THE ALGEBRAIC GEOMETRY ALGEBRAIC TOPOLOGY DUMP

ERNEST YEUNG ERNESTYALUMNI@GMAIL.COM

From the beginning of 2016, I decided to cease all explicit crowdfunding for any of my materials on physics, math. I failed to raise *any* funds from previous crowdfunding efforts. I decided that if I was going to live in *abundance*, I must lose a scarcity attitude. I am committed to keeping all of my material **open-sourced**. I give all my stuff *for free*.

In the beginning of 2017, I received a very generous donation from a reader from Norway who found these notes useful, through *PayPal*. If you find these notes useful, feel free to donate directly and easily through PayPal, which won't go through a 3rd. party such as indiegogo, kickstarter, patreon. Otherwise, under the *open-source MIT license*, feel free to copy, edit, paste, make your own versions, share, use as you wish.

gmail : ernestyalumni
linkedin : ernestyalumni
twitter : ernestyalumni

Contents

Date: 5 mars 2017.
Key words and phrases. Algebraic Geometry, Algebraic Topology.

Abstract. Everything about Algebraic Geometry, Algebraic Topology

Part 1. Algebra; Groups, Rings, R-Modules, Categories

We should know some algebra. I will follow mostly Rotman (2010) [29].

1. Prime numbers, GCD (greatest common denominator), integers, Euler's totient, Chinese Remainder Theorem, integer divison, modulus, remainders; Euclid's Lemma

Definition 1 (natural numbers $\mathbb{N}$). *natural numbers* $\mathbb{N}$

$$\mathbb{N} = \{ \text{ integers } n | n \geq 0 \} \tag{1}$$

i.e. $\mathbb{N}$ *is set of all nonnegative integers.*

Definition 2 (prime). *natural number p is* ***prime*** *if $p \geq 2$, and $\nexists$ factorization $p = ab$, where $a < p$, $b < p$ are natural numbers.*

Definition 3. $a, b \in \mathbb{Z}$ ***relatively prime*** *if* $\gcd(a,b) = 1$

Axiom 1. ***Least Integer Axiom*** $\exists$ *smallest integer in every $C \subset \mathbb{N}$, $C \neq \emptyset$*

cf. pp. 1, Ch. 1 Things Past of Rotman (2010) [29]

Theorem 1 (Division Algorithm). *$\forall a, b \in \mathbb{Z}$, $a \neq 0$, $\exists ! \, q, r \in \mathbb{Z}$ s.t.*

$$b = qa + r \text{ and } 0 \leq r < |a|$$

Proof. Consider $n \in \mathbb{Z}$, $b - na \in \mathbb{Z}$
Let $C = \{ b - na | n \in \mathbb{Z} \} \bigcap \mathbb{N}$.
$C \neq \emptyset$ (otherwise, consider $b - na < 0$, $b < na$, then contradiction)
By Least Integer Axiom, $\exists$ smallest $r \in C$, $r = b - na$.
define $q = n$ when $r = b - na$.
Suppose

$$\begin{aligned} & qa + r = q'a + r' \\ & (q - q')a = r' - r \\ & |(q - q')a| = |r' - r| \end{aligned} ,$$

$0 \leq r' < |a|$. Now $0 \leq |r' - r| < |a|$
if $|q - q'| \neq 0$, $|(q - q')a| \geq |a|$

$$\Longrightarrow q = q', r = r'$$

Conclude both sides are 0 (by contradiction) □

cf. pp. 2, Thm. 1.4, Ch. 1 Things Past of Rotman (2010) [29]

Definition 4 (divisor). *$a, b \in \mathbb{Z}$, a* ***divisor*** *of b if $\exists \, d \in \mathbb{Z}$ s.t. $b = ad$.*
a ***divides*** *b or b multiple of a, denote*

$$a | b$$

$a|b$ iff b has remainder $r = 0$ after dividing by a

cf. pp. 3, Ch. 1 Things Past of Rotman (2010) [29]

1.1. Greatest Common Denominator (GCD); Euclid's Lemma.

Definition 5 (common divisor). ***common divisor*** *of integers a and b, is integer c, s.t.* $c|a$ and $c|b$.
greatest common divisor *or* ***gcd*** *of a and b, denoted $(a,b) \equiv gcd(a,b)$ defined by*

$$(a,b) \equiv gcd(a,b) = \begin{cases} 0 & \text{if } a = 0 = b \\ & \text{the largest common divisor of } a \text{ and } b \text{ otherwise} \end{cases}$$

cf. pp. 3, Ch. 1 Things Past of Rotman (2010) [29]

Theorem 2. *If $a, b \in \mathbb{Z}$, then $gcd(a,b) \equiv (a,b) = d$ is linear combination of a and b, i.e. $\exists \, s, t \in \mathbb{Z}$ s.t.*

$$d = sa + tb \tag{2}$$

cf. pp.4, Thm. 1.7, Ch. 1 Things Past of Rotman (2010) [29]

Proof. Let $I :=$

$$I := \{ sa + tb | s, t \in \mathbb{Z} \}$$

If $I \neq \{0\}$, let d be smallest positive integer in I.
$d \in I$, so $d = sa + tb$ for some $s, t \in \mathbb{Z}$.
Claim: $I = (d) \equiv \{ kd | k \in \mathbb{Z} \} =$ set of all multiples of d.
Clearly $(d) \subseteq I$, since $kd = k(sa + tb) = (ks)a + (kt)b \in I$.
Let $c \in I$.
By division algorithm, $c = qd + r$, $0 \leq r < d$

$$r = c - qd = s'a + t'b - qsa - qtb = (s' - sq)a + (t' - qt)b \in I$$

If $r \in I$, but $r < d$, contradiction that $\min_{\substack{i \in I \\ i > 0}} i = d$.
So $r = 0$, and $d | c = c/d$.

$$c \in (d), \text{ so } I \subseteq (d) \Longrightarrow I = (d)$$

□

Theorem 3 (**Euclid's Lemma**; 1.10 of Rotman (2010) [29]). *If p prime and $p|ab$, then $p|a$ or $p|b$.*
More generally,
if prime p divides product $a_1 a_2 \dots a_n$,
then it must divide at least 1 of the factors a_i.
i.e. (notation),
If prime p, and $ab/p \in \mathbb{Z}$,
then $a/p \in \mathbb{Z}$ or $b/p \in \mathbb{Z}$.
More generally,
if prime p, s.t. $a_1 a_2 \dots a_n / p \in \mathbb{Z}$,
then $\exists$ 1 a_i s.t. $a_i / p \in \mathbb{Z}$

