mathbb{R}^m \to W$ is a vector space ismorphism. Let $V := (\mathbb{R}^m, |\bullet|_{m,1})$. By (1) of Theorem 46.1, $L : V \to W$ is bounded, so, by Lemma 45.9, $L^{-1} : W \to V$ is bounded below. By (2) of Theorem 46.1, $L : V \to W$ is bounded below, so, by HW#11-5, $L^{-1} : W \to V$ is bounded.

Proof of (1): By (1) of Theorem 46.1, $T \circ L : V \to X$ is bounded. As

$$L^{-1} : W \to V \qquad \text{and} \qquad T \circ L : V \to X$$

are both bounded, it follows that $T \circ L \circ L^{-1} : W \to X$ is bounded. That is, $T : W \to X$ is bounded. *End of proof of (1).*

Proof of (2): Assume: T is 1-1. Want: $T : W \to X$ is bdd below.

Since $L : V \to W$ is a vector space isomorphism, L is 1-1. As

$$L : V \to W \qquad \text{and} \qquad T : W \to X$$

are both 1-1, it follows that $T \circ L : V \to X$ is 1-1. Then, by (2) of Theorem 46.1, $T \circ L : V \to X$ is bounded below. As

$$L^{-1} : W \to V \qquad \text{and} \qquad T \circ L : V \to X$$

are both bounded below, it follows that $T \circ L \circ L^{-1} : W \to X$ is bounded below. That is, $T : W \to X$ is bounded below. *End of proof of (2).* $\square$

THEOREM 46.3. *Let V be a finite dimensional vector space and let $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$. Then $|\bullet| \approx \|\bullet\|$.*

Proof. Let $I := \mathrm{id}_V : V \to V$. Let $V' := (V, |\bullet|)$ and let $V'' := (V, \|\bullet\|)$. Since V'' is finite dimensional, we conclude, by (1) of Theorem 46.2, that $I : V'' \to V'$ is bounded. Then, by $\Leftarrow$ of Remark 43.4, we get: $|\bullet| << \|\bullet\|$. It remains to show: $\|\bullet\| << |\bullet|$.

Since V' is finite dimensional, we see, by (1) of Theorem 46.2, that $I : V' \to V''$ is bounded. Then, by $\Leftarrow$ of Remark 43.4, $\|\bullet\| << |\bullet|$. □

REMARK 46.4. *Let V be a finite dimensional vector space. Then $\mathcal{N}(V) \neq \varnothing$.*

Proof. Let $m := \dim V$. Since $|\bullet|_{m,2} \in \mathcal{N}(\mathbb{R}^m)$, we get $\mathcal{N}(\mathbb{R}^m) \neq \varnothing$. Choose an ordered basis B of V. Then $L_B : \mathbb{R}^m \to V$ is a vector space isomorphism, so $\mathbb{R}^m$ and V are vector space isomorphic. So, since $\mathcal{N}(\mathbb{R}^m) \neq \varnothing$, we see that $\mathcal{N}(V) \neq \varnothing$. □

Another way of looking at the preceding proof: Let $m := \dim V$, and choose an ordered basis B of V. Then $(|\bullet|_{m,2}) \circ (L_B^{-1}) : V \to [0, \infty)$ is a norm on V, and so we have: $\mathcal{N}(V) \neq \varnothing$, as desired. QED

It turns out that, using fancy set theory (specifically, two results, one called "Zorn's Lemma", and another called "the Well-Ordering Principle"), one can prove that infinite dimensional vector spaces also have ordered bases, although the meaning of "ordered" needs to be made precise. Then, using this, one can prove that, $\forall$vector space V, $\mathcal{N}(V) \neq \varnothing$. In this exposition, we will only need this result for finite dimensional vector spaces, so we are content with Remark 46.4.

REMARK 46.5. *Let V be a finite dimensional vector space. Then $\#\{\mathcal{T}_{|\bullet|} \text{ s.t. } |\bullet| \in \mathcal{N}(V)\} = 1$.*

Proof. Let $\mathcal{A} := \{\mathcal{T}_{|\bullet|} \text{ s.t. } |\bullet| \in \mathcal{N}(V)\}$. We wish to show: $\#\mathcal{A} = 1$. By Remark 46.4, choose $|\bullet|_0 \in \mathcal{N}(V)$. Then $\mathcal{T}_{|\bullet|_0} \in \mathcal{A}$. Then $\mathcal{A} \neq \varnothing$, so $\#\mathcal{A} \geqslant 1$. We wish to show: $\#\mathcal{A} \leqslant 1$. It suffices to show: $\forall \mathcal{T}', \mathcal{T}'' \in \mathcal{A}$, $\mathcal{T}' = \mathcal{T}''$. Let $\mathcal{T}', \mathcal{T}'' \in \mathcal{A}$ be given. Want: $\mathcal{T}' = \mathcal{T}''$.

Since $\mathcal{T}', \mathcal{T}'' \in \mathcal{A}$, choose $|\bullet|, \|\bullet\| \in \mathcal{N}(V)$ such that $\mathcal{T}' = \mathcal{T}_{|\bullet|}$ and $\mathcal{T}'' = \mathcal{T}_{\|\bullet\|}$. By Theorem 46.3, we have $|\bullet| \approx \|\bullet\|$. Then, by Corollary 44.6, $\mathcal{T}_{|\bullet|} = \mathcal{T}_{\|\bullet\|}$. Then $\mathcal{T}' = \mathcal{T}_{|\bullet|} = \mathcal{T}_{\|\bullet\|} = \mathcal{T}''$, as desired. □

DEFINITION 46.6. *Let V be a finite dimensional vector space. Then $\mathcal{T}_V := \mathrm{ELT}\{\mathcal{T}_{|\bullet|} \text{ s.t. } |\bullet| \in \mathcal{N}(V)\}$.*

Let V be a finite dimensional vector space. Then, by Remark 46.5, we see that $\mathcal{T}_V \neq ☹$, so $\mathcal{T}_V$ is a topology on V. The **standard topology** on V is the topology $\mathcal{T}_V$ of Definition 46.6.

The upshot of Theorem 46.3 is that, while a finite dimensional vector space V will have many norms, they are all equivalent, and so they all give the same topology, and this one topology is the standard topology $\mathcal{T}_V$ on V. In infinite dimensions, there are even more norms, and they often give quite different topologies from one another, and so the infinite dimesional world is richer and more complicated. Fortunately (or unfortunately, depending on point of view), this infinite dimensional world is typically not a focus of undergraduate mathematics.

In the sequel, every finite dimensional vector space is given its standard topology, turning it into a topological space.

We record that any norm on a finite dimensional vector space gives rise to the standard topology:

REMARK 46.7. *Let V be a finite dimensional vector space and let $|\bullet| \in \mathcal{N}(V)$. Then $\mathcal{T}_{|\bullet|} = \mathcal{T}_V$.*

Proof. Let $\mathcal{A} := \{\mathcal{T}_{\|\bullet\|} \text{ s.t. } \|\bullet\| \in \mathcal{N}(V)\}$. Then $\mathcal{T}_V = \mathrm{ELT}\mathcal{A}$.

As $|\bullet| \in \mathcal{N}(V)$, we get $\mathcal{T}_{|\bullet|} \in \mathcal{A}$. By Remark 46.5, $\#\mathcal{A} = 1$. So, since $\mathcal{T}_{|\bullet|} \in \mathcal{A}$, we see that $\mathcal{A} = \{\mathcal{T}_{|\bullet|}\}$, and so $\mathrm{ELT}\mathcal{A} = \mathcal{T}_{|\bullet|}$. Then $\mathcal{T}_{|\bullet|} = \mathrm{ELT}\mathcal{A} = \mathcal{T}_V$, as desired. $\square$

THEOREM 46.8. *Let V be a finite dimensional vector space and let W be a normed vector space. Let $T \in L(V,W)$. Then $T : V \to W$ is continuous.*

Proof. By Remark 46.4, choose $|\bullet| \in \mathcal{N}(V)$. By Remark 46.7, $T_{|\bullet|} = \mathcal{T}_V$.

Let $V' := (V, |\bullet|)$. Then the standard topology on V' is given by: $\mathcal{T}_{V'} = \mathcal{T}_{|\bullet|}$. Since V' is finite dimensional, by (1) of Theorem 46.2, we see that $T : V' \to W$ is bounded. Then, by (2)⇒(5) of Theorem 42.13, $T : V' \to W$ is continuous. So, since $\mathcal{T}_{V'} = \mathcal{T}_{|\bullet|} = \mathcal{T}_V$, we conclude that $T : V \to W$ is continuous. $\square$

COROLLARY 46.9. *Let V and W be finite dimensional vector spaces. Let $T \in L(V,W)$. Then $T : V \to W$ is continuous.*

Proof. By Remark 46.4, choose $|\bullet| \in \mathcal{N}(W)$. By Remark 46.7, $T_{|\bullet|} = \mathcal{T}_W$.

Let $W' := (W, |\bullet|)$. Then the standard topology on W' is given by: $\mathcal{T}_{W'} = \mathcal{T}_{|\bullet|}$. Since V is finite dimensional, Theorem 46.8, we see that $T : V \to W'$ is bounded. Then, by (2)⇒(5) of Theorem 42.13, $T : V \to W'$ is continuous. So, since $\mathcal{T}_{W'} = \mathcal{T}_{|\bullet|} = \mathcal{T}_W$, we conclude that $T : V \to W$ is continuous. □

COROLLARY 46.10. *Let V, W be finite dimensional vector spaces. Let $T : V \hookrightarrow\!\!> W$ be a vector space isomorphism. Then $T : V \hookrightarrow\!\!> W$ is a homeomorphism.*

Proof. Since $T \in L(V,W)$, by Corollary 46.9, $T : V \to W$ is continuous. It remains to show: $T^{-1} : W \to V$ is continuous.

By Fact 44.4, $T^{-1} \in L(W,V)$. So, by Corollary 46.9, $T^{-1} : W \to V$ is continuous, as desired. □

DEFINITION 46.11. *Let $m \in \mathbb{N}$ and let $v, w \in \mathbb{R}^m$. We define the* **dot product** *of v and w as $v \bullet w := v_1w_1 + \cdots + v_mw_m$.*

For example, $(2,5,7) \cdot (0,1,-1) = (2)(0) + (5)(1) + (7)(-1) = -2$.

We will sometimes write $v \bullet w = \sum_j v_jw_j$. It will be the reader's responsibility to figure out, from the context, that $\sum_j$ means $\sum_{j=1}^m$.

We will NOT be following the **Einstein Summation Convention**, which asserts that, whenever one sees a repeated index (in any term in any formula), a summation on that index is understood. So, for example, those who follow that convention might write $v \bullet w = v_jw_j$, with the understanding that, because the index j appears twice in the term "v_jw_j", the "$\sum_j$" is understood, without being written.

Recall that, for any vector space S, for any set Z, the set S^Z (of functions $Z \to S$) has a standard vector space structure, given by:

- $\forall f, g \in S^Z$, $f + g \in S^Z$ is defined by $(f+g)(z) = [f(z)] + [g(z)]$,
- $\forall c \in \mathbb{R}$, $\forall f \in S^Z$, $cf \in S^Z$ is defined by $(cf)(z) = c \cdot [f(z)]$.

Recall that, for any $\ell \in \mathbb{N}$, $[1..\ell] = \{1,\ldots,m\}$. So, for example, $[1..5] = \{1,2,3,4,5\}$. Also,

$$[1..2] \times [1..3] \quad = \quad \{\,(1,1)\,,\,(1,2)\,,\,(1,3)\,,\,(2,1)\,,\,(2,2)\,,\,(2,3)\,\}.$$

Recall (Definition 39.11) that, for any $\ell, m \in \mathbb{N}$, for any set S,

$$S^{\ell\times m} \quad := \quad S^{[1..\ell]\times[1..m]};$$

if S is a vector space, then $S^{\ell\times m}$ has a standard vector space structure.

DEFINITION 46.12. *Let $\ell, m, n \in \mathbb{N}$ and let S be a set. Then*

$$S^{\ell\times m\times n} \quad := \quad S^{[1..\ell]\times[1..m]\times[1..n]}.$$

For all $\ell, m, n \in \mathbb{N}$, for any vector space S, the set $S^{\ell\times m\times n}$ has a standard vector space structure.

More generally,

DEFINITION 46.13. *Let $d \in \mathbb{N}$, let $\ell_1, \ldots, \ell_d\mathbb{N}$ and let S be a set. Then*

$$S^{\ell_1\times\cdots\times\ell_d} \quad := \quad S^{[1..\ell_1]\times\cdots\times[1..\ell_d]}.$$

For all $d \in \mathbb{N}$, for all $\ell_1, \ldots, \ell_d\mathbb{N}$, for any vector space S, For all $\ell, m, n \in \mathbb{N}$, for any vector space S, the set $S^{\ell_1\times\cdots\times\ell_d}$ has a standard vector space structure.

Let $A \in \mathbb{R}^{2\times 7\times 9}$. Sometimes we say that A is a three dimensional tensor with entries in $\mathbb{R}$. Or, to be more specific we might say that A is a $2 \times 7 \times 9$ tensor with entries in $\mathbb{R}$. Note that while each element of $\mathbb{R}^{2\times 7\times 9}$ is three dimensional, the full vector space $\mathbb{R}^{2\times 7\times 9}$ is $2 \cdot 7 \cdot 9 = 126$ dimensional. Because A is three dimensional, we cannot display it easily on a two dimensional space like a page of paper, but there are ways of "flattening" A: For example, there's a vector space isomorphism $\mathbb{R}^{2\times 7\times 9} \to (\mathbb{R}^2)^{7\times 9}$, and the image of A in $(\mathbb{R}^2)^{7\times 9}$ can be displayed as a 7×9 matrix, with entries in $\mathbb{R}^2$.

NOTE: A zero dimensional tensor is a scalar. A one dimensional tensor is sometimes called a vector, although this is not to be confused with an element of an abstract vector space. To avoid this confusion, a one dimensional tensor is sometimes called a "tuple". A two dimensional tensor is a matrix.

For the rest of this class, we fix

$$X \quad := \quad \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{bmatrix} \quad \in \quad \mathbb{R}^{3\times 4}.$$

Then $X : \{1, 2, 3\} \times \{1, 2, 3, 4\} \to \mathbb{R}$, and, for example, $X(2, 3) = 6$. Howoever, usually, when tensors are involved, we use subscripts for the input values, writing $X_{23} = 6$, or, sometimes, $X_{2,3} = 6$. We have $X_{2\bullet} : \{1, 2, 3, 4\} \to \mathbb{R}$, *i.e.*, $X_{2\bullet} \in \mathbb{R}^4$. We have $X_{2\bullet} = (4, 5, 6, 0)$; this is the second row of X. Similarly, we have $X_{\bullet 3} = (3, 6, 9) \in \mathbb{R}^3$; this is the third column of X.

This notational convention continues to three dimensional tensors. So, for example, for any $Z \in \mathbb{R}^{3\times 4\times 2}$, we have both $Z_{\bullet\bullet 1} \in \mathbb{R}^{3\times 4}$ and $Z_{\bullet\bullet 2} \in \mathbb{R}^{3\times 4}$; because of the "2" in "$3 \times 4 \times 2$", we have $Z_{\bullet\bullet 3} = ☹$. If I have in mind some three dimesional tensor $Z \in \mathbb{R}^{3\times 4\times 2}$, it's hard to display it on a page of paper. Instead, I could show you $Z_{\bullet\bullet 1}$ and $Z_{\bullet\bullet 2}$.

DEFINITION 46.14. *For all $m \in \mathbb{N}$, we define e^m to be the standard ordered basis of $\mathbb{R}^m$.*

Thus, for all $m \in \mathbb{N}$, we have $e^m \in (\mathbb{R}^m)^m$ and

$$e_1^m \;=\; (1,0,\ldots,0) \,\in\, \mathbb{R}^m \quad , \quad \ldots \quad , \quad e_m^m \;=\; (0,\ldots,0,1) \,\in\, \mathbb{R}^m.$$

We will sometimes be sloppy and omit the superscript m, writing $e_1,\ldots,e_m$. Recall X above. Recall that $X_{2\bullet} = (4,5,6,0)$. Then

$$X_{2\bullet} \quad = \quad 4e_1 \,+\, 5e_2 \,+\, 6e_3 \,+\, 0e_4,$$

or, being more careful about superscripts, $X_{2\bullet} = 4e_1^4 + 5e_2^4 + 6e_3^4 + 0e_4^4$. Of course $0e_4^4 = 0_4$, so we also have $X_{2\bullet} = 4e_1^4 + 5e_2^4 + 6e_3^4$. This kind of equation will sometimes be called a **reproducing equation**. In this example, we "reproduced" the second row of X as a linear combination that uses its (nonzero) entries as coefficients.

The next fact records three reproducing formulas. It reproduces (1) a vector, (2) a row of a matrix, and (3) a column of a matrix:

FACT 46.15. *The following are all true:*

(1) $\forall m \in \mathbb{N}$, $\forall v \in \mathbb{R}^m$, $\quad v = \displaystyle\sum_j v_j e_j$,

(2) $\forall \ell, m \in \mathbb{N}$, $\forall A \in \mathbb{R}^{\ell\times m}$, $\forall j$, $\quad A_{j\bullet} = \displaystyle\sum_k A_{jk} e_k$, $\qquad$ *and*

(3) $\forall \ell, m \in \mathbb{N}$, $\forall A \in \mathbb{R}^{\ell\times m}$, $\forall k$, $\quad A_{\bullet k} = \displaystyle\sum_j A_{jk} e_j$.

Proof. Omitted. □

Much is expected of the reader in understanding (1), (2) and (3) of Fact 46.15. For example, in (2), we write "$\forall j$", and the reader is supposed to understand, from context, that this means "$\forall j \in [1..\ell]$. Note that, if we were following the Einstein Summation Convention, then the equation in (2) could be simplified to read: "$A_{j\bullet} = A_{jk}e_k$", and "$\displaystyle\sum_k$" would be understood, because k is the only repeated index in the term "$A_{jk}e_k$". Being more careful, we would rewrite Fact 46.15:

FACT 46.16. *The following are all true:*

(1) $\forall m \in \mathbb{N}$, $\forall v \in \mathbb{R}^m$, $\quad v = \sum_{j=1}^{m} v_j e_j^m$,

(2) $\forall \ell, m \in \mathbb{N}$, $\forall A \in \mathbb{R}^{\ell \times m}$, $\forall j \in [1..\ell]$, $\quad A_{j\bullet} = \sum_{k=1}^{m} A_{jk} e_k^m$,

and (3) $\forall \ell, m \in \mathbb{N}$, $\forall A \in \mathbb{R}^{\ell \times m}$, $\forall k \in [1..m]$, $\quad A_{\bullet k} = \sum_{j=1}^{\ell} A_{jk} e_j^\ell$.

Proof. Omitted. ☐

47. Class 21 on 10 April 2018, Tu of Week 12

We define $0^0 = 1$. Then, $\forall x \in \mathbb{R}$, $x^0 = 1$. Define

$$\ln \;:=\; \exp^{-1} \;:\; (0,\infty) \;\hookrightarrow\!\!> \;\mathbb{R}.$$

For all $p > 0$, we define $0^p = 0$. For all $x > 0$, for all $p \in \mathbb{R}$, we define $x^p = \exp(p \cdot [\ln x])$. Finally, we define

$$\tan \;:=\; \sin / \cos \qquad \text{and} \qquad \cot \;:=\; \cos / \sin .$$

Let V be a vector space. Let $|\bullet| \in \mathcal{N}(V)$. For any $p \geqslant 0$, define $|\bullet|^p : V \to [0,\infty)$ by $(|\bullet|^p)(x) = |x|^p$. Then $|\bullet|^0 = C_V^1$ and $|\bullet|^1 = |\bullet|$.

Let $|\bullet| : \mathbb{R} \to [0,\infty)$ denote the usual absolute value function. Then, since $|\bullet| \in \mathcal{N}(\mathbb{R})$, we get $|\bullet|^0 = C_{\mathbb{R}}^1$ and $|\bullet|^1 = |\bullet|$. For all $p \in \mathbb{N}_0$, let $(\bullet)^p : \mathbb{R} \to \mathbb{R}$ denote the pth power function, defined by $(\bullet)^p(x) = x^p$. Then $(\bullet)^0 = C_{\mathbb{R}}^1$ and $(\bullet)^1 = \mathrm{id}_{\mathbb{R}}$.

DEFINITION 47.1. *Let V and W be finite dimensional vector spaces. Then, by $\breve{\mathcal{O}}(V,W)$, we mean the set of functions $\varepsilon : V \dashrightarrow W$ such that*

(1) *$dom[\varepsilon] \;\in\; \mathcal{N}_V(0_V)$,*
(2) *ε is continuous at 0_V and*
(3) *$\varepsilon(0_V) = 0_W$.*

We will sometimes omit (V,W) and write $\breve{\mathcal{O}}$ in place of $\breve{\mathcal{O}}(V,W)$, provided V and W are clear. Also, we may sometimes $\mathcal{O}$ instead of $\breve{\mathcal{O}}$.

Condition (1) in Definition 47.1, is expressed by saying, "ε is defined near zero." Conditions (2) and (3) in Definition 47.1 are expressed by saying, "ε vanishes continuously at zero."

Focusing on partial functions $\mathbb{R} \dashrightarrow \mathbb{R}$, we have:

- $(\bullet)^0 \quad \notin \quad \breve{\mathcal{O}}(\mathbb{R},\mathbb{R})$,
- $(\bullet)^1, (\bullet)^2, (\bullet)^3, \ldots \quad \in \quad \breve{\mathcal{O}}(\mathbb{R},\mathbb{R})$,

- $|\bullet|^0 \quad \notin \quad \breve{o}(\mathbb{R},\mathbb{R})$,
- $\forall p > 0, \qquad |\bullet|^p \quad \in \quad \breve{o}(\mathbb{R},\mathbb{R})$,
- $\sin, \tan \quad \in \quad \breve{o}(\mathbb{R},\mathbb{R})$ and
- $\cos, \cot \quad \notin \quad \breve{o}(\mathbb{R},\mathbb{R})$.

DEFINITION 47.2. *Let V and W be finite dimensional vector spaces. Let $|\bullet| \in \mathcal{N}(W)$. Then, by $\widehat{\mathcal{O}}(V,W,|\bullet|)$, we mean the set of functions $\alpha : V \dashrightarrow W$ such that: $\exists U \in \mathcal{N}_V(0_V)$ s.t.*

(1) *$U \subseteq dom[\alpha]$ and*
(2) *$\sup |\alpha_*(U)|_W < \infty$.*

Condition (1) in Definition 47.2 is expressed by saying, "α is defined on U." Condition (2) in Definition 47.2 is expressed by saying, "α is bounded on U." Conditions (1) and (2) in Definition 47.2 are expressed by saying "α is defined and bounded near zero."

Assigned HW#12-1.

DEFINITION 47.3. *Let V and W be finite dimensional vector spaces. Then $\widehat{\mathcal{O}}(V,W) := \mathrm{ELT}\{\widehat{\mathcal{O}}(V,W,|\bullet|)$ s.t. $|\bullet| \in \mathcal{N}(W)\}$.*

We will sometimes omit (V,W) and write $\widehat{\mathcal{O}}$ in place of $\widehat{\mathcal{O}}(V,W)$, provided V and W are clear. Also, we may sometimes write "$\mathcal{O}$" instead of "$\widehat{\mathcal{O}}$".

Assigned HW#12-2.

COROLLARY 47.4. *Let V and W be finite dimensional vector spaces. Then $\breve{o}(V,W) \subseteq \widehat{\mathcal{O}}(V,W)$.*

Proof. We wish to show: $\forall \varepsilon \in \breve{o}(V,W)$, $\varepsilon \in \widehat{\mathcal{O}}(V,W)$. Let $\varepsilon \in \breve{o}(V,W)$ be given. We wish to show: $\varepsilon \in \widehat{\mathcal{O}}(V,W)$.

By (1) and (2) of Definition 47.1, we know that $\mathrm{dom}[\varepsilon] \in \mathcal{N}_V(0_V)$ and that ε is continuous at 0_V. Then, by HW#12-2, $\varepsilon \in \widehat{\mathcal{O}}(V,W)$. □

Focusing on partial functions $\mathbb{R} \dashrightarrow \mathbb{R}$, we have:

- $(\bullet)^0, (\bullet)^1, (\bullet)^2, (\bullet)^3, \ldots \quad \in \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$,
- $\forall p \geqslant 0, \qquad |\bullet|^p \quad \in \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$,
- $\sin, \cos, \tan \quad \in \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$,
- $\cot \quad \notin \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$,
- $\mathrm{adj}_0^{1000}(\cot) \quad \notin \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$,
- $\sqrt{\bullet} \quad \notin \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$,

- $\mathrm{adj}_0^2(C^1_{\mathbb{R}}) \quad \in \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$ and
- $\chi_{\mathbb{R}}^{(0,\infty)} \quad \in \quad \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$.

DEFINITION 47.5. *Let V and W be finite dimensional vector spaces. Let $|\bullet| \in \mathcal{N}(V)$ and let $p \geqslant 0$. Then*

$$\begin{aligned} \breve{o}_p(V,W,|\bullet|) &:= [|\bullet|^p]\cdot[\breve{o}(V,W)], \qquad \textit{and} \\ \widehat{\mathcal{O}}_p(V,W,|\bullet|) &:= [|\bullet|^p]\cdot[\widehat{\mathcal{O}}(V,W)]. \end{aligned}$$

Assigned HW#12-3.

DEFINITION 47.6. *Let V and W be finite dimensional vector spaces. Let $p \geqslant 0$. Then*

$$\begin{aligned} \breve{o}_p(V,W) &:= \mathrm{ELT}\{\breve{o}_p(V,W,|\bullet|) \ \textit{s.t.}\ |\bullet| \in \mathcal{N}(W)\}, \qquad \textit{and} \\ \widehat{\mathcal{O}}_p(V,W) &:= \mathrm{ELT}\{\widehat{\mathcal{O}}_p(V,W,|\bullet|) \ \textit{s.t.}\ |\bullet| \in \mathcal{N}(W)\}. \end{aligned}$$

Let V and W be finite dimensional vector spaces. Let $|\bullet| \in \mathcal{N}(W)$. Then we have both $\breve{o}_0(V,W) = \breve{o}(V,W)$ and $\widehat{\mathcal{O}}_0(V,W) = \widehat{\mathcal{O}}(V,W)$. Also, for all $p \geqslant 0$, we have $\breve{o}_p(V,W) \subseteq \widehat{\mathcal{O}}_p(V,W)$.

Given a function f, the next remark gives us criteria for determining, for each $p > 0$, whether f in $\breve{o}_p$ or $\widehat{\mathcal{O}}_p$ or both or neither.

REMARK 47.7. *Let V and W be finite dimensional vector spaces, let $|\bullet| \in \mathcal{N}(V)$, let $p > 0$ and let $f : V \dashrightarrow W$. Let $g := f/[|\bullet|^p]$, and let $h := \mathrm{adj}_{0_V}^{0_W}(g)$. Then both of the following hold:*

$$\begin{aligned} &(1)\ [\ f \in \breve{o}_p(V,W)\ ] \ \Leftrightarrow\ [\ (\,h \in \breve{o}(V,W)\,)\,\&\,(\,f(0_V) = 0_W\,)\ ] \\ \&\,&(2)\ [\ f \in \widehat{\mathcal{O}}_p(V,W)\ ] \ \Leftrightarrow\ [\ (\,h \in \widehat{\mathcal{O}}(V,W)\,)\,\&\,(\,f(0_V) = 0_W\,)\ ]. \end{aligned}$$

Proof. Let $V^* := V_{0_V}^{\times}$. That is, $V^* = V\backslash\{0_V\}$. By definition of h, we have: both ($g = h$ on V^*) and ($h(0_V) = 0_W$). Since $p > 0$, it follows that $0^p = 0$. Then we have: $|0_V|_V^p = 0^p = 0$.

Proof of $\Rightarrow$ of (1): Assume that $f \in \breve{o}_p(V,W)$. We wish to show both that $h \in \breve{o}(V,W)$ and that $f(0_V) = 0_W$.

Since $f \in \breve{o}_p(V,W) = [|\bullet|^p]\cdot[\breve{o}(V,W)]$, choose $\varepsilon \in \breve{o}(V,W)$ such that $f = [|\bullet|^p]\cdot\varepsilon$. Since $\varepsilon \in \breve{o}(V,W)$, we get $\varepsilon(0_V) = 0_W$. Then

$$f(0_V) \ =\ [\,|0_V|_V^p\,]\cdot[\,\varepsilon(0_V)\,] \ =\ 0\cdot 0_W \ =\ 0_W.$$

Want: $h \in \breve{o}(V,W)$. Since $\varepsilon \in \breve{o}(V,W)$, it suffices to show: $h = \varepsilon$.

Since $f = [|\bullet|^p]\cdot\varepsilon$, we get: $f/[|\bullet|^p] = \varepsilon$ on V^*. That is, we get: $g = \varepsilon$ on V^*. So, since $h = g$ on V^*, we get: $h = \varepsilon$ on V^*. So, as $h(0_V) = 0_W = \varepsilon(0_V)$, we get $h = \varepsilon$. *End of proof of $\Rightarrow$ of (1).*

Proof of $\Leftarrow$ of (1): Assume that both $h \in \breve{\mathcal{o}}(V,W)$ and $f(0_V) = 0_W$. We wish to prove: $f \in \breve{\mathcal{o}}_p(V,W)$. We define $\phi := [|\bullet|^p] \cdot h$. Then we have $\phi \in [|\bullet|^p] \cdot [\breve{\mathcal{o}}(V,W)] = \breve{\mathcal{o}}_p(V,W)$, so it suffices to show: $f = \phi$.

As $f/[|\bullet|^p] = g$, we get: $f = [|\bullet|^p] \cdot g$ on V^*. So, as $g = h$ on V^*, we get: $f = [|\bullet|^p] \cdot h$ on V^*. That is, $f = \phi$ on V^*. It remains to show: $f(0_V) = \phi(0_V)$. Since $f(0_V) = 0_W$, we wish to show: $\phi(0_V) = 0_W$.

We calculate: $\phi(0_V) = [|0_V|^p] \cdot [h(0_V)] = 0 \cdot 0_W = 0_W$, as desired. *End of proof of $\Leftarrow$ of (1).*

Proof of $\Rightarrow$ of (2): Assume that $f \in \widehat{\mathcal{O}}_p(V,W)$. We wish to show both that $h \in \widehat{\mathcal{O}}(V,W)$ and that $f(0_V) = 0_W$.

Since $f \in \widehat{\mathcal{O}}_p(V,W) = [|\bullet|^p] \cdot [\widehat{\mathcal{O}}(V,W)]$, choose $\alpha \in \widehat{\mathcal{O}}(V,W)$ such that $f = [|\bullet|^p] \cdot \alpha$. Then

$$f(0_V) \;=\; [\,|0_V|_V^p\,] \cdot [\,\alpha(0_V)\,] \;=\; 0 \cdot [\,\alpha(0_V)\,] \;=\; 0_W.$$

Want: $h \in \widehat{\mathcal{O}}(V,W)$. Since $\alpha \in \widehat{\mathcal{O}}(V,W)$, it suffices to show: $h = \alpha$.

Since $f = [|\bullet|^p] \cdot \alpha$, we get: $f/[|\bullet|^p] = \alpha$ on V^*. That is, we get: $g = \alpha$ on V^*. So, since $h = g$ on V^*, we get: $h = \alpha$ on V^*. So, as $h(0_V) = 0_W = \alpha(0_V)$, we get $h = \alpha$. *End of proof of $\Rightarrow$ of (2).*

Proof of $\Leftarrow$ of (2): Assume that $h \in \widehat{\mathcal{O}}(V,W)$ and that $f(0_V) = 0_W$. We wish to show that $f \in \widehat{\mathcal{O}}_p(V,W)$.

Let $\phi := [|\bullet|^p] \cdot h$. Then $\phi \in [|\bullet|^p] \cdot [\widehat{\mathcal{O}}(V,W)] = \widehat{\mathcal{O}}_p(V,W)$, so it suffices to show: $f = \phi$.

As $f/[|\bullet|^p] = g$, we get: $f = [|\bullet|^p] \cdot g$ on V^*. So, as $g = h$ on V^*, we get: $f = [|\bullet|^p] \cdot h$ on V^*. That is, $f = \phi$ on V^*. It remains to show: $f(0_V) = \phi(0_V)$. Since $f(0_V) = 0_W$, we wish to show: $\phi(0_V) = 0_W$.

We calculate: $\phi(0_V) = [|0_V|^p] \cdot [h(0_V)] = 0 \cdot 0_W = 0_W$, as desired. *End of proof of $\Leftarrow$ of (2).* $\square$

Using Remark 47.7, $\forall p \in \mathbb{N}_0$, one can show:

(1) $(\bullet)^p \in \widehat{\mathcal{O}}_p(\mathbb{R},\mathbb{R})$,
(2) $\forall \delta > 0, \quad (\bullet)^p \notin \widehat{\mathcal{O}}_{p+\delta}(\mathbb{R},\mathbb{R})$,
(3) $(\bullet)^p \notin \breve{\mathcal{o}}_p(\mathbb{R},\mathbb{R})$ and
(4) $\forall \delta \in (0,p], \quad (\bullet)^p \in \breve{\mathcal{o}}_{p-\delta}(\mathbb{R},\mathbb{R})$.

We summarize (1) and (2) above as saying, $\forall p \in \mathbb{N}_0$, that $(\bullet)^p$ is just barely in $\widehat{\mathcal{O}}_p(\mathbb{R},\mathbb{R})$. We summarize (3) and (4) above as saying, $\forall p \in \mathbb{N}_0$, that $(\bullet)^p$ is just barely outside $\breve{\mathcal{o}}_p(\mathbb{R},\mathbb{R})$.