Proof. If $p \nmid a$, i.e. $a/p \notin \mathbb{Z}$, then $\gcd(p,a) \equiv (p,a) = 1$.
From Thm. 2,

$$1 = sp + ta$$
$$\Longrightarrow b = spb + tab = p(sb + td)$$

ab/p and so $ab = pd$, so $b = spb + tdp$, i.e. b is a multiple of p ($b/p \in \mathbb{Z} \equiv p|b$). □

Corollary 1 (1.11 of Rotman (2010) [29])**.** *Let $a,b,c \in \mathbb{Z}$.*
If c,a relatively prime, i.e. $\gcd(c,a) = 1$, and if $c|ab \equiv ab/c \in \mathbb{Z}$, then $c|b \equiv b/c \in \mathbb{Z}$

Proof.
$$\gcd(c,a) = 1 = sc + ta \Longrightarrow b = sbc + tab = sbc + t(qc) = c(sb + tq) \Longrightarrow b/c = sb + tq$$
□

Theorem 4 (Euclidean Algorithm)**.** *Let $a,b \in \mathbb{Z}^+$.*
$\exists$ algorithm that finds $d = \gcd a,b$

cf. pp. 5, Thm. 1.14 (Euclidean Algorithm), Ch. 1 Things Past of Rotman (2010) [29].

Proof. □

Definition 6. *Let fixed $m \geq 0$. Then $a,b \in \mathbb{Z}$ are* ***congruent modulo*** *m, denoted by*

$$a \equiv b \bmod m$$

if $m|(a-b)$, i.e. $(a-b)/m \in \mathbb{Z}$, i.e. if $(a-b)/m \in \mathbb{Z}$, i.e. $(a-b)$ integer multiple of m

Proposition 1. *If $m \geq 0$ is fixed, $m \in \mathbb{Z}$, then $\forall\, a,b,c \in \mathbb{Z}$*
(1) $a \equiv a \bmod m$
(2) *if* $a \equiv b \bmod m$*, then* $b \equiv a \bmod m$
(3) *if* $a \equiv b \bmod m$*, and* $b \equiv c \bmod m$*, then* $a \equiv c \bmod m$

cf. Prop. 1.18 of Rotman (2010) [29]

Proof. (1) $(a-a)/m = 0/m = 0$
(2) $(b-a)/m = (-1)(a-b)/m \in \mathbb{Z}$
(3) $(a-c)/m = (a-b+b-c)/m = (a-b)/m + (b-c)/m \in \mathbb{Z}$
□

EY : 20171225 to recap,

$$\boxed{\begin{gathered} a \equiv b \bmod n \\ \text{meaning} \\ \frac{a-b}{n} \in \mathbb{Z} \text{ or } a - b = kn,\ k \in \mathbb{Z} \text{ or } a = b + kN \text{ but rather} \\ a = pn + r \\ b = qn + r \end{gathered}} \qquad (3)$$

for $a = b + kn$, but b need not be a remainder of division of a by n. More precisely, $a = b \bmod n$ asserts that a,b have the same remainder when divided by n, i.e.

$$a = pn + r$$
$$b = qn + r$$

So $a \sim b$ or $[a] = [b]$ is an equivalence relation since
$a \sim a$ since $a \equiv a \bmod N$, since $a = a + 0N$,
if $a \sim b$, then $b \sim a$, since $a - b = kN$, then $b = a - kN$

if $a \sim b$, $b \sim c$, then $a \sim c$, since $a - b = kN$, then $a - c = (k+l)N$.
$$b - c = lN$$
cf. Prop. 1.19 of Rotman (2010) [29]

Proposition 2. *Let $m \geq 0$ be fixed*
(1) *If $a = qm + r$, then $a \equiv r \bmod m$*
(2) *If $0 \leq r' < r < m$, then $r \not\equiv \bmod m$ i.e. r and r' aren't congruent $\bmod m$*
(3) *$a \equiv b \bmod m$ iff a,b leave same remainder after dividing by m*
(4) *If $m \geq 2$, $\forall\, a \in \mathbb{Z}$, $a \equiv b \bmod m$ for some $b \in 0,1,\dots m-1$*

Proof. (1) If $a = qm + r$, then $a \equiv r \bmod m$
$$\frac{a-r}{m} = q \in \mathbb{Z}$$
(2) *Want*: If $0 \leq r' < r < m$, then $r \not\equiv \bmod m$.
Suppose $\frac{r-r'}{m} = k \in \mathbb{Z}$. Then $r - r' = km$ or $r = r' + km$.
$$m > r > r' \leq 0$$
$$m > r' + km > r' \leq 0$$
$$m - r' > km > 0$$
But $k > 0$ (since $m > 0$ and $r - r' = km > 0$) and $m - r' > km > 0$ is a contradiction.
(3) *Want*: $a \equiv b \bmod m$ iff a,b leave same remainder after dividing by m. By
By Division Algorithm, this is true:
$$a = q_a m + r_a$$
$$b = q_b m + r_b$$
$$\frac{a-b}{m} = q_a + \frac{r_a}{m} - q_b - \frac{r_b}{m} = k = q_a - q_b + \frac{r_a - r_b}{m} \in \mathbb{Z}$$
Now
$$|m| > r_a \leq 0$$
$$|m| > r_b \leq 0$$
$2|m| > r_a + r_b$.
And if $r_a > r_b$, $|m| > r_a > r_a - r_b > 0$.
In both cases, $r_a = r_b$ since $q_a - q_b + \frac{r_a - r_b}{m} \in \mathbb{Z}$ needs to be enforced.
(4) *Want*: If $m \geq 2$, $\forall\, a \in \mathbb{Z}$, $a \equiv b \bmod m$ for some $b \in 0,1,\dots m-1$.
By Division Algorithm, $a = q_a m + r_a$, $0 \leq r_a < |m|$. $\frac{a - r_a}{m} = q_a \in \mathbb{Z}$ so let $b = r_a$. □

Theorem 5 (1.26 of Rotman (2010) [29])**.** *If $\gcd(a,m) \equiv (a,m) = 1$, then $\forall\, b \in \mathbb{Z}$, $\exists\, x$ s.t.*
$$ax \equiv b \bmod m$$
In fact, $x = sb$, where $sa \equiv 1 \bmod m$ is 1 solution. Moreover, any 2 solutions are congruent $\bmod m$.
i.e.
If $\gcd a,b = 1$, then $\forall\, y \in \mathbb{Z}$, $\exists\, x$ s.t. $ax \equiv y \bmod b$, $x = sy$, where $sa \equiv 1 \bmod b$ is 1 solution. Moreover, any 2 solutions are congruent $\bmod m$. This implies that

$ax \equiv y \bmod b$ or $\frac{Ax - y}{b} \in \mathbb{Z}$, and $\frac{(as-1)y}{b} \in \mathbb{Z}$.
$sa \equiv 1 \bmod b$ or $\frac{sa-1}{b} \in \mathbb{Z}$, which implies that $sa - 1 = b(-t)$ or

$$sa + tb = 1$$

for some $s, t \in \mathbb{Z}$.

Proof. $\gcd(a,m) = 1 = sa + tm$, by Thm. 2

Then $b = b \cdot 1 = b(sa+tm) = sab + tmb$ or $b = tbm + sab$ or $a(sb) = -tbm + b$.

So $a(sb) \bmod m \equiv b$.

Let $x := sb$ and so $ax \bmod m = b$.

Now suppose $x \neq sb$ s.t. $ax \bmod m = b$. Then $ax = qm + b$. From $a(sb) \bmod m = b$, we also get $a(sb) = q'm + b$. Then $a(x-sb) \bmod m = 0$, so $m | a(x-sb) \equiv a(x-sb)/m \in \mathbb{Z}$.

By Corollary 1 (which says, if $\gcd(c,a) = 1$ and if $ab/c \in \mathbb{Z}$, then $b/c \in \mathbb{Z}$), since $\gcd(m,a) = (m,a) = 1$, and since $a(x-sb)/m \in \mathbb{Z}$, then $(x-sb)/m \in \mathbb{Z}$. So $(x-sb) = qm$ or $(sb) \bmod m = x$. □

Proposition 3 (3.1 of Scheinerman (2006) [31])**.** *Let* $a, b \in \mathbb{Z}$, *let* $c = a \bmod b$, *i.e.* $a = qb + c$ *s.t.* $0 \leq c < b$.

Then

$$gcd(a,b) = gcd(b,c) \tag{4}$$

cf. Sec. 3.3 Euclid's method of Scheinerman (2006) [31]

Proof. If d common divisor of a, b, i.e. $a/d, b/d \in \mathbb{Z} \equiv d|a, d|b$.
$c/d \in \mathbb{Z} \equiv d|c$ since $c = a - qb$.

If d is common divisor of b, c, i.e. $d|b, d|c \equiv c/d, b/d \in \mathbb{Z}$,
then $d|a \equiv a/d \in \mathbb{Z}$ since $a = qb + c$. So set of common divisors of a, b same as set of common divisors of b and c.

Then $\gcd(a,b) = \gcd(b,c)$. □

1.2. Euler's totient; relatively prime.

cf. Ch. 5 Arrays, Sec. 5.1 Euler's totient of Scheinerman (2006) [31]

For

$$\begin{aligned} &\varphi : \mathbb{Z}^+ \to \mathbb{Z}^+ \\ &\varphi : n \mapsto \varphi(n) := \text{ number of elements of } \{1, 2, \dots n\} \\ \text{that are relative prime to } \quad & \\ &n = |\{i | i \in \{1, 2, \dots n\}, (n,i) = 1 \text{ or equivalently } n \propto i\}| \end{aligned}$$

e.g. $\varphi(10) = 4$ since $\varphi(10) = |\{1, 3, 7, 9\}|$.

we want $|(a,b)| 1 \leq a, b, \leq n, \gcd(a,b) \equiv (a,b) = 1|$.

$$\begin{aligned} p_n = \frac{1}{n^2}\left[-1 + 2\sum_{i=1}^n \varphi(k) \right] = \\ = \text{ probability that 2 integers, chosen uniformly and independently from } \{1, 2, \dots n\} \text{ are relatively prime} \end{aligned}$$

If p is prime, $\forall\, i \in \{1, 2, \dots p\}$, $(p,i) \equiv \gcd(p,i) = 1$, i.e. relatively prime to p, except 1 $i \in \{1, 2, \dots p\}$.

Therefore

$$\varphi(p) = p - 1$$

Consider $\varphi(p^2)$.

$\{1, 2, \dots p^2\}$, only numbers *not* relatively prime to p^2 are multiples of p since $p, 2p, 3p, \dots p^2$ all divide p^2, i.e. $p|p^2, 2p|p^2 \dots (p-1)p|p^2 \equiv p^2/p, p^2/2p, \dots p^2/p(1-p)$.

Assume $\varphi(p^n) = p^2 - p^{n-1} = p^{n-1}(p-1)$.

$$\varphi(p^{n+1}) = \varphi(pp^n) = p^n\varphi(p) = p^n(p-1)$$

Therefore,

Proposition 4 (5.1)**.** *Let* p *prime,* $n \in \mathbb{Z}^+$

e.g. $\varphi(77)$.

$\forall\, n$ s.t. $1 \leq n \leq 77$.

$$\begin{aligned} \gcd(n, 77) &= 1 \\ \gcd(n, 7) &= 1 \\ \gcd(n, 11) &= 1 \end{aligned}$$

By Prop. 3,

$$\begin{aligned} \gcd(n, 7) &= \gcd(7, n \mod 7) \\ \gcd(n, 11) &= \gcd(11, n \mod 11) \end{aligned}$$

cf. Example (10) of Dummit and Foote [2].

To recap,

Definition 7 (Euler φ-function)**.** $\forall\, n \in \mathbb{Z}^+$,
let $\varphi(n) :=$ *number of positive integers* $a \leq n$ *with a relatively prime to* n, *i.e.* $\gcd(a,n) = 1 \equiv (a,n)$

e.g. $\varphi(12) = 4$, since $1, 5, 7, 11$ are only positive integers less than or equal to 12.