Using Remark 47.7, $\forall p \geqslant 0$, one can show:

(1) $|\bullet|^p \in \widehat{\mathcal{O}}_p(\mathbb{R},\mathbb{R})$,

(2) $\forall \delta > 0, \quad |\bullet|^p \notin \widehat{\mathcal{O}}_{p+\delta}(\mathbb{R},\mathbb{R})$,
(3) $|\bullet|^p \notin \breve{o}_p(\mathbb{R},\mathbb{R})$ and
(4) $\forall \delta \in (0,p], \quad |\bullet|^p \in \breve{o}_{p-\delta}(\mathbb{R},\mathbb{R})$.

We summarize (1) and (2) above as saying, $p \geqslant 0$, that $|\bullet|^p$ is just barely in $\widehat{\mathcal{O}}_p(\mathbb{R},\mathbb{R})$. We summarize (3) and (4) above as saying, $p \geqslant 0$, that $|\bullet|^p$ is just barely outside $\breve{o}_p(\mathbb{R},\mathbb{R})$.

The remarks of the preceding two paragraphs motivate:

DEFINITION 47.8. *Let V and W be finite dimensional vector spaces and let $f : V \dashrightarrow W$. Then:*

- *By f is* **constant order***, we mean $f \in \widehat{\mathcal{O}}_0(V,W)$.*
- *By f is* **linear order***, we mean $f \in \widehat{\mathcal{O}}_1(V,W)$.*
- *By f is* **quadratic order***, we mean $f \in \widehat{\mathcal{O}}_2(V,W)$.*
- *By f is* **cubic order***, we mean $f \in \widehat{\mathcal{O}}_3(V,W)$.*
- *By f is* **quartic order***, we mean $f \in \widehat{\mathcal{O}}_4(V,W)$.*
- *$\forall p \geqslant 0$, by f is* **order** *p, we mean $f \in \widehat{\mathcal{O}}_p(V,W)$.*
- *By f is* **subconstant***, we mean $f \in \breve{o}_0(V,W)$.*
- *By f is* **sublinear***, we mean $f \in \breve{o}_1(V,W)$.*
- *By f is* **subquadratic***, we mean $f \in \breve{o}_2(V,W)$.*
- *By f is* **subcubic***, we mean $f \in \breve{o}_3(V,W)$.*
- *By f is* **subquartic***, we mean $f \in \breve{o}_4(V,W)$.*
- *$\forall p \geqslant 0$, by f is* **sub-***p, we mean $f \in \breve{o}_p(V,W)$.*

For all $p \in \mathbb{N}_0$, $(\bullet)^p : \mathbb{R} \to \mathbb{R}$ is order p, but not sub-p. In particular, $(\bullet)^2 : \mathbb{R} \to \mathbb{R}$ is quadratic order, but not subquadratic.

For all $p \geqslant 0$, $|\bullet|^p : \mathbb{R} \to \mathbb{R}$ is order p, but not sub-p. In particular, $|\bullet|^2 : \mathbb{R} \to \mathbb{R}$ is quadratic order, but not subquadratic.

Assigned HW#12-4.

Let V and W be finite dimensional vector spaces. According to HW#12-4, for any $p \geqslant 0$, for any $\delta > 0$, any sub-p function $V \dashrightarrow W$ is an order q function. This implies

$$\widehat{\mathcal{O}}_0 \supseteq \breve{o}_0 \supseteq \widehat{\mathcal{O}}_1 \supseteq \breve{o}_1 \supseteq \widehat{\mathcal{O}}_2 \supseteq \breve{o}_2 \supseteq \cdots .$$

More generally, we assert: "the collection of $\widehat{\mathcal{O}}$ and $\breve{o}$ function spaces are totally ordered by inclusion". The precise meaning of this is: $\forall p,q \geqslant 0$, $\forall \mathcal{F} \in \{\breve{o}_p, \widehat{\mathcal{O}}_p\}$, $\forall \mathcal{G} \in \{\breve{o}_q, \widehat{\mathcal{O}}_q\}$, either $\mathcal{F} \subseteq \mathcal{G}$ or $\mathcal{G} \subseteq \mathcal{F}$.

FACT 47.9. *Let V and W be finite dimensional vector spaces. We define $\breve{o} := \breve{o}(V,W)$ and $\widehat{\mathcal{O}} := \widehat{\mathcal{O}}(V,W)$. Then:*

(1A) $\breve{\mathcal{O}} + \breve{\mathcal{O}} \subseteq \breve{\mathcal{O}}$, *i.e.*, $\forall f, g \in \breve{\mathcal{O}}$, $f + g \in \breve{\mathcal{O}}$,
(1B) $\mathbb{R}\breve{\mathcal{O}} \subseteq \breve{\mathcal{O}}$, *i.e.*, $\forall a \in \mathbb{R}$, $\forall f \in \breve{\mathcal{O}}$, $af \in \breve{\mathcal{O}}$,
(2A) $\widehat{\mathcal{O}} + \widehat{\mathcal{O}} \subseteq \widehat{\mathcal{O}}$, *i.e.*, $\forall f, g \in \widehat{\mathcal{O}}$, $f + g \in \widehat{\mathcal{O}}$, *and*
(2B) $\mathbb{R}\widehat{\mathcal{O}} \subseteq \widehat{\mathcal{O}}$, *i.e.*, $\forall a \in \mathbb{R}$, $\forall f \in \widehat{\mathcal{O}}$, $af \in \widehat{\mathcal{O}}$.

Proof. Unassigned HW. □

FACT 47.10. *We define* $\breve{\mathcal{O}} := \breve{\mathcal{O}}(\mathbb{R}, \mathbb{R})$. *and* $\widehat{\mathcal{O}} := \widehat{\mathcal{O}}(\mathbb{R}, \mathbb{R})$. *Then:*

(1) $\breve{\mathcal{O}} \cdot \breve{\mathcal{O}} \subseteq \breve{\mathcal{O}}$, *i.e.*, $\forall f, g \in \breve{\mathcal{O}}$, $fg \in \breve{\mathcal{O}}$, *and*
(2) $\widehat{\mathcal{O}} \cdot \widehat{\mathcal{O}} \subseteq \widehat{\mathcal{O}}$, *i.e.*, $\forall f, g \in \widehat{\mathcal{O}}$, $fg \in \widehat{\mathcal{O}}$.

Proof. Unassigned HW. □

FACT 47.11. *Let V, W and X be finite dimensional VSs. Then:*

(1) $[\breve{\mathcal{O}}(W, X)] \circ [\breve{\mathcal{O}}(V, W)] \subseteq \breve{\mathcal{O}}(V, X)$, *and*
(2) $[\widehat{\mathcal{O}}(W, X)] \circ [\breve{\mathcal{O}}(V, W)] \subseteq \widehat{\mathcal{O}}(V, X)$.

Proof. Unassigned HW. □

WARNING: Let $\breve{\mathcal{O}} := \breve{\mathcal{O}}(\mathbb{R}, \mathbb{R})$ and let$\widehat{\mathcal{O}} := \widehat{\mathcal{O}}(\mathbb{R}, \mathbb{R})$. Then, unforunately, we have $\widehat{\mathcal{O}} \circ \widehat{\mathcal{O}} \nsubseteq \widehat{\mathcal{O}}$ and $\breve{\mathcal{O}} \circ \widehat{\mathcal{O}} \nsubseteq \widehat{\mathcal{O}}$. For example, define $f : \mathbb{R} \to \mathbb{R}$ and $\gamma : \mathbb{R}\backslash\{1\} \to \mathbb{R}$ by $f(x) = x + 1$ and $\gamma(x) = x/(x - 1)$. Let $g := \mathrm{adj}_1^2(\gamma)$. Then $f \in \widehat{\mathcal{O}}$ and $g \in \breve{\mathcal{O}} \subseteq \widehat{\mathcal{O}}$. Note that we have both $\mathrm{dom}[f] = \mathrm{dom}[g] = \mathbb{R}$ and $g(f(0)) = g(1) = 2$. However, for all $x \in \mathbb{R}\backslash\{0\}$, we calculate $(g(f(x)) = (x + 1)/x$. Then $g \circ f$ is not bounded near 0, so $g \circ f \notin \widehat{\mathcal{O}}$.

As a general rule, if you see a constant order function f on the *RIGHT* of a composite sign, then there's no way to control the composite, unless you somehow know that f vanishes at zero. This problem doesn't arise if you know, for some $p > 0$, that f is order p. It also doesn't arise if you know, for some $p \geqslant 0$, that f is sub-p.

FACT 47.12. *Let V and W be finite dimensional vector spaces and let $\| \bullet \| \in \mathcal{N}(W)$. Let $p > 0$. Then*

$$\|\breve{\mathcal{O}}(V, W)\|^p \subseteq \breve{\mathcal{O}}(V, \mathbb{R}) \qquad \textit{and} \qquad \|\widehat{\mathcal{O}}(V, W)\|^p \subseteq \widehat{\mathcal{O}}(V, \mathbb{R}).$$

Proof. By (1) of Fact 47.11, $(\breve{\mathcal{O}}(V, W)) \circ (\breve{\mathcal{O}}(V, W)) \subseteq \breve{\mathcal{O}}(V, \mathbb{R})$. Then

$$\begin{aligned} \|\breve{\mathcal{O}}(V, W)\|^p &= \| \bullet \|^p \circ (\breve{\mathcal{O}}(V, W)) \\ &\subseteq (\breve{\mathcal{O}}(V, W)) \circ (\breve{\mathcal{O}}(V, W)) \subseteq \breve{\mathcal{O}}(V, \mathbb{R}). \end{aligned}$$

It remains to show: $\|\widehat{\mathcal{O}}(V,W)\|^p \subseteq \widehat{\mathcal{O}}(V,\mathbb{R})$. Want: $\forall f \in \widehat{\mathcal{O}}(V,W)$, $\|f\|^p \in \widehat{\mathcal{O}}(V,\mathbb{R})$. Given $f \in \widehat{\mathcal{O}}(V,W)$. We wish to show: $\|f\|^p \in \widehat{\mathcal{O}}(V,\mathbb{R})$. Let $g := \|f\|^p$. We wish to show: $g \in \widehat{\mathcal{O}}(V,\mathbb{R})$.

Since $f \in \widehat{\mathcal{O}}(V,W)$, choose $U \in \mathcal{N}_V(0_V)$ s.t.

$$[\ U \subseteq \mathrm{dom}[f]\ ] \qquad \text{and} \qquad [\ \sup \|f_*(U)\| < \infty\ ].$$

Let $|\bullet| \in \mathcal{N}(\mathbb{R})$ be absolute value. It suffices to show:

$$[\ U \subseteq \mathrm{dom}[g]\ ] \qquad \text{and} \qquad [\ \sup |g_*(U)| < \infty\ ].$$

We have $\mathrm{dom}[g] = \mathrm{dom}[f]$. Then $U \subseteq \mathrm{dom}[f] = \mathrm{dom}[g]$. It remains to show: $\sup |g_*(U)| < \infty$. Let $M := \sup \|f_*(U)\|$. Then $0 \leqslant M < \infty$, so $M^p < \infty$. It therefore suffices to show: $|g_*(U)| \leqslant M^p$. We wish to show: $\forall x \in U$, $|g(x)| \leqslant M^p$. Let $x \in U$ be given. Want: $|g(x)| \leqslant M^p$.

Since $x \in U \subseteq \mathrm{dom}[f]$, we get $f(x) \in f_*(U)$. Then

$$\|f(x)\| \ \in\ \|f_*(U)\| \ \leqslant\ \sup \|f_*(U)\| \ =\ M.$$

As $0 \leqslant \|f(x)\| \leqslant M$, we get $0 \leqslant \|f(x)\|^p \leqslant M^p$. Then $0 \leqslant g(x) \leqslant M^p$. Since $0 \leqslant g(x)$, we get $|g(x)| = g(x)$. Then $|g(x)| = g(x) \leqslant M^p$. $\square$

THEOREM 47.13. *For all $s > 0$, let $\widehat{\mathcal{O}}_s := \widehat{\mathcal{O}}_s(\mathbb{R},\mathbb{R})$. Let $p > 0$ and $q \geqslant 0$. Then:* $\widehat{\mathcal{O}}_p \cdot \widehat{\mathcal{O}}_q \subseteq \widehat{\mathcal{O}}_{p+q}$ *and* $\widehat{\mathcal{O}}_q \circ \widehat{\mathcal{O}}_p \subseteq \widehat{\mathcal{O}}_{qp}$.

Proof. We wish to show: $\forall f \in \widehat{\mathcal{O}}_p$, $\forall g \in \widehat{\mathcal{O}}_q$,

$$fg \in \widehat{\mathcal{O}}_{p+q} \qquad \text{and} \qquad g \circ f \in \widehat{\mathcal{O}}_{qp}.$$

Let $f \in \widehat{\mathcal{O}}_p$ and $g \in \widehat{\mathcal{O}}_q$ be given. We wish to show:

$$fg \in \widehat{\mathcal{O}}_{p+q} \qquad \text{and} \qquad g \circ f \in \widehat{\mathcal{O}}_{qp}.$$

Let $\widehat{\mathcal{O}} := \widehat{\mathcal{O}}(\mathbb{R},\mathbb{R})$. By (2) of Fact 47.10, we have $\widehat{\mathcal{O}} \cdot \widehat{\mathcal{O}} \subseteq \widehat{\mathcal{O}}$. Let $|\bullet| \in \mathcal{N}(\mathbb{R})$ be absolute value. Then $f \in \widehat{\mathcal{O}}_q = |\bullet|^p \cdot \widehat{\mathcal{O}}$ and $g \in \widehat{\mathcal{O}}_q = |\bullet|^q \cdot \widehat{\mathcal{O}}$, so choose $\alpha, \beta \in \widehat{\mathcal{O}}$ s.t.

$$(\ f = |\bullet|^p \cdot \alpha\ ) \qquad \text{and} \qquad (\ g = |\bullet|^q \cdot \beta\ ).$$

Then

$$\begin{aligned} fg \ &=\ |\bullet|^p \cdot \alpha \cdot |\bullet|^q \cdot \beta \ =\ |\bullet|^{p+q} \cdot \alpha \cdot \beta \\ &\in\ |\bullet|^{p+q} \cdot \widehat{\mathcal{O}} \cdot \widehat{\mathcal{O}} \subseteq |\bullet|^{p+q} \cdot \widehat{\mathcal{O}}. \end{aligned}$$

It remains to show $g \circ f \in \widehat{\mathcal{O}}_{qp}$.

For all $x \in \mathbb{R}$, we have

$$\begin{aligned} g(f(x)) &= (|\bullet|^q \cdot \beta)(f(x)) = |f(x)|^q \cdot [\beta(f(x))] \\ &= \big|\, (|\bullet|^p \cdot \alpha)(x) \,\big|^q \cdot [\beta(f(x))] \\ &= \big|\, |x|^p \cdot [\alpha(x)] \,\big|^q \cdot [\beta(f(x))] \\ &= |x|^{qp} \cdot |\alpha(x)|^q \cdot [\beta(f(x))]. \end{aligned}$$

Then $g \circ f = |\bullet|^{qp} \cdot |\alpha|^q \cdot [\beta \circ f]$.

By Fact 47.12, $|\widehat{\mathcal{O}}|^q \subseteq \widehat{\mathcal{O}}$. Then $|\alpha|^q \in |\widehat{\mathcal{O}}|^q \subseteq \widehat{\mathcal{O}}$. Since $p > 0$, we have $\widehat{\mathcal{O}}_p \subseteq \breve{o}$ By (2) of Fact 47.11, we have $\widehat{\mathcal{O}} \circ \breve{o} \subseteq \widehat{\mathcal{O}}$. Then $\beta \circ f \in \widehat{\mathcal{O}} \circ \widehat{\mathcal{O}}_p \subseteq \widehat{\mathcal{O}} \circ \breve{o} \subseteq \widehat{\mathcal{O}}$. Recall that $\widehat{\mathcal{O}} \cdot \widehat{\mathcal{O}} \subseteq \widehat{\mathcal{O}}$. Then

$$\begin{aligned} g \circ f &= |\bullet|^{qp} \cdot |\alpha|^q \cdot [\beta \circ f] \\ &\in |\bullet|^{qp} \cdot \widehat{\mathcal{O}} \cdot \widehat{\mathcal{O}} = |\bullet|^{qp} \cdot \widehat{\mathcal{O}} = \widehat{\mathcal{O}}_{qp}, \end{aligned}$$

as desired. $\square$

Assigned HW#12-5.

THEOREM 47.14. *For all $s \geqslant 0$, let $\widehat{\mathcal{O}}_s := \widehat{\mathcal{O}}_s(\mathbb{R}, \mathbb{R})$. Let $p, q \geqslant 0$. Then:*

(1) $\widehat{\mathcal{O}}_p \cdot \widehat{\mathcal{O}}_q \subseteq \widehat{\mathcal{O}}_{p+q}$.
(2) $\widehat{\mathcal{O}}_p \cdot \breve{o}_q \subseteq \breve{o}_{p+q}$.
(3) $\breve{o}_p \cdot \widehat{\mathcal{O}}_q \subseteq \breve{o}_{p+q}$.
(4) $\breve{o}_p \cdot \breve{o}_q \subseteq \breve{o}_{p+q}$.

Proof. Unassigned HW. $\square$

In Theorem 47.14, simply remember that, when $p, q > 0$,

- if you mix $\widehat{\mathcal{O}}$ and $\widehat{\mathcal{O}}$, you get $\widehat{\mathcal{O}}$,
- if you mix $\widehat{\mathcal{O}}$ and $\breve{o}$, you get $\breve{o}$,
- if you mix $\breve{o}$ and $\widehat{\mathcal{O}}$, you get $\breve{o}$ and
- if you mix $\breve{o}$ and $\breve{o}$, you get $\breve{o}$.

Also remember that multiplication adds exponents.

THEOREM 47.15. *Let V, W and X be finite dimensional vector spaces. Then:*

(5) $\forall p > 0$, $\forall q \geqslant 0$, $\quad [\widehat{\mathcal{O}}_q(W, X)] \circ [\widehat{\mathcal{O}}_p(V, W)] \subseteq \widehat{\mathcal{O}}_{qp}(V, X)$.
(6) $\forall p \geqslant 0$, $\forall q > 0$, $\quad [\widehat{\mathcal{O}}_q(W, X)] \circ [\breve{o}_p(V, W)] \subseteq \breve{o}_{qp}(V, X)$.
(7) $\forall p > 0$, $\forall q \geqslant 0$, $\quad [\breve{o}_q(W, X)] \circ [\widehat{\mathcal{O}}_p(V, W)] \subseteq \breve{o}_{qp}(V, X)$.

(8) $\forall p \geqslant 0$, $\forall q \geqslant 0$, $\qquad [\breve{o}_q(W,X)] \circ [\breve{o}_p(V,W)] \subseteq \breve{o}_{qp}(V,X)$.

Proof. Unassigned HW. □

In Theorem 47.15, the boundary cases (where either $p = 0$ or $q = 0$) are generally not important. Simply remember that, when $p, q > 0$,

- if you mix $\widehat{\mathcal{O}}$ and $\widehat{\mathcal{O}}$, you get $\widehat{\mathcal{O}}$,
- if you mix $\widehat{\mathcal{O}}$ and $\breve{o}$, you get $\breve{o}$,
- if you mix $\breve{o}$ and $\widehat{\mathcal{O}}$, you get $\breve{o}$ and
- if you mix $\breve{o}$ and $\breve{o}$, you get $\breve{o}$.

Also remember that, whereas multiplication adds exponents, composition multiplies them.

If you're worried about learning the boundary cases, keep in mind that if you have $\widehat{\mathcal{O}}_0$ on the right side of $\circ$, then there is no inclusion. On the other hand, if you have $\widehat{\mathcal{O}}_0$ on the left hand side, then, with two exceptions, the answer is as expected. The first exception is $\widehat{\mathcal{O}}_0 \circ \widehat{\mathcal{O}}_0$, where there is no inclusion. The second exception is $\widehat{\mathcal{O}}_0 \circ \breve{o}_p$, which is not contained in $\breve{o}_0$, but *IS* contained in $\widehat{\mathcal{O}}_0$.

Specifically, let V, W, X be finite dimensional vector spaces. Then:

(5') $\forall q \geqslant 0$, $\qquad [\widehat{\mathcal{O}}_q(W,X)] \circ [\widehat{\mathcal{O}}_0(V,W)] \not\subseteq \widehat{\mathcal{O}}_0(V,X)$.
(6') $\forall p \geqslant 0$, $\qquad [\widehat{\mathcal{O}}_0(W,X)] \circ [\breve{o}_p(V,W)] \not\subseteq \breve{o}_0(V,X)$.
(6") $\forall p \geqslant 0$, $\qquad [\widehat{\mathcal{O}}_0(W,X)] \circ [\breve{o}_p(V,W)] \subseteq \widehat{\mathcal{O}}_0(V,X)$.
(7') $\forall q \geqslant 0$, $\qquad [\breve{o}_q(W,X)] \circ [\widehat{\mathcal{O}}_0(V,W)] \not\subseteq \breve{o}_0(V,X)$.
(7") $\forall q \geqslant 0$, $\qquad [\breve{o}_q(W,X)] \circ [\widehat{\mathcal{O}}_0(V,W)] \not\subseteq \widehat{\mathcal{O}}_0(V,X)$.

We leave all these as exercises. For (6") the argument is similar to the argument in Theorem 47.13. For (5'), (6'), (7') and (7"), see the WARNING following Fact 47.11. In (5') and (7"), keep in mind that $\widehat{\mathcal{O}}_0$ contains all the other function spaces. So, if a composition of functions spaces is not contained in $\widehat{\mathcal{O}}_0$, then it's not contained in any of the others. Thus, (7') is a consequence of (7").

48. Class 22 on 12 April 2018, Th of Week 12

Recall that, for any vector space V, for any set S, V^S has a standard vector space structure. In particular, since $\mathbb{R}^{3\times 4} = \mathbb{R}^{[1..3]\times[1..4]}$, we see that $\mathbb{R}^{3\times 4}$ has a standard vector space structure. We will practice vector addition and scalar multiplication in $\mathbb{R}^{3\times 4}$ soon. In the vector space $\mathbb{R}^{3\times 4}$, the vector addition is sometimes called matrix addition. Our mathematical world starts is scalars, proceeds to vectors, then to matrices, and then to what are called 3-dimensional tensors, to be discussed soon. From there we go to 4-dimensional tensors, and on and on.

Recall: $\forall m \in \mathbb{N}$, e^m is the standard ordered basis of $\mathbb{R}^m$, so

$$e_1^m = (1,0,\ldots,0) \in \mathbb{R}^m \quad , \quad \ldots \quad , \quad e_m^m = (0,\ldots,0,1) \in \mathbb{R}^m.$$

For all $m,n \in \mathbb{N}$, we let $E^{mn} \in (\mathbb{R}^{m\times n})^{m\times n}$ be a matrix of matrices whose ij entry is the matrix $E_{ij}^{mn} \in \mathbb{R}^{m\times n}$ is an $m \times n$ matrix with a one in the (i,j) entry and with 0s in all the other entries. So, for example, setting m to 3 and n to 4, we have

$$\begin{array}{ccc}
E_{11}^{34} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, & \ldots, & E_{14}^{34} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
\vdots & \vdots & \vdots \\
E_{31}^{34} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & \ldots, & E_{34}^{34} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
\end{array}$$

Note that $\{E_{ij}^{34} \,|\, i \in [1..3], j \in [1..4]\}$ is a set of 12 matrices; it is a basis of the 12-dimensional vector space $\mathbb{R}^{3\times 4}$. Let

$$X \quad := \quad \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{bmatrix} \quad \in \quad \mathbb{R}^{3\times 4}.$$

This variable X will be bound for this entire class. The $(2,3)$-entry of X is $X_{23} = 6$. The 2nd row of X is $X_{2\bullet} = (4,5,6,0) \in \mathbb{R}^4$. The 3rd column of X is $X_{\bullet 3} = (3,6,9) \in \mathbb{R}^3$. The (horizontal) row vector corresponding to $X_{2\bullet}$ is $(X_{2\bullet})^H = \begin{bmatrix} 4 & 5 & 6 & 0 \end{bmatrix} \in \mathbb{R}^{1\times 4}$. The (vertical) column vector corresponding to $X_{\bullet 3}$ is

$$(X_{\bullet 3})^V \quad = \quad \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \quad \in \quad \mathbb{R}^{3\times 1}.$$

We can reproduce $X_{\bullet 3}$ from the entries $3,6,9$ in the 3rd column of X:

$$X_{\bullet 3} \quad = \quad (3,6,9) \quad = \quad 3e_1 + 6e_2 + 9e_3 \quad = \quad \sum_{j=1}^{3} X_{j3} e_j^3.$$

The lazy among us would simply write $X_{\bullet 3} = \sum X_{j3} e_j$, expecting the reader to understand that, because the index j is repeated, the $\sum$ is on

j, and expecting the reader to figure out that j ranges from 1 to 3, and expecting the reader to understand that, here, e_j means e_j^3. The super lazy among us follow the Einstein convention that any repeated index in a term is automatically summed, and would write $X_{\bullet 3} = X_{j3}e_j$. We can also reproduce X via the super lazy formula $X = X_{ij}E_{ij}$. Less lazy is $X = \sum X_{ij}E_{ij}^{34}$. Better is $X = \sum_{i,j} X_{ij}E_{ij}^{34}$. Still better: $X = \sum_i \sum_j X_{ij}E_{ij}^{34}$. Still better: $X = \sum_{i=1}^{3} \sum_{j=1}^{4} X_{ij}E_{ij}^{34}$. Note that, by commutativity of addition of matrices, you can interchange the order of summation, and write $X = \sum_{j=1}^{4} \sum_{i=1}^{3} X_{ij}E_{ij}^{34}$.

Define $Y \in \mathbb{R}^{3\times 4}$ by

$$\begin{aligned} Y_{\bullet 1} &= (2, 0, 0) \\ Y_{\bullet 2} &= (0, 4, 0) \\ Y_{\bullet 3} &= (0, 0, 6) \\ Y_{\bullet 4} &= (9, 0, 8). \end{aligned}$$

Then, in a more conventional fomat, we have:

$$Y \quad = \quad \begin{bmatrix} 2 & 0 & 0 & 9 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 6 & 8 \end{bmatrix} \quad \in \quad \mathbb{R}^{3\times 4}.$$

This variable Y will be bound for this entire class.

Since $\mathbb{R}^{3\times 4}$ is a vector space, we should be able to calculate $10Y$ and $X + Y$. They are:

$$\begin{aligned} 10Y \quad &= \quad \begin{bmatrix} 20 & 0 & 0 & 90 \\ 0 & 40 & 0 & 0 \\ 0 & 0 & 60 & 80 \end{bmatrix} \qquad \text{and} \\ X + Y \quad &= \quad \begin{bmatrix} 3 & 2 & 3 & 9 \\ 4 & 9 & 6 & 0 \\ 7 & 8 & 15 & 8 \end{bmatrix}. \end{aligned}$$

Recall that $\mathbb{R}^{3\times 4\times 2} = \mathbb{R}^{[1..3]\times[1..4]\times[1..2]}$. Because $\mathbb{R}$ is a vector space, there is a standard vector space structure on $\mathbb{R}^{3\times 4\times 2}$.

Define $Z \in \mathbb{R}^{3\times 4\times 2}$ by $Z_{\bullet\bullet 1} = X$ and $Z_{\bullet\bullet 2} = Y$. There is no conventional format for displaying Z on a 2-dimensional page of paper, as Z

is, by its nature, 3-dimensional. It is an example of a 3-dimensional tensor. Be careful, however: $\dim(\mathbb{R}^{3\times 4\times 2}) = 3\cdot 4\cdot 2 = 24$, but each element of $\mathbb{R}^{3\times 4\times 2}$ is a 3-dimensional tensor. Since Z is 3-dimensional array of numbers, we have to look at various ways of “flattening” Z, so as to display it on a page. One approach is

$$Z_{\bullet\bullet 1} = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{bmatrix} \qquad \text{and}$$

$$Z_{\bullet\bullet 2} = \begin{bmatrix} 2 & 0 & 0 & 9 \\ 0 & 4 & 0 & 0 \\ 0 & 8 & 0 & 8 \end{bmatrix}.$$

Given a 3-dimensional tensor like Z, we would say that we “know” Z if, $\forall i \in [1..3]$, $\forall j \in [1..4]$, $\forall k \in [1..2]$, we can compute Z_{ijk}. So, for example, we should be able to compute Z_{321} and Z_{142}. We have

$$Z_{321} = X_{32} = 8 \qquad \text{and} \qquad Z_{142} = Y_{14} = 9.$$

This variable Z will be bound for this entire class.

Recall that $L(\mathbb{R}^3, \mathbb{R}^4)$ denotes the set of linear maps $R^3 \to \mathbb{R}^4$; it is a vector subspace of the vector space $(\mathbb{R}^4)^{\mathbb{R}^3}$, which means that $L(\mathbb{R}^3, \mathbb{R}^4)$ is closed under vector addition (*a.k.a.* function addition) and scalar multiplication. That is:

- $\forall F, G \in L(\mathbb{R}^3, \mathbb{R}^4), \qquad F + G \in L(\mathbb{R}^3, \mathbb{R}^4),$ and
- $\forall a \in \mathbb{R}, \forall F \in L(\mathbb{R}^3, \mathbb{R}^4), \qquad aF \in L(\mathbb{R}^3, \mathbb{R}^4).$

In other words,

- the sum of any two linear maps $\mathbb{R}^3 \to \mathbb{R}^4$ is again linear and
- any multiple of a linear map $\mathbb{R}^3 \to \mathbb{R}^4$ is again linear.

There is a standard vector space isomorphism

$$\mathbb{R}^{3\times 4} \quad \hookrightarrow\!> \quad L(\mathbb{R}^3, \mathbb{R}^4).$$

We begin by describing the image T of X under this isomorphism. Define $T \in L(\mathbb{R}^3, \mathbb{R}^4)$ by

$$\begin{aligned} T(e_1^3) &= X_{1\bullet}, \\ T(e_2^3) &= X_{2\bullet}, \\ T(e_3^3) &= X_{3\bullet}. \end{aligned}$$

This variable T will be bound for the entire class.

Because T is linear, we can now, for any $x, y, z \in \mathbb{R}$, compute $T(x, y, z)$. For example,

$$\begin{aligned}
T(2,4,5) &= T(2e_1^3 + 4e_2^3 + 5e_3^3) \\
&= 2 \cdot [T(e_1^3)] + 4 \cdot [T(e_2^3)] + 5 \cdot [T(e_3^3)] \\
&= 2 \cdot [X_{1\bullet}] + 4 \cdot [X_{2\bullet}] + 5 \cdot [X_{3\bullet}] \\
&= 2 \cdot (1,2,3,0) + 4 \cdot (4,5,6,0) + 5 \cdot (7,8,9,0) \\
&= (\,2\,,\,4\,,\,6\,,\,0) + (\,16\,,\,20\,,\,24\,,\,0\,) + (\,35\,,\,40\,,\,45\,,\,0\,) \\
&= (\,53\,,\,64\,,\,75\,,\,0\,).
\end{aligned}$$

More generally, for any $p \in \mathbb{R}^3$, we have

$$\begin{aligned}
Tp &= T\left(\sum_j p_j e_j^3\right) = \sum_j T\left(p_j e_j^3\right) \\
&= \sum_j p_j \cdot [T(e_j^3)] = \sum_j p_j \cdot [X_{j\bullet}] \\
&= \sum_j p_j \cdot \left[\sum_k X_{jk} e_k^4\right] = \sum_j \sum_k p_j X_{jk} e_k^4 \\
&= \sum_{j,k} p_j X_{jk} e_k^4
\end{aligned}$$

We used the matrix X to construct the linear map T. One can also go backwards: If someone knows T and can, for any $p \in \mathbb{R}^3$ compute Tp and tell it to us, then we can figure out X because the rows of X are $T(e_1^3)$, $T(e_2^3)$ and $T(e_3^3)$. That is, for all $j \in [1..3]$, we have $X_{j\bullet} = T(e_j^3)$. Recall the dot product $\bullet : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ given by $v \bullet w = v_1 w_1 + v_2 w_2 + v_3 w_3 + v_4 w_4$. For all $v \in \mathbb{R}^4$, for all $k \in [1..4]$, we have $v_k = v \bullet e_k$. Then, for all $j \in [1..3]$, for all $k \in [1..4]$, we have $X_{jk} = (T(e_j^3))_k = (T(e_j^3)) \cdot e_k^4$. This last formula allows us to recover all the entries of X from the linear map T.

We have now described a way of going from a matrix $X \in \mathbb{R}^{3\times 4}$ to a linear map $T \in L(\mathbb{R}^3, \mathbb{R}^4)$, and a way of going back from T to X. We generalize this to $\mathbb{R}^{m\times n}$ and $L(\mathbb{R}^m, \mathbb{R}^n)$ as follows:

DEFINITION 48.1. *Let $m, n \in \mathbb{N}$. Then*

(1) $\forall A \in \mathbb{R}^{m\times n}$, *$L_A \in L(\mathbb{R}^m, \mathbb{R}^n)$ is defined by* $L_A(p) = \sum_{j,k} p_j A_{jk} e_k^n$.

(2) $\forall F \in L(\mathbb{R}^m, \mathbb{R}^n)$, *$[F] \in \mathbb{R}^{m\times n}$ is defined by* $[F]_{jk} = [F(e_j^m)] \bullet e_k^n$.