If p prime, $\varphi(p) = p - 1$.

More generally,
$\forall\, a \geq 1$,

$$\boxed{\varphi(p^a) = p^a - p^{a-1} = p^{a-1}(p-1)} \tag{5}$$

φ is multiplicative in the sense that

$$\varphi(ab) = \varphi(a)\varphi(b) \text{ if } \gcd(a,b) = 1 \tag{6}$$

$\Longrightarrow$ general formula.

If $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_s^{\alpha_s}$ (Fumdanetal Thm. of Arithmetic, $\forall\, n \in \mathbb{Z}, n > 1$), then

$$\boxed{\begin{aligned} &\varphi(n) = \varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2}) \dots \varphi(p_s^{\alpha_s}) \\ &\quad p_1^{\alpha_1 - 1}(p_1 - 1)p_2^{\alpha_2 - 1}(p_2 - 1) \dots p_s^{\alpha_s - 1}(p_s - 1) \end{aligned}} \tag{7}$$

cf. pp. 69 Thm. 5.4 (Chinese Remainder) of Scheinerman (2006) [31].

Theorem 6. *Let* $n \in \mathbb{Z}^+$,
let $p_1, p_2, \dots p_t$ *be distinct prime divisors of* n *(i.e.* $\forall\, p_i$, $\frac{n}{p_i^{k_i}} \in \mathbb{Z}$ *for some* $k_i \geq 1$*)*
Then

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_t}\right) \tag{8}$$

Proof. By Fundamental Thm. of Arithmetic,

$$n = p_1^{e_1} p_2^{e_2} \dots p_t^{e_t}$$

where p_j are distinct primes, and e_j are positive integers.

From Eqns. 5, 6, i.e. where

$$\begin{aligned} \varphi(p^a) &= p^a - p^{a-1} = p^{a-1}(p-1) \\ \varphi(ab) &= \varphi(a)\varphi(b) \text{ if } \gcd(a,b) = 1 \end{aligned}$$

$$\begin{aligned} \varphi(n) = \varphi(p_1^{e_1} p_2^{e_2} \dots p_t^{e_t}) = \varphi(p_1^{e_1})\varphi(p_2^{e_2}) \dots \varphi(p_t^{e_t}) = \\ = p_1^{e_1}(1 - \frac{1}{p_1})p_2^{e_2}(1 - \frac{1}{p_2}) \dots p_t^{e_t}(1 - \frac{1}{p_t}) = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \dots (1 - \frac{1}{p_t}) \end{aligned}$$

□

Exercise 10. cf. pp. 7 Exercise 10 Dummit and Foote [2].

Prove: $\forall$ given $N \in \mathbb{Z}^+$ (positive number),
$\exists$ only finite many integers n with $\varphi(n) = N$, where φ denotes Euler's φ-function.

EY, Indeed, by definition,

$$\begin{gathered} \varphi(n) = N \\ a_1, a_2 \dots a_N \text{ s.t. } a_i \leq n \\ \gcd(a_i, n) = 1 \text{ i.e. } 1 = s_i a_i + t_i n \end{gathered}$$

Given $N \in \mathbb{Z}^+$, let $n \in \mathbb{Z}$, s.t. $\varphi(n) = N$ (given hypothesis).

Let $p =$ least (i.e. smallest) prime s.t. $p > N + 1$.

If $q \geq p$ is a prime divisor of n, i.e.

$$n = q^k m$$

for some $k \geq 1$, and m with q not dividing m.

Then

$$\varphi(n) = \varphi(q^k)\varphi(m) = q^{k-1}(q-1)\varphi(m) \geq q - 1 \geq p - 1 > N$$

Contradiction.

Thus, $\nexists$ prime divisor of n greater than $N + 1$.

Particularly, distinct prime divisors of n belong to a finite set, say these primes are $p_1, p_2 \dots p_m$.

Definition 8. *prime divisor q of n if q is prime and*

$$\frac{n}{q} \in \mathbb{Z} \text{ i.e. } n = q^k m \text{ for some } k \geq 1 \text{ and } \frac{m}{q} \notin \mathbb{Z}^+ \tag{9}$$

Now

$$n = p_1^{a_1} p_2^{a_2} \dots p_m^{a_m}$$

for some $0 < a_i$, so

$$\varphi(n) = \varphi(p_1^{a_1})\varphi(p_2^{a_2}) \dots \varphi(p_m^{a_m}), \text{ so } \varphi(n) = \prod_{i=1}^m p_i^{a_i - 1}(p_i - 1)$$

Note, $\forall$ prime p_i, $\varphi(n) \geq p_i^{a_i-1}(p_i - 1) \geq p_i - 1 > N$ for sufficiently large a_i.

Thus, $\forall p_i$, $\exists$ only finitely many permissible choices for exponents a_i.
So set of all n with $\varphi(n) = N$ is subset of finite set, hence finite.

$\forall N \in \mathbb{Z}^+$, $\exists$ largest integer n with $\varphi(n) = N$.
Thus, as $n \to \infty$, $\varphi(n) \to \infty$.

Scheinerman (2006) [31]

cf. Ex. 1.19, pp. 13, Sec. 1.1 Some Number Theory of Rotman (2010) [29] **Exercise 1.19.** If a and b are relatively prime and if each divides an integer n, then their product ab also divides n, i.e.

Theorem 7. *If $\gcd a, b = 1$, and if $n/a \in \mathbb{Z} \equiv a|n$, and $n/b \in \mathbb{Z} \equiv b|n$, then $n/ab \in \mathbb{Z} \equiv ab|n$.*

Proof. $\gcd a, b = 1$, so $sa + tb = 1$ for some $s, t \in \mathbb{Z}$ (Thm. 5).
$\frac{n}{a}, \frac{n}{b} \in \mathbb{Z}$, so $n = au$, $n = bv$
$n = n \cdot 1 = n(sa + tb) = bvsa + autb = ab(vs + ut)$, so $\frac{n}{ab} = vs + ut \in \mathbb{Z}$. $\square$

1.2.1. *Chinese Remainder Theorem.*

Theorem 8. *If m, m' relatively prime (i.e. $\gcd(m, m') = 1$), then for*

$$\begin{gathered} x \equiv b \bmod m \\ x \equiv b' \bmod m' \end{gathered}$$

i.e. given $b, b' m, m'$, and wanting to find x, $\exists x$ and $\forall 2x$'s, $x = x' \mod mm'$, i.e.

Let m, n relatively prime positive integers (i.e. $\gcd m, n = 1$),
$\forall a, b \in \mathbb{Z}$,
then pair of congruences
$x \equiv a \bmod m$
$x \equiv b \bmod n$
has a solution (x), and this solution x is uniquely determined, modulo mn.

Proof. cf. The Chinese Remainder Theorem by Keith Conrad

Suppose
$(x - a)/m \in \mathbb{Z}$ or $x - a = my$
$(x - b)/n \in \mathbb{Z}$ or $x - b = nz$ or $a + my - b = nz$
$\gcd m, n = 1$, so then $\forall b \in \mathbb{Z}$, $\exists w$ s.t. $mw \equiv b \bmod n$ i.e. $\frac{mw-b}{n} \in \mathbb{Z}$, in fact, $w = sb$, where $sm \equiv 1 \bmod n$, or $\frac{sm-1}{n} \in \mathbb{Z}$, is 1 solution (Thm. 5).

$$\begin{aligned} &my = b - a + nz \\ &smy = sb - sa + snz = (1 + nv)y = s(b - a) + snz \text{ or } y = s(b - a) + n(sz - vy) \\ &\text{or } y \equiv s(b - a) \bmod n \end{aligned}$$

$$\begin{aligned} x = a + my = a + m(s(b - a) + n(sz - vy)) = a + ms(b - a) + mn(sz - vy) \equiv a + ms(b - a) + mnu \\ x - a = m(s(b - a) + nu) \implies x = a \bmod m \\ x - b = a + ms(b - a) + mnu - b = a + (1 + m)(b - a) + mnu - b = m(b - a) + mnu \implies x \equiv b \bmod n \end{aligned}$$

Uniqueness: Suppose $x, y \in \mathbb{Z}$ s.t.

$$\begin{aligned} x &\equiv a \bmod m & y &\equiv a \bmod m \\ x &\equiv b \bmod n & y &\equiv b \bmod n \end{aligned}$$

Given $\gcd m, n = 1$, $sm + tn = 1$.

Since $\frac{x-a}{m}, \frac{y-a}{m} \in \mathbb{Z}$, $\frac{x-y}{m} \in \mathbb{Z}$, likewise, $\frac{x-a}{n}, \frac{y-a}{n} \in \mathbb{Z}$, $\frac{x-y}{n} \in \mathbb{Z}$
Since $\frac{x-y}{m}, \frac{x-y}{n} \in \mathbb{Z}$, $\frac{x-y}{mn} \in \mathbb{Z}$ by Thm. 7.

Thus, $x - y = mnk$ for some $k \in \mathbb{Z}$. For instance, $k = 0$, $x = y$.

This shows any 2 solutions are the same, modulo mn. $\square$

cf. Ch. 1 Things Past, Thm. 1.28 of Rotman (2010) [29], pp. 68 Thm. 5.2 (Chinese Remainder) of Scheinerman (2006) [31].

2. Groups

cf. pp. 16 Chapter 1 Introduction to Groups. Dummit and Foote (2004) [2]

Definition 9 (binary operation)**.** (1) *binary operation $*$ on set G is a function $* : G \times G \to G$. $\forall a, b \in G$, $a * b \equiv *(a, b)$*
(2) *binary operation $*$ on set G is* associative*: if $\forall a, b, c \in G$, $a * (b * c) = (a * b) * c$*
(3) *If $*$ is binary operation on set G, a, b of G commut if $a * b = b * a$.*
*$*$ (or G) is **commutative** if $\forall a, b \in G$ $a * b = b * a$.*

cf. pp. 16. Sec. 1.1. Basic Axioms and Examples, Dummit and Foote (2004) [2]

Definition 10 (Group)**.** (1) *Group is an ordered pair $(G, *)$ where G is a set, $*$ is a binary operation on G s.t.*
(a) $(a * b) * c = a * (b * c)$, $\forall a, b, c \in G$, *i.e.* $*$ associative
(b) $\exists e \in G$, *s.t.* $\forall a \in G$, $a * e = e * a = a$ *($\exists$* identity *e)*
(c) $\forall a \in G$, $\exists a^{-1} \in G$, *called an* inverse *of a, s.t.* $a * a^{-1} = a^{-1} * a = e$
(2) *(optional; abelian or* commutative*) $(G, *)$ abelian (or commutative) if $a * b = b * a$, $\forall a, b \in G$.*

e.g.

(1) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are groups under $+$ with $e = 0$ and $a^{-1} = -a$, $\forall a$.
(2) $\mathbb{Q} - \{0\}, \mathbb{R} - \{0\}, \mathbb{C} - \{0\}, \mathbb{Q}^+, \mathbb{R}^+$ groups under $\times$ with $e = 1$, $a^{-1} = \frac{1}{a}$

(3) (**direct product of groups**)If $(A,*), (B,\circ)$ are groups, we can form new group $A\times B$ called **direct product** s.t.

$$A\times B = \{(a,b) | a\in A, b\in B\}$$

and $(a_1,b_1)(a_2,b_2) = (a_1 * a_2, b_1\circ b_2)$ cf. Example 6, Sec. 1.1 Dummit and Foote (2004) [2]

Proposition 5. *If G group under operation $*$, then*

(1) *identity of G is unique*
(2) *$\forall\, a\in G$, a^{-1} uniquely determined.*
(3) $(a^{-1})^{-1} = a \quad \forall\, a\in G$
(4) $(a*b)^{-1} = (b^{-1}) * (a^{-1})$
(5) *$\forall\, a_1, a_2, \dots a_n\in G$, $a_1, a_2\dots a_n$ independent of how expression is bracketed (generalized associative law)*

cf. Prop. 1, Sec. 1.1 Dummit and Foote (2004)[2]

3. Groups; normal subgroups

Definition 11 (normal subgroup $K\triangleleft G$)**.**
normal subgroup *K of $G\equiv K\triangleleft G$ -*
subgroup $K\subset G$, if $\forall\, k\in K$, $\forall\, g\in G$,

$$gkg^{-1}\in K$$

Definition 12 (quotient group)**.**
quotient group *$G \mod K\equiv G/K$ -*
if G/K = family of all left cosets of subgroups $K\subset G =$

$$= \{gK | g\in G, gK = \{gk|k\in K\}$$

and
K = normal subgroup of G, i.e. $K\triangleleft G$, and so

$$aKbK = abK \qquad \forall\, a,b\in G,$$

so G/K group.