With this notation, we have $L_X = T$ and $[T] = X$.

REMARK 48.2. *Let $m, n \in \mathbb{N}$. Then the two maps*

$$\begin{array}{ccc} A & \mapsto & L_A \\ \mathbb{R}^{m\times n} & \leftrightarrow & L(\mathbb{R}^m, \mathbb{R}^n) \\ [F] & \leftarrow\!\shortmid & F \end{array}$$

are both linear, and they are inverses of one another.

Proof. Unassigned HW. □

By Remark 48.2, for all $m, n \in \mathbb{N}$, the two vector spaces $\mathbb{R}^{m\times n}$ and $L(\mathbb{R}^m, \mathbb{R}^n)$ are isomorphic to one another, and, in fact, Remark 48.2 displays isomorphisms in each direction. In this isomorphism, $X \leftrightarrow T$.

For $m, n \in \mathbb{N}$, many people prefer the to set up an isomorphism $\mathbb{R}^{n\times m} \leftrightarrow L(\mathbb{R}^m, \mathbb{R}^n)$. Note that, on the left of $\leftrightarrow$, the n appears before the m, whereas, on the right, it's the other way around. Under this new isomorphism, the matrix corrsponding to $T \in L(\mathbb{R}^3, \mathbb{R}^4)$ would not be X, but rather, the transpose of X:

$$X^t \quad = \quad \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \\ 0 & 0 & 0 \end{bmatrix} \quad \in \quad \mathbb{R}^{4\times 3}.$$

For this exposition, we prefer the isomorphism of Remark 48.2.

We now define multiplication of matrices, in terms of composition:

DEFINITION 48.3. *Let $\ell, m, n \in \mathbb{N}$, $A \in \mathbb{R}^{\ell\times m}$, $B \in \mathbb{R}^{m\times n}$. Then $AB := [L_B \circ L_A]$.*

The next result asserts: the (i, k) entry of AB is $(A_{i\bullet}) \bullet (B_{\bullet j})$, *i.e.*, the dot product of (the ith row of A) and (the jth column of B).

REMARK 48.4. *Let $\ell, m, n \in \mathbb{N}$, $A \in \mathbb{R}^{\ell\times m}$, $B \in \mathbb{R}^{m\times n}$. Let $i \in [1..\ell]$ and $k \in [1..n]$. Then $(AB)_{ik} = \sum_{j=1}^{m} A_{ij}B_{jk}$.*

Proof. Since $L_\bullet$ and $[\bullet]$ are inverses, we see that $[L_A] = A$ and that $[L_B] = B$. Then, for all $j \in [1..m]$, we have $A_{ij} = [L_A]_{ij} = (L_A(e_i^\ell))\bullet e_j^m$ and $B_{jk} = [L_B]_{jk} = (L_B(e_j^m)) \bullet e_k^n$.

We have $(AB)_{ik} = [L_B\circ L_A]_{ik} = ((L_B\circ L_A)(e_i^\ell))\bullet e_k^n$. Let $v := L_A(e_i^\ell)$. Then $(L_B \circ L_A)(e_i^\ell) = L_B(L_A(e_i^\ell)) = L_B(v)$. Also, $\forall j \in [1..m]$, we have: $v_j = v \bullet e_j^m = (L_A(e_i^\ell)) \bullet e_j^m = A_{ij}$. Since $v = \sum_{j=1}^{m} v_j e_j^m$, we get $L_B(v) = \sum_{j=1}^{m} v_j \cdot [L_B(e_j^m)]$. Then

$$\begin{aligned}
(AB)_{ik} &= ((L_B \circ L_A)(e_i^\ell)) \bullet e_k^n \\
&= [L_B(v)] \bullet e_k^n \\
&= \left(\sum_{j=1}^{m} v_j \cdot [L_B(e_j^m)] \right) \bullet e_k^n \\
&= \sum_{j=1}^{m} v_j \cdot \big([L_B(e_j^m)] \bullet e_k^n \big) \\
&= \sum_{j=1}^{m} A_{ij} B_{jk},
\end{aligned}$$

as desired. $\square$

Next on the agenda: bilinear maps.

Let U, V and W be sets and let $F : U \times V \to W$. Then, $\forall v \in V$, $F(\bullet, v) : U \to W$ is defined by $(F(\bullet, v))(u) = F(u, v)$. Also, $\forall u \in U$, $F(u, \bullet) : V \to W$ is defined by $(F(u, \bullet))(v) = F(u, v)$.

DEFINITION 48.5. *Let U, V and W be vector spaces. Then $B(U, V, W)$ denotes the set of all $F : U \times V \to W$ such that both*

(1) $\forall v \in V$, $\quad F(\bullet, v) \in L(U, W)$ *and*
(2) $\forall u \in U$, $\quad F(u, \bullet) \in L(V, W)$.

Condition (1) of Definition 48.5 is expressed by saying that F is "linear in its first variable". Condition (2) is expressed by saying that F is "linear in its second variable". Conditions (1) and (2) together are expressed by saying that F is "**bilinear**".

Let U, V and W be vector spaces. Then $W^{U \times V}$ has a standard vector space structure, and we leave it as an unassigned exercise to verify that $B(U, V, W)$ is a vector subspace of $W^{U \times V}$. That is: (a sum of two bilinear maps is again bilinear) AND (a scalar multiple of a bilinear map is again bilinear).

Let U, V and W be vector spaces and let $* \in B(U, V, W)$. For all $u \in U$, $v \in V$, it is traditional to use the notation $u * v$ to denote $*(u, v)$. Similar remarks apply when $*$ is replaced by any special character, not in an alphabet. When we get to *tri*linear maps, there's no such convention, and special characters are not typically used to denote trilinear maps.

For all $m \in \mathbb{N}$, there is a standard example of a bilinear function called dot product: $\bullet \in B(\mathbb{R}^m, \mathbb{R}^m, \mathbb{R})$. That is, for any $m \in \mathbb{N}$, the dot product is a bilinear functional on $\mathbb{R}^m \times \mathbb{R}^m$. (The word "functional" in place of "function" indicates that the target vector space is $\mathbb{R}$.)

To give just one example of a bilinear functional $*$ on $\mathbb{R}^3 \times \mathbb{R}^4$, recall:

$$X \quad = \quad \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{bmatrix} \quad \in \quad \mathbb{R}^{3\times 4}.$$

Define $* \in B(\mathbb{R}^3, \mathbb{R}^4, \mathbb{R})$ by

$$\forall i \in [1..3], \;\; \forall j \in [1..4], \qquad e_i^3 * e_j^4 = X_{ij}.$$

The reader may expect that, to define a bilinear $* : \mathbb{R}^3 \times \mathbb{R}^4 \to \mathbb{R}$, we need, $\forall p \in \mathbb{R}^3$, $\forall q \in \mathbb{R}^4$, to give a formula for $p * q$. However, if we know every $e_i^3 * e_j^4$, then we can use bilinearity to compute $p * q$. For example, say we want to compute $(5, 2, 1) * (4, 6, 1, 2)$. First, we have

$$(5, 2, 1) * (4, 6, 1, 2) \;\; = \;\; (5e_1^3 + 2e_2^3 + e_3^3) * (4e_1^4 + 6e_2^4 + e_3^4 + 2e_4^4),$$

and via bilinearity of $*$, we cna expand the right hand side into 12 terms, the first of which would be $5 \cdot 4 \cdot (e_1^3 * e_1^4)$, which is equal to $5 \cdot 4 \cdot X_{11} = 5 \cdot 4 \cdot 1 = 20$. The others are all easy to compute, and we leave it as an exercise for the reader to compute them and add them; the answer is 284. In class, we gave a description of how to get at this same 12 term computation by writing $(5, 2, 1)$ vertically on the right of X, writing $(4, 6, 1, 2)$ horizontally across the top, and the multiplying each entry of X by the numbers above it and to its right. Those 12 products are the same as the terms in the expansion of $(5, 2, 1) * (4, 6, 1, 2)$. Generalizing, $\forall p \in \mathbb{R}^3$, $\forall q \in \mathbb{R}^3$, we compute

$$\begin{aligned} p * q \;\; &= \;\; \left(\sum_i p_i e_i^3 \right) * \left(\sum_j q_j e_j^4 \right) \\ &= \;\; \sum_{i,j} p_i \cdot q_j \cdot (e_i^3 * e_j^4) \\ &= \;\; \sum_{i,j} p_i q_j X_{ij}, \end{aligned}$$

and it is understood that in each sum, i ranges over $[1..3]$, while j ranges over $[1..4]$. Note that the final sum has 12 terms.

The variable $*$, defined in the last paragraph, will be bound for the rest of this class. Note: A person who forgets X (but remembers $*$) can recover each of the 12 entries of X from the 12 equations:

$$\forall i \in [1..3],\ \ \forall j \in [1..4], \qquad X_{ij} = e_i^3 * e_j^4.$$

The preceding discussion motivates:

DEFINITION 48.6. *Let $\ell, m \in \mathbb{R}$. Then*

(1) $\forall A \in \mathbb{R}^{\ell\times m}$, $B_A \in B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R})$ *is defined by* $B_A(e_i^\ell, e_j^m) = A_{ij}$.
(2) $\forall F \in B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R})$, $[F] \in \mathbb{R}^{\ell\times m}$ *is defined by* $[F]_{ij} = F(e_i^\ell, e_j^m)$.

For the reader who, in (1) of Definition 48.6, would prefer to see a general formula for $p * q$, we have: For all $\ell, m \in \mathbb{N}$, for all $A \in \mathbb{R}^{\ell\times m}$,

$$\forall p \in \mathbb{R}^\ell,\ \ \forall q \in \mathbb{R}^m, \qquad B_A(p,q) = \sum_{i,j} p_i q_j A_{ij},$$

and, in the sum, i ranges over $[1..\ell]$ and j over $[1..m]$; the sum has ℓm terms. (When $\ell = 3$ and $m = 4$, we get the expected 12 terms.)

In the notation of Definition 48.6, $B_X = *$ and $[*] = X$.

REMARK 48.7. *Let $\ell, m \in \mathbb{N}$. Then the two maps*

$$\begin{array}{ccc} A & \mapsto & B_A \\ \mathbb{R}^{\ell\times m} & \leftrightarrow & B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R}) \\ [F] & \leftarrow\!\shortmid & F \end{array}$$

are both linear, and they are inverses of one another.

Proof. Unassigned HW. □

By Remark 48.7, for all $\ell, m \in \mathbb{N}$, the two vector spaces $\mathbb{R}^{\ell\times m}$ and $B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R})$ are isomorphic to one another, and, in fact, Remark 48.2 displays isomorphisms in each direction. In this isomorphism, $X \leftrightarrow *$.

We've been doing bilinear functionals. Now let's approach vector valued bilinear functions. First, recall: $X, Y \in \mathbb{R}^{3\times 4}$ are given by

$$X = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{bmatrix}, \qquad Y = \begin{bmatrix} 2 & 0 & 0 & 9 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 6 & 8 \end{bmatrix}.$$

Also, recall: $Z \in \mathbb{R}^{3\times 4\times 2}$ is defined by $Z_{\bullet\bullet 1} = X$ and $Z_{\bullet\bullet 2} = Y$. Then $Z_{131} = X_{13} = 7$ and $Z_{132} = Y_{13} = 0$. For all $i \in [1..3]$, for all $j \in [1..4]$,

$Z_{ij\bullet} \in \mathbb{R}^2$ is defined by $(Z_{ij\bullet})_k = Z_{ijk}$. So, since $Z_{131} = X_{13} = 7$ and $Z_{132} = Y_{13} = 0$, we see that $Z_{13\bullet} = (7, 0)$. Note that we can reproduce $Z_{13\bullet}$ as $Z_{13\bullet} = Z_{131}e_1^2 + Z_{132}e_2^2$. In general,

$$\forall i \in [1..3],\ \forall j \in [1..4], \qquad Z_{ij\bullet} = \sum_{k=1}^{2} Z_{ijk}e_k^2.$$

Note: $\forall i \in [1..3]$, $\forall j \in [1..4]$, $\forall k \in [1..2]$, $Z_{ijk} = Z_{ij\bullet} \bullet e_k^2$.

There is a standard VS isomorphism $\mathbb{R}^{3\times 4\times 2} \hookrightarrow\!\!> B(\mathbb{R}^3, \mathbb{R}^4, \mathbb{R}^2)$, and we begin by describing the image $\circledast \in B(\mathbb{R}^3, \mathbb{R}^4, \mathbb{R}^2)$ of Z. Define $\circledast \in B(\mathbb{R}^3, \mathbb{R}^4, \mathbb{R}^2)$ by

$$\forall i \in [1..3],\ \forall j \in [1..4], \qquad e_i^3 \circledast e_j^4 = Z_{ij\bullet},$$

or, equivalently by

$$\forall i \in [1..3],\ \forall j \in [1..4], \qquad e_i^3 \circledast e_j^4 = \sum_{k=1}^{2} Z_{ijk}e_k^2,$$

or, equivalently by

$$\forall p \in \mathbb{R}^3,\ \forall q \in \mathbb{R}^4, \qquad p \circledast q = \sum_{i,j,k} p_i q_j Z_{ijk} e_k^2,$$

and, in the sum, i ranges over $[1..3]$, j over $[1..4]$ and k over $[1..2]$.

The variable $\circledast$, defined in the last paragraph, will be bound for the rest of this class. Note: A person who forgets Z (but remembers $\circledast$) can recover each of the 24 entries of Z from the 24 equations:

$$\forall i \in [1..3],\ \forall j \in [1..4],\ \forall k \in [1..2], \qquad Z_{ijk} = (e_i^3 \circledast e_j^4) \bullet e_k^2.$$

The preceding discussion motivates:

DEFINITION 48.8. *Let $\ell, m, n \in \mathbb{R}$. Then*

(1) $\forall A \in \mathbb{R}^{\ell\times m\times n}$, $B_A \in B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R}^n)$ *is defined by*

$$B_A(e_i^\ell, e_j^m) \quad = \quad \sum_{k=1}^{n} A_{ijk}e_k^n.$$

(2) $\forall F \in B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R}^n)$, $[F] \in \mathbb{R}^{\ell\times m\times n}$ *is defined by*

$$[F]_{ijk} \quad = \quad [F(e_i^\ell, e_j^m)] \bullet e_k^n.$$

For the reader who, in (1) of Definition 48.8, would prefer to see a general formula for $p{*}q$, we have: For all $\ell, m, n \in \mathbb{N}$, for all $A \in \mathbb{R}^{\ell\times m\times n}$,

$$\forall p \in \mathbb{R}^\ell,\ \forall q \in \mathbb{R}^m, \qquad B_A(p, q) = \sum_{i,j,k} p_i q_j A_{ij} e_k^n,$$

and, in the sum, i ranges over $[1..\ell]$, j over $[1..m]$ and k over $[1..n]$; the sum has ℓmn terms. (When $\ell = 3$, $m = 4$, $n = 2$, we get 24 terms.)

In the notation of Definition 48.8, $B_Z = \circledast$ and $[\circledast] = Z$.

REMARK 48.9. *Let $\ell, m, n \in \mathbb{N}$. Then the two maps*

$$\begin{array}{ccc} A & \mapsto & B_A \\ \mathbb{R}^{\ell\times m\times n} & \leftrightarrow & B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R}^n) \\ [F] & \leftarrow\!\shortmid & F \end{array}$$

are both linear, and they are inverses of one another.

Proof. Unassigned HW. □

By Remark 48.9, for all $\ell, m, n \in \mathbb{N}$, the two vector spaces $\mathbb{R}^{\ell\times m\times n}$ and $B(\mathbb{R}^\ell, \mathbb{R}^m, \mathbb{R}^n)$ are isomorphic to one another. Remark 48.9 displays isomorphisms in each direction. In this isomorphism, $Z \leftrightarrow \circledast$.

We've been doing bilinear. Now let's do trilinear.

Let U, V W and X be sets and let $F : U \times V \times W \to X$. Then, $\forall v \in V$, $\forall w \in W$, $F(\bullet, v, w) : U \to X$ is defined by

$$(F(\bullet, v, w))(u) \quad = \quad F(u, v, w).$$

Also, $\forall u \in U$, $\forall w \in W$, $F(u, \bullet, w) : V \to X$ is defined by

$$(F(u, \bullet, w))(v) \quad = \quad F(u, v, w).$$

Also, $\forall u \in U$, $\forall v \in V$, $F(u, v, \bullet) : W \to X$ is defined by

$$(F(u, v, \bullet))(w) \quad = \quad F(u, v, w).$$

DEFINITION 48.10. *Let U, V, W and X be vector spaces. Then $T(U, V, W, X)$ denotes the set of all $F : U \times V \times W \to X$ such that*

(1) $\forall v \in V$, $\forall w \in W$, $\quad F(\bullet, v, w) \in L(U, X)$ *and*
(2) $\forall u \in U$, $\forall w \in W$, $\quad F(u, \bullet, w) \in L(V, X)$.
(3) $\forall u \in U$, $\forall v \in V$, $\quad F(u, v, \bullet) \in L(W, X)$.

Condition (1) of Definition 48.5 is expressed by saying that F is "linear in its first variable". Condition (2) is expressed by saying that F is "linear in its second variable". Condition (3) is expressed by saying that F is "linear in its third variable". Conditions (1), (2) and (3) together are expressed by saying that F is "**trilinear**".

Let U, V, W and X be vector spaces. Then $X^{U\times V\times W}$ has a standard vector space structure, and we leave it as an unassigned exercise to verify that $T(U, V, W, X)$ is a vector subspace of $X^{U\times V\times W}$. That is, we have: both (a sum of two trilinear maps is again trilinear) AND (a scalar multiple of a trilinear map is again trilinear).

DEFINITION 48.11. *Let $m_1, m_2, m_3 \in \mathbb{R}$. Then*

(1) $\forall A \in \mathbb{R}^{m_1 \times m_2 \times m_3}$, $B_A \in B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R})$ *is defined by*
$$B_A(e_{j_1}^{m_1}, e_{j_2}^{m_2}, e_{j_3}^{m_3}) \quad = \quad A_{j_1 j_2 j_3}.$$
(2) $\forall F \in B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R})$, $[F] \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ *is defined by*
$$[F]_{j_1 j_2 j_3} \quad = \quad F(e_{j_1}^{m_1}, e_{j_2}^{m_2}, e_{j_3}^{m_3}).$$

For the reader who, in (1) of Definition 48.11, would prefer to see a general formula for $F(p^1, p^2, p^3)$, we have:

For all $m_1, m_2, m_3 \in \mathbb{N}$, for all $A \in \mathbb{R}^{m_1 \times m_2 \times m_3}$,
$$\forall p^1 \in \mathbb{R}^{m_1},\ \forall p^2 \in \mathbb{R}^{m_2}\ \forall p^3 \in \mathbb{R}^{m_3},$$
$$B_A(p^1, p^2, p^3) = \sum_{j_1, j_2, j_3} p_{j_1}^1 p_{j_2}^2 p_{j_3}^3 q_j A_{j_1 j_2 j_3},$$
and, in the sum, j_1 ranges over $[1..m_1]$, j_2 over $[1..m_2]$ and $j_3$over $[1..m_3]$; the sum has $m_1 m_2 m_3$ terms.

REMARK 48.12. *Let $m_1, m_2, m_3 \in \mathbb{N}$. Then the two maps*
$$\begin{array}{ccc} A & \mapsto & B_A \\ \mathbb{R}^{m_1 \times m_2 \times m_3} & \leftrightarrow & B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R}) \\ [F] & \leftarrow\!\shortmid & F \end{array}$$
are both linear, and they are inverses of one another.

Proof. Unassigned HW. □

By Remark 48.7, for all $m_1, m_2, m_3 \in \mathbb{N}$, the two vector spaces $\mathbb{R}^{m_1 \times m_2 \times m_3}$ and $B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R})$ are isomorphic to one another, and, in fact, Remark 48.2 displays isomorphisms in each direction.

We've been doing trilinear functionals. Now let's approach vector valued trilinear functions.

DEFINITION 48.13. *Let $m_1, m_2, m_3, n \in \mathbb{N}$, Then*

(1) $\forall A \in \mathbb{R}^{m_1 \times m_2 \times m_3 \times n}$, $T_A \in B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R}^n)$ *is defined by*
$$T_A(e_{j_1}^{m_1}, e_{j_2}^{m_2}, e_{j_3}^{m_3}) \quad = \quad \sum_{k=1}^{n} A_{j_1 j_2 j_3 k} e_k^n.$$
(2) $\forall F \in B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R}^n)$, $[F] \in \mathbb{R}^{m_1 \times m_2 \times m_3 \times n}$ *is defined by*
$$[F]_{j_1 j_2 j_3 k} \quad = \quad [F(e_{j_1}^{m_1}, e_{j_2}^{m_2}, e_{j_3}^{m_3})] \bullet e_k^n.$$

For the reader who, in (1) of Definition 48.13, would prefer to see a general formula for $F(p^1, p^2, p^3)$, we have:

For all $m_1, m_2, m_3 \in \mathbb{N}$, for all $A \in \mathbb{R}^{m_1 \times m_2 \times m_3}$,

$$\forall p^1 \in \mathbb{R}^{m_1},\ \forall p^2 \in \mathbb{R}^{m_2}\ \forall p^3 \in \mathbb{R}^{m_3},$$
$$T_A(p^1, p^2, p^3) = \sum_{j_1, j_2, j_3, k} p^1_{j_1} p^2_{j_2} p^3_{j_3} q_j A_{j_1 j_2 j_3 k} e^n_k,$$

and, in the sum, j_1 ranges over $[1..m_1]$, j_2 ranges $[1..m_2]$ and j_3 over $[1..m_3]$, k over $[1..n]$; the sum has $m_1 m_2 m_3 n$ terms.

REMARK 48.14. *Let $m_1, m_2, m_3, n \in \mathbb{N}$. Then the two maps*

$$\begin{array}{ccc} A & \mapsto & T_A \\ \mathbb{R}^{m_1 \times m_2 \times m_3 \times n} & \leftrightarrow & B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R}^n) \\ [F] & \leftarrow\!\shortmid & F \end{array}$$

are both linear, and they are inverses of one another.

Proof. Unassigned HW. $\square$

By Remark 48.7, for all $m_1, m_2, m_3, n \in \mathbb{N}$, the two vector spaces $\mathbb{R}^{m_1 \times m_2 \times m_3 \times n}$ and $B(\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \mathbb{R}^{m_3}, \mathbb{R}^n)$ are isomorphic to one another, and, in fact, Remark 48.2 displays isomorphisms in each direction.

We've been doing trilinear. Next: general multilinear functions.

DEFINITION 48.15. *Let $d \in \mathbb{N}$. Let $V_1, \ldots, V_d$ be vector spaces. Let X be a vector space. Then $M^d(V_1, \ldots, V_d, X)$ denotes the set of all $F : V_1 \cdots V_d \to X$ such that*

(1) $\forall v_2 \in V_2, \ldots, \forall v_d \in V_d, \quad F(\bullet, v_2, \ldots, v_d) \in L(V_1, X)$ *and*

$\qquad \cdots\cdots\cdots\cdots \qquad$ *and*

(d) $\forall v_1 \in V_1, \ldots, \forall v_{d-1} \in V_{d-1}, \quad F(v_1, \ldots, v_{d-1}, \bullet) \in L(V_d, X)$.

In Definition 48.15, $\forall c \in [1..d]$, Condition (c) is expressed by saying that F is "linear in its cth variable". Conditions (1), ..., (d) together are expressed by saying that F is "**multilinear**" or "d**-multilinear**".

Let $d \in \mathbb{N}$. Let $V_1, \ldots, V_d$ be vector spaces. Let X be a vector space. Then $X^{V_1 \times \cdots \times V_d}$ has a standard vector space structure, and we leave it as an unassigned exercise to verify that $M^d(V_1, \ldots, V_d, X)$ is a vector subspace of $X^{V_1 \times \cdots \times V_d}$. That is, we have: both (a sum of two multilinear maps is again multilinear) AND (a scalar multiple of a multilinear map is again multilinear).

DEFINITION 48.16. *Let $d \in \mathbb{N}$. Let $m_1, \ldots, m_d \in \mathbb{R}$. Then*

(1) $\forall A \in \mathbb{R}^{m_1 \times \cdots \times m_d}$, $M_A^d \in M^d(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R})$ *is defined by*

$$M_A^d(e_{j_1}^{m_1}, \ldots, e_{j_d}^{m_d}) \quad = \quad A_{j_1 \cdots j_d}.$$

(2) $\forall F \in B(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R})$, $[F] \in \mathbb{R}^{m_1 \times \cdots \times m_d}$ *is defined by*

$$[F]_{j_1 \cdots j_d} \quad = \quad F(e_{j_1}^{m_1}, \ldots, e_{j_d}^{m_d}).$$

For the reader who, in (1) of Definition 48.16 would prefer to see a general formula for $F(p^1, \ldots, p^d)$, we have:

For all $d \in \mathbb{N}$, for all $m_1, \ldots, m_d \in \mathbb{N}$, for all $A \in \mathbb{R}^{m_1 \times \cdots \times m_d}$,

$$\forall p^1 \in \mathbb{R}^{m_1}, \ \ldots, \ \forall p^d \in \mathbb{R}^{m_d},$$
$$M_A^d(p^1, \ldots, p^d) = \sum_{j_1, \ldots, j_d} p_{j_1}^1 \cdots p_{j_d}^d A_{j_1 \cdots j_d},$$

and, in the sum, j_1 ranges over $[1..m_1]$, $\ldots$, j_d over $[1..m_d]$; the sum has $m_1 \cdots m_d$ terms.

REMARK 48.17. *Let $d \in \mathbb{N}$. Let $m_1, \ldots, m_d \in \mathbb{N}$. Then*

$$\begin{array}{ccc} A & \mapsto & M_A^d \\ \mathbb{R}^{m_1 \times \cdots \times m_d} & \leftrightarrow & B(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R}) \\ [F] & \leftarrow\!\shortmid & F \end{array}$$

are both linear, and they are inverses of one another.

Proof. Unassigned HW. □

By Remark 48.7, for all $d \in \mathbb{N}$, for all $m_1, \ldots, m_d \in \mathbb{N}$, the two vector spaces $\mathbb{R}^{m_1 \times \cdots \times m_d}$ and $B(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R})$ are isomorphic to one another. Remark 48.7 displays isomorphisms in each direction.

We've been doing multilinear functionals. Finally, let's approach vector valued multilinear functions.

DEFINITION 48.18. *Let $d \in \mathbb{N}$, $m_1, \ldots, m_d \in \mathbb{N}$, $n \in \mathbb{N}$. Then*

(1) $\forall A \in \mathbb{R}^{m_1 \times \cdots \times m_d \times n}$, $M_A^d \in B(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R}^n)$ *is defined by*

$$M_A^d(e_{j_1}^{m_1}, \ldots, e_{j_d}^{m_d}) \quad = \quad \sum_{k=1}^{n} A_{j_1 \cdots j_d k} e_k^n.$$

(2) $\forall F \in B(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R}^n)$, $[F] \in \mathbb{R}^{m_1 \times \cdots \times m_d \times n}$ *is defined by*

$$[F]_{j_1 \cdots j_d k} \quad = \quad [F(e_{j_1}^{m_1}, \ldots, e_{j_d}^{m_d})] \bullet e_k^n.$$

For the reader who, in (1) of Definition 48.18 would prefer to see a general formula for $F(p^1, \ldots, p^d)$, we have:

For all $d \in \mathbb{N}$, for all $m_1, \ldots, m_d \in \mathbb{N}$, for all $A \in \mathbb{R}^{m_1 \times \cdots \times m_d}$,

$$\forall p^1 \in \mathbb{R}^{m_1},\ \ldots,\ \forall p^d \in \mathbb{R}^{m_d},$$
$$M_A^d(p^1, \ldots, p^d) = \sum_{j_1, \ldots, j_d, k} p^1_{j_1} \cdots p^d_{j_d} A_{j_1 \cdots j_d k} e^n_k,$$

and, in the sum, j_1 ranges over $[1..m_1]$, $\ldots$, j_d over $[1..m_d]$, k over $[1..n]$; the sum has $m_1 \cdots m_d n$ terms.

REMARK 48.19. *Let $d \in \mathbb{N}$. Let $m_1, \ldots, m_d \in \mathbb{N}$. Let $n \in \mathbb{N}$. Then*

$$\begin{array}{ccc} A & \mapsto & M_A^d \\ \mathbb{R}^{m_1 \times \cdots \times m_d \times n} & \leftrightarrow & B(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R}^n) \\ [F] & \leftarrow\!\shortmid & F \end{array}$$

are both linear, and they are inverses of one another.

Proof. Unassigned HW. □

By Remark 48.19, for all $d \in \mathbb{N}$, for all $m_1, \ldots, m_d, n \in \mathbb{N}$, the two vector spaces $\mathbb{R}^{m_1 \times \cdots \times m_d \times n}$ and $B(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}, \mathbb{R}^n)$ are isomorphic to one another. Remark 48.19 displays isomorphisms in each direction.

49. Class 23 on 17 April 2018, Tu of Week 13

Assigned HW#13-1, HW#13-2 and HW#13-3.

THEOREM 49.1. *Let V and W be finite dimensional normed vector spaces. Let Z be a normed vector space. Let $* \in B(V, W, Z)$. Then: $\exists K \geqslant 0$ such that, $\forall v \in V$, $\forall w \in W$, $\quad |v * w|_Z \leqslant K \cdot |v|_V \cdot |w|_W$.*

Proof. Let $m := \dim V$, $n := \dim W$. Choose $B \in \mathrm{OB}(V)$, $C \in \mathrm{OB}(W)$. Then $L_B : \mathbb{R}^m \to V$ and $L_C : \mathbb{R}^n \to W$ are vector space isomorphisms. Let $V_1 := (\mathbb{R}^m, |\bullet|_{m,1})$, $W_1 := (\mathbb{R}^n, |\bullet|_{n,1})$. By (1) of Theorem 46.1, $L_B^{-1} : V \to V_1$ and $L_C^{-1} : W \to W_1$ are both bounded. We define $S := (L_B^{-1})^{\wedge}_{V,V_1}$ and $T := (L_C^{-1})^{\wedge}_{W,W_1}$. Then $L_B^{-1} : V \to V_1$ is S-bounded and $L_C^{-1} : W \to W_1$ is T-bounded.

Define $\circledast \in B(V_1, W_1, Z)$ by $x \circledast y = [L_B(x)] * [L_C(y)]$. By HW#13-2, choose $K_1 \geqslant 0$ such that, $\forall x \in V_1$, $\forall y \in W_1$, $|x \circledast y|_Z \leqslant K_1 \cdot |x|_{V_1} \cdot |y|_{W_1}$. Let $K := K_1 S T$. Want: $\forall v \in V$, $\forall w \in W$, $|v * w|_Z \leqslant K \cdot |v|_V \cdot |w|_W$. Let $v \in V$ and $w \in W$ be given. We wish to show: $|v * w|_Z \leqslant K \cdot |v|_V \cdot |w|_W$.

Let $x := L_B^{-1}(v)$ and let $y := L_C^{-1}(w)$. Then

$$x \circledast y \quad = \quad [L_B(x)] * [L_C(y)] \quad = \quad v * w.$$

Since $L_B^{-1} : V \to V_1$ is S-bounded, we get $|L_B^{-1}(v)|_{V_1} \leqslant S|v|_V$. Since $L_C^{-1} : W \to W_1$ is T-bounded, we get $|L_C^{-1}(w)|_{W_1} \leqslant T|w|_W$. Then

$$\begin{aligned} |v * w|_Z &= |x \circledast y|_Z \leqslant K_1 \cdot |x|_{V_1} \cdot |y|_{V_2} \\ &= K_1 \cdot |L_B^{-1}(v)|_{V_1} \cdot |L_C^{-1}(w)|_{W_1} \\ &\leqslant K_1 \cdot S \cdot |v|_V \cdot T \cdot |w|_W \\ &= K \cdot |v|_V \cdot |w|_W, \end{aligned}$$

as desired. $\square$

Assigned HW#13-4.

DEFINITION 49.2. *Let S, V, W, Z be vector spaces. $* \in B(V, W, Z)$. Let $f : S \dashrightarrow V$ and let $g : S \dashrightarrow W$. Then $f \underset{S}{*} g : S \dashrightarrow Z$ is defined by $(f \underset{S}{*} g)(x) = [f(x)] * [g(x)]$.*

We will sometimes omit the "S" in "$\underset{S}{*}$", provided S is clear. Note that, in Definition 49.2, $\mathrm{dom}[f \underset{S}{*} g] = (\mathrm{dom}[f]) \bigcap (\mathrm{dom}[g])$.

THEOREM 49.3. *Let S, V, W, Z be finite dimensional vector spaces. Let $* \in B(V, W, Z)$. Then $[\widehat{\mathcal{O}}(S, V)] \underset{S}{*} [\breve{o}(S, W)] \subseteq \breve{o}(S, Z)$.*

Proof. We wish to show: $\forall \alpha \in \widehat{\mathcal{O}}(S, V)$, $\forall \beta \in \breve{o}(S, W)$, we have: $\alpha \underset{S}{*} \beta \in \breve{o}(S, Z)$. Let $\alpha \in \widehat{\mathcal{O}}(S, V)$ and $\beta \in \breve{o}(S, W)$ be given. We wish to show: $\alpha \underset{S}{*} \beta \in \breve{o}(S, Z)$.