Definition 13 (exact sequence of groups)**.** ***exact sequence*** *if $imf_{n+1} = kerf_n \qquad \forall\, n$ for sequence of group homomorphisms and groups*

$$G_{n+1}\xrightarrow{f_{n+1}} G_n \xrightarrow{f_n} G_{n-1} \tag{10}$$

Theorem 9. (1)

$$1 \qquad A\xrightarrow{f} B$$

(2)

$$B\xrightarrow{g} C \qquad 1$$

(3)

$$1 \qquad A\xrightarrow{h} B \qquad 1$$

Proof. (1) $\text{im}(1\to A) = 1$, since $1\to A$ is a group homomorphism $((1\to A)(1) = 1_A)$.
if $1\to A\overset{f}{\mapsto} B$ exact, $\ker f = \text{im}(1\to A) = 1$, so if $f(x) = 1$, $x=1$, f injective.
If f injective, $\ker f = 1$. $1 = \text{im}(1\to A)$. $1\to A\overset{f}{\mapsto} B$, exact.

(2) $\ker(C\to 1) = C$, by def. of $C\to 1$
if $B\overset{g}{\mapsto} C\to 1$ exact, $\text{im} g = g(B) = \ker(C\to 1) = C$. $g(B) = C$ implies g surjective.
If g surjective, $g(B) = C = \ker(C\to 1)$. $B\overset{g}{\mapsto} C\to 1$ exact.

(3) From (i), $1\to A\overset{h}{\mapsto} B$ exact iff h injective. From (ii), $A\overset{h}{\mapsto} B\to 1$, exact iff h surjective.
h isomorphism.

□

3.1. 1st, 2nd, 3rd Isomorphism Theorems.

Theorem 10 (1st Isomorphism Theorem (Modules) Thm. 7.8 of Rotman (2010) [29])**.** *If $f: M\to N$ is R-map of modules, then $\exists\, R$-isomorphism s.t.*

$$\begin{tikzcd} M \arrow[r, "f"] \arrow[d, "\pi"] & N \\ M/kerf \arrow[ur, "\varphi = \cong"'] & \end{tikzcd}$$

$$\begin{aligned} & \varphi : M/kerf \to imf \\ & \varphi : m + kerf \mapsto f(m) \end{aligned} \tag{11}$$

Proof. View M, N as abelian groups.

Recall natural map $\pi : M\to M/N$
$m\mapsto m+N$
Define φ s.t. $\varphi\pi = f$.
(φ well-defined). Let $m + \ker f = m' + \ker f$, $m, m'\in M$, then $\exists\, n\in\ker f$ s.t. $m = m'+n$.

$$\varphi(m+\ker f) = \varphi\pi(m) = f(m) = f(m'+n) = f(m') + f(n) = \varphi\pi(m') + 0 = \varphi(m'+\ker f)$$

$\Longrightarrow \varphi$ well-defined.
(φ surjective). Clearly, $\text{im}\varphi\subseteq\text{im} f$.
Let $y\in\text{im} f$. So $\exists\, m\in M$ s.t. $y = f(m)$. $f(m) = \varphi\pi(m) = \varphi(m+\ker f) = y$. So $y\in\text{im}\varphi$. $\text{im} f\subseteq\text{im}\varphi$.
$\Longrightarrow \varphi$ surjective.
(φ injective) If $\varphi(a+\ker f) = \varphi(b+\ker f)$, then

$$\varphi\pi(a) = \varphi\pi(b) \text{ or } f(a) = f(b) \text{ or } 0 = f(a) - f(b) = f(a-b) \text{ so } a-b\in\ker f(a-b) + \ker f = \ker f \text{ so } a+\ker f = b+\ker f$$

φ isomorphism.
φ R-map. $\varphi(r(m+N)) = \varphi(rm+N) = f(rm)$.
Since f R-map, $f(rm) = rf(m) = r\varphi(m+N)$. φ is R-map indeed.

□

Theorem 11 (2nd Isomorphism Theorem (Modules) Thm. 7.9 of Rotman (2011) [29])**.** *If S,T are submodules of module M, i.e. $S,T\in M$, then $\exists$ $R-$isomorphism*

$$\begin{tikzcd} S \arrow[r, "h"] \arrow[d, "\pi|_S"] & (S+T)/T = imh \\ S/(S\cap T) = S/kerh \arrow[ur, "\cong"'] & \end{tikzcd}$$

$$S/(S\cap T)\to (S+T)/T \tag{12}$$

Proof. Let natural map $\pi : M \to M/T$.

So $\ker \pi = T$.

Define $h := \pi|_S$, so $h : S \to M/T$, so $\ker h = S \cap T$,

$$(S+T)/T = \{(s+t)+T | a \in S+T, s \in S, t \in T\}$$

i.e. $(S+T)/T$ consists of all those cosets in M/T having a representation in S.

By 1st. isomorphism theorem,

$$S/S\cap T \xrightarrow{\cong} (S+T)/T$$

□

Theorem 12 (3rd Isomorphism Theorem (Modules) Thm. 7.10 of Rotman (2011) [29]). *If $T \subseteq S \subseteq M$ is a tower of submodules, then $\exists$ R-isomorphism*

$$\begin{tikzcd} M/T \arrow[r, "g"] \arrow[d, "\pi"] & M/S \\ (M/T)/(S/T) = (M/T)/kerg \arrow[ur, "\cong"] & \end{tikzcd}$$

(13) $$(M/T)/(S/T) \to M/S$$

Proof. Define $g : M/T \to M/S$ to be **coset enlargement**, i.e.