Choose $|\bullet|_S \in \mathcal{N}(S)$ and $|\bullet|_V \in \mathcal{N}(V)$ and $|\bullet|_W \in \mathcal{N}(W)$. Choose $\|\bullet\| \in \mathcal{N}(Z)$. By Theorem 49.1, choose $K \geqslant 0$ such that, $\forall v \in V$, $\forall w \in W$, $\|v * w\| \leqslant K \cdot |v|_V \cdot |w|_W$. Let $|\bullet|_Z := [\|\bullet\|]/[K+1]$. Then: $\forall v \in V$, $\forall w \in W$, we have $|v * w|_Z \leqslant |v|_V \cdot |w|_W$.

Since $\alpha \in \widehat{\mathcal{O}}(S, V)$, choose $U \in \mathcal{N}_S(0_S)$ such that

$$U \subseteq \mathrm{dom}[\alpha] \qquad \text{and} \qquad \sup |\alpha_*(U)|_V < \infty.$$

Since $\mathrm{dom}[\alpha] \supseteq U \in \mathcal{N}_S(0_S)$, we conclude that $\mathrm{dom}[\alpha] \in \mathcal{N}_S(0_S)$.

Since $\beta \in \breve{o}(S, W)$, we know

- $\mathrm{dom}[\beta] \in \mathcal{N}_S(0_S)$,
- $\beta(0_S) = 0_W$ and
- $\beta : S \dashrightarrow W$ is continuous at 0_S.

Let $\gamma := \alpha *_S \beta$. We wish to show: $\alpha \in \breve{\mathcal{O}}(S,Z)$. We wish to show:

- $\mathrm{dom}[\gamma] \in \mathcal{N}_S(0_S)$,
- $\gamma(0_S) = 0_Z$ and
- $\gamma : S \dashrightarrow Z$ is continuous at 0_S.

Because

$$\mathrm{dom}[\alpha] \in \mathcal{N}_S(0_S) \qquad \text{and} \qquad \mathrm{dom}[\beta] \in \mathcal{N}_S(0_S) \qquad \text{and}$$
$$\mathrm{dom}[\gamma] = (\mathrm{dom}[\alpha]) \cap (\mathrm{dom}[\beta]),$$

it follows that $\mathrm{dom}[\gamma] \in \mathcal{N}_S(0_S)$. Also, we have

$$\gamma(0_S) \quad = \quad [\alpha(0_S)] * [\beta(0_S)] \quad = \quad [\alpha(0_S)] * [0_W] \quad = \quad 0_Z.$$

It remains to show: $\gamma : S \dashrightarrow Z$ is continuous at 0_S. We wish to show: $\forall \varepsilon > 0$, $\exists \delta > 0$ such that, $\forall x \in \mathrm{dom}[\gamma]$,

$$[\ |x|_S < \delta\ ] \quad \Rightarrow \quad [\ |\gamma(x)|_Z < \varepsilon\ ].$$

Let $\varepsilon > 0$ be given. We wish to show: $\exists \delta > 0$ such that, $\forall x \in \mathrm{dom}[\gamma]$,

$$[\ |x|_S < \delta\ ] \quad \Rightarrow \quad [\ |\gamma(x)|_Z < \varepsilon\ ].$$

Since $U \in \mathcal{N}_S(0_S)$ and since $\mathcal{B}_S(0_S)$ is a neighborhood base at 0_S in S, choose $\rho > 0$ such that $B_S(0_S, \rho) \subseteq U$. Let $M := (\sup |\alpha_*(U)|_V) + 1$. Then $M > 0$ and $|\alpha_*(U)|_V < M$. Since $\beta : S \dashrightarrow W$ is continuous at 0_S and since $\beta(0_S) = 0_W$, choose $\tau > 0$ such that, for all $x \in \mathrm{dom}[\beta]$,

$$[\ |x|_S < \tau\ ] \quad \Rightarrow \quad [\ |\beta(x)|_W < \varepsilon / M\ ].$$

Let $\delta := \min\{\rho, \tau\}$. We wish to show: $\forall x \in \mathrm{dom}[\gamma]$,

$$[\ |x|_S < \delta\ ] \quad \Rightarrow \quad [\ |\gamma(x)|_Z < \varepsilon\ ].$$

Let $x \in \mathrm{dom}[\gamma]$ be given. We wish to show:

$$[\ |x|_S < \delta\ ] \quad \Rightarrow \quad [\ |\gamma(x)|_Z < \varepsilon\ ].$$

Assume: $|x|_S < \delta$. We wish to show: $|\gamma(x)|_Z < \varepsilon$.

Recall: $\mathrm{dom}[\gamma] = (\mathrm{dom}[\alpha]) \cap (\mathrm{dom}[\beta])$. So, since $x \in \mathrm{dom}[\gamma]$, we get $x \in \mathrm{dom}[\alpha]$ and $x \in \mathrm{dom}[\beta]$.

Since $|x|_S < \delta \leqslant \rho$, we get $x \in B_S(0_S, \rho)$. So, since $B_S(0_S, \rho) \subseteq U$, we get $x \in U$. So, since $x \in \mathrm{dom}[\alpha]$, we get $\alpha(x) \in \alpha_*(U)$. Then $|\alpha(x)|_V \in |\alpha_*(U)|_V < M$. Since $x \in \mathrm{dom}[\beta]$ and $|x|_S < \delta \leqslant \tau$, by choice of τ, we see that $|\beta(x)|_W < \varepsilon/M$. Let $v := \alpha(x)$ and let $w := \beta(x)$. Then $|v * w|_Z \leqslant |v|_V \cdot |w|_W$. Then

$$\begin{aligned} |\gamma(x)|_Z &= |(\alpha *_S \beta)(x)|_Z = |[\alpha(x)] * [\beta(x)]|_Z \\ &= |v * w|_Z \leqslant |v|_V \cdot |w|_W \\ &= |\alpha(x)|_V \cdot |\beta(x)|_W \\ &< M \cdot [\varepsilon/M] = \varepsilon, \end{aligned}$$

as desired. $\square$

COROLLARY 49.4. *Let S, V, W, Z be finite dimensional vector spaces. Let $\circledast \in B(W, V, Z)$. Then $[\widehat{\mathcal{O}}(S, W)] \underset{S}{\circledast} [\breve{o}(S, V)] \subseteq \breve{o}(S, Z)$.*

Proof. Let $V_1 := W$ and $W_1 := V$. By Theorem 49.3 (with V replaced by V_1, W by W_1 and $*$ by $\circledast$), we have

$$[\widehat{\mathcal{O}}(S, V_1)] \underset{S}{\circledast} [\breve{o}(S, W_1)] \subseteq \breve{o}(S, Z).$$

Then $[\widehat{\mathcal{O}}(S, W)] \underset{S}{\circledast} [\breve{o}(S, V)] = [\widehat{\mathcal{O}}(S, V_1)] \underset{S}{\circledast} [\breve{o}(S, W_1)] \subseteq \breve{o}(S, Z)$. $\square$

COROLLARY 49.5. *Let S, V, W, Z be finite dimensional vector spaces. Let $* \in B(V, W, Z)$. Then $[\breve{o}(S, V)] \underset{S}{*} [\widehat{\mathcal{O}}(S, W)] \subseteq \breve{o}(S, Z)$.*

Proof. Define $\circledast \in B(W, V, Z)$ by $w \circledast v = v * w$. Then

$$[\breve{o}(S, V)] \underset{S}{*} [\widehat{\mathcal{O}}(S, W)] = [\widehat{\mathcal{O}}(S, W)] \underset{S}{\circledast} [\breve{o}(S, V)].$$

By Corollary 49.4, we have $[\widehat{\mathcal{O}}(S, W)] \underset{S}{\circledast} [\breve{o}(S, V)] \subseteq \breve{o}(S, Z)$. Then $[\breve{o}(S, V)] \underset{S}{*} [\widehat{\mathcal{O}}(S, W)] = [\widehat{\mathcal{O}}(S, W)] \underset{S}{\circledast} [\breve{o}(S, V)] \subseteq \breve{o}(S, Z)$. $\square$

COROLLARY 49.6. *Let S, V, W, Z be finite dimensional vector spaces. Let $* \in B(V, W, Z)$. Then $[\breve{o}(S, V)] \underset{S}{*} [\breve{o}(S, W)] \subseteq \breve{o}(S, Z)$.*

Proof. By HW#12-4, $\breve{o}(S, W) \subseteq \widehat{\mathcal{O}}(S, W)$. By Corollary 49.5,

$$[\breve{o}(S, V)] \underset{S}{*} [\widehat{\mathcal{O}}(S, W)] \subseteq \breve{o}(S, Z).$$

Then $[\breve{o}(S, V)] *_S [\breve{o}(S, W)] \subseteq [\breve{o}(S, V)] *_S [\widehat{\mathcal{O}}(S, W)] \subseteq \breve{o}(S, Z)$. $\square$

Assigned HW#13-5.

THEOREM 49.7. *Let S, V, W, Z be finite dimensional vector spaces. Let $* \in B(V, W, Z)$. Let $p, q \geqslant 0$. Then*

(1) $[\widehat{\mathcal{O}}_p(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}_q(S,W)] \subseteq \widehat{\mathcal{O}}_{p+q}(S,Z)$,
(2) $[\widehat{\mathcal{O}}_p(S,V)] \underset{S}{*} [\breve{o}_q(S,W)] \subseteq \breve{o}_{p+q}(S,Z)$,
(3) $[\breve{o}_p(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}_q(S,W)] \subseteq \breve{o}_{p+q}(S,Z)$ *and*
(4) $[\breve{o}_p(S,V)] \underset{S}{*} [\breve{o}_q(S,W)] \subseteq \breve{o}_{p+q}(S,Z)$.

Proof. Choose $|\bullet| \in \mathcal{N}(S)$.

Proof of (1): By HW#13-5, $[\widehat{\mathcal{O}}(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}(S,W)] \subseteq \widehat{\mathcal{O}}(S,Z)$. Then

$$\begin{aligned}
&[\widehat{\mathcal{O}}_p(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}_q(S,W)] \\
&\quad= [(|\bullet|^p)\cdot(\widehat{\mathcal{O}}(S,V))] \underset{S}{*} [(|\bullet|^q)\cdot(\widehat{\mathcal{O}}(S,W))] \\
&\quad= (|\bullet|^p)\cdot(|\bullet|^q)\cdot(\,[\widehat{\mathcal{O}}(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}(S,W)]\,) \\
&\quad\subseteq (|\bullet|^{p+q})\cdot(\,\widehat{\mathcal{O}}(S,Z)\,) = \widehat{\mathcal{O}}_{p+q}(S,Z),
\end{aligned}$$

as desired. *End of proof of (1).*

Proof of (2): By Theorem 49.3, $[\widehat{\mathcal{O}}(S,V)] \underset{S}{*} [\breve{o}(S,W)] \subseteq \breve{o}(S,Z)$. Then

$$\begin{aligned}
&[\widehat{\mathcal{O}}_p(S,V)] \underset{S}{*} [\breve{o}_q(S,W)] \\
&\quad= [(|\bullet|^p)\cdot(\widehat{\mathcal{O}}(S,V))] \underset{S}{*} [(|\bullet|^q)\cdot(\breve{o}(S,W))] \\
&\quad= (|\bullet|^p)\cdot(|\bullet|^q)\cdot(\,[\widehat{\mathcal{O}}(S,V)] \underset{S}{*} [\breve{o}(S,W)]\,) \\
&\quad\subseteq (|\bullet|^{p+q})\cdot(\,\breve{o}(S,Z)\,) = \breve{o}_{p+q}(S,Z),
\end{aligned}$$

as desired. *End of proof of (2).*

Proof of (3): By Corollary 49.5, $[\breve{o}(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}(S,W)] \subseteq \breve{o}(S,Z)$. Then

$$\begin{aligned}
&[\breve{o}_p(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}_q(S,W)] \\
&\quad= [(|\bullet|^p)\cdot(\breve{o}(S,V))] \underset{S}{*} [(|\bullet|^q)\cdot(\widehat{\mathcal{O}}(S,W))] \\
&\quad= (|\bullet|^p)\cdot(|\bullet|^q)\cdot(\,[\breve{o}(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}(S,W)]\,) \\
&\quad\subseteq (|\bullet|^{p+q})\cdot(\,\breve{o}(S,Z)\,) = \breve{o}_{p+q}(S,Z),
\end{aligned}$$

as desired. *End of proof of (3).*

Proof of (4): By Corollary 49.6, $[\breve{\mathcal{O}}(S,V)] \underset{S}{*} [\breve{\mathcal{O}}(S,W)] \subseteq \breve{\mathcal{O}}(S,Z)$. Then

$$\begin{aligned}
[\breve{\mathcal{O}}_p(S,V)] & \underset{S}{*} [\breve{\mathcal{O}}_q(S,W)] \\
&= [(|\bullet|^p) \cdot (\breve{\mathcal{O}}(S,V))] \underset{S}{*} [(|\bullet|^q) \cdot (\breve{\mathcal{O}}(S,W))] \\
&= (|\bullet|^p) \cdot (|\bullet|^q) \cdot (\,[\breve{\mathcal{O}}(S,V)] \underset{S}{*} [\breve{\mathcal{O}}(S,W)]\,) \\
&\subseteq (|\bullet|^{p+q}) \cdot (\,\breve{\mathcal{O}}(S,Z)\,) \;=\; \breve{\mathcal{O}}_{p+q}(S,Z),
\end{aligned}$$

as desired. *End of proof of (4).* □

We also recall the main results from Theorem 47.15:

THEOREM 49.8. *Let V,W,X be finite dimensional vector spaces. Let $p,q>0$. Then*

(5) $[\widehat{\mathcal{O}}_q(W,X)] \circ [\widehat{\mathcal{O}}_p(V,W)] \subseteq \widehat{\mathcal{O}}_{qp}(V,X)$,
(6) $[\widehat{\mathcal{O}}_q(W,X)] \circ [\breve{\mathcal{O}}_p(V,W)] \subseteq \breve{\mathcal{O}}_{qp}(V,X)$,
(7) $[\breve{\mathcal{O}}_q(W,X)] \circ [\widehat{\mathcal{O}}_p(V,W)] \subseteq \breve{\mathcal{O}}_{qp}(V,X)$ *and*
(8) $[\breve{\mathcal{O}}_q(W,X)] \circ [\breve{\mathcal{O}}_p(V,W)] \subseteq \breve{\mathcal{O}}_{qp}(V,X)$.

Proof. Pf of (5): Follows from (5) of Theorem 47.15. *End of pf of (5).*
Pf of (6): Follows from (6) of Theorem 47.15. *End of pf of (6).*
Pf of (7): Follows from (7) of Theorem 47.15. *End of pf of (7).*
Pf of (8): Follows from (8) of Theorem 47.15. *End of pf of (8).* □

HW#13-1 asserts that every linear function has linear order. We next prove (in Theorem 49.9, below) that a function that is both linear and sublinear must be identically zero.

Let S be a set and let W be a vector space. We define $\mathbf{0}_{SW} := C_S^{0_W}$. Then $\mathbf{0}_{SW} : S \to W$. For all $x \in S$, we have $\mathbf{0}_{SW}(x) = 0_W$. Recall that W^S has a standard vector space structure. Then $\mathbf{0}_{SW} = 0_{W^S}$. That is, $\mathbf{0}_{SW}$ is the zero element of W^S.

Let V and W be vector spaces. Then $\mathbf{0}_{VW} = 0_{L(V,W)}$.

Let V and W be finite dimensional vector spaces. Then, for all $p \geqslant 0$, we have both $\mathbf{0}_{VW} = 0_{\widehat{\mathcal{O}}_p(V,W)}$ and $\mathbf{0}_{VW} = 0_{\breve{\mathcal{O}}_p(V,W)}$.

THEOREM 49.9. *Let V and W be finite dimensional vector spaces. Then $[L(V,W)] \cap [\breve{\mathcal{O}}_1(V,W)] = \{\mathbf{0}_{VW}\}$.*

Proof. Since $[L(V,W)] \cap [\breve{\mathcal{O}}_1(V,W)] \supseteq \{\mathbf{0}_{VW}\}$, it suffices to prove: $[L(V,W)] \cap [\breve{\mathcal{O}}_1(V,W)] \subseteq \{\mathbf{0}_{VW}\}$. We wish to show:

$$\forall T \in [L(V,W)] \cap [\breve{\mathcal{O}}_1(V,W)], \qquad T \in \{\mathbf{0}_{VW}\}.$$

Given $\forall T \in [L(V,W)] \cap [\breve{\mathcal{O}}_1(V,W)]$. Want: $T \in \{\mathbf{0}_{VW}\}$.

We have: both $T \in L(V,W)$ and $T \in \breve{\mathcal{O}}_1(V,W)$. We wish to show: $T = \mathbf{0}_{VW}$. Assume that $T \neq \mathbf{0}_{VW}$. We aim for a contradiction.

Since $T \neq \mathbf{0}_{VW}$, choose $x \in V$ s.t. $Tx \neq 0_W$. Let $y := Tx$. Then $y \neq 0_W$. Since $Tx \neq 0_W$, it follows that $x \neq 0_V$.

Choose $|\bullet|_V \in \mathcal{N}(V)$ and $|\bullet|_W \in \mathcal{N}(W)$. Since $x \neq 0_V$, $|x|_V > 0$. Since $y \neq 0_W$, $|y|_W > 0$. Let $\varepsilon := [|y|_W]/[|x|_V]$. Then $\varepsilon > 0$.

Since $T \in \breve{\mathcal{O}}_1(V,W) = [|\bullet|_V] \cdot [\breve{\mathcal{O}}(V,W]$, choose $\alpha \in \breve{\mathcal{O}}(V,W)$ such that $T = |\bullet|_V \cdot \alpha$. We have

$$\begin{aligned} \mathrm{dom}[T] &= (\mathrm{dom}[|\bullet|_V]) \cap (\mathrm{dom}[\alpha]) \\ &= V \cap (\mathrm{dom}[\alpha]) = \mathrm{dom}[\alpha]. \end{aligned}$$

Since $T \in L(V,W)$, we get $\mathrm{dom}[T] = V$.

For all $s \in \mathbb{R}$, we have both

$$\begin{aligned} T(sx) &= s \cdot [Tx] = sy \qquad \text{and} \\ T(sx) &= (|\bullet|_V \cdot \alpha)(sx) \\ &= |sx|_V \cdot [\alpha(sx)] \\ &= |s| \cdot |x|_V \cdot [\alpha(sx)], \end{aligned}$$

so

$$\begin{aligned} |T(sx)|_W &= |sy|_W = |s| \cdot |y|_W \qquad \text{and} \\ |T(sx)|_W &= |s| \cdot |x|_V \cdot |\alpha(sx)|_W, \end{aligned}$$

so $|s| \cdot |y|_W = |s| \cdot |x|_V \cdot |\alpha(sx)|_W$. Then, $\forall s \in \mathbb{R}_0^\times$, since $|s| \neq 0$, we get $|y|_W = |x|_V \cdot |\alpha(sx)|_W$, so $|\alpha(sx)|_W = \varepsilon$.

Because $\alpha \in \breve{\mathcal{O}}(V,W)$, it follows both that $\alpha(0_V) = 0_W$ and that $\alpha : V \dashrightarrow W$ is continuous at 0_V. Then choose $\delta > 0$ s.t., $\forall v \in \mathrm{dom}[\alpha]$,

$$[\ |v|_V < \delta\ ] \quad \Rightarrow \quad [\ |\alpha(v)|_W < \varepsilon\ ].$$

Let $s := \delta/(2 \cdot |x|_V)$. Since $\delta > 0$ and $|x|_V > 0$, we get $s > 0$. Then $s \neq 0$, so $|\alpha(sx)|_W = \varepsilon$. Let $v := sx$. Then $v \in V = \mathrm{dom}[T] = \mathrm{dom}[\alpha]$ and $|v|_V = s \cdot |x|_V = \delta/2 < \delta$, so, by choice of δ, we get $|\alpha(v)|_W < \varepsilon$. Then $\varepsilon = |\alpha(sx)|_W = |\alpha(v)|_W < \varepsilon$, so $\varepsilon < \varepsilon$. Contradiction. □

50. Class 24 on 19 April 2018, Th of Week 13

DEFINITION 50.1. *Let V and W be vector spaces. Then we define*

(1) $B(V,W) := B(V,V,W)$,

(2) $T(V,W) := T(V,V,V,W)$ *and*
(3) $\forall d \in \mathbb{N}$, $M^d(V,W) := M^d(V,\ldots,V,W)$.

DEFINITION 50.2. *Let V and W be vector spaces, let $d \in \mathbb{N}$ and let $F \in M^d(V,W)$. Then we define $\Delta_F : V \to W$ by $\Delta_F(u) = F(u,\ldots,u)$.*

In Definition 50.2, the mapping $\Delta_F : V \to W$ is called the **diagonal restriction** of F.

We work three examples: First, let

$$A \quad := \quad \begin{bmatrix} 1 & 4 \\ 0 & 3 \end{bmatrix}.$$

Let $G := B_A$. Then $G \in B(\mathbb{R}^2,\mathbb{R}^2,\mathbb{R}) = B(\mathbb{R}^2,\mathbb{R})$, and so we have $G : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$. For all $p,q,x,y \in \mathbb{R}$, we have

$$G(\,(p,q)\,,\,(x,y)\,) \quad = \quad px \,+\, 4py \,+\, 3qy.$$

Let $Q := \Delta_G$. Then $Q : \mathbb{R}^2 \to \mathbb{R}$. For all $x,y \in \mathbb{R}$, we have

$$Q(x,y) \quad = \quad G(\,(x,y)\,,\,(x,y)\,) \quad = \quad x^2 \,+\, 4xy \,+\, 3y^2.$$

Note that $Q : \mathbb{R}^2 \to \mathbb{R}$ is a homogeneous quadratic polynomial. Second, the transpose of A is

$$A^t \quad := \quad \begin{bmatrix} 1 & 0 \\ 4 & 3 \end{bmatrix}.$$

For all $j,k \in [1..2]$, we have $A^t_{jk} = A_{kj}$. Let $G_1 := B_{A^t}$. We leave it as an exercise to verify, for all $p,q,x,y \in \mathbb{R}$, we have

$$G_1(\,(p,q)\,,\,(x,y)\,) \quad = \quad G(\,(x,y)\,,\,(p,q)\,).$$

Thus, for all $u,v \in \mathbb{R}^2$, we have $G_1(u,v) = G(v,u)$. Then, for all $u \in \mathbb{R}^2$, we have $\Delta_{G_1}(u) = G_1(u,u) = G(u,u) = \Delta_G(u)$. Then $\Delta_{G_1} = \Delta_G = Q$. Third, the symmetrization of A is $C := (A + A^t)/2$. We compute

$$C \quad := \quad \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

For all $j,k \in [1..2]$, $C_{jk} = (A_{jk} + A_{kj})/2$. Then, for all $j,k \in [1..2]$, we get $C_{kj} = C_{jk}$. This is expressed by saying that C is a *symmetric* matrix. Let $H := B_C$. Then $H = (B_A + B_{A^t})/2 = (G + G_1)/2$. Then, for all $u,v \in \mathbb{R}^2$, we have

$$H(u,v) \quad = \quad \frac{[G(u,v)] + [G(v,u)]}{2}.$$

Then, for all $u, v \in \mathbb{R}^2$, we bet $H(v, u) = H(u, v)$. This is expressed by saying that H is a *symmetric* bilinear function. We have $\Delta_H = (\Delta_G + \Delta_{G_1})/2 = (Q + Q)/2 = Q$. Thus the three matrices A, A^t and C give rise to three different bilinear functions G, G_1 and H, but those bilinear functions all give rise to the same quadratic polynomial Q.

Generally, for any quadratic polynomial $Q : \mathbb{R}^2 \to \mathbb{R}$, there are many bilinear functions F such that $Q = B_F$, but exactly one of them is symmetric. This generalizes to higher dimensions, and also, to cubics, quartics and higher degree. We explain those generalizations next. We start by defining what it means for a function $P : V \to W$ to be a (homogeneous) polynomial.

DEFINITION 50.3. *Let V and W be vector spaces. Then we define*

(0) $C(V, W) := \{C_V^x \,|\, x \in W\}$,
(0') $P^0(V, W) := C(V, W)$,
(1) $P^1(V, W) := L(V, W)$,
(2) $Q(V, W) := \{\Delta_F \,|\, F \in B(V, W)\}$,
(3) $K(V, W) := \{\Delta_F \,|\, F \in T(V, W)\}$,
(∗) $\forall d \in [2..\infty)$, $P^d(V, W) := \{\Delta_F \,|\, F \in M^d(V, W)\}$,

Let V and W be vector spaces. Then

(0) $P^0(V, W) = C(V, W)$,
(1) $P^1(V, W) = L(V, W)$,
(2) $P^2(V, W) = Q(V, W)$,
(3) $P^3(V, W) = K(V, W)$.

Also, for all $F : V \to W$,

(0) F is constant means $F \in C(V, W)$,
(1) F is (homogeneous) linear means $F \in L(V, W)$,
(2) F is (homogeneous) quadratic means $F \in Q(V, W)$,
(3) F is (homogeneous) cubic means $F \in K(V, W)$,
(∗) $\forall d \in \mathbb{N}_0$, F is (homogeneous) of degree d means $F \in P^d(V, W)$.

Let V and W be vector spaces, $d \in [2..\infty)$, $F \in P^d(V, W)$. Then there are many $F \in M^d(V, W)$ such that $\Delta_F = P$, but only one such F is *symmetric*, a term that we define next.

DEFINITION 50.4. *Let V and W be vector spaces. Then*

(1) $SM^1(V, W) := L(V, W)$,
(2) $SB(V, W) := \{F \in B(V, W) \,|\, \forall x, y \in V,\ F(x, y) = F(y, x)\}$,

(3) $ST(V,W) := \{F \in B(V,W) \mid \forall x,y,z \in V,$
$$F(x,y,z) = F(y,z,x) = F(z,x,y) = F(y,x,z) = F(x,z,y) = F(z,y,x) \quad \},$$

$(*)$ $\forall d \in [2..\infty)$, $SM^d(V,W) :=$
$$\{F \in M^d(V,W) \mid \forall x_1,\ldots,x_d \in V,\ \forall \sigma : [1..d] \hookrightarrow\!\!> [1..d],\ F(x_1,\ldots,x_d) = F(x_{\sigma(1)},\ldots,x_{\sigma(d)}) \quad \}.$$

Let V and W be vector spaces. Then $SM^2(V,W) = SB(V,W)$ and $SM^3(V,W) = ST(V,W)$. Also, we define **symmetric multilinar function** as follows:

(2) $\forall F \in B(V,W)$, F is **symmetric** means $F \in SB(V,W)$,
(3) $\forall F \in T(V,W)$, F is **symmetric** means $F \in ST(V,W)$,
$(*)$ $\forall d \in [2..\infty)$, $\forall F \in M^d(V,W)$,
F is **symmetric** means $F \in SM^d(V,W)$,

Let V and W be vector spaces and let $d \in [2..\infty)$. The mapping $F \mapsto \Delta_F : M^d(V,W) \to P^d(V,W)$ is linear and surjective. It is a theorem that the restriction of that mapping to $SM^d(V,W)$ is a vector space isomorphism. That is, $F \mapsto \Delta_F : SM^d(V,W) \to P^d(V,W)$ is a vector space isomorphism. Let $\Pi^d_{VW} : P^d(V,W) \to SM^d(V,W)$ be the inverse isomorphism. Then, for all $P \in P^d(V,W)$, by the **polarization** of P we mean $\Pi^d_{VW}(P)$. Therefore, the polarization of

a homogeneous polynomial of degree d from V to W

is a symmetric d-multilinear function on V^d taking values in W. Polarization allows us to convert

questions about homogeneous polynomials

into questions about symmetric multilinear algebra.

Let V be a vector space and $d \in [2..\infty)$. Then $\exists$vector space Z (called the dth **symmetric power** of V) such that, $\forall$vector space W, $\exists$an isomorphism $\Phi^d_{VW} : SM^d(V,W) \hookrightarrow\!\!> L(Z,W)$. Since SM is on one side and L is on the other, this isomorphism allows us to convert

questions about symmetric multilinear algebra

into questions about linear algebra.

THE BIG IDEA: Let V and W be vector spaces, let $f : V \dashrightarrow W$, let $p \in V$ and let $d \in \mathbb{N}$. We attempt to approximate f near p by a polynomial, as follows. Recall (Definition 25.21): for all $h \in V$, $f^T_p(h) = [f(p+h)] - [f(p)]$. Approximating f near p is equivalent to approximating f_p^T near 0_V. Assume that we are able to find homogeneous polynomials $P_1 \in P^1(V,W), \ldots, P_d \in P^d(V,W)$ such that $f_p^T - P_1 - \cdots - P_d \in \breve{\mathcal{o}}_d(V,W)$. As $P_1 \in P^1(V,W) = L(V,W)$, use linear algebra to study P_1. For each $j \in [2..d]$, let $F_j := \Pi_{VW}^j(P_j)$. For each $j \in [2..d]$, let Z_j be the jth symmetric power of V and let $L_j := \Phi_{VW}^j(F_j)$. As $L_j \in L(Z_j,W)$, use linear algebra to study L_j. Finally, the remainder $f_p^T - P_1 - \cdots - P_d$ is sub-d. Assuming d is very large, we try to show that, in some sense (to be determined later), any sub-d function is small, and can be neglected. The upshot of this discussion: LINEAR ALGEBRA IS THE STUDY OF EVERYTHING!

We now begin to implement this big idea, step by step.

REMARK 50.5. *Let V and W be finite dimensional vector spaces and let $\alpha \in \widehat{\mathcal{O}}(V,W)$. Then dom$[\alpha] \in \mathcal{N}_V(0_V)$.*

Proof. Since $\alpha \in \widehat{\mathcal{O}}(V,W)$, choose $U \in \mathcal{N}_V(0_V)$ such that $U \subseteq \mathrm{dom}[\alpha]$. Since $\mathrm{dom}[\alpha] \supseteq U \in \mathcal{N}_V(0_V)$, it follows that $\mathrm{dom}[\alpha] \in \mathcal{N}_V(0_V)$. $\square$

Let V and W be finite dimensional vector spaces and let $p \geqslant 0$. Then $\breve{\mathcal{o}}_p(V,W) \subseteq \widehat{\mathcal{O}}_p(V,W) \subseteq \widehat{\mathcal{O}}(V,W)$. So, by Remark 50.5, we see

- $\forall \alpha \in \breve{\mathcal{o}}_p(V,W), \quad \mathrm{dom}[\alpha] \in \mathcal{N}_V(0_V)$ and
- $\forall \alpha \in \widehat{\mathcal{O}}_p(V,W), \quad \mathrm{dom}[\alpha] \in \mathcal{N}_V(0_V)$.

REMARK 50.6. *Let V and W be vector spaces, let $f : V \dashrightarrow W$ and let $p \in V$. Assume that dom$[f] \notin \mathcal{N}_V(p)$. Then dom$[f_p^T] \notin \mathcal{N}_V(0_V)$.*

Proof. Unassigned HW. $\square$

DEFINITION 50.7. *Let V and W be finite dimensional vector spaces, let $f : V \dashrightarrow W$ and let $p \in V$. Then we define*

$$\mathrm{LINS}_p^{V,W} f \quad := \quad \{\, L \in L(V,W) \mid f_p^T \, - \, L \ \in \ \breve{\mathcal{o}}_1(V,W) \,\}.$$

In Definition 50.7, when V and W are clear, we will omit them from the notation, and write $\mathrm{LINS}_p f$ instead of $\mathrm{LINS}_p^{V,W} f$. The functions L in $\mathrm{LINS}_p f$ are called **linearizations** of f at p. They are those linear functions that approximate f_p^T at 0_V, in the sense that the remainder, $f_p^T - L$, is sublinear. We'll show in a moment (Remark 50.12) that, if a linearization exists, then it's unique.

There's a strong connection between differentiation and linearization. For example, recall that the absolute value function $|\bullet| : \mathbb{R} \to \mathbb{R}$ is *NOT* differentiable at 0. Unassigned HW: Let $|\bullet| : \mathbb{R} \to \mathbb{R}$ denote the absolute value function. Show that $\mathrm{LINS}_0|\bullet| = \varnothing$, *i.e.* show that $|\bullet|$ has *NO* linearization at 0.

For functions $f : \mathbb{R} \to \mathbb{R}$, if L is a linearization of f at a point $p \in \mathbb{R}$, then there's a tangent line T to the graph of f at $(p, f(p))$, and the graph of L is the line passing through the origin $(0,0)$ that is parallel to T. We'll eventually see that this always works, provided the domain of f is a neighborhood of p. However, if the domain is not a neighborhood, then we may be able to find a tangent line, even though no linearization exists:

Let $X_0 := \{1, 1/2, 1/3, \ldots\} \cup \{0\}$, let $f_0 := (\mathrm{id}_{\mathbb{R}})|X_0 : X_0 \to \mathbb{R}$ and let $p_0 := 0$. The graph of f_0 consists of the origin $(0,0)$, together with countably many points approaching the origin. Then we have $f_0'(p_0) = 1$, and, also, $\mathrm{dom}[f_0] = X_0 \notin \mathcal{N}_{\mathbb{R}}(p_0)$. So the line $y = x$ is, in some sense, a tangent line to f_0 at $(0,0)$, even though the domain of f_0 is not a neighborhood of p_0. Because $\mathrm{dom}[f_0] \notin \mathcal{N}_{\mathbb{R}}(p_0)$, by the next remark (Remark 50.9), we see that $\mathrm{LINS}_0 f_0 = \varnothing$, *i.e.*, there is *NO* linearization to f_0 at p_0.

REMARK 50.8. *Let V and W be finite dimensional vector spaces, $f : V \dashrightarrow W$ and $p \in V$. Then:*

$$[\ dom[f] \in \mathcal{N}_V(p)\ ] \qquad \Leftrightarrow \qquad [\ dom[f_p^T] \in \mathcal{N}_V(0_V)\ ].$$

Proof. Unassigned HW. □

REMARK 50.9. *Let V and W be finite dimensional vector spaces, $f : V \dashrightarrow W$ and $p \in V$. Assume* $\mathrm{LINS}_p f \neq \varnothing$. *Then $dom[f] \in \mathcal{N}_V(p)$.*

Proof. Choose $L \in \mathrm{LINS}_p f$. Then $L \in L(V,W)$ and $f_p^T - L \in \breve{\mathcal{O}}_1(V,W)$. Since $L \in L(V,W)$, we have $\mathrm{dom}[L] = V$. Let $R := f_p^T - L$. Then $\mathrm{dom}[R] = (\mathrm{dom}[f_p^T) \cap (\mathrm{dom}[L])$. So, since $\mathrm{dom}[f_p^T] \subseteq V = \mathrm{dom}[L]$, we conclude that $\mathrm{dom}[R] = \mathrm{dom}[f_p^T]$.

We have $R = f_p^T - L \in \breve{\mathcal{O}}_1(V,W)$, so, by definition of $\breve{\mathcal{O}}_1$, we see that $\mathrm{dom}[R] \in \mathcal{N}_V(0_V)$. Then $\mathrm{dom}[f_p^T] = \mathrm{dom}[R] \in \mathcal{N}_V(0_V)$. Then, by $\Leftarrow$ of Remark 50.8, $\mathrm{dom}[f] \in \mathcal{N}_V(p)$. □

Let X be a topological space, let Y be a set, let $\alpha, \beta : X \dashrightarrow Y$ and let $S \subseteq X$. Recall that $\alpha = \beta$ **on** S means: $\forall z \in S$, $\alpha(z) = \beta(z)$.

DEFINITION 50.10. *Let X be a topological space, let Y be a set, let $\alpha, \beta : X \dashrightarrow Y$ and let $p \in X$. Then $\alpha = \beta$* **near** *p means: there exists $U \in \mathcal{N}_X(p)$ such that $\alpha = \beta$ on U.*

FACT 50.11. *Let V and W be finite dimensional vector spaces and let $\alpha, \beta : V \dashrightarrow W$. Assume that $\alpha = \beta$ near 0_V, and that $\alpha \in \breve{\mathcal{o}}_1(V,W)$. Then we have: $\beta \in \breve{\mathcal{o}}_1(V,W)$.*

Proof. Unassigned HW. I may assign this next week. $\square$

Let $p \geqslant 0$. Unassigned HW: Show that Fact 50.11 continues to be true if $\breve{\mathcal{o}}_1$ is replaced by $\breve{\mathcal{o}}_p$. Unassigned HW: Show that Fact 50.11 continues to be true if $\breve{\mathcal{o}}_1$ is replaced by $\widehat{\mathcal{O}}_p$.

REMARK 50.12. *Let V and W be finite dimensional vector spaces, let $f : V \dashrightarrow W$ and let $p \in V$. Then $\#(\mathrm{LINS}_p\, f) \leqslant 1$.*

Proof. We wish to show:

$$\forall L, M \in \mathrm{LINS}_p f, \qquad L = M.$$

Let $L, M \in \mathrm{LINS}_p f$ be given. We wish to show: $L = M$.

By Lemma 50.9, we have $\mathrm{dom}[f] \in \mathcal{N}_V(p)$. So, by $\Rightarrow$ of Remark 50.8, we have $\mathrm{dom}[f_T^p] \in \mathcal{N}_V(0_V)$. Let $U := \mathrm{dom}[f_p^T]$. Then $U \in \mathcal{N}_V(0_V)$.

Let $R := f_p^T - L$ and $S := f_p^T - M$. Then, as $L, M \in \mathrm{LINS}_p f$, we see that $R, S \in \breve{\mathcal{o}}_1(V,W)$. So, since $\breve{\mathcal{o}}_1(V,W)$ is a vector space of functions, we get $S - R \in \breve{\mathcal{o}}_1(V,W)$. By assumption, $L, M \in L(V,W)$, So, since $L(V,W)$ is a vector space of functions, we get $L - M \in L(V,W)$. We have $S - R = L - M$ on U, so, since $U \in \mathcal{N}_V(0_V)$, we get $S - R = L - M$ near 0_V. So, since $S - R \in \breve{\mathcal{o}}_1(V,W)$, by Lemma 50.11, we see that $L - M \in \breve{\mathcal{o}}_1(V,W)$. By Theorem 49.9, $[L(V,W)] \cap [\breve{\mathcal{o}}_1(V,W)] = \{\mathbf{0}_{VW}\}$. So, since $L - M \in L(V,W)$ and since $L - M \in \breve{\mathcal{o}}_1(V,W)$, we conclude that $L - M = \mathbf{0}_{VW}$. Then $L = M$. $\square$

Let $A := [6] \in \mathbb{R}^{1\times 1}$. Then A is a 1×1 matrix, and the only entry of A is 6. That entry of A is A_{11}, and so we have: $A_{11} = 6$. To get at the unique entry of a 1×1 matrix, we make the following definition:

DEFINITION 50.13. *For any set S, for any $A \in S^{1\times 1}$, we define* $\mathrm{ENT}\, S := S_{11}$

Let $A := [6] \in \mathbb{R}^{1\times 1}$ and let

$$B := \begin{bmatrix} 7 & 5 & 3 \\ 2 & 4 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3}.$$

Then $\mathrm{ENT}\, A = A_{11} = 6$, while $\mathrm{ENT}\, B = ☹$ and $B_{11} = 7$.

DEFINITION 50.14. *Let V and W be finite dimensional vector spaces, let $f : D \dashrightarrow W$ and let $p \in V$. Then $D_p f := \mathrm{ELT}(\mathrm{LINS}_p f)$.*

In Definition 50.14, we have $D_pf \in L(V,W)$, and D_pf is called the **total derivative** of f at p.

Let V and W be finite dimensional vector spaces, let $f : D \dashrightarrow W$ and let $p \in V$. By Remark 50.12, either f has no linearization at p, or it has exactly one. If it has none, then $D_pf = ☹$. If there is exactly one linearization L of f at p, then:

$$[\,D_pf = L\,] \quad \text{and} \quad [\,L \in L(V,W)\,] \quad \text{and} \quad [\,f_p^T - L \in \mathcal{O}_1(V,W)\,].$$

Let $m \in \mathbb{N}$, let $p \in \mathbb{R}^m$, let W be a finite dimensional vector space and let $f : \mathbb{R}^m \dashrightarrow W$. Then the partial derivatives of f at p are the vectors: $(\partial_1 f)(p), \ldots, (\partial_m f)(p) \in W$. The matrix of the total derivative is: $[D_pf] \in \mathbb{R}^{m\times n}$. The rows of this matrix are the vectors: $[D_pf]_{\bullet 1}, \ldots, [D_pf]_{\bullet m} \in W$. We will see later that, for all $j \in [1..m]$, we have: $[D_pf]_{\bullet j} = (\partial_j f)(p)$. In this sense, the total derivative is assembled from the partial derivatives.

REMARK 50.15. *Let V and W be finite dimensional vector spaces, $f : D \dashrightarrow W$ and $p \in V$. Assume $D_pf \neq ☹$. Then $dom[f] \in \mathcal{N}_V(p)$.*

Proof. Since $\mathrm{ELT}(\mathrm{LINS}_pf) = D_pf \neq ☹$, we see that $\mathrm{LINS}_pf \neq \varnothing$. Then, by Remark 50.9, $\mathrm{dom}[f] \in \mathcal{N}_V(p)$. $\square$

Let $X_0 := \{1, 1/2, 1/3, \ldots\} \cup \{0\}$, let $f_0 := (\mathrm{id}_{\mathbb{R}})|X_0 : X_0 \to \mathbb{R}$ and let $p_0 := 0$. Then $\mathrm{dom}[f_0] \notin \mathcal{N}_V(p_0)$, so, by Remark 50.15, $D_{p_0}f_0 = ☹$. On the other hand, we have $f_0'(p_0) = 1$. So, in this case, there's no clear connection between $D_{p_0}f_0$ and $f_0'(p_0)$. We will eventually show that this disconnect is exactly caused by the fact that $\mathrm{dom}[f_0] \notin \mathcal{N}_V(p_0)$.

Let's next look at an example where the domain is a neighborhood of the point: Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Let $p := 3$. Then $f'(p) = 6$. We wish to compute D_pf, and to compare it to $f'(p)$. We begin by computing f_p^T: For all $h \in \mathbb{R}$, we have

$$\begin{aligned} f_p^T(h) &= [f(p+h)] - [f(p)] \;=\; (p+h)^2 - p^2 \\ &= p^2 + 2ph + h^2 - p^2 \;=\; 2ph + h^2 \\ &= 2\cdot 3\cdot h + h^2 \;=\; 6h + h^2. \end{aligned}$$

Next we try to break f_p^T into a sum of two functions, one in $L(\mathbb{R},\mathbb{R})$ and the other in $\breve{\mathcal{O}}_1(\mathbb{R},\mathbb{R})$. Define $L \in L(\mathbb{R},\mathbb{R})$ by $L(h) = 6h$. Define $R : \mathbb{R} \to \mathbb{R}$ by $R(h) = h^2$. Then $f_p^T = L + R$ and $L \in L(\mathbb{R},\mathbb{R})$. Unassigned HW: Prove the quadratic analogue of HW#13-2. That is, prove that, for all finite dimensional vector spaces V and W, $\mathcal{Q}(V,W) \subseteq \widehat{\mathcal{O}}_2(V,W)$.

In particular, $\mathcal{Q}(\mathbb{R},\mathbb{R}) \subseteq \widehat{\mathcal{O}}_2(\mathbb{R},\mathbb{R})$. Recall: $B(\mathbb{R},\mathbb{R}) = B(\mathbb{R},\mathbb{R},\mathbb{R})$ is the set of bilinear maps $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. As R is the diagonal restriction of the bilinear mapping

$$(x,y) \;\mapsto\; 6xy \quad \in \quad B(\mathbb{R},\mathbb{R}),$$

we get $R \in \mathcal{Q}(\mathbb{R},\mathbb{R})$. Then $R \in \mathcal{Q}(\mathbb{R},\mathbb{R}) \subseteq \widehat{\mathcal{O}}_2(\mathbb{R},\mathbb{R}) \subseteq \breve{o}_1(\mathbb{R},\mathbb{R})$. Since $L \in L(\mathbb{R},\mathbb{R})$ and $f_p^T - L = R \in \breve{o}_1(\mathbb{R},\mathbb{R})$, we get $D_p f = L$. Then $[D_p f] = [L] = [6] \in \mathbb{R}^{1\times 1}$, so $\mathrm{ENT}[D_p f] = 6 = f'(p)$.

Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. We next pursue two goals:

(1) Show: $f'(p) =^* \mathrm{ENT}[D_p f]$.

(2) Show: $(\,\mathrm{dom}[f] \in \mathcal{N}_{\mathbb{R}}(p)\,) \;\Rightarrow\; (\,f'(p) = \mathrm{ENT}[D_p f]\,)$.

Let $X_0 := \{1, 1/2, 1/3, \ldots\} \cup \{0\}$, let $f_0 := (\mathrm{id}_{\mathbb{R}})|X_0 \,:\, X_0 \to \mathbb{R}$ and let $p_0 := 0$. Keep in mind that $f_0'(p_0) = 1 \neq ☹ = \mathrm{ENT}[D_{p_0} f_0]$. So, the requirement, in (2), that $\mathrm{dom}[f] \in \mathcal{N}_{\mathbb{R}}(p)$ cannot be eliminated.

51. Class 25 on 24 April 2018, Tu of Week 14

Unassigned HW: Let V be a normed vector space, let W be a vector space, let $f : V \dashrightarrow W$ and let $p \in V$. Show:

$$[\;\mathrm{dom}[f] \in \mathcal{N}_V(p)\;] \quad \Leftrightarrow \quad [\;\mathrm{dom}[f_p^T] \in \mathcal{N}_V(0_V)\;].$$

Hint: For $\Rightarrow$, use the fact that $\mathrm{dom}[f_p^T] = (\mathrm{dom}[f]) - p$. For $\Leftarrow$, use the fact that $\mathrm{dom}[f] = (\mathrm{dom}[f_p^T]) + p$.

Assigned HW#14-1, HW#14-2, HW#14-3 and HW#14-4.

Unassigned HW: Let V, W be finite dimensional vector spaces, let $\alpha, \beta : V \dashrightarrow W$ and let $p \geqslant 0$. Assume both that $\alpha = \beta$ near 0_V and that $\alpha \in \widehat{\mathcal{O}}_p(V,W)$. Show that $\beta \in \widehat{\mathcal{O}}_p(V,W)$.

DEFINITION 51.1. *For all VSs V and W, let $L_V W := L(V,W)$.*

Let V and W be vector spaces, let $G \in L_V L_V L_V W$ and let $x, y, z \in V$. We have $L_V L_V L_V W = L(V, L(V, L(V,W)))$. Since $G \in L_V L_V L_V W$ and $x \in V$, we get $G(x) \in L_V L_V W$. Since $G(x) \in L_V L_V W$ and $y \in V$, we get $(G(x))(y) \in L_V W$. Since $(G(x))(y) \in L_V W$ and $z \in V$, we get $((G(x))(y))(z) \in W$. Note: In the notation we will set up below, we will the simpler $G(x)(y)(z)$ to mean $((G(x))(y))(z)$.

DEFINITION 51.2. *Let V and W be vector spaces and let $k \in \mathbb{N}$. Then $L_V^k W := L_V \cdots L_V W$, with, on the RHS, L_V repeated k times.*

DEFINITION 51.3. *Let V and W be vector spaces and let $k \in \mathbb{N}$. Let $G \in L_V^k W$ and let $x_1, \ldots, x_k \in V$. Then we define*

$$G(x_1) \cdots (x_k) \qquad := \qquad (\cdots((G(x_1))(x_2))\cdots)(x_k).$$

Recall that, for any vector spaces V and W, we defined

$$M^k(V,W) \qquad := \qquad M^k(V,\ldots,V,W),$$

with, on the RHS, V repeated k times.

REMARK 51.4. *Let V and W be vector spaces, $k \in \mathbb{N}$. Define*

$$\Phi : M^k(V,W) \to L_V^k W \qquad \textit{and} \qquad \Psi : L_V^k W \to M^k(V,W)$$

by

$$\begin{array}{rcl} (\Phi(F))(x_1)\cdots(x_k) & = & F(x_1,\ldots,x_k) \qquad \textit{and} \\ (\Psi(G))(x_1,\ldots,x_k) & = & G(x_1)\cdots(x_k). \end{array}$$

Then Φ and Ψ are both linear, and they are inverses of one another.

Proof. Unassigned HW. □

By Remark 51.4 $\forall$vector spaces V and W, $\forall k \in \mathbb{N}$, the two related vector spaces $M^k(V,W)$ and $L_V^k W$ are isomorphic to one another. Remark 51.4 displays isomorphisms in each direction.

DEFINITION 51.5. *Let V and W be finite dimensional vector spaces. Let $f : V \dashrightarrow W$. Then $Df : V \to L_V W$ is defined by $(Df)(p) = D_p f$.*

DEFINITION 51.6. *Let V and W be finite dimensional vector spaces. Let $f : V \dashrightarrow W$. Then*

(1) $D_\circ^0 f := f$,
(2) $D_\circ^1 f := Df$ *and*
(3) $\forall k \in [2..\infty)$, $D_\circ^k f := D \cdots Df$, *with D is repeated k times.*

In Definition 51.6, we have

$$\begin{array}{rcl} D^0_\circ f \;=\; f & : & V \;\to\; W, \\ D^1_\circ f \;=\; Df & : & V \;\to\; L_V W, \\ D^2_\circ f & : & V \;\to\; L_V L_V W, \\ D^3_\circ f & : & V \;\to\; L_V L_V L_V W, \\ D^4_\circ f & : & V \;\to\; L_V^4 W, \\ & \vdots & \\ D^k_\circ f & : & V \;\to\; L_V^k W, \\ & \vdots & \end{array}$$

DEFINITION 51.7. *Let V and W be finite dimensional vector spaces. Let $f : V \dashrightarrow W$. Then*

(1) $D^0 f := f$,
(2) $D^1 f := Df$ *and*
(3) $\forall k \in [2..\infty)$, $D^k f : V \to M^k(V, W)$ *is defined by*

$$((D^k f)(p))(x_1, \ldots, x_k) \quad = \quad ((D^k_\circ f)(p))(x_1) \cdots (x_k).$$

DEFINITION 51.8. *Let V and W be finite dimensional vector spaces, let $p \in V$ and let $k \in \mathbb{N}_0$. Let $f : V \dashrightarrow W$. Then $D^k_p f := (D^k f)(p)$.*

Let V and W be fdVSs and let $p \in V$. Let $f : V \dashrightarrow W$. Note that $D^0_p f = f(p) \in W$ and that $D^1_p f = (Df)(p) = D_p f \in L_V W$.

Let V and W be fdVSs, let $p \in V$, let $k \in \mathbb{N}_0$ and let $u_1, \ldots, u_k \in V$. Let $f : V \dashrightarrow W$. Then

$$\begin{aligned} (D^k_p f)(u_1, \ldots, u_k) &= ((D^k f)(p))(u_1, \ldots, u_k) \\ &= ((D^k_\circ f)(p))(u_1) \cdots (u_k). \end{aligned}$$

Also, it turns out that

$$(D^k_p f)(u_1, \ldots, u_k) \;\;{}^*{=}\;\; (\partial_{u_1} \cdots \partial_{u_k} f)(p).$$

Also, it turns out that, for any $\sigma : [1..k] \hookrightarrow\!\!> [1..k]$,

$$(D^k_p f)(u_1, \ldots, u_k) \;\;{}^*{=}\;\; (\partial_{u_{\sigma(1)}} \cdots \partial_{u_{\sigma(k)}} f)(p).$$

REMARK 51.9. *Let V and W be finite dimensional vector spaces, let $f, g : V \dashrightarrow W$, let $p \in V$ and let $k \in \mathbb{N}_0$. Then*

$$[\ D^k f = D^k g\ ] \quad \Leftrightarrow \quad [\ D^k_\circ f = D^k_\circ g\ ].$$

Proof. Unassigned HW. □

Assigned HW#14-5.

COROLLARY 51.10. *Let V and W be finite dimensional vector spaces. Let $f, g : V \dashrightarrow W$. Let U be an open subset of V. Assume that $f = g$ on U. Then $Df = Dg$ on U.*

Proof. We wish to show: $\forall p \in U$, $(Df)(p) = (Dg)(p)$. Let $p \in U$ be given. We wish to show: $(Df)(p) = (Dg)(p)$.

Let $S := \mathrm{LINS}_p f$ and $T := \mathrm{LINS}_p g$. By HW#14-5, $S \subseteq T$. By HW#14-5 (with f and g interchanged), $T \subseteq S$. As $S \subseteq T \subseteq S$, $S = T$. Then $(Df)(p) = D_p f = \mathrm{ELT}\, S = \mathrm{ELT}\, T = D_p g = (Dg)(p)$. □

Openness of U in Corollary 51.10 is needed: Define $f, g : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = |x| \qquad \text{and} \qquad g(x) = x.$$

Then $f = g$ on $[0, \infty)$, but $(Df)(0) = ☹ \neq (Dg)(0)$.

COROLLARY 51.11. *Let V and W be finite dimensional vector spaces. Let $f, g : V \dashrightarrow W$. Let U be an open subset of V. Assume that $f = g$ on U. Then: $\forall k \in \mathbb{N}$, $D^k f = D^k g$ on U.*

Proof. For all $k \in \mathbb{N}$, define $P_k := [\, D^k f = D^k g \text{ on } U\,]$. We wish to show: $\forall k \in \mathbb{N}$, P_k. By Corollary 51.10, P_1. By the Principle of Mathematical Induction, we want: $\forall k \in \mathbb{N}$, $(P_k \Rightarrow P_{k+1})$. Let $k \in \mathbb{N}$ be given. We want to show: $P_k \Rightarrow P_{k+1}$. Assume: P_k. We want to show: P_{k+1}. We know: $D^k f = D^k g$ on U. We want: $D^{k+1} f = D^{k+1} g$ on U.

Since $D^k f = D^k g$ on U, by Remark 51.9, $D_\circ^k f = D_\circ^k g$ on U. Then, by Corollary 51.10, $DD_\circ^k f = DD_\circ^k g$ on U. That is, $D_\circ^{k+1} f = D_\circ^{k+1} g$ on U. Then, by Remark 51.9, $D^{k+1} f = D^{k+1} g$ on U, as desired. □

REMARK 51.12. *Let V and W be finite dimensional vector spaces and let $p \geqslant 0$. Then*

(1) $\forall \alpha \in \widehat{\mathcal{O}}_p(V, W)$, *dom*$[\alpha] \in \mathcal{N}_V(0_V)$ *and*
(2) $\forall \alpha \in \breve{o}_p(V, W)$, *dom*$[\alpha] \in \mathcal{N}_V(0_V)$.

Proof. Unassigned HW. □

REMARK 51.13. *Let V and W be finite dimensional vector spaces. Let $f : V \dashrightarrow W$. Then dom$[Df] \subseteq \mathrm{IntD}_V f$.*

Proof. We wish to show: $\forall p \in \mathrm{dom}[Df]$, $p \in \mathrm{IntD}_V f$. Let $p \in \mathrm{dom}[Df]$ be given. We wish to show: $p \in \mathrm{IntD}_V f$.

Since $p \in \mathrm{dom}[Df]$, we get $(Df)(p) \neq ☹$. Let $L := D_p f$. Then $L = (Df)(p) \neq ☹$, so $L \in L(V, W)$. Then $\mathrm{dom}[L] = V$.

Let $R := f_p^T - L$. Then $f_p^T = L + R$ and $R \in \breve{\circ}_1(V, W)$. Then, by Remark 51.12, $\mathrm{dom}[R] \subseteq \mathcal{N}_V(0_V)$. We have

$$\mathrm{dom}[L + R] \quad = \quad (\mathrm{dom}[L]) \;\cap\; (\mathrm{dom}[R]),$$

so, since $\mathrm{dom}[R] \subseteq V = \mathrm{dom}[L]$, we get $\mathrm{dom}[L + R] = \mathrm{dom}[R]$. Then $\mathrm{dom}[f_p^T] = \mathrm{dom}[L + R] = \mathrm{dom}[R] \in \mathcal{N}_V(0_V)$. Then $\mathrm{dom}[f] \in \mathcal{N}_V(p)$. Let $S := \mathrm{dom}[f]$. Since $S \in \mathcal{N}_V(p)$, by HW#3-1, we get $p \in \mathrm{Int}_V S$. Then $p \in \mathrm{Int}_V(\mathrm{dom}[f]) = \mathrm{IntD}_V f$, as desired. $\square$

Let V be a nonzero finite dimensional vector space. Then V has no open points, so, for all $S \subseteq V$, we have $\mathrm{Int}_V S \subseteq \mathrm{LP}_V S$. It follows, for any finite dimensional vector space W, for any $f : V \dashrightarrow W$, that $\mathrm{IntD}_V f \subseteq \mathrm{LPD}_V f$, so, by Remark 51.13, $\mathrm{dom}[Df] \subseteq \mathrm{LPD}_V f$.

REMARK 51.14. *Let V and W be finite dimensional vector spaces, let $f : V \dashrightarrow W$ and let $p \in V$. Then:*

$$[\ f \textit{ is continuous at } p\ ] \quad \Leftrightarrow \quad [\ f_p^T \textit{ is continuous at } 0_V\ ].$$

Proof. Unassigned HW. $\square$

REMARK 51.15. *Let V and W be finite dimensional vector spaces, let $p \geqslant 0$ and let $\alpha \in \breve{\circ}_p(V, W)$. Then α is continuous at 0_V.*

Proof. Since $\alpha \in \breve{\circ}_p(V, W) \subseteq \breve{\circ}_0(V, W) = \breve{\circ}(V, W)$, it follows, by (2) of Definition 47.1, that α is continuous at 0_V. $\square$

REMARK 51.16. *Let V and W be finite dimensional vector spaces, let $p > 0$ and let $\alpha \in \widehat{\mathcal{O}}_p(V, W)$. Then α is continuous at 0_V.*

Proof. Since $\alpha \in \widehat{\mathcal{O}}_p(V, W) \subseteq \breve{\circ}_0(V, W) = \breve{\circ}(V, W)$, it follows, by (2) of Definition 47.1, that α is continuous at 0_V. $\square$

THEOREM 51.17. *Let V and W be finite dimensional vector spaces, let $f : V \dashrightarrow W$ and let $p \in \mathrm{dom}[Df]$. Then f is continuous at p.*

Proof. By Remark 51.14, we wish to show: f_p^T is continuous at 0_V.

Since $p \in \mathrm{dom}[Df]$, we get $(Df)(p) \neq ☹$. Let $L := D_p f$. Then $L = (Df)(p) \neq ☹$, so $L \in L(V, W)$. By Theorem 46.8, $L : V \to W$ is continuous. In particular, L is continuous at 0_V. Let $R := f_p^T - L$.

Then $R \in \breve{\mathcal{O}}_1(V,W)$, so, by Remark 51.15, that R is continuous at 0_V. As L and R are both continuous at 0_V, $L+R$ is continuous at 0_V. So, since $f_p^T = L+R$, we see that f_p^T is continuous at 0_V, as desired. $\square$

Let V and W be finite dimensional vector spaces and $f : V \dashrightarrow W$. Then, by Theorem 51.17, we have: $\mathrm{dom}[Df] \subseteq \mathrm{dct}[f]$.

LEMMA 51.18. *Let V and W be finite dimensional vector spaces, $L \in L(V,W)$, $p,u \in V$. Assume: V is nonzero. Then $(\partial_u L)(p) = Lu$.*

Proof. For all $h \in \mathbb{R}$, we have

$$\begin{aligned}(SS_L^{p,u})(h) &= \frac{[L(p+hu)] - [Lp]}{h} \\ &= \frac{[Lp] + h \cdot [Lu] - [Lp]}{h} \\ &= \frac{h \cdot [Lu]}{h}.\end{aligned}$$

Therefore, for all $h \in \mathbb{R}_0^\times$, we have $(SS_L^{p,u})(h) = Lu = C_{\mathbb{R}}^{Lu}(h)$. Then $SS_L^{p,u} = C_{\mathbb{R}}^{Lu}$ on $\mathbb{R}_0^\times$, so $\lim\limits_0 SS_L^{p,u} = \lim\limits_0 C_{\mathbb{R}}^{Lu}$. Then $(\partial_u L)(p) = Lu$. $\square$

REMARK 51.19. *Let V and W be finite dimensional vector spaces and let $\alpha \in \breve{\mathcal{O}}(V,W)$. Assume: V is nonzero. Then $\lim\limits_{0_v} \alpha = 0_W$.*

Proof. Since V is nonzero, it follows that V has no open points. Then, for all $S \subseteq V$, we have: $\mathrm{Int}_V S \subseteq \mathrm{LP}_V S$. In particular, we have $\mathrm{IntD}_V \alpha \subseteq \mathrm{LPD}_V \alpha$. Since $\alpha \in \breve{\mathcal{O}}(V,W)$, we see, by (1) of Definition 47.1, that $\mathrm{dom}[\alpha] \in \mathcal{N}_V(0_V)$. Then, by HW#3-1, $0_V \in \mathrm{Int}_V(\mathrm{dom}[\alpha])$ *i.e.*, $0_V \in \mathrm{IntD}_V \alpha$. So, by Remark 25.19, we want: $\alpha \to 0_W$ near 0_V.

As $\alpha \in \breve{\mathcal{O}}(V,W)$, by (2) and (3) of Definition 47.1, we have both ($\alpha \to \alpha(0_V)$ near 0_V) and ($\alpha(0_V) = 0_W$). Then $\alpha \to 0_W$ near 0_V. $\square$

LEMMA 51.20. *Let V and W be finite dimensional vector spaces, $u \in V$ and $R \in \breve{\mathcal{O}}_1(V,W)$. Assume: V is nonzero. Then $(\partial_u R)(0_V) = 0_W$.*

Proof. Let $\psi := SS_R^{0_V,u}$. Then $(\partial_u R)(0_V) = \lim\limits_0 \psi$. Want: $\lim\limits_0 \psi = 0_W$.

We have $R \in \breve{\mathcal{O}}_1(V,W) \subseteq \breve{\mathcal{O}}_0(V,W) = \breve{\mathcal{O}}(V,W)$, so, by (3) of Definition 47.1, $R(0_V) = 0_W$. Then, $\forall h \in \mathbb{R}$,

$$\begin{aligned}\psi(h) &= \frac{[R(0_V + hu)] - [R(0_V)]}{h} \\ &= \frac{[R(hu)] - [0_W]}{h} = \frac{R(hu)}{h}.\end{aligned}$$

Define $\iota : \mathbb{R} \to V$ by $\iota(h) = hu$. Then, $\forall h \in \mathbb{R}$,

$$\psi(h) \quad = \quad \frac{R(\iota(h))}{h} \quad = \quad \frac{(R \circ \iota)(h)}{\mathrm{id}_{\mathbb{R}}(h)} \quad = \quad \left(\frac{R \circ \iota}{\mathrm{id}_R} \right)(h).$$

Then $\psi = \dfrac{R \circ \iota}{\mathrm{id}_R}$.

We have $\iota \in L(\mathbb{R}, V) \subseteq \widehat{\mathcal{O}}_1(\mathbb{R}, V)$ and $R \in \breve{o}_1(V, W)$. Then

$$\begin{array}{rcl} R \circ \iota & \in & [\breve{o}_1(V, W)] \circ [\widehat{\mathcal{O}}_1(\mathbb{R}, V)] \\ & \subseteq & \breve{o}_1(\mathbb{R}, W) \quad = \quad [|\bullet|][\breve{o}(\mathbb{R}, W)]. \end{array}$$

Choose $\alpha \in \breve{o}(\mathbb{R}, W)$ such that $R \circ \iota = |\bullet|\alpha$. Then

$$\psi \quad = \quad \frac{R \circ \iota}{\mathrm{id}_R} \quad = \quad \left[\frac{|\bullet|}{\mathrm{id}_{\mathbb{R}}} \right] \alpha.$$

Let $\beta := \mathrm{adj}_0^0([|\bullet|]/[\mathrm{id}_{\mathbb{R}}])$. then $\psi = \beta\alpha$ on $\mathbb{R}_0^\times$. Then $\lim\limits_0 \psi = \lim\limits_0 \beta\alpha$.

Since $\mathrm{im}[\beta] \subseteq \{-1, 0, 1\}$ and since $\mathrm{dom}[\beta] = \mathbb{R}$, we conclude that $\beta \in \widehat{\mathcal{O}}(\mathbb{R}, \mathbb{R})$. Define $* \in B(\mathbb{R}, W, W)$ by $c * w = cw$. That is, let $*$ denote scalar multiplication in W. Then $\beta\alpha = \beta \underset{\mathbb{R}}{*} \alpha$. Since

$$\begin{array}{rcl} \beta \underset{\mathbb{R}}{*} \alpha & \in & [\widehat{\mathcal{O}}(\mathbb{R}, \mathbb{R})] \underset{\mathbb{R}}{*} [\breve{o}(\mathbb{R}, W)] \\ & = & [\widehat{\mathcal{O}}_0(\mathbb{R}, \mathbb{R})] \underset{\mathbb{R}}{*} [\breve{o}_0(\mathbb{R}, W)] \\ & \subseteq & \breve{o}_0(\mathbb{R}, W) \quad = \quad \breve{o}(\mathbb{R}, W), \end{array}$$

it follows, from Remark 51.19, that $\lim\limits_0 (\beta \underset{\mathbb{R}}{*} \alpha) = 0_W$.

Then $\lim\limits_0 \psi = \lim\limits_0 \beta\alpha = \lim\limits_0 (\beta \underset{\mathbb{R}}{*} \alpha) = 0_W$, as desired. □

52. Class 26 on 26 April 2018, Th of Week 14

Recall: Let V and W be finite dimensional vector spaces and let $u \in V$. Assume V is nonzero. Then

(1) $\forall L \in L(V, W)$, $(\partial_u L)(0_V) = L(u)$ and
(2) $\forall R \in \breve{o}_1(V, W)$, $(\partial_u R)(0_V) = 0_W$.

For (1), see Lemma 51.18. For (2), see Lemma 51.20.

REMARK 52.1. *Let V and W be finite dimensional vector spaces, let $f : V \dashrightarrow W$ and let $u \in V$. Let $p \in V$ and let $g := f_p^T$. Then $(\partial_u g)(0_V) = (\partial_u f)(p)$.*

Proof. Unassigned HW. □

REMARK 52.2. *Let V and W be finite dimensional vector spaces, let $\phi, \psi : V \dashrightarrow W$ and let $u \in V$. Let $p \in \mathrm{LPD}_V(\phi + \psi)$. Then $(\partial_u(\phi + \psi))(p) \ {=^*}\ [(\partial_u \phi)(p)] + [(\partial_u \psi)(p)]$.*

Proof. Unassigned HW. □

Recall (Theorem 51.17): Let V and W be fdVSs, let $f : V \dashrightarrow W$ and let $p \in \mathrm{dom}[Df]$. Then f is continuous at p.

Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. We have been pursuing two goals:

(1) Show: $f'(p) \ {=^*}\ \mathrm{ENT}[D_p f]$.
(2) Show: $(\,\mathrm{dom}[f] \in \mathcal{N}_{\mathbb{R}}(p)\,) \ \Rightarrow\ (\,f'(p) = \mathrm{ENT}[D_p f]\,)$.

We will omit (2), for lack of time this semester. We will eventually show that (1) is a corollary of the next theorem.