(14) $$g : M + T \mapsto m + S$$

g well-defined: if $m + T = m' + T$, then $m - m' \in T \subseteq S$, and $m + S = m' + S \Longrightarrow g(m+T) = g(m'+T)$

$\ker g = S/T$ since

$$\begin{aligned} g(s+T) &= s + S = S \qquad (S/T \subseteq \ker g) \\ g(m+T) &= m + S = 0 = S = s + S, \text{ so } m = s \Longrightarrow \ker g \subseteq S/T \end{aligned}$$

$\mathrm{im} g = M/S$ since

$$\begin{aligned} g(m+T) &= m + S \Longrightarrow \mathrm{im} g \subseteq M/S \\ m + S &= g(m+T) \end{aligned}$$

Then by 1st isomorphism, and commutative diagram, done. □

4. Rings

cf. Ch. 7 "Introduction to Rings" pp. 223, Dummit and Foote (2014)[2]

Definition 14 (Ring). *ring R is a set, together with 2 binary operations $+, \times$ (addition and multiplication, $\times \equiv \cdot$) s.t.*

1. (a) $(R,+)$ *abelian group*
   (b) $\times$ *associative:* $a(bc) = (ab)c \quad \forall a,b,c \in R$
   (c) *distributivity in R:* $\forall a,b,c \in R$
   $$(a+b)c = ac + bc \text{ and } a(b+c) = ab + ac$$
2. R commutative *if multiplication commutative*
3. *R has an identity* 1 *if* $\exists 1 \in R$ *s.t.*
   $$1a = a1 = a \quad \forall a \in R$$

Definition 15 (division ring). *ring R with identity 1, where $1 \neq 0$ is a* ***division ring*** *(or skew field) if $\forall a \in R$, $a \neq 0$, $\exists$ multiplicative inverse $1/a$, i.e. $\exists b \in R$ s.t. $ab = ba = 1$*

e.g.

1. rational numbers $\mathbb{Q}$
   real numbers $\mathbb{R}$
   complex numbers $\mathbb{C}$
   are commutative rings with identity (in fact, they're fields)
   Ring axioms for each follow ultimately from ring axioms for $\mathbb{Z}$
   (verified when $\mathbb{Z}$ constructed from $\mathbb{Z}$ (Sec. 7.5)), $\mathbb{C}$ from $\mathbb{R}$ (Example 1, Sec. 13.1).
   Construction of $\mathbb{R}$ from $\mathbb{Z}$ carried out in basic analysis texts
2. **quotient group** $\mathbb{Z}/n\mathbb{Z}$ is a commutative ring with identity (element 1) under operations of addition and multiplication of residue classes (frequently referred to as "modular arithmetic").
   We saw additive abelian groups axioms followed from general principles of theory of quotient groups ($\mathbb{Z}/n\mathbb{Z}$) was prototypical quotient group. cf. Example 4, pp. 224, Dummit and Foote (2014)[2]
3. **the (real) Hamiltonian Quaternions**.

Definition 16 ((real) Hamiltonian Quaternions). *Let $\mathbb{H} = \{a + bi + cj + dk | a,b,c,d \in \mathbb{R}\}$ s.t. "componentwise" addition is defined as*

(15) $$(a+bi+cj+dk) + (a'+b'i+c'j+d'k) = (a+a') + (b+b')i + (c+c')j + (d+d')k$$

and multiplication defined by expanding using distributive laws

$$(a+bi+cj+dk)(a'+b'i+c'j+d'k)$$

using

(16) $$\begin{aligned} i^2 = j^2 = k^2 &= -1 \\ ij = -ji &= k \\ jk = -kj &= i \\ ki = -ik &= j \end{aligned}$$

Working out the multiplication

$$\begin{aligned} &(a+bi+cj+dk)(a'+b'i+c'j+d'k) = \\ &= \begin{aligned} &aa' + ab'i + ac'j + ad'k + ba'i - bb' + bc'k - bd'j + \\ &ca'j - cb'k - cc' + cd'i + da'k + db'j - dc'i - dd' = \end{aligned} \\ &= aa' - bb' - cc' - dd' + (ab' + ba' + cd' - dc')i + (ac' - bd' + ca' + db')j + (ad' + bc' - cb' + da')k \end{aligned}$$

Hamiltonian Quaternions are noncommutative ring with identity $(1 = 1 + 0i + 0j + 0k)$.

Similarly define *rational* Hamiltonian Quaternions ring by taking $a,b,c,d \in \mathbb{Q}$.

real and rational Hamiltonian Quaternions both are divison rings, where inverse of nonzero element defined as

(17) $$(a+bi+cj+dk)^{-1} = \frac{a - bi - cj - dk}{a^2+b^2+c^2+d^2}$$

cf. Example 5, pp. 224, Dummit and Foote (2014)[2]

4. **rings of functions** (important class)
   Let X be any nonempty set.
   Let A be any ring.

Definition 17 (function ring). *collection $R = \{f : X \to A\}$ is a ring under pointwise addition and multiplication of functions s.t.*

(18) $$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ (fg)(x) &= f(x)g(x) \end{aligned}$$

cf. Example 6, pp. 225, Dummit and Foote (2014)[2]

(5) cf. Example 7 of p. 225, Dummit and Foote (2014)[2] *ring which doesn't have an identity*:
- ring $2\mathbb{Z}$ of even integers (sum and product of even integers is an even integer)

- function $f:\mathbb{R}\to\mathbb{R}$ has compact support if $\exists\, a,b\in\mathbb{R}$, (depending on f) s.t. $f(x)=0, \quad \forall\, x\notin[a,b]$
 $\{f:\mathbb{R}\to\mathbb{R}|f$ has compact support $\}$ is a commutative ring without identity (since an identity couldn't have compact support),
 Similarly,
 $\{$ cont. $f:\mathbb{R}\to\mathbb{R}|f$ has compact support $\}$ is a commutative ring without identity.

Proposition 6 (Dummit and Foote (2014)[2], Prop. 1, pp. 226)**.** *Let R be a ring. Then*

(1) $0a=a0=0 \quad \forall\, a\in R$
(2) $(-a)0=a(-b)=-(ab) \quad \forall\, a,b\in R$ *(recall $-a$ is additive inverse of a)*
(3) $(-a)(-b)=ab \qquad \forall\, a,b\in R$
(4) *if R has identity 1, identity unique and $-a=(-1)a$*

Proof. Use distributivity and additive existence inverse for abelian group $(R,+)$.