THEOREM 52.3. *Let V and W be fdVSs, let $f : V \dashrightarrow W$, and let $p, u \in V$. Assume that V is nonzero. Then $(D_p f)(u) \ {^*{=}}\ (\partial_u f)(p)$.*

Proof. We wish to show:

$$[\ (D_p f)(u) \neq ☹\ ] \quad \Rightarrow \quad [\ (D_p f)(u) \ = \ (\partial_u f)(p)\ ].$$

Assume: $(D_p f)(u) \neq ☹$. Want: $(D_p f)(u) = (\partial_u f)(p)$.

Let $L := D_p f$. Then $L(u) = (D_p f)(u) \neq ☹$, so $L \neq ☹$. Then $L \in \mathcal{L}(V, W)$. Let $g := f_p^T$ and let $R := g - L$. Then $R \in \breve{\mathcal{O}}_1(V, W)$ and $g = L + R$. By Remark 52.1, $(\partial_u g)(0_V) = (\partial_u f)(p)$. We wish to show: $(D_p f)(u) = (\partial_u g)(0_V)$. Equivalently, we want: $(\partial_u g)(0_V) = L(u)$.

By Lemma 51.18 and Lemma 51.20, we have:

(1) $(\partial_u L)(0_V) = L(u)$ and
(2) $(\partial_u R)(0_V) = 0_W$.

Since $(Df)(p) = D_p f \neq ☹$, we get $p \in \mathrm{dom}[Df]$. By Remark 51.13, $\mathrm{dom}[Df] \subseteq \mathrm{IntD}_V f$. Since $g = f_T^p$, we get $\mathrm{dom}[g] = (\mathrm{dom}[f]) - p$, and so $\mathrm{IntD}_V g = (\mathrm{IntD}_V f) - p$. Then $0_V = p - p \in (\mathrm{IntD}_V f) - p = \mathrm{IntD}_V g$.

Since V is nonzero, V has no open points, and so we have: $\forall S \subseteq V$, $\mathrm{Int}_V S \subseteq \mathrm{LP}_V S$. Then $\mathrm{IntD}_V g \subseteq \mathrm{LPD}_V g$. Then

$$0_V \quad \in \quad \mathrm{IntD}_V g \quad \subseteq \quad \mathrm{LPD}_V g \quad = \quad \mathrm{LPD}_V(L + R).$$

So, by Remark 52.2, $(\partial_u(L + R))(0_V) \ {=^*}\ [(\partial_u L)(0_V)] + [(\partial_u R)(0_V)]$. Therefore, since $g = L + R$, by (1) and (2) above, we conclude that $(\partial_u g)(0_V) \ {=^*}\ [L(u)] + [0_W]$. So, since $[L(u)] + [0_W] = L(u) \neq ☹$, it follows that $(\partial_u g)(0_V) = L(u)$, as desired. □

In Theorem 52.3, we cannot replace "$\overset{*}{=}$" by "$=$", and a counterexample is as follows: Let $S := \{(x,x^2)\,|\,x > 0\}$. Let $f := \chi_{\mathbb{R}^2}^S : \mathbb{R}^2 \to \mathbb{R}$. Let $p := 0_2$. Then f is not continuous at p, so, by Theorem 51.17, $D_p f = ☹$. Unassigned exercise: Show, $\forall u \in \mathbb{R}^2$, that $(\partial_u f)(p) = 0$. Then, for all $u \in \mathbb{R}^2$, we have $(D_p f)(u) = ☹ \neq 0 = (\partial_u f)(p)$.

COROLLARY 52.4. *Let $m \in \mathbb{N}$, let W be a finite dimensional vector space, let $f : \mathbb{R}^m \dashrightarrow W$, let $p \in \mathbb{R}^m$ and let $j \in [1..m]$. Then $(D_p f)(e_j^m) \;\overset{*}{=}\; (\partial_j f)(p)$.*

Proof. Let $u := e_j^m$. By definition of ∂_j, we have: $(\partial_j f)(p) = (\partial_u f)(p)$. By Theorem 52.3, we have: $(D_p f)(u) \;\overset{*}{=}\; (\partial_u f)(p)$.

Then $(D_p f)(e_j^m) = (D_p f)(u) \;\overset{*}{=}\; (\partial_u f)(p) = (\partial_j f)(p)$. □

For all $n \in \mathbb{N}$, for all $k \in [1..n]$, define $\pi_k^n : \mathbb{R}^n \to \mathbb{R}$ by $\pi_k^n(x) = x_k$.

Let $m, n \in \mathbb{N}$ and let $L \in L(\mathbb{R}^m, \mathbb{R}^n)$. Recall:

$$\forall j \in [1..m],\, \forall k \in [1..n], \quad [L]_{jk} \;=\; (L(e_j^m))_k.$$

Equivalently:

$$\forall j \in [1..m],\, \forall k \in [1..n], \quad [L]_{jk} \;=\; \pi_k^n(L(e_j^m)).$$

COROLLARY 52.5. *Let $m, n \in \mathbb{N}$, let $f : \mathbb{R}^m \dashrightarrow \mathbb{R}^n$, let $p \in \mathbb{R}^m$. For all $k \in [1..n]$, let $f_k := \pi_k^n \circ f : \mathbb{R}^m \to \mathbb{R}$. Then:*

$$\forall j \in [1..m],\, \forall k \in [1..n], \qquad [D_p f]_{jk} \;\overset{*}{=}\; (\partial_j f_k)(p).$$

Proof. Let $j \in [1..m]$ and $k \in [1..n]$ be given. We wish to show: $[D_p f]_{jk} \;\overset{*}{=}\; (\partial_j f_k)(p)$.

Let $u := e_j^m$. By definition of ∂_j, $(\partial_j f_k)(p) = (\partial_u f_k)(p)$. By Theorem 52.3, $(D_p f)(u) \;\overset{*}{=}\; (\partial_u f)(p)$, so $\pi_k^n((D_p f)(u)) \;\overset{*}{=}\; \pi_k^n((\partial_u f)(p))$. Also, we have $[D_p f]_{jk} = \pi_k^n((D_p f)(e_j^m)) = \pi_k^n((D_p f)(u))$.

For all $x \in \mathbb{R}^m$, $f(x) = (f_1(x), \ldots, f_n(x))$. Unassigned exercise: Show, for all $x \in \mathbb{R}^m$, that

$$(\partial_u f)(x) \quad = \quad (\ (\partial_u f_1)(x)\ ,\ \ldots\ ,\ (\partial_u f_k)(x)\ ),$$

so $\pi_k^n((\partial_u f)(x)) = (\partial_u f_k)(x)$. Then $\pi_k^n((\partial_u f)(p)) = (\partial_u f_k)(p)$. Then

$$\begin{aligned} [D_p f]_{jk} &= \pi_k^n((D_p f)(u)) \;\overset{*}{=}\; \pi_k^n((\partial_u f)(p)) \\ &= (\partial_u f_k)(p) \;=\; (\partial_j f_k)(p), \end{aligned}$$

as desired. □

We have $\mathbb{R}^1 = \mathbb{R}^{\{1\}} \neq \mathbb{R}$. Also, $\pi_1^1 : \mathbb{R}^1 \to \mathbb{R}$ is a VS isomorphism.

Note: For any $L \in L(\mathbb{R}, \mathbb{R})$, we have $\mathrm{ENT}[L] = L(1)$. (*Proof:* Choose $m \in \mathbb{R}$ such that, for all $x \in \mathbb{R}$, $L(x) = mx$. Then $L(1) = m$ and $[L] = [m]$. Then $\mathrm{ENT}[L] = \mathrm{ENT}[m] = m = L(1)$. QED)

We now complete our goal:

COROLLARY 52.6. *Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$ and let $p \in \mathbb{R}$. Then we have: $f'(p) \;=^*\; \mathrm{ENT}[D_p f]$.*

Proof. Let $u := 1$. Then, by definition of f' and $\partial_u f$, we have $\partial_u f = f'$. Let $L := D_p f$. Then $L(1) = (D_p f)(u)$. We want: $f'(p) \;=^*\; \mathrm{ENT}[L]$.

By Theorem 52.3, $(D_p f)(u) \;\;^*\!\!=\;\; (\partial_u f)(p)$. We have $\mathrm{ENT}[L] = L(1)$. Then $f'(p) = (\partial_u f)(p) \;=^*\; (D_p f)(u) = L(1) = \mathrm{ENT}[L]$, as desired. $\square$

We now practice differentiation of polynomials, in the multivariable setting. We start with differentiation of constants.

REMARK 52.7. *Let V and W be finite dimensional vector spaces, let $x \in V$ and let $C \in C(V, W)$. Then $D_x C = \mathbf{0}_{VW}$.*

Proof. For all $h \in V$, since $C \in C(V, W)$, we get $C(x+h) = C(x)$; then $C_x^T(h) = [C(x+h)] - [C(x)] = 0_W = \mathbf{0}_{VW}(h)$. Then $C_x^T = \mathbf{0}_{VW}$. Let $L := \mathbf{0}_{VW}$, $R := \mathbf{0}_{VW}$. Then $L \in L(V, W)$, $R \in \breve{\mathcal{O}}_1(V, W)$, $C_x^T = L + R$.

Since $L \in L(V, W)$ and $C_x^T - L = R \in \breve{\mathcal{O}}_1(V, W)$, we conclude that $L \in \mathrm{LINS}_x C$, so $L = D_x C$. Then $D_x C = L = \mathbf{0}_{VW}$, as desired. $\square$

Let V and W be finite dimensional vector spaces and let $C \in C(V, W)$. By Remark 52.7, $\forall x \in V$, $D_x C = \mathbf{0}_{VW}$. Let $Y := L(V, W)$ and let $Z := L(V, Y)$. Then $Y = L_V W$ and $Z = L_V Y = L_V L_V W$. Also, we have $0_Y = C_V^{0_W}$ and $0_Z = C_V^{0_Y}$. For all $x \in V$,

$$(DC)(x) \;=\; D_x C \;=\; \mathbf{0}_{VW} \;=\; C_V^{0_W} \;=\; 0_Y \;=\; C_V^{0_Y}(x) \;=\; 0_Z(x).$$

Then $DC = 0_Z = 0_{L_V L_V W}$.

We move on to differentiation of linear functions.

REMARK 52.8. *Let V and W be finite dimensional vector spaces, let $x \in V$ and let $L \in L(V, W)$. Then $D_x L = L$.*

Proof. For all $h \in V$, we have

$$\begin{aligned} L_x^T(h) \;&=\; [L(x+h)] \,-\, [L(x)] \\ &=\; [L(x)] \,+\, [L(h)] \,-\, [L(x)] \;\;=\;\; L(h). \end{aligned}$$

Then $L_x^T = L$. Let $L := L$ and $R := \mathbf{0}_{VW}$. Then $L \in L(V,W)$ and $R \in \breve{\mathcal{O}}_1(V,W)$ and $L_x^T = L + R$.

Since $L \in L(V,W)$ and $L_x^T - L = R \in \breve{\mathcal{O}}_1(V,W)$, we conclude that $L \in \mathrm{LINS}_x L$, so $L = D_x L$. Then $D_x L = L$, as desired. $\square$

Let V and W be finite dimensional vector spaces and let $L \in L(V,W)$. By Remark 52.8, $\forall x \in V$, $D_x L = L$. So, for all $x \in V$,

$$(DL)(x) \quad = \quad D_x L \quad = \quad L \quad = \quad C_V^L(x).$$

Let $Y := L(V,W)$. Then $DL = C_V^L \in C(V,Y)$.

We move on to differentiation of quadratic functions.

Let V and W be finite dimensional vector spaces and let $Q \in Q(V,W)$. Let $* := \Pi_{VW}^2 Q$ denote the polarization of Q. Recall that $* \in SB(V,W)$ and that $\Delta_* = Q$. For all $x \in V$, we have $Q(x) = \Delta_* x = x * x$.

REMARK 52.9. *Let V and W be finite dimensional vector spaces, let $x \in V$ and let $Q \in Q(V,W)$. Let $* := \Pi_{VW}^2 Q$, and define $L \in L(V,W)$ by $L(h) = 2 \cdot [x * h]$. Then $D_x Q = L$.*

Proof. For all $h \in V$, we have

$$\begin{aligned} Q_x^T(h) &= [Q(x+h)] - [Q(x)] \\ &= [(x+h)*(x+h)] - [x*x] \\ &= [x*x] + 2\cdot[x*h] + [h*h] - [x*x] \\ &= 2\cdot[x*h] + [h*h] \\ &= [L(h)] + [Q(h)]. \end{aligned}$$

Then $Q_x^T = L + Q$. We have $Q \in Q(V,W) \subseteq \widehat{\mathcal{O}}_2(V,W) \subseteq \breve{\mathcal{O}}_1(V,W)$.

Since $L \in L(V,W)$ and $Q_x^T - L = Q \in \breve{\mathcal{O}}_1(V,W)$, we conclude that $L \in \mathrm{LINS}_x Q$, so $L = D_x Q$. Then $D_x Q = L$, as desired. $\square$

Let V and W be finite dimensional vector spaces and let $Q \in Q(V,W)$. We seek a formula for DQ, but it requires some preliminary definitions.

DEFINITION 52.10. *Let V and W be vector spaces. Then we define $M^0(V,W) := W$ and $SM^0(V,W) = W$.*

Let V and W be vector spaces. Recall:

$$\begin{aligned} &M^1(V,W) = L(V,W), \quad SM^1(V,W) = L(V,W), \\ &M^2(V,W) = B(V,W), \quad SM^2(V,W) = SB(V,W), \\ &M^3(V,W) = T(V,W) \quad \text{and} \quad SM^3(V,W) = ST(V,W). \end{aligned}$$

DEFINITION 52.11. *Let V and W be vector spaces, let $n \in \mathbb{N}_0$ and let $F \in M^n(V,W)$. Then*

(1) *we define* $F_0^* \ := \ F \ \in \ M^n(\,V\,,\,M^0(V,W)\,)$,
(2) *we define* $F_n^* \ := \ F \ \in \ M^0(\,V\,,\,M^n(V,W)\,)$ *and*
(3) $\forall j \in (0..n)$, *we define* $F_j^* \ \in \ M^{n-j}\,(\,V\,,\,M^j(V,W)\,)$ *by*

$$(F_j^*(x_1,\ldots,x_{n-j}))(y_1,\ldots,y_j) \ = \ F(x_1,\ldots,x_{n-j},y_1,\ldots,y_j).$$

Let V be a vector space, let $x \in V$ and let $n \in \mathbb{N}$. Then the notation "$x^{(n)}$" will be used to denote "$x,x,\ldots,x$", with x repeated n times. We sometimes omit the parentheses and simply write x^n for $x^{(n)}$.

Let V and W be vector spaces, let $n \in [2..\infty)$ and let $F \in M^n(V,W)$. Then, for all $x,y \in V$, we have: $((F_1^*)(x^{(n-1)}))(y) = F(x^{(n-1)},y)$. Also, for all $x \in V$, we have $\Delta_F(x) = F(x^{(n)})$ and $\Delta_{F_1^*}(x) = F_1^*(x^{(n-1)})$.

REMARK 52.12. *Let V and W be vector spaces, $n \in \mathbb{N}_0$, $j \in [0..n]$ and $F \in SM^n(V,W)$. Then $F_j^* \ \in \ SM^{n-j}\,(\,V\,,\,SM^j(V,W)\,)$.*

Proof. Unassigned HW. ☐

We use the new notation to restate Remark 52.9:

REMARK 52.13. *Let V and W be finite dimensional vector spaces, $x \in V$ and $Q \in Q(V,W)$. Let $B := \Pi_{VW}^2 Q$. Then $D_xQ = 2 \cdot [B_1^*(x)]$.*

Proof. Define $L \in L(V,W)$ by $L(h) = 2 \cdot [B(x,h)]$. By Remark 52.9 (with $*$ replaced by B), we have $D_xQ = L$. It therefore suffices to show that $2 \cdot [B_1^*(x)] = L$. We wish to show: $\forall h \in V$, $2 \cdot [(B_1^*(x))(h)] = L(h)$. Let $h \in V$ be given. We want to prove: $2 \cdot [(B_1^*(x))(h)] = L(h)$.

By definition of B_1^*, $2 \cdot [(B_1^*(x))(h)] = 2 \cdot [B(x,h)]$. By definition of L, $L(h) = 2 \cdot [B(x,h)]$. Then $2 \cdot [(B_1^*(x))(h)] = 2 \cdot [B(x,h)] = L(h)$. ☐

Let V and W be finite dimensional vector spaces and let $Q \in Q(V,W)$. Let $B := \Pi_{VW}^2 Q$. We now have enough notation to develop a formula for DQ: By Remark 52.13, $\forall x \in V$, $D_xQ = B_1^*(x)$. So, $\forall x \in V$,

$$(DQ)(x) \quad = \quad D_xQ \quad = \quad B_1^*(x).$$

Let $Y := L(V,W)$. Then our formula reads: $DQ = B_1^* \in L(V,Y)$.

We move on to differentiation of cubic functions.

Let V and W be finite dimensional vector spaces and let $K \in K(V,W)$. Let $T := \Pi_{VW}^3 K$ denote the polarization of K. Recall that $T \in ST(V,W)$ and that $\Delta_T = K$. For all $x \in V$, we have $K(x) = \Delta_T x = T(x^{(3)})$.

REMARK 52.14. *Let V and W be finite dimensional vector spaces, $x \in V$ and $K \in K(V,W)$. Let $T := \Pi^3_{VW} K$. Then $D_x K = 3 \cdot [T_1^*(x^{(2)})]$.*

Proof. Let $L := 3 \cdot [T_1^*(x^{(2)})]$, $B := 3 \cdot [T_2^*(x)]$ and $R := \Delta_B + K$. Then, for all $h \in V$, we have $L(h) = 3 \cdot [T(x^{(2)}, h)]$ and

$$\begin{aligned} R(h) &= [B(h,h)] \; + \; [K(h)] \\ &= 3 \cdot [T(x,h,h)] \; + \; [T(h,h,h)] \\ &= 3 \cdot [T(x,h^{(2)})] \; + \; [T(h^{(3)})] \end{aligned}$$

For all $h \in V$, we have

$$\begin{aligned} K_x^T(h) &= [K(x+h)] \; - \; [K(x)] \\ &= [T((x+h)^{(3)})] \; - \; [T(x^{(3)})] \\ &= [T(x^{(3)})] \; + \; 3 \cdot [T(x^{(2)}, h)] \; + \\ &\qquad\qquad 3 \cdot [T(x, h^{(2)})] \; + \; [T(h^{(3)})] \; - \; [T(x^{(3)})] \\ &= 3 \cdot [T(x^{(2)}, h)] + 3 \cdot [T(x, h^{(2)})] + [T(h^{(3)})] \\ &= [L(h)] + [R(h)] \quad = \quad (L+R)(h). \end{aligned}$$

Then $Q_x^T = L + R$. Since $B \in B(V,W)$, we get

$$\Delta_B \quad \in \quad Q(V,W) \quad \subseteq \quad \widehat{\mathcal{O}}_2(V,W) \quad \subseteq \quad \breve{o}_1(V,W).$$

Also,

$$K \quad \in \quad K(V,W) \quad \subseteq \quad \widehat{\mathcal{O}}_3(V,W) \quad \subseteq \quad \breve{o}_1(V,W).$$

Then $R = \Delta_B + K \in \breve{o}_1(V,W)$.

Since $L \in L(V,W)$ and $K_x^T - L = R \in \breve{o}_1(V,W)$, we conclude that $L \in \mathrm{LINS}_x K$, so $L = D_x K$. Then $D_x K = L = 3 \cdot [T_1^*(x^{(2)})]$. ☐

In the notation of Remark 52.14, we have $K(x) = T(x^{(3)})$, and so K is a multivariable analogue for x^3 in freshman calculus. In freshman calculus, we have $(d/dx)(x^3) = 3x^2$, which we prove using

$$(d/dx)(x^3) \quad = \quad \lim_{h \to 0} \frac{(x+h)^3 \; - \; x^3}{h},$$

followed by expanding $(x+h)^3$. The analogue in multivariable analysis is $D_x K = 3 \cdot [T_1^*(x^{(2)})]$, which we prove using

$$K_x^T(h) \quad = \quad [T((x+h)^{(3)})] \; - \; [T(x^{(3)})],$$

followed by expanding $T((x+h)^{(3)})$.

Let V and W be finite dimensional vector spaces and let $K \in K(V, W)$. Let $T := \Pi^3_{VW} K$. We now have enough notation to develop a formula for DK: By Remark 52.14, $\forall x \in V$, $D_x K = 3 \cdot [T_1^*(x^{(2)})]$. So, $\forall x \in V$,

$$(DK)(x) \quad = \quad D_x K \quad = \quad 3 \cdot [T_1^*(x^{(2)})] = \quad (3 \cdot \Delta_{T_1^*})(x).$$

Let $Y := L(V, W)$. Then our formula reads: $DK = 3 \cdot [\Delta_{T_1^*}]$. We have $T_1^* \in B(V, Y)$, so $\Delta_{T_1^*} \in Q(V, Y)$, and so $3 \cdot [\Delta_{T_1^*}] \in Q(V, Y)$. Then $DK = 3 \cdot [\Delta_{T_1^*}] \in Q(V, Y)$.

We now differentiate arbitrary homogeneous polynomials:

REMARK 52.15. *Let V and W be finite dimensional vector spaces, $x \in V$, let $n \in \mathbb{N}$ and $P \in P^n(V, W)$. Let $F := \Pi^n_{VW} P$. Then $D_x P = n \cdot [F_1^*(x^{(n-1)})]$.*

Proof. Unassigned HW. $\square$

In the notation of Remark 52.15, we have $P(x) = F(x^{(n)})$, and so P is a multivariable analogue for x^n in freshman calculus. In freshman calculus, we have $(d/dx)(x^n) = nx^{n-1}$, which we prove using

$$(d/dx)(x^n) \quad = \quad \lim_{h \to 0} \frac{(x + h)^n \, - \, x^n}{h},$$

followed by expanding $(x+h)^n$. The analogue in multivariable analysis is $D_x P = n \cdot [F_1^*(x^{(n-1)})]$, which we prove using

$$K_x^T(h) \quad = \quad [F((x + h)^{(n)})] \, - \, [F(x^{(n)})],$$

followed by expanding $F((x + h)^{(n)})$.

Let V and W be fdVSs, $n \in \mathbb{N}$, $P \in P^n(V, W)$. Let $F := \Pi^n_{VW} P$. We now have enough notation to develop a formula for DP: By Remark 52.15, $\forall x \in V$, $D_x P = n \cdot [F_1^*(x^{(n-1)})]$. So, $\forall x \in V$,

$$(DP)(x) \quad = \quad D_x P \quad = \quad n \cdot [T_1^*(x^{(n-1)})] = \quad (n \cdot \Delta_{T_1^*})(x).$$

Let $Y := L(V, W)$. Then our formula reads: $DP = n \cdot [\Delta_{F_1^*}]$. We have $F_1^* \in M^{n-1}(V, Y)$, so $\Delta_{F_1^*} \in P^{n-1}(V, Y)$, and so $n \cdot [\Delta_{F_1^*}] \in P^{n-1}(V, Y)$. Then $DP = n \cdot [\Delta_{F_1^*}] \in P^{n-1}(V, Y)$.

Next time we will work on second (total) derivatives of homogeneous polynomials. Let V and W be finite dimensional vector spaces, let $n \in \mathbb{N}$ and let $P \in P^n(V, W)$. Let $F := \Pi^n_{VW} P$. We will prove: $\forall x \in V$, $D_x^2 P = n(n-1) \cdot [F_2^*(x^{(n-2)})]$. This is the analogue of the freshman calculus formula: $(d/dx)^2(x^n) = n(n-1) \cdot x^{n-2}$.

53. Class 27 on 1 May 2018, Tu of Week 15

Recall: Let V and W be fdVSs. Let $f : V \dashrightarrow W$. Then

$$\begin{array}{rcll} Df & : & V \dashrightarrow L_V W & \text{and} \\ DDf & : & V \dashrightarrow L_V L_V W. & \end{array}$$

There is a vector space isomorphism $\Psi : L_V L_V W \to B(V,W)$ given by $(\Psi(G))(y,z) = G(y)(z)$. We defined

$$D^2 f \quad := \quad \Psi \circ (DDf) \quad : \quad V \dashrightarrow B(V,W).$$

It is a fact that $\mathrm{im}[D^2 f] \subseteq SB(V,W)$, but we omit the proof, for lack of time. For any $L \in L_V W$, we showed: $DL = C_V^L \in C(V, L_V W)$.

Let V and W be fdVSs, $f : V \dashrightarrow W$ and $x \in V$. We defined:

$$D_x f := (Df)(x) \in L_V W \quad \text{and} \quad D_x^2 f := (D^2 f)(x) \in B(V,W).$$

Let $L := D_x f \in L_V W$. Then $DL = C_V^L \in C(V, L_V W)$. That is, $DD_x f = C_V^{D_x f} \in C(V, L_V W)$. By contrast, we have

$$D_x Df \quad = \quad (DDf)(x) \quad \in \quad L_V L_V W \quad = \quad L(V, L_V W).$$

Let $\Psi : L_V L_V W \to B(V,W)$ be given by $(\Psi(G))(y,z) = G(y)(z)$, as in the preceding paragraph. Since $D_x Df = (DDf)(x)$, we conclude that $\Psi(D_x Df) = \Psi((DDf)(x))$. Then

$$\begin{array}{rcccl} \Psi(D_x Df) & = & \Psi((DDf)(x)) & = & (\Psi \circ (DDf))(x) \\ & = & (D^2 f)(x) & = & D_x^2 f. \end{array}$$

Then, $\forall y, z \in V$, $(D_x Df)(y)(z) = (\Psi(D_x Df))(y,z) = (D_x^2 f)(y,z)$.

Recall: Let V and W be vector spaces, let $n \in [2..\infty)$ and let $F \in \mathrm{SM}^n(V,W)$. Then

(1) $F_0^* = F = F_n^*$ and

(2) $F_1^* \in SM^{n-1}(V, L(V,W))$ is def'd by: $\forall x_1, \ldots, x_{n-1}, y \in V$,

$$(F_1^*(x_1, \ldots, x_{n-1}))(y) \quad = \quad F(x_1, \ldots, x_{n-1}, y).$$

(3) $F_2^* \in SM^{n-2}(V, SB(V,W))$ is def'd by: $\forall x_1, \ldots, x_{n-2}, y, z \in V$,

$$(F_1^*(x_1, \ldots, x_{n-2}))(y,z) \quad = \quad F(x_1, \ldots, x_{n-2}, y, z).$$

For all $j \in \mathbb{N}_0$, for all x, we used $x^{(n)}$ as an abbreviation of $x, x, \ldots, x$, with x repeated j times. With this notation, we have:

$$\begin{array}{rrcll} \forall x, y \in V, & (F_1^*(x^{(n-1)}))(y) & = & F(x^{(n-1)}, y), & \text{and} \\ \forall x, y, z \in V, & (F_2^*(x^{(n-2)}))(y,z) & = & F(x^{(n-2)}, y, z). & \end{array}$$

Recall that $\Delta_{F_1^*} : V \to L(V,W)$ is the diagonal restriction of F_1^*, defined by $\Delta_{F_1^*}(x) = F_1^*(x^{(n-1)})$. Then

$$\forall x, y \in V, \qquad (\Delta_{F_1^*}(x))(y) \;=\; F(x^{(n-1)}, y).$$

Recall that $\Delta_{F_2^*} : V \to SB(V,W)$ is the diagonal restriction of F_2^*, defined by $\Delta_{F_2^*}(x) = F_2^*(x^{(n-2)})$. Then

$$\forall x, y, z \in V, \qquad (\Delta_{F_2^*}(x))(y,z) \;=\; F(x^{(n-2)}, y, z).$$

Recall: Let V and W be finite dimensional vector spaces and let $n \in \mathbb{N}$. By definition,

$$P^n(V,W) \quad := \quad \{\ \Delta_F \mid F \in M^n(V,W)\ \}.$$

Moreover, for any $F \in M^n(V,W)$ the diagonal restrictions of F and of the symmetrization $SF \in SM^n(V,W)$ are the same, and so

$$P^n(V,W) \quad = \quad \{\ \Delta_F \mid F \in SM^n(V,W)\ \}.$$

Moreover, for any $P \in P^n(V,W)$, $\exists! F \in SM^n(V,W)$ s.t. $P = \Delta_F$. This unque F is called the **polarization** of P, and is denoted $\Pi^n_{VW}P$.

We proved the multivariable analogue of $(d/dx)(x^n) = nx^{n-1}$. It reads as follows:

THEOREM 53.1. *Let V and W be finite dimensional vector spaces and let $n \in \mathbb{N}$. Let $P \in P^n(V,W)$. Let $F := \Pi^n_{VW}P$. Then:*

(1) $\forall x \in V, \quad D_xP \;=\; n \cdot [F_1^*(x^{(n-1)})],$ *and*
(2) $DP \;=\; n \cdot \Delta_{F_1^*}$.

Let V and W be finite dimensional vector spaces and let $n \in \mathbb{N}$. Let $P \in P^n(V,W)$. Let $F := \Pi^n_{VW}P$. If we want to compute DDP, we'll need the polarization $\Pi^{n-1}_{V,L_VW}(DP)$. Since $n \cdot F_1^* \in SM^{n-1}(V, L_VW)$ and since $\Delta_{n \cdot F_1^*} = DP$, we conclude that $\Pi^{n-1}_{V,L_VW}(DP) = n \cdot F_1^*$. Using this, we get our next result, which is the multivariable analogue of the formula $(d/dx)^2(x^n) = n(n-1) \cdot x^{n-2}$.

THEOREM 53.2. *Let V and W be finite dimensional vector spaces, let $n \in [2..\infty)$ and let $P \in P^n(V,W)$. Let $F := \Pi^n_{VW}P$. Then*

(1) $\forall x \in V, \quad D_x^2P \;=\; n(n-1) \cdot [F_2^*(x^{(n-2)})],$ *and*
(2) $D^2P \;=\; n(n-1) \cdot \Delta_{F_2^*}$.

Proof. Proof of (1): Given $x \in V$. Want: $D_x^2P = n(n-1) \cdot [F_2^*(x^{(n-2)})]$. We wish to show: $\forall y, z \in V$, $(D_x^2P)(y,z) = (n(n-1) \cdot [F_2^*(x^{(n-2)})])(y,z)$.

Given $y, z \in V$. Want: $(D_x^2 P)(y,z) = (n(n-1) \cdot [F_2^*(x^{(n-2)})])(y,z)$. We compute

$$\begin{aligned}(n(n-1) \cdot [F_2^*(x^{(n-2)})])(y,z) &= n(n-1) \cdot [(F_2^*(x^{(n-2)}))(y,z)] \\ &= n(n-1) \cdot [F(x^{(n-2)}, y, z)].\end{aligned}$$

We therefore want: $(D_x^2 P)(y,z) = n(n-1) \cdot [F(x^{(n-2)}, y, z)]$. Recall:

$$(D_x^2 P)(y,z) = (D_x DP)(y)(z).$$

We therefore want: $(D_x DP)(y)(z) = n(n-1) \cdot [F(x^{(n-2)}, y, z)]$.

Let $\overline{W} := L_V W$ and $\overline{P} := DP$. Then $\overline{P} \in P^{n-1}(V, \overline{W})$. By (2) of Theorem 53.1, we have $\overline{P} = n \cdot \Delta_{F_1^*}$. Let $\overline{F} := \Pi_{VW}^{n-1} \overline{P}$. Then $\overline{F} \in SM^{n-1}(V, \overline{W})$. Since $\overline{P} = n \cdot \Delta_{F_1^*} = \Delta_{n \cdot F_1^*}$, we get $\overline{F} = n \cdot F_1^*$. By (1) of Theorem 53.1 (with n replaced by $n-1$, P by $\overline{P}$, F by $\overline{F}$), we have $D_x \overline{P} = (n-1) \cdot [\overline{F}_1^*(x^{(n-2)})]$. Then

$$\begin{aligned}(D_x DP)(y) &= (D_x \overline{P})(y) \\ &= ((n-1) \cdot [\overline{F}_1^*(x^{(n-2)})])(y) \\ &= (n-1) \cdot [(\overline{F}_1^*(x^{(n-2)}))(y)] \\ &= (n-1) \cdot [\overline{F}(x^{(n-2)}, y)] \\ &= (n-1) \cdot [(n \cdot F_1^*)(x^{(n-2)}, y)] \\ &= (n-1) \cdot [n \cdot (F_1^*(x^{(n-2)}, y))] \\ &= n(n-1) \cdot [F_1^*(x^{(n-2)}, y)].\end{aligned}$$

Then

$$\begin{aligned}(D_x DP)(y)(z) &= (n(n-1) \cdot [F_1^*(x^{(n-2)}, y)])(z) \\ &= n(n-1) \cdot [(F_1^*(x^{(n-2)}, y))(z)] \\ &= n(n-1) \cdot [F(x^{(n-2)}, y, z)],\end{aligned}$$

as desired. *End of proof of (1).*

Proof of (2): We wish to show:

$$\forall x \in V, \qquad (D^2 P)(x) = (n(n-1) \cdot \Delta_{F_2^*})(x).$$

Let $x \in V$ be given. We wish to show: $(D^2 P)(x) = (n(n-1) \cdot \Delta_{F_2^*})(x)$. By definition of $D_x^2 P$, we have $D_x^2 P = (D^2 P)(x)$. We therefore wish to show: $D_x^2 P = (n(n-1) \cdot \Delta_{F_2^*})(x)$. We have

$$(n(n-1) \cdot \Delta_{F_2^*})(x) = n(n-1) \cdot [\Delta_{F_2^*}(x)].$$

We therefore wish to show: $D_x^2 P = n(n-1)\cdot[\Delta_{F_2^*}(x)]$.

By (1), we know that $D_x^2 P = n(n-1)\cdot[F_2^*(x^{(n-2)})]$. We have

$$F_2^* \quad \in \quad SM^{n-2}(\,V\,,\,SB(V,W)\,),$$

so, by definition of $\Delta_{F_2^*}$, we have $\Delta_{F_2^*}(x) = F_2^*(x^{(n-2)})$. Then

$$D_x^2 P \quad = \quad n(n-1)\cdot[F_2^*(x^{(n-2)})] \quad = \quad n(n-1)\cdot[\Delta_{F_2^*}(x)],$$

as desired. *End of proof of (2).* $\square$

Recall: Let S, V, W and Z be vector spaces. Let $* \in B(V,W,Z)$. Let $f : S \dashrightarrow V$ and $g : S \dashrightarrow W$. Then $f \underset{S}{*} g : S \dashrightarrow Z$ is defined by: $\forall x \in S$, $(f \underset{S}{*} g)(x) = [f(x)] * [g(x)]$.

LEMMA 53.3. *Let S, V, W and Z be vector spaces. Let $* \in B(V,W,Z)$. Let $L \in L(S,V)$ and $M \in L(S,W)$. Then $L \underset{S}{*} M \in Q(S,Z)$.*

Proof. We wish to show: $\exists B \in B(S,Z)$ s.t. $L \underset{S}{*} M = \Delta_B$.

Define $B \in B(S,Z)$ by $B(x,y) = [L(x)] * [M(y)]$. We wish to show: $L \underset{S}{*} M = \Delta_B$. We wish to show: $\forall x \in S$, $(L \underset{S}{*} M)(x) = \Delta_B(x)$. Let $x \in S$ be given. We wish to show: $(L \underset{S}{*} M)(x) = \Delta_B(x)$.

We have $(L \underset{S}{*} M)(x) = [L(x)] * [M(x)] = B(x,x) = \Delta_B(x)$. $\square$

DEFINITION 53.4. *Let S, V, W, Z be VSs, $* \in B(V,W,Z)$. Then*

(1) $\forall f : S \dashrightarrow V$, $\forall w \in W$, $\quad f \underset{S}{*} w := f \underset{S}{*} C_S^w$, *and*
(2) $\forall v \in V$, $\forall g : S \dashrightarrow W$, $\quad v \underset{S}{*} g := C_S^v \underset{S}{*} g$.

In Definition 53.4, as usual, we may sometimes omit the "S" in "$\underset{S}{*}$".

The next result is the multivariable **Product Rule**.

THEOREM 53.5. *Let S, V, W and Z be VSs. Let $* \in B(V,W,Z)$. Let $f : S \dashrightarrow V$ and $g : S \dashrightarrow W$. Let $x \in S$. Then*

$$D_x(f \underset{S}{*} g) \quad =^* \quad [\,(D_x f) \underset{S}{*} (g(x))\,] \;+\; [\,(f(x)) \underset{S}{*} (D_x g)\,].$$

Proof. We want: $(\;[\,(D_x f) \underset{S}{*} (g(x))\,] \;+\; [\,(f(x)) \underset{S}{*} (D_x g)\,] \neq ☹\;) \quad \Rightarrow$

$$(\; D_x(f \underset{S}{*} g) \quad = \quad [\,(D_x f) \underset{S}{*} (g(x))\,] \;+\; [\,(f(x)) \underset{S}{*} (D_x g)\,]\;).$$

Assume: $[\,(D_x f) \underset{S}{*} (g(x))\,] \;+\; [\,(f(x)) \underset{S}{*} (D_x g)\,] \neq ☹$. We want:

$$D_x(f \underset{S}{*} g) \quad = \quad [\,(D_x f) \underset{S}{*} (g(x))\,] \;+\; [\,(f(x)) \underset{S}{*} (D_x g)\,].$$

Let $h := f \underset{S}{*} g$, $L := D_x f$, $M := D_x g$, $v := f(x)$, $w := g(x)$. Then $h(x) = v * w$. We wish to show: $D_x h \;=\; [\, L \underset{S}{*} w \,] \,+\, [\, v \underset{S}{*} M \,]$.

Let $A := C_S^v$, $B := C_S^w$. Then

$$f_x^T = [f(x + \bullet)] - A \qquad \text{and} \qquad g_x^T = [g(x + \bullet)] - B$$

Also, $A \underset{S}{*} B = C_S^{v*w}$. Then $h_x^T = [h(x + \bullet)] - [A \underset{S}{*} B]$. Also,

$$L \underset{S}{*} w = L \underset{S}{*} B \qquad \text{and} \qquad v \underset{S}{*} M = A \underset{S}{*} M.$$

We want: $D_x h \;=\; [\, L \underset{S}{*} B \,] \,+\, [\, A \underset{S}{*} M \,]$. It therefore suffices to show:

$$h_x^T \;-\; [\, L \underset{S}{*} B \,] \;-\; [\, A \underset{S}{*} M \,] \quad \in \quad \breve{\mathcal{o}}_1(S, Z).$$

So, since $h_x^T = [h(x + \bullet)] - [A \underset{S}{*} B]$, we wish to show:

$$[h(x + \bullet)] \;-\; [A \underset{S}{*} B] \;-\; [\, L \underset{S}{*} B \,] \;-\; [\, A \underset{S}{*} M \,] \quad \in \quad \breve{\mathcal{o}}_1(S, Z).$$

Let $\delta := f_x^T - L$ and let $\varepsilon := g_x^T - M$. Since $D_x f = L$, we see that $\delta \in \breve{\mathcal{o}}_1(S, V)$. Since $D_x g = M$, we see that $\varepsilon \in \breve{\mathcal{o}}_1(S, W)$. We have

$$\begin{aligned} \delta \;&=\; f_x^T - L \;=\; [f(x + \bullet)] - A - L \\ \varepsilon \;&=\; g_x^T - M \;=\; [g(x + \bullet)] - B - M, \end{aligned}$$

so $f(x + \bullet) = A + L + \delta$ and $g(x + \bullet) = B + M + \varepsilon$. Then

$$h(x + \bullet) \quad = \quad (A + L + \delta) \underset{S}{*} (B + M + \varepsilon).$$

Expanding this out, and bringing three terms to the LHS, we get

$$\begin{aligned} &[h(x + \bullet)] \;-\; [A \underset{S}{*} B] \;-\; [\, L \underset{S}{*} B \,] \;-\; [\, A \underset{S}{*} M \,] \\ =\quad &[\, A \underset{S}{*} \varepsilon \,] \;+\; [\, L \underset{S}{*} M \,] \;+\; [\, \delta \underset{S}{*} B \,] \\ &+ [\, L \underset{S}{*} \varepsilon \,] \;+\; [\, \delta \underset{S}{*} M \,] \;+\; [\, \delta \underset{S}{*} \varepsilon \,]. \end{aligned}$$

It therefore suffices to show:

$$A \underset{S}{*} \varepsilon,\; L \underset{S}{*} M,\; \delta \underset{S}{*} B,\; L \underset{S}{*} \varepsilon,\; \delta \underset{S}{*} M,\; \delta \underset{S}{*} \varepsilon \;\in\; \breve{\mathcal{o}}_1(S, Z).$$

We have

$$\begin{array}{lcl} A \in C(S, V) \subseteq \widehat{\mathcal{O}}_0(S, V) & \text{and} & B \in C(S, W) \subseteq \widehat{\mathcal{O}}_0(S, W), \\ L \in L(S, V) \subseteq \widehat{\mathcal{O}}_1(S, V) & \text{and} & M \in L(S, W) \subseteq \widehat{\mathcal{O}}_1(S, W), \\ \delta \in \breve{\mathcal{o}}_1(S, V) & \text{and} & \varepsilon \in \breve{\mathcal{o}}_1(S, W). \end{array}$$

Then

$$\begin{aligned}
&A \underset{S}{*} \varepsilon \in [\widehat{\mathcal{O}}_0(S,V)] \underset{S}{*} [\breve{o}_1(S,W)] \subseteq \breve{o}_1(S,Z),\\
&L \underset{S}{*} M \in [\widehat{\mathcal{O}}_1(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}_1(S,W)] \subseteq \widehat{\mathcal{O}}_2(S,Z) \subseteq \breve{o}_1(S,Z),\\
&\delta \underset{S}{*} B \in [\breve{o}_1(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}_0(S,W)] \subseteq \breve{o}_1(S,Z),\\
&L \underset{S}{*} \varepsilon \in [\widehat{\mathcal{O}}_1(S,V)] \underset{S}{*} [\breve{o}_1(S,W)] \subseteq \breve{o}_2(S,Z) \subseteq \breve{o}_1(S,Z) \quad \text{and}\\
&\delta \underset{S}{*} M \in [\breve{o}_1(S,V)] \underset{S}{*} [\widehat{\mathcal{O}}_1(S,W)] \subseteq \breve{o}_2(S,Z) \subseteq \breve{o}_1(S,Z).
\end{aligned}$$

It remains to show: $\delta \underset{S}{*} \varepsilon \in \breve{o}_1(S,Z)$. We have

$$\delta \underset{S}{*} \varepsilon \in [\breve{o}_1(S,V)] \underset{S}{*} [\breve{o}_1(S,W)] \subseteq \breve{o}_2(S,Z) \subseteq \breve{o}_1(S,Z),$$

as desired. □

COROLLARY 53.6. *Let V and W be finite dimensional vector spaces, $x \in V$, $a \in \mathbb{R}$ and $g : V \dashrightarrow W$. Then $D_x(ag) \;=^*\; a \cdot [D_x g]$.*

Proof. Define $* \in B(\mathbb{R}, W, W)$ by $s * w = sw$. Let $f := C_V^a$. Then

$$\begin{aligned}
&f(x) = a, \qquad D_x f \;=\; \mathbf{0}_{VW},\\
&ag \;=\; a \underset{V}{*} g \;=\; f \underset{V}{*} g \qquad \text{and}\\
&a \cdot [D_x g] \;=\; a \underset{V}{*} [D_x g] \;=\; f \underset{V}{*} [D_x g].
\end{aligned}$$

By Theorem 53.5,

$$D_x(f \underset{V}{*} g) \;=^*\; [\,(D_x f) \underset{V}{*} (g(x))\,] \;+\; [\,(f(x)) \underset{V}{*} (D_x g)\,].$$

Then

$$\begin{aligned}
D_x(ag) &= D_x(f \underset{V}{*} g)\\
&=^* [\,(D_x f) \underset{V}{*} (g(x))\,] \;+\; [\,(f(x)) \underset{V}{*} (D_x g)\,]\\
&= [\,(\mathbf{0}_{VW}) \underset{V}{*} (g(x))\,] \;+\; [\,a \underset{V}{*} (D_x g)\,]\\
&= [\,\mathbf{0}_{VW}\,] \;+\; [\,a \cdot (D_x g)\,] \;=\; a \cdot (D_x g),
\end{aligned}$$

as desired. □

LEMMA 53.7. *Let V, W be finite dimensional vector spaces, $x \in V$, $f : V \dashrightarrow W$ and $g : V \dashrightarrow W$. Then $D_x(f+g) \;=^*\; [D_x f] + [D_x g]$.*

Proof. Unassigned HW. □

The next result is the multivariable **Chain Rule**.

THEOREM 53.8. *Let U, V and W be finite dimensional vector spaces. Let $x \in V$, $f : U \dashrightarrow V$ and $g : V \dashrightarrow W$. Let $y := f(x)$. Then $D_x(g \circ f) \;=^*\; (D_y g) \circ (D_x f)$.*

Proof. We wish to prove:

$$[\ (D_y g) \circ (D_x f) \neq ☹\ ] \ \Rightarrow\ [\ D_x(g \circ f) \ = \ (D_y g) \circ (D_x f)\ ].$$

Assume: $(D_y g) \circ (D_x f) \neq ☹$. We want: $D_x(g \circ f) \ = \ (D_y g) \circ (D_x f)$.

Let $h := g \circ f$. Then $h : U \dashrightarrow W$. Let $L := D_x f$ and $M := D_y g$. Then $L \in L(U, V)$ and $M \in L(V, W)$. We want: $D_x h = M \circ L$. It therefore suffices to show: $h_x^T - (M \circ L) \in \breve{o}_1(U, W)$.

Let $\delta := f_x^T - L$ and let $\varepsilon := g_y^T - M$. Since $D_x f = L$, we see that $\delta \in \breve{o}_1(U, V)$. Since $D_y g = M$, we see that $\varepsilon \in \breve{o}_1(V, W)$.

Claim: $h_x^T = (g_y^T) \circ (f_x^T)$. *Proof of Claim:* We want: $\forall u \in U$, $(h_x^T)(u) = ((g_y^T) \circ (f_x^T))(u)$. Let $u \in U$ be given. We want to prove that: $(h_x^T)(u) = ((g_y^T) \circ (f_x^T))(u)$.

Let $v := (f_x^T)(u)$. Adding $y = f(x)$ to $v = [f(x + u)] - [f(x)]$ gives $y + v = f(x + u)$. Then $g(y + v) = g(f(x + u))$. Also, $g(y) = g(f(x))$. So, since $h = g \circ f$, we get $g(y + v) = h(x + u)$ and $g(y) = h(x)$. Then $(g_y^T)(v) = [g(y + v)] - [g(y)] = [h(x + u)] - [h(x)] = (h_x^T)(u)$. Then $(h_x^T)(u) = (g_y^T)(v) = (g_y^T)((f_x^T)(u)) = ((g_y^T) \circ (f_x^T))(u)$, as desired. *End of proof of Claim.*

By definition of δ and ε, we have $f_x^T = L + \delta$ and $g_y^T = M + \varepsilon$. By the claim, $h_x^T = (g_y^T) \circ (f_x^T)$. Then

$$\begin{aligned} h_x^T &= (g_y^T) \circ (f_x^T) \\ &= (M + \varepsilon) \circ (L + \delta) \\ &= (M \circ (L + \delta)) + (\varepsilon \circ (L + \delta)). \end{aligned}$$

Since M is linear, we get $M \circ (L + \delta) = (M \circ L) + (M \circ \delta)$. Then

$$h_x^T \ = \ (M \circ L) \ + \ (M \circ \delta) \ + \ (\varepsilon \circ (L + \delta)),$$

so $h_x^T - (M \circ L) = (M \circ \delta) + (\varepsilon \circ (L + \delta))$, It therefore suffices to show: $M \circ \delta,\ \varepsilon \circ (L + \delta) \ \in \ \breve{o}_1(U, W)$.

We have

$$\begin{array}{lll} L \in L(U, V) \subseteq \widehat{\mathcal{O}}_1(U, V), & & M \in L(V, W) \subseteq \widehat{\mathcal{O}}_1(V, W), \\ \delta \in \breve{o}_1(U, V) & \text{and} & \varepsilon \in \breve{o}_1(V, W). \end{array}$$

Then $M \circ \delta \in [\widehat{\mathcal{O}}_1(V, W)] \circ [\breve{o}_1(U, V)] \subseteq \breve{o}_1(U, W)$. It remains to show: $\varepsilon \circ (L + \delta) \in \breve{o}_1(U, W)$.

We have

$$\begin{aligned} L + \delta &\in [\widehat{\mathcal{O}}_1(U, V)] + [\breve{o}_1(U, V)] \\ &\subseteq [\widehat{\mathcal{O}}_1(U, V)] + [\widehat{\mathcal{O}}_1(U, V)] \ \subseteq \ \widehat{\mathcal{O}}_1(U, V). \end{aligned}$$

Then $\varepsilon \circ (L+\delta) \in [\breve{\mathcal{O}}_1(V,W)] \circ [\widehat{\mathcal{O}}_1(U,V)] \subseteq \breve{\mathcal{O}}_1(U,W)$, as desired. $\square$

Let $\ell, m, n \in \mathbb{N}$, $f : \mathbb{R}^\ell \dashrightarrow \mathbb{R}^m$, $g : \mathbb{R}^m \dashrightarrow \mathbb{R}^n$, $x \in \mathbb{R}^\ell$, $y := f(x)$. By Theorem 53.8, $D_x(g \circ f) \;=^*\; (D_y g) \circ (D_x f)$. Taking matrices,

$$[D_x(g \circ f)] \;=^*\; [D_x f] \cdot [D_y g].$$

Then, replacing y by $f(x)$ and taking transposes,

$$[D_x(g \circ f)]^t \;=^*\; [D_{f(x)} g]^t \cdot [D_x f]^t.$$

This is similar to the one-variable Chain Rule: Let $f : \mathbb{R} \dashrightarrow \mathbb{R}$, $g : \mathbb{R} \dashrightarrow \mathbb{R}$, $x \in \mathrm{LPD}_{\mathbb{R}}(g \circ f)$. Then

$$(g \circ f)'(x) \;=^*\; [g'(f(x))] \cdot [f'(x)].$$

54. Class 28 on 3 May 2018, Th of Week 15

I will continue to hold office hours as usual until the final exam.

DEFINITION 54.1. *Let A and B be sets. Then*

$$A \coprod B \quad := \quad \begin{cases} A \bigcup B, & \text{if } A \bigcap B = \varnothing \\ ☹, & \text{if } A \bigcap B \neq \varnothing. \end{cases}$$

DEFINITION 54.2. *Let $\mathcal{S}$ be a set of sets.*
Then $\mathcal{S}$ is **pairwise-disjoint** *means:*

$$\forall A, B \in \mathcal{S}, \quad (\ [\ A = B\ ] \text{ or } [\ A \bigcap B = \varnothing\ ]\ ).$$

We use "pw-dj" as an abbreviation for "pairwise-disjoint". A pw-dj collection of sets is called a **partition**.

Let X be a set. Let $\mathcal{P}$ be a set of sets. By $\mathcal{P}$ is a **partition of** X, we mean: both ($\mathcal{P}$ is pw-dj) and ($\bigcup \mathcal{P} = X$).

DEFINITION 54.3. *Let $\mathcal{S}$ be a set of sets. Then*

$$\coprod \mathcal{S} \quad := \quad \begin{cases} \bigcup \mathcal{S}, & \text{if } \mathcal{S} \text{ is pairwise-disjoint} \\ ☹, & \text{if } \mathcal{S} \text{ is not pairwise-disjoint.} \end{cases}$$

Let f and g be two functions. Recall: $f \subseteq (\mathrm{dom}[f]) \times (\mathrm{im}[f])$. Similarly, $g \subseteq (\mathrm{dom}[g]) \times (\mathrm{im}[g])$. Note that $g \supseteq f$ is equivalent to:

$$(\ \mathrm{dom}[g] \supseteq \mathrm{dom}[f]\ ) \qquad \text{and} \qquad (\ g|(\mathrm{dom}[f]) = f\ )$$

By g is an **extension** of f, we mean: $g \supseteq f$.

Let f be a sequence of functions. Assume that $f_1 \subseteq f_2 \subseteq \cdots$. Let $g := f_1 \cup f_2 \cup \cdots$. Then g is a function and

- $\mathrm{dom}[g] \;=\; (\mathrm{dom}[f_1]) \;\cup\; (\mathrm{dom}[f_2]) \;\cup\; \cdots$ and

- $\forall j \in \mathbb{N}, \quad g \supseteq f_j$.

DEFINITION 54.4. *A function μ is called a* **set function** *if:*

$$(\text{ } dom[\mu] \text{ is a set of sets }) \qquad \text{and} \qquad (\text{ } im[\mu] \subseteq [0,\infty] \text{ }).$$

Let μ be a set function. Intuitively, we think of μ as assigning a size or "measure" to each set in its domain. The sets in the domain of μ are said to be "μ-measurable", and, for any μ-measurable set A, the "μ-measure" of A is, by definition, $\mu(A)$. By the "μ-total space", we mean: the union of all μ-measurable sets, *i.e.*, $\bigcup \mathrm{dom}[\mu]$.

For example: Let $\mathcal{I} \subseteq 2^{\mathbb{R}}$ denote the set of all intervals in $\mathbb{R}$. Define $\ell : \mathcal{I} \to [0,\infty]$ by $\ell(I) = (\sup I) - (\inf I)$. Then ℓ is a set function; the ℓ-total space is $\bigcup \mathrm{dom}[\ell] = \bigcup \mathcal{I} = \mathbb{R}$. We call ℓ the "length" function.

[The variables $\mathcal{I}$ and ℓ are bound for the remainder of this class.]

Let $\mathcal{I}_2 := \{A \times B \,|\, A, B \in \mathcal{I}\}$. Define $\ell_2 : \mathcal{I}_2 \to [0,\infty]$ by

$$\ell_2(A \times B) \quad = \quad \begin{cases} [\ell(A)] \cdot [\ell(B)], & \text{if } \ell(A) \neq 0 \neq \ell(B) \\ 0, & \text{otherwise.} \end{cases}$$

Then ℓ_2 is a set function; the ℓ_2-total space is $\bigcup \mathrm{dom}[\ell_2] = \bigcup \mathcal{I}_2 = \mathbb{R}^2$. We call ℓ_2 the "area" function.

[The variables $\mathcal{I}_2$, ℓ_2 are bound for the remainder of this class.]

DEFINITION 54.5. *Let μ be a set function. Then μ is* **finitely additive** *means:* $\quad \forall$*finite* $\mathcal{R} \subseteq dom[\mu], \quad \mu\left(\coprod \mathcal{R}\right) \;\; {}^*\!\!= \sum_{A \in \mathcal{R}} \mu(A).$

DEFINITION 54.6. *Let μ be a set function. Then μ is* **countably additive** *means:* $\forall$*countable* $\mathcal{R} \subseteq dom[\mu], \; \mu\left(\coprod \mathcal{R}\right) \;\; {}^*\!\!= \sum_{A \in \mathcal{R}} \mu(A).$

DEFINITION 54.7. *Let μ be a set function. Assume $\bigcup dom[\mu] \subseteq \mathbb{R}$. Then, by μ is* **translation invariant***, we mean: $\forall S \subseteq \mathbb{R}$, $\forall a \in \mathbb{R}$,*

$$\mu(S+a) \quad = \quad \mu(S).$$

We generalize Definition 54.7 to Euclidean spaces of any dimension:

DEFINITION 54.8. *Let μ be a set function and let $d \in \mathbb{N}$. Assume that $\bigcup dom[\mu] \subseteq \mathbb{R}^d$. Then, by μ is* **translation invariant***, we mean: $\forall S \subseteq \mathbb{R}^d$, $\forall a \in \mathbb{R}^d$,* $\qquad \mu(S+a) \quad = \quad \mu(S).$

Recall "area" function ℓ_2. This set function ℓ_2 is translation invariant. At the dawn of measure theory, a basic question was: Is there a "good" way of measuring the area of every subset of $\mathbb{R}^2$? More precisely: Does there exist a countably additive, translation invariant set function $\mu \supseteq \ell_2$ s.t. $\mathrm{dom}[\mu] = 2^{\mathbb{R}^2}$? The answer turned out to be no, but mathematicians eventually found a countably additive, translation invariant set function $\lambda_2 \supseteq \ell_2$ s.t. $\mathrm{dom}[\lambda_2]$ is so large, or "broad", that it has, as elements, all the subsets of $\mathbb{R}^2$ that anyone would ever care about. This extension λ_2 is called "Lebesgue measure on $\mathbb{R}^2$", and we will spend much of this class describing it, as well as its one-dimensional cousin, known as "Lebesgue measure on $\mathbb{R}$".

We first talk about ways in which a collection $\mathcal{S}$ of sets can be made larger (or "broadened"). We already know two:

Let $\mathcal{S}$ be a set of sets. Recall that

$$\begin{aligned} \langle\mathcal{S}\rangle_\cup &:= \{ \bigcup \mathcal{R} \mid \mathcal{R} \subseteq \mathcal{S} \} \qquad \text{and} \\ \langle\mathcal{S}\rangle_\cap &:= \{ \bigcap \mathcal{R} \mid \varnothing \neq \mathcal{R} \subseteq \mathcal{S} \}. \end{aligned}$$

We make several similar definitions (with "pw-dj" = "pairwise-disjoint"):

DEFINITION 54.9. *Let $\mathcal{S}$ be a set of sets. Then*

$$\begin{aligned} \langle\mathcal{S}\rangle_{\mathrm{fin}\,\cup} &:= \{ \bigcup \mathcal{R} \,|\, (\mathcal{R} \subseteq \mathcal{S}) \,\&\, (\mathcal{R} \textit{ is finite}) \}, \\ \langle\mathcal{S}\rangle_{\mathrm{fin}\,\amalg} &:= \{ \coprod \mathcal{R} \,|\, (\mathcal{R} \subseteq \mathcal{S}) \,\&\, (\mathcal{R} \textit{ is finite, pw-dj}) \}, \\ \langle\mathcal{S}\rangle_{\mathrm{fin}\,\cap} &:= \{ \bigcap \mathcal{R} \,|\, (\varnothing \neq \mathcal{R} \subseteq \mathcal{S}) \,\&\, (\mathcal{R} \textit{ is finite}) \}, \\ \langle\mathcal{S}\rangle_{\mathrm{ctbl}\,\cup} &:= \{ \bigcup \mathcal{R} \,|\, (\mathcal{R} \subseteq \mathcal{S}) \,\&\, (\mathcal{R} \textit{ is countable}) \}, \\ \langle\mathcal{S}\rangle_{\mathrm{ctbl}\,\amalg} &:= \{ \coprod \mathcal{R} \,|\, (\mathcal{R} \subseteq \mathcal{S}) \,\&\, (\mathcal{R} \textit{ is countable, pw-dj}) \}, \\ \langle\mathcal{S}\rangle_{\mathrm{ctbl}\,\cap} &:= \{ \bigcap \mathcal{R} \,|\, (\varnothing \neq \mathcal{R} \subseteq \mathcal{S}) \,\&\, (\mathcal{R} \textit{ is countable}) \}. \end{aligned}$$

LEMMA 54.10. *Let $\mathcal{S}$ be a set of sets. Then:*

$$\begin{aligned} &\text{(i)} && [\,(\varnothing \in \mathcal{S}) \,\&\, (\forall A, B \in \mathcal{S},\ A \cup B \in \mathcal{S})\,] \;\Leftrightarrow\; [\,\langle\mathcal{S}\rangle_{\mathrm{fin}\,\cup} \subseteq \mathcal{S}\,] \\ &\&\text{(ii)} && [\,\forall A, B \in \mathcal{S},\ A \cap B \in \mathcal{S}\,] \;\Leftrightarrow\; [\,\langle\mathcal{S}\rangle_{\mathrm{fin}\,\cap} \subseteq \mathcal{S}\,]. \end{aligned}$$

Proof. Unassigned HW. *Hint:* Use Mathematical Induction. □

Certain collections of sets are, by their nature "broad". For example, if a collection of sets is closed under various set-theoretic operations, then it might be considered a broad collection of sets. The most basic set-theoretic operations are: union, intersection and set subtraction. This kind of thinking motivates the next two definitions.

DEFINITION 54.11. *Let $\mathcal{S}$ be a set of sets.*
Then $\mathcal{S}$ is a **ring of sets** *means:*

(1) $\langle\mathcal{S}\rangle_{\mathrm{fin}\,\cup} \subseteq \mathcal{S}$,
(2) $\langle\mathcal{S}\rangle_{\mathrm{fin}\,\cap} \subseteq \mathcal{S}$ *and*

(3) $\forall A, B \in \mathcal{S}, \quad A \backslash B \in \mathcal{S}$.

Let $\mathcal{R}$ be a set of sets. Let $\mathcal{S}$ be a ring of sets. By $\mathcal{R}$ is a **subring** of $\mathcal{S}$, we mean: both ($\mathcal{R} \subseteq \mathcal{S}$) and ($\mathcal{R}$ is a ring of sets).

DEFINITION 54.12. *Let $\mathcal{S}$ be a set of sets.*
Then $\mathcal{S}$ is a **σ-ring of sets** *means:*

(1) $\langle \mathcal{S} \rangle_{\mathrm{ctbl}\,\cup} \subseteq \mathcal{S}$,
(2) $\langle \mathcal{S} \rangle_{\mathrm{ctbl}\,\cap} \subseteq \mathcal{S}$ *and*
(3) $\forall A, B \in \mathcal{S}, \quad A \backslash B \in \mathcal{S}$.

Let $\mathcal{R}$ be a set of sets. Let $\mathcal{S}$ be a σ-ring of sets. By $\mathcal{R}$ is a **σ-subring** of $\mathcal{S}$, we mean: both ($\mathcal{R} \subseteq \mathcal{S}$) and ($\mathcal{R}$ is a σ-ring of sets).

The next lemma is a quantified equivalence for a ring of sets.

LEMMA 54.13. *Let $\mathcal{S}$ be a nonempty set of sets.*
Then: (*$\mathcal{S}$ is a ring of sets*) $\Leftrightarrow$
(*$\forall A, B \in \mathcal{S}$, both ($A \amalg B \;{}^*\!\!\in \mathcal{S}$) and ($A \backslash B \in \mathcal{S}$)*).

Proof. Proof of $\Rightarrow$: Unassigned HW. *End of proof of $\Rightarrow$.*

Proof of $\Leftarrow$: Assume

$\forall A, B \in \mathcal{S}$, both ($A \amalg B \;{}^*\!\!\in \mathcal{S}$) and ($A \backslash B \in \mathcal{S}$).

Want: $\mathcal{S}$ is a ring of sets. We know:

(i) $\forall A, B \in \mathcal{S}$, $A \amalg B \;{}^*\!\!\in \mathcal{S}$ and
(ii) $\forall A, B \in \mathcal{S}$, $A \backslash B \in \mathcal{S}$.

By Definition 54.11, we wish to verify:

(1) $\langle \mathcal{S} \rangle_{\mathrm{fin}\,\cup} \subseteq \mathcal{S}$,
(2) $\langle \mathcal{S} \rangle_{\mathrm{fin}\,\cap} \subseteq \mathcal{S}$ and
(3) $\forall A, B \in \mathcal{S}, \quad A \backslash B \in \mathcal{S}$.

Since $\mathcal{S} \neq \varnothing$, choose $Z \in \mathcal{S}$. Then, by (ii), $Z \backslash Z \in \mathcal{S}$. That is, $\varnothing \in \mathcal{S}$. For any two sets A and B, $A \cup B = A \amalg (B \backslash A)$. Thus, from (i) and (ii), we see: $\forall A, B \in \mathcal{S}$, we have $A \cup B \in \mathcal{S}$. Then, by (i) of Lemma 54.10, (1) holds. By (ii), (3) holds. It remains to show that (2) holds.

For any two sets A and B, we have $A \backslash (A \backslash B) = A \cap B$. Thus, from (ii), we see: $\forall A, B \in \mathcal{S}$, we have $A \cap B \in \mathcal{S}$. Then, by (ii) of Lemma 54.10, we see that (2) holds. *End of proof of $\Leftarrow$.* $\square$

Let $\mathcal{B} := \langle \mathcal{I} \rangle_\sigma$ and $\mathcal{B}_2 := \langle \mathcal{I}_2 \rangle_\sigma$. Elements of $\mathcal{B}$ are called **Borel sets** in $\mathbb{R}$, and elements of $\mathcal{B}_2$ are called **Borel sets** in $\mathbb{R}^2$.

[The variables $\mathcal{B}$, $\mathcal{B}_2$ are bound for the remainder of this class.]

We seek to extend the length function, $\ell : \mathcal{I} \to [0, \infty]$, to a collection of sets so broad that it has, as elements, "all the subsets of $\mathbb{R}$ that anyone would ever care about". It turns out that $\mathcal{B}$ is such a collection, but the broadness required isn't described in precise mathematical terms, so we can't prove it with mathematical rigor. However, I would issue the following challenge: Try to find a subset of $\mathbb{R}$ that is *not* an element of $\mathcal{B}$. Similarly, see if you can find a subset of $\mathbb{R}^2$ that is *not* an element of $\mathcal{B}_2$. It turns out that, if we accept the Axiom of Choice, then it's possible to prove that such subsets of $\mathbb{R}$ and of $\mathbb{R}^2$ do exist, but good luck describing any specific one. In fact, from the perspective of any conventional applications, if we are able to extend ℓ and ℓ_2 to countably additive, translation invariant functions $\lambda : \mathcal{B} \to [0, \infty]$ and $\lambda_2 : \mathcal{B}_2 \to [0, \infty]$, then we deserve to celebrate . . . responsibly!

Recall that $\mathcal{B} := \langle \mathcal{I} \rangle_\sigma$, *i.e.*, $\mathcal{B}$ is obtained by closing $\mathcal{I}$ under countable union, countable intersection and set differences. When you close a collection $\mathcal{S}$ of sets under, say, countable union, the definition is straightforward: Form all possible countable unions of sets in $\mathcal{S}$; this gives, a new collection $\mathcal{T} := \langle \mathcal{S} \rangle_{\mathrm{ctbl}\,\cup}$. However $\mathcal{T}$ may not be closed under, say, set differences. Next, we might close $\mathcal{T}$ under set differences, yielding $\mathcal{U} := \{A \backslash B \,|\, A, B \in \mathcal{T}\}$. Unfortunately, this last collection $\mathcal{U}$ may *not* be closed under countable union. So, we might then look at $\langle \mathcal{U} \rangle_{\mathrm{ctbl}\,\cup}$, which *is* closed under countable union, but may not be closed under set differences or countable intersection. And so on.

If we start with $\mathcal{I}$ and repeat various closure operations repeatedly, we eventually end up with $\mathcal{B}$, the smallest σ-ring containing $\mathcal{I}$. This way of thinking of the $\mathcal{B}$ is very intuitive, but it turns out that, to implement it, we would need to perform uncountably many of these closure operations, and so we would need a version of recursion called "transfinite" recursion. To avoid this, we opt for a different approach:

Note: $2^{\mathbb{R}}$ is a σ-ring of sets that contains $\mathcal{I}$. Find *every* σ-subring of $2^{\mathbb{R}}$ that contains $\mathcal{I}$, and then form the set $\mathcal{B}$ consisting of those subsets A of $\mathbb{R}$ s.t. A is an element of each of those σ-subrings. It turns out that that collection of sets A is the smallest σ-ring containing $\mathcal{I}$. We now implement that idea, first for rings (see Definition 54.14), then for σ-rings (see Definition 54.15).

DEFINITION 54.14. *Let $\mathcal{S}$ be a set of sets and let $X := \bigcup \mathcal{S}$. Then*

$$r(\mathcal{S}) \;\; := \;\; \{\; \mathcal{T} \subseteq 2^X \;|\; (\mathcal{T} \supseteq \mathcal{S}) \;\&\; (\mathcal{T} \textit{ is a ring of sets}) \;\},$$

$$\langle \mathcal{S} \rangle_r \ := \ \bigcap [r(\mathcal{S})].$$

In Definition 54.14, $\mathcal{S}$ is a set of sets, so $r(\mathcal{S})$ is a set of sets of sets.

DEFINITION 54.15. *Let $\mathcal{S}$ be a set of sets and let $X := \bigcup \mathcal{S}$. Then*

$$\begin{aligned} \sigma(\mathcal{S}) \ &:= \ \{ \ \mathcal{T} \subseteq 2^X \mid (\mathcal{T} \supseteq \mathcal{S}) \ \& \ (\mathcal{T} \textit{ is a } \sigma\textit{-ring of sets}) \ \}, \\ \langle \mathcal{S} \rangle_\sigma \ &:= \ \bigcap [\sigma(\mathcal{S})]. \end{aligned}$$

In Definition 54.15, $\mathcal{S}$ is a set of sets, so $\sigma(\mathcal{S})$ is a set of sets of sets.

Let $\mathcal{S}$ be a set of sets. The process of computing $\langle \mathcal{S} \rangle_\sigma$ can be difficult: We might use transfinite recursion, but that's beyond the scope of this course, and not easy in any case. The alternative is to try identify $\sigma(\mathcal{S})$, which is also difficult. Note that, according to Lemma 55.2 below, we only need *FOUR* steps to go from $\mathcal{S}$ to $\langle \mathcal{S} \rangle_r$. By contrast, going from $\mathcal{S}$ to $\langle \mathcal{S} \rangle_\sigma$ can take uncountably many steps. So understanding $\langle \mathcal{S} \rangle_r$ is much easier, and we discuss that first.

It helps to weaken slightly the definition of "ring of sets" obtaining what we will call a "pre-ring of sets":

DEFINITION 54.16. *Let $\mathcal{S}$ be a set of sets.*
Then $\mathcal{S}$ is a **pre-ring of sets** *means:* $\quad \forall A, B \in \mathcal{S},\ A \backslash B \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}$.

Note that any partition is a pre-ring of sets.

Unassigned HW: Show that, while the set $\mathcal{I}$ of intervals is not a ring of sets, it *is* a pre-ring of sets. Show: $\mathcal{I}_2$ is also a pre-ring of sets.

For a pre-ring $\mathcal{S}$ of sets, computing $\langle \mathcal{S} \rangle_r$ is made relatively easy by:

LEMMA 54.17. *Let $\mathcal{S}$ be a pre-ring of sets. Then $\langle \mathcal{S} \rangle_r = \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}$.*

Proof. Unassigned HW. *Hint:* Show, in order, First, show

$$\forall A \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}, \forall B \in \mathcal{S}, \quad A \backslash B \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}.$$

Then, by induction, prove:

$$\forall n \in \mathbb{N}, [\, \forall A \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}, \forall B_1, \ldots, B_n \in \mathcal{S}, \quad A \backslash (B_1 \cup \cdots \cup) B_n \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}].$$

Then prove

$$\forall A \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}, \forall B \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}, \quad A \backslash B \in \langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}.$$

Then, using Lemma 54.13, show that $\langle \mathcal{S} \rangle_{\mathrm{fin}\amalg}$ is a ring of sets. $\square$

Let $\mathcal{A} := \langle \mathcal{I} \rangle_r$. An example of a set in $\mathcal{A} \backslash \mathcal{I}$ is

$$[1, 2] \ \cup \ (3, 4) \ \cup \ [5, 7) \ \cup \ (20, 100].$$

Since $\mathcal{I}$ is a pre-ring of sets, by Lemma 54.17, we have $\mathcal{A} = \langle\mathcal{I}\rangle_{\mathrm{fin}\amalg}$. That is to say: a set is in $\mathcal{A}$ iff it is a finite disjoint union of intervals (like the set above). Think of elements of $\mathcal{I}$ as "Lego blocks" in one dimension, and think of the elements of $\mathcal{A}$ as sets that we can assemble from these blocks, in finite time. (The phrase "in finite time" is our colloquial way of indicating that each set in $\mathcal{A}$ is a *finite* union of sets in $\mathcal{I}$. Eventually, we'll look at constructions that can be accomplished "in countable time", *i.e.*, by countable unions.) Since we know the size (*i.e.*, the length) of each block, there's only one reasonable way to assign a size to each such assembly. Thus, extending ℓ to a finitely additive $\kappa : \mathcal{A} \to [0, \infty]$ should be straightforward.

[The variable $\mathcal{A}$ is bound for the remainder of this class.]

There are many subsets of $\mathbb{R}$ that we might care about that are *NOT* in $\mathcal{A}$. For example, let

$$\begin{aligned} Z \quad := \quad & [1, 1 + (1/2)] \ \cup \ [2, 2 + (1/4)] \ \cup \\ & \qquad [3, 3 + (1/8)] \ \cup \ [4, 4 + (1/16)] \ \cup \ \cdots . \end{aligned}$$

Then Z is a *countable* union of intervals, but is not equal to any finite union of intervals. Then $Z \notin \langle\mathcal{I}\rangle_{\mathrm{fin}\amalg} = \mathcal{A}$. Nevertheless, whatever we might mean by measure, the measure of Z should be

$$(1/2) \quad + \quad (1/4) \quad + \quad (1/8) \quad + \quad (1/16) \quad + \quad \cdots ,$$

which is equal to 1. So, if our goal is to extend ℓ to a broad collection of sets, we would hope that Z would be an element of that collection. While $\mathcal{A}$ is insufficient, note that $Z \in \langle\mathcal{I}\rangle_{\mathrm{ctbl}\amalg} \subseteq \langle\mathcal{I}\rangle_\sigma = \mathcal{B}$.

We can also do "Legos in two dimensions": Let $\mathcal{A}_2 := \langle\mathcal{I}_2\rangle_r$. Then, by Lemma 54.17, $\mathcal{A}_2 = \langle\mathcal{I}_2\rangle_{\mathrm{fin}\amalg}$. Our "Lego blocks" are now rectangles, and every element of $\mathcal{A}_2$ is a finite-time assembly of these two dimensional blocks. However, if you want to assemble a good approximation to something round, like a disk, then that finite amount of time may be very large. (In three dimensions, it gets even worse: my son worked on his Lego Death Star for weeks.) Since we know the size (*i.e.*, the area) of each block, there's only one reasonable way to assign a size to each such assembly. Thus, extending ℓ_2 to a finitely additive, translation invariant $\kappa_2 : \mathcal{A}_2 \to [0, \infty]$ should be straightforward.

[The variable $\mathcal{A}_2$ is bound for the remainder of this class.]

In two dimensions, we have $\mathcal{A}_2 = \langle\mathcal{I}_2\rangle_{\mathrm{fin}\amalg}$. This collection of sets is even more obviously *NOT* broad enough to include, as elements, "all the subsets of $\mathbb{R}^2$ that anyone would ever care about". For example, a disk is not a finite union of rectangles, and, if we cannot develop a formula for the area of a disk, then we will not earn the respect of the ancient Greeks, who are, even now, the Guardians of Mathematics.

By contrast, when we look at σ-rings like $\mathcal{B}$ and $\mathcal{B}_2$, the problem goes away: As we mentioned earlier, any subset of $\mathbb{R}$ that anyone cares about is an element of $\mathcal{B}$, and any subset of $\mathbb{R}^2$ that anyone cares about is an element of $\mathcal{B}_2$. In this sense, $\mathcal{B}$ and $\mathcal{B}_2$ are both very broad. If we are somehow able to extend ℓ and ℓ_2 to countably additive, translation invariant functions $\lambda : \mathcal{B} \to [0, \infty]$ and $\lambda_2 : \mathcal{B}_2 \to [0, \infty]$ then we deserve to celebrate . . . responsibly!

DEFINITION 54.18. *Let $\mathcal{P}$ and $\mathcal{Q}$ be partitions. Then $\mathcal{P} \lessdot \mathcal{Q}$ means: $\mathcal{P} \subseteq \langle \mathcal{Q} \rangle_{\text{fin}\sqcup}$.*

To indicate that $\mathcal{P} \lessdot \mathcal{Q}$, we say that $\mathcal{Q}$ is a **finite refinement** of $\mathcal{P}$.

Intuitively, if $\mathcal{P}$ is a box of Legos, and $\mathcal{Q}$ is another box of Legos, then $\mathcal{P} \lessdot \mathcal{Q}$ means that each Lego in $\mathcal{P}$ can be build out of Legos from $\mathcal{Q}$, in finite time. Consequently, if some assembly can be built with $\mathcal{P}$-Legos in finite time, then it can also be built with $\mathcal{Q}$-Legos in a finite (but typically longer) period of time. Therefore:

LEMMA 54.19. *Let $\mathcal{P}$ and $\mathcal{Q}$ be partitions. Assume that $\mathcal{P} \lessdot \mathcal{Q}$. Then we have $\langle \mathcal{P} \rangle_r \subseteq \langle \mathcal{Q} \rangle_r$.*

Proof. Unassigned HW. Hint: Think Legos. □

DEFINITION 54.20. *Let μ be a set function and let $\mathcal{Q} := dom[\mu]$. Let $\mathcal{R} := \langle \mathcal{Q} \rangle_r$. Then* FAES μ *is the set of all $\nu : \mathcal{R} \to [0, \infty]$ such that*

$$\text{both} \quad (\ \nu \text{ is finitely additive }) \quad \text{and} \quad (\ \nu|\mathcal{Q} = \mu\ ).$$

In Definition 54.20, "FAES" = "Finitely Additive ExtensionS".

DEFINITION 54.21. *Let μ be a set function. By μ is a* **partition measure***, we mean: dom$[\mu]$ is a partition.*

DEFINITION 54.22. *Let μ be a set function, X a set. By μ is a* **partition measure on** *X, we mean: dom$[\mu]$ is a partition of X.*

Our basic result about FAES is:

THEOREM 54.23. *Let μ be a partition measure.*
Then $\#(\text{FAES}\,\mu) = 1$.

Proof. Unassigned HW. Hint: Think Legos. $\square$

DEFINITION 54.24. *Let μ be a partition measure. We define $\mu^* := \mathrm{ELT}(\mathrm{FAES}\,\mu)$.*

DEFINITION 54.25. *Let μ and ν be partition measures. Define $\mathcal{P} := dom[\mu]$ and $\mathcal{Q} := dom[\nu]$. Then $\mu \lessdot \nu$ means:*

$$[\,\mathcal{P} \lessdot \mathcal{Q}\,] \quad \textit{and} \quad \left[\forall P \in \mathcal{P}, \quad \mu(P) = \sum_{\substack{Q \in \mathcal{Q} \\ Q \subseteq P}} \nu(Q) \right].$$

To indicate that $\mu \lessdot \nu$, we say that ν is a **finite refinement** of μ.

Intuitively, $\mathcal{P}$ is a box of Legos, and μ tells us the size of each of the $\mathcal{P}$-Legos. Intuitively, $\mathcal{Q}$ is another box of Legos, and ν tells us the size of each of the $\mathcal{Q}$-Legos. The condition $\mathcal{P} \lessdot \mathcal{Q}$ tells us that each $\mathcal{P}$-Lego can be built out of $\mathcal{Q}$-Legos, in finite time. The condition

$$\forall P \in \mathcal{P}, \quad \mu(P) = \sum_{\substack{Q \in \mathcal{Q} \\ Q \subseteq P}} \nu(Q)$$

tells us: the size of any $\mathcal{P}$-Lego is the same as the size of the $\mathcal{Q}$-Legos needed to build it. These two conditions together imply that,

if we build an assembly out of $\mathcal{P}$-Legos,
and measure its total size,
and then rebuild it out of $\mathcal{Q}$-Legos,
and measure its total size once more,

then the two total sizes are the same. That is:

LEMMA 54.26. *Let μ and ν be partition measures. Assume $\mu \lessdot \nu$. Then $\mu^* \subseteq \nu^*$.*

Proof. Unassigned HW. Hint: Think Legos. $\square$

The inclusion $\mu^* \subseteq \nu^*$ says that the set function $\nu^* : \mathcal{Q}^* \to [0, \infty]$ is an extension of the set function $\mu^* : \mathcal{P}^* \to [0, \infty]$.

Let μ and ν be partition measures. Note that

$$(\ \mu \lessdot \nu\ ) \quad \Rightarrow \quad \left(\bigcup \mathrm{dom}[\mu] \;=\; \bigcup \mathrm{dom}[\nu] \right).$$

That is, if $\mu \lessdot \nu$, then the μ-total space is equal to the ν-total space.

DEFINITION 54.27. *Let μ be a sequence of set functions. By $\mu_\bullet$ is a* **partition measure scheme***, we mean:*

(1) $\forall j \in \mathbb{N}$, *μ_j is a partition measure.*
(2) $\forall j \in \mathbb{N}$, $\mu_j \lessdot \mu_{j+1}$.

Let μ be a partition measure scheme. Note that

$$\bigcup \operatorname{dom}[\mu_1] \quad = \quad \bigcup \operatorname{dom}[\mu_2] \quad = \quad \bigcup \operatorname{dom}[\mu_3] \quad = \quad \cdots.$$

That is, $\forall j \in \mathbb{N}$, μ_j-total space is equal to the μ_1-total space.

DEFINITION 54.28. *Let μ be a sequence of set functions and let X be a set. By $\mu_\bullet$ is a* **partition measure scheme on** X*, we mean:*

(1) *$\mu_\bullet$ is a partition measure scheme,* *and*
(2) $\bigcup$ *dom*$[\mu_1] \;=\; X$.

Let μ be a partition measure scheme. Then $\mu_1 \lessdot \mu_2 \lessdot \mu_3 \lessdot \cdots$.

Intuitively, we have countably many boxes of Legos. For all $j \in \mathbb{N}$, the jth box contains the sets in $\operatorname{dom}[\mu_j]$. For all $j \in \mathbb{N}$, μ_j tells us the size of the Legos in the jth box. The condition $\mu_1^* \subseteq \mu_2^* \subseteq \mu_3^* \subseteq \cdots$ ensures: for all $j, k \in \mathbb{N}$, if $j \leqslant k$,

if we build an assembly out of Legos from the jth box,
and measure its total size,
and then rebuild it out of Legos from the kth box,
and measure its total size once more,

then the two total sizes are the same. Or, to put it another way: By Lemma 54.26, we have $\mu_1^* \subseteq \mu_2^* \subseteq \mu_3^* \subseteq \cdots$.

DEFINITION 54.29. *Let μ be a partition measure scheme. Then we define* $\mu_\bullet^* \;:=\; \mu_1^* \bigcup \mu_2^* \bigcup \mu_3^* \bigcup \cdots$.

Intuitively, we take all the Legos from all the boxes, throw them all into one big box. The function $\mu_\bullet^*$ tells us the size of any finite-time assembly that can be built from Legos from that big box.

For example: Define a sequence $\mathcal{D}$ of partitions of $\mathbb{R}$ by

$$\mathcal{D}_j \quad = \quad \{\, (a, a + 2^{1-j}] \mid a \in 2^{j-1}\mathbb{Z} \,\}.$$

Then $\mathcal{D}_1$ is called the "first dyadic partition" of $\mathbb{R}$, and is equal to

$$\{\, \ldots,\ (-4,-3],\ (-3,-2],\ (-2,-1],\ (-1,0],\ (0,1],\ (1,2],\ (2,3],\ (3,4],\ \ldots \,\}.$$

Also, $\mathcal{D}_2$ is called the "second dyadic partition" of $\mathbb{R}$, and is equal to

$$\{\, \ldots,\ (-4/2,-3/2],\ (-3/2,-2/2],\ (-2/2,-1/2],\ (-1/2,0/2],\ (0/2,1/2],\ (1/2,2/2],\ (2/2,3/2],\ (3/2,4/2],\ \ldots \,\}.$$

Also, $\mathcal{D}_3$ is called the "third dyadic partition" of $\mathbb{R}$, and is equal to

$$\{\ \ldots,\ (-4/4,-3/4],\ (-3/4,-2/4],\ (-2/4,-1/4],\ (-1/4,0/4],\\ (0/4,1/4],\ (1/4,2/4],\ (2/4,3/4],\ (3/4,4/4],\ \ldots\ \}.$$

Also, $\mathcal{D}_4$ is called the "fourth dyadic partition" of $\mathbb{R}$, and is equal to

$$\{\ \ldots,\ (-4/8,-3/8],\ (-3/8,-2/8],\ (-2/8,-1/8],\ (-1/8,0/8],\\ (0/8,1/8],\ (1/8,2/8],\ (2/8,3/8],\ (3/8,4/8],\ \ldots\ \}.$$

Note: $\mathcal{D}_1 \lessdot \mathcal{D}_2 \lessdot \mathcal{D}_3 \lessdot \mathcal{D}_4 \lessdot \cdots$. Define a sequence m of partition measures by $m_j = \ell|\mathcal{D}_j$. Note: $m_1 \lessdot m_2 \lessdot m_3 \lessdot m_4 \lessdot \cdots$. We conclude that $m_\bullet$ is a partition measure scheme.

[The sequences $\mathcal{D}_\bullet$ and $m_\bullet$ are bound for the remainder of this class.]

For another example: Define a sequence $\mathcal{D}^2$ of partitions of $\mathbb{R}^2$ by

$$\mathcal{D}_j^2 \quad = \quad \{\ A \times B \mid A, B \in \mathcal{D}_j\ \}.$$

Define a sequence m^2 of partition measures by $m_j^2 = \ell_2|\mathcal{D}_j^2$. Note: $m_1^2 \lessdot m_2^2 \lessdot m_3^2 \lessdot m_4^2 \lessdot \cdots$. We conclude that $m_\bullet^2$ is a partition measure scheme.

[The sequences $\mathcal{D}_\bullet^2$ and $m_\bullet^2$ are bound for the remainder of this class.]

Intuitively, for all $j \in \mathbb{N}$, $\mathcal{D}_j^2$ is a box of two-dimensional Legos, each one a square (containing part, but not all, of its boundary). Each box is a refinement of the preceding box, and so can build more assemblies than the preceding box. We now take all the Legos from all these boxes and put them into one big box. The function $(m_\bullet^2)^*$ tells us the total size of any assembly that can be built, in finite time, from that big box.

Imagine we want to build the open disk of radius $\sqrt{2}$ about $(0,0)$. We begin by choosing four Legos from $\mathcal{D}_1$ that cover the inscribed square $(-1,1] \times (-1,1]$. Then, with smaller Legos, we partially fill in the part of the disk not already not covered by those first four Legos. We continue, and get better and better approximations to the disk, but, unfortunately, in finite time, we can't exactly cover the disk.

We now switch from finite time to countable time. In *countable* time, we can exactly cover our disk, and then compute its area as the sum of the areas of the countably many Legos used to cover it.

DEFINITION 54.30. *Let μ be a set function and let $\mathcal{Q} := dom[\mu]$. Let $\mathcal{S} := \langle \mathcal{Q} \rangle_\sigma$. Then* CAES μ *is the set of all $\nu : \mathcal{S} \to [0,\infty]$ such that*

$$\textit{both} \quad (\ \nu \textit{ is countably additive }) \quad \textit{and} \quad (\ \nu|\mathcal{Q} = \mu\ ).$$

In Definition 54.30, "CAES" = "Countably Additive ExtensionS".

DEFINITION 54.31. *Let ν be a set function. Then ν is a* **measure** *means: ($dom[\nu]$ is a σ-ring) and (ν is countably additive).*

Let μ be a set function. Then, $\forall \nu \in \mathrm{CAES}\,\mu$, ν is a measure.

DEFINITION 54.32. *Let μ be a partition measure. Then μ is* **σ-finite** *means: both ($dom[\mu]$ is countable) and ($\forall A \in dom[\mu]$, $mu(A) < \infty$).*

You are responsible for knowing the statement of Theorem 54.33 below, but not for knowing its proof, which was not covered.

THEOREM 54.33. *Let μ be a σ-finite partition measure scheme. Then $\#(\mathrm{CAES}\,\mu_\bullet^*) \leqslant 1$.*

Proof. TO BE ADDED LATER ☐

DEFINITION 54.34. *Let X be a topological space and let μ be a partition measure scheme on X. Then $\mu_\bullet$ is a* **pre-measure** *on X means: μ_1 is regular and σ-finite.*

Unassigned HW: Show both that $m_\bullet$ is a pre-measure on $\mathbb{R}$ and that $m_\bullet^2$ is a pre-measure on $\mathbb{R}^2$.

DEFINITION 54.35. *Let X be a topological space and let μ be a partition measure scheme on X. Then*

(1) *$\mu_\bullet$ is* **inner regular** *means: $\forall j \in \mathbb{N}$, $\forall B \in dom[\mu_j]$, $\forall \varepsilon > 0$, $\exists A \in dom[\mu_\bullet^*]$ s.t. $[\ \mathrm{Cl}_X A \subseteq B\ ]$ and $[\ (\mu_\bullet^*)(B \backslash A) < \varepsilon\ ]$,*

(2) *$\mu_\bullet$ is* **outer regular** *means: $\forall j \in \mathbb{N}$, $\forall B \in dom[\mu_j]$, $\forall \varepsilon > 0$, $\exists C \in dom[\mu_\bullet^*]$ s.t. $[\ B \subseteq \mathrm{Int}_X C\ ]$ and $[\ (\mu_\bullet^*)(C \backslash B) < \varepsilon\ ]$,*

&(3) *$\mu_\bullet$ is* **regular** *means: (μ is both inner regular and outer regular) and ($\forall A \in dom[\mu_1]$, $\mathrm{Cl}_X A$ is sequentially compact).*

You are responsible for knowing the statement of Theorem 54.36 below, but not for knowing its proof, which was not covered.

THEOREM 54.36. *Let X be a topological space. Let μ be a pre-measure on X. Then $\#(\mathrm{CAES}\,\mu_\bullet^*) = 1$.*

Proof. TO BE ADDED LATER ☐

DEFINITION 54.37. *Let X be a topological space, and let μ be a pre-measure on X. Then $\widehat{\mu_\bullet} := \mathrm{ELT}(\mathrm{CAES}\,\mu_\bullet^*)$.*

Define $\lambda := \widehat{m_\bullet}$ and $\lambda_2 := \widehat{m_\bullet^2}$. Then λ is called **Lebesgue measure** on $\mathbb{R}$, and λ_2 is called **Lebesgue measure** on $\mathbb{R}^2$.

[The variables λ and λ_2 are bound for the remainder of this class.]

You are responsible for knowing the statement of Theorem 54.38 below, but not for knowing its proof, which was not covered.

THEOREM 54.38. *We have:*

(1) $dom[\lambda] = \mathcal{B}$,
(2) $\lambda|\mathcal{I} = \ell$,
(3) $dom[\lambda_2] = \mathcal{B}_2$ *and*
(2) $\lambda_2|\mathcal{I}_2 = \ell_2$.

Proof. TO BE ADDED LATER □

DEFINITION 54.39. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$. *Then*

$$\mathcal{O}_f \quad := \quad \{\ (x, y) \in \mathbb{R}^2 \mid (\ x \in dom[f]\ ) \,\&\, (\ 0 < y < f(x)\ )\ \}.$$

In Definition 54.39, the set $\mathcal{O}_f$ is called the **ordinate set** of f.

DEFINITION 54.40. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$. *Then*

$$\int f \quad := \quad [\lambda_2(\mathcal{O}_f)] \ - \ [\lambda_2(\mathcal{O}_{-f})].$$

DEFINITION 54.41. *Define* $\operatorname{sgn} : \mathbb{R} \to \{-1, 0, 1\}$ *by*

$$\operatorname{sgn} x \quad = \quad \begin{cases} -1, & \text{if } x < 0 \\ 0, & \text{if } x = 0 \\ 1, & \text{if } x > 0. \end{cases}$$

DEFINITION 54.42. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$ *and let* $a, b \in \mathbb{R}$. *Then*

$$\int_a^b f \quad := \quad [\operatorname{sgn}(b - a)] \cdot \left[\int \left(\left(\chi_{[a|b]}^{\mathbb{R}} \right) \cdot f \right) \right].$$

The next result is the **Fundamental Theorem of Calculus**. You are responsible for knowing its statement, but not for knowing its proof, which was not covered.

THEOREM 54.43. *Let* $f : \mathbb{R} \dashrightarrow \mathbb{R}$ *and let* $a, b \in \mathbb{R}$. *Assume that* $a < b$ *and that* f *is continuous on* $[a, b]$. *Define* $g : \mathbb{R} \dashrightarrow \mathbb{R}$ *by* $g(x) = \int_a^x f$. *Then: (* g *is c/d on* $[a, b]$ *) and (* $g' = f$ *on* (a, b) *).*

Proof. TO BE ADDED LATER □

55. Appendix: Generating a ring of sets (NOT ON FINAL)

This section will NOT be covered on the final.

LEMMA 55.1. *Let $\mathcal{C}$ be a set of sets. Assume both*

$$\langle\mathcal{C}\rangle_{\mathrm{fin}\,\cap} \subseteq \mathcal{C} \qquad \textit{and} \qquad \langle\mathcal{C}\rangle_{\mathrm{fin}\,\cup} \subseteq \mathcal{C}.$$

Let $\mathcal{D} := \{C\backslash C_0 \,|\, C, C_0 \in \mathcal{C}\}$. Then $\mathcal{D}$ is a pre-ring of sets.

Proof. By Definition 54.16, we want: $\forall D, D' \in \mathcal{D}$, $D\backslash D' \in \langle\mathcal{D}\rangle_{\mathrm{fin}\amalg}$. Let $D, D' \in \mathcal{D}$ be given. We want: $D\backslash D' \in \langle\mathcal{D}\rangle_{\mathrm{fin}\amalg}$.

Since $D, D' \in \mathcal{D}$, choose $C, C_0, C', C_0' \in \mathcal{C}$ such that

$$\text{both} \quad [\ D = C\backslash C_0\ ] \quad \text{and} \quad [\ D' = C'\backslash C_0'\ ].$$

Then $C_0 \cup C' \in \langle\mathcal{C}\rangle_{\cup} \subseteq \mathcal{C}$ and $C \cap C' \cap C_0' \in \langle\mathcal{C}\rangle_{\cap} \subseteq \mathcal{C}$. Let

$$\begin{aligned} X &:= C\backslash(C_0 \cup C') \qquad \text{and} \\ Y &:= (C \cap C' \cap C_0')\backslash C_0. \end{aligned}$$

Then $X, Y \in \mathcal{D}$. Then $X \amalg Y \overset{*}{\in} \langle\mathcal{D}\rangle_{\mathrm{fin}\amalg}$. It therefore suffices to show that $D\backslash D' = X \coprod Y$. We therefore wish to show:

(1) $X \cap Y = \varnothing$,
(2) $D\backslash D' \subseteq X \cup Y$, and
(3) $D\backslash D' \supseteq X \cup Y$.

Proof of (1): Since $X \subseteq C\backslash C'$ and $Y \subseteq C'$, we conclude that $X \cap Y = \varnothing$, as desired. *End of proof of (1).*

Proof of (2): We wish to show: $\forall q \in D\backslash D'$, $q \in X \cup Y$. Let $q \in D\backslash D'$ be given. We wish to show: $q \in X \cup Y$. It suffices to show: $(q \notin X) \Rightarrow (q \in Y)$. Assume $q \notin X$. Want: $q \in Y$. We want to prove: $q \in C$ and $q \in C'$ and $q \in C_0'$ and $q \notin C_0$.

Since $q \in D = C\backslash C_0$, we get $q \in C$ and $q \notin C_0$. Since $q \in C$ and $q \notin X = C\backslash(C_0 \cup C')$, we see that $q \in C_0 \cup C'$. So, as $q \notin C_0$, we see that $q \in C'$ It remains to show: $q \in C_0'$.

Since $q \in D\backslash D'$, we have $q \notin D'$. So, since $D' = C'\backslash C_0'$ and $q \in C'$, it follows that $q \in C_0'$, as desired. *End of proof of (2).*

Proof of (3): We have $X = C\backslash(C_0 \cup C') \subseteq C\backslash C_0 = D$. So, since $X \cap D' \subseteq (C\backslash C') \cap C' = \varnothing$, we get $X \subseteq D\backslash D'$. It remains to show: $Y \subseteq D\backslash D'$.

We have $Y = (C \cap C' \cap C_0')\backslash C_0 \subseteq C\backslash C_0 = D$. So, since $Y \cap D' \subseteq C_0' \cap (C'\backslash C_0') = \varnothing$, we get $Y \subseteq D\backslash D'$, as desired. *End of proof of (3).*

□

The next lemma shows that, given a set $\mathcal{A}$ of sets, we can, in FOUR steps, generate a ring of sets. This is in contrast to σ-rings, where the generation might take uncountably many steps, and therefore might require transfinite recursion to implement.

LEMMA 55.2. *Let $\mathcal{A}$ be a set of sets, $\mathcal{B} := \langle \mathcal{A} \rangle_{\text{fin}\,\cap}$, $\mathcal{C} := \langle \mathcal{B} \rangle_{\text{fin}\,\cup}$ and $\mathcal{D} := \{C \backslash C_0 \,|\, C, C_0 \in \mathcal{C}\}$ and $\mathcal{E} := \langle \mathcal{D} \rangle_{\text{fin}\,\amalg}$. Then $\mathcal{E} = \langle \mathcal{A} \rangle_r$.*

Proof. Since $\langle \mathcal{A} \rangle_r = \bigcap [r(\mathcal{A})]$, we wish to show: $\mathcal{E} = \bigcap [r(\mathcal{A})]$.

For any $\mathcal{R} \in r(\mathcal{A})$, $\mathcal{A} \subseteq \mathcal{R}$, so $\mathcal{B} \subseteq \mathcal{R}$, so $\mathcal{C} \subseteq \mathcal{R}$, so $\mathcal{D} \subseteq \mathcal{R}$, so $\mathcal{E} \subseteq \mathcal{R}$. Then $\mathcal{E} \subseteq \bigcap [r(\mathcal{A})]$. Want: $\mathcal{E} \supseteq \bigcap [r(\mathcal{A})]$. It suffices to show: $\mathcal{E} \in r(\mathcal{A})$.

Let $X := \bigcup \mathcal{A}$. We have $\mathcal{A} \subseteq 2^X$, so $\mathcal{B} \subseteq 2^X$, so $\mathcal{C} \subseteq 2^X$, so $\mathcal{D} \subseteq 2^X$, so $\mathcal{E} \subseteq 2^X$. It remains to show: $\mathcal{E}$ is a ring of sets.

We have $\mathcal{B} = \langle \mathcal{A} \rangle_{\text{fin}\,\cap}$, and so $\langle \mathcal{B} \rangle_{\text{fin}\,\cap} \subseteq \mathcal{B}$. We have $\mathcal{C} = \langle \mathcal{B} \rangle_{\text{fin}\,\cup}$, and so $\langle \mathcal{C} \rangle_{\text{fin}\,\cup} \subseteq \mathcal{C}$. Because $\cap$ distributes over $\cup$, we get

$$\langle \langle \mathcal{B} \rangle_{\text{fin}\,\cup} \rangle_{\text{fin}\,\cap} \quad \subseteq \quad \langle \langle \mathcal{B} \rangle_{\text{fin}\,\cap} \rangle_{\text{fin}\,\cup}.$$

So, since $\langle \mathcal{B} \rangle_{\text{fin}\,\cup} = \mathcal{C}$ and $\langle \mathcal{B} \rangle_{\text{fin}\,\cap} = \mathcal{B}$, this gives $\langle \mathcal{C} \rangle_{\text{fin}\,\cap} = \langle \mathcal{B} \rangle_{\text{fin}\,\cup}$. Then $\langle \mathcal{C} \rangle_{\text{fin}\,\cap} = \langle \mathcal{B} \rangle_{\text{fin}\,\cup} = \mathcal{C}$. So, since $\langle \mathcal{C} \rangle_{\text{fin}\,\cup} \subseteq \mathcal{C}$, by Lemma 55.1, we conclude that $\mathcal{D}$ is a pre-ring of sets. Then, by Lemma 54.17, we conclude that $\langle \mathcal{D} \rangle_r = \mathcal{E}$, and so $\mathcal{E}$ is a ring of sets, as desired. □

NOTE TO SELF: Next year, let's do several things a little differently. First, let's rename "partition measure scheme" to "refinement sequence". Second, add construction of Lebesgue measure on any Euclidean space $\mathbb{R}^d$. Third, add Lebesgue-Stieltjes measure. Fourth, describe product measure on $\{0, 1\}^{\mathbb{N}}$. Fifth, add the Kolmogorov Extension Theorem. END OF NOTE TO SELF.

Index of terms

FOR OFFICE USE ONLY: Corollary 34.10, p. 260