(1)
$$0a=(0+0)a=0a+0a\to 0a=0a0=a(0+0)=a0+a0 \text{ so } a0=0$$
(2)
$$ab+(-a)b=(a+(-a))b=0b=0\Longrightarrow(-a)b=-(ab)$$
$$ab+a(-b)=a(b+(-b))=a0=0\Longrightarrow a(-b)=-(ab)$$
(3)
$$(-a)(b+(-b))=(-a)b+(-a)(-b)=-(ab)+(-a)(-b)=0\Longrightarrow(-a)(-b)=ab$$
(4)
$$(-1)a+a=(-1)a+1a=(-1+1)a=0a=0 \text{ so } -a=(-1)a$$
Suppose $\exists\, 1'$ s.t. $1'a=a1'=a$ and $1'\neq 1$
$$-a+a=0=(-1)a+1'a=(-1+1')a=0a\Longrightarrow 0=-1+1'$$
Then $1'=1$. Contradiction. □

Definition 18. *Let R be a ring.*

(1) ***zero divisor*** *a, $a\neq 0$, $a\in R$ if $\exists\, b\in R$, $b\neq 0$ so either $ab=0$ or $ba=0$.*
(2) *Assume R has identity $1\neq 0$.*
 unit *in R, $u\in R$ if $\exists$ some $v\in R$ s.t. $uv=vu=1$.*
 $R^x\equiv$ set of units in R.
$$\forall\, u,v\in R^x,\ (uv)(v^{-1}u^{-1})=u(1)u^{-1}=1$$
 $1\in R^x$ so, $1u=u1=u$
 *if $u\in R^x$, $u^{-1}\in R^x$ since $uu^{-1}=u^{-1}u=1$ R^x is a **group of units***

Thus a *field* = commutative ring F with identity $1\neq 0$ s.t. $\forall\, a\in$ field, $a\neq 0$ is a unit, i.e. $F^x=F-\{0\}$.

Proposition 7. *zero divisor can never be a unit.*

Proof. Suppose a unit in R, and $ab=0$ for some $b\neq 0$, $b\neq R$.
Then $va=1$ s.t. some $v\in R$ and
$$b=1b=vab=v(ab)=v(0)=0$$
Contradiction. □

5. Commutative Rings

cf. Ch. 3 "Commutative Rings I" of Rotman (2010) [29]

Definition 19. *commutative ring R is a set with 2 binary operations, addition and multiplication, s.t.*

(i) *R abelian group under addition*
(ii) *(commutativity) $ab=ba \quad \forall\, a,b\in R$ (this isn't there for noncommutativity)*
(iii) *(associativity) $a(bc)=(ab)c \quad \forall\, a,b,c\in R$*
(iv) *$\exists\, 1\in R$ s.t. $1a=a \quad \forall\, a\in R \qquad$ (many names used: one, unit, identity)*
(v) *(distributivity) $a(b+c)=ab+ac \qquad a,b,c\in R$ (this splits up into 2 distributivity laws for noncommutativity)*

To reiterate, abelian group under addition R (is defined as)
(1) associative $\forall\, x,y,z\in R$, $x+(y+z)=(x+y)+z$
(2) $\exists\, 0\in R$, $0+x=x+0$, $\forall\, x\in R$
(3) $\forall\, x\in R$, $\exists\,(-x)\in R$ s.t. $x+(-x)=0=(-x)+x$

abelian, if commutativity: $x+y=y+x$.

5.1. Linear Algebra; Linear Algebra with commutative rings as fields.

5.1.1. *Linear Algebra.*

Definition 20 (subspace)**.** *If V vector space over field k, then **subspace** of V is subset U of V s.t.*

(1) *$0\in U$*
(2) *$u,u'\in U$ imply $u+u'\in U$*
(3) *$u\in U$, and $a\in k$ imply $au\in U$*

proper subspace *of $V\equiv U\subsetneq V$ is subspace $U\subseteq V$ with $U\neq V$.*

$U=V$, $U=\{0\}$ are always subspaces of a vector space V.
Examples (Example 3.70 Rotman (2010) [29])

(ii) If $V=(a_1,\dots a_n)$, $v\neq 0$, $v\in\mathbb{R}^n$,
 line through origin $l=\{av|a\in\mathbb{R}\}$ is a subspace of $\mathbb{R}^n$.
 plane through origin $=\{av_1+bv_2|v_1,v_2$ fixed pair of noncollinear vectors, $a,b\in\mathbb{R}\}$ are subspaces of $\mathbb{R}^n$
(iii) If $m\leq n$, $\mathbb{R}^m$ regarded as set of all vectors in $\mathbb{R}^n$ s.t. last $n-m$ coordinates are 0, then $\mathbb{R}^m$ subspace of $\mathbb{R}^n$. e.g. $\mathbb{R}^2=\{(x,y,0)\in\mathbb{R}^3\}\subsetneq\mathbb{R}^3$
(iv) If k field, **homogeneous linear system over** k of m equations in n unknowns is a set of equations
$$a_{11}x_1+\dots+a_{1n}x_n=0$$
$$a_{21}x_1+\dots+a_{2n}x_n=0$$
$$\vdots$$
$$a_{m1}x_1+\dots+a_{mn}x_n=0$$
 where $a_{ji}\in k$.
 solution of this system is vector $(c_1\dots c_n)\in k^n$ s.t. $\sum_i a_{ji}c_i=0$, $\forall\, j$.
 solution $(c_1\dots c_n)$ **nontrivial** if $\exists$ some $c_i\neq 0$.
 solution space (or null space) of system = set of all solutions.
 solution space also a subspace of k^n

e.g. $k=\mathbb{I}_p$,
$$3x-2y+z\equiv 1 \bmod 7$$
$$x+y-2z\equiv 0 \bmod 7$$
$$-x+2y+z\equiv 4 \bmod 7$$

Definition 21 (list). *list := vector space V is ordered set $v_1 \dots v_n$ of vectors in V, i.e. $\exists$ some $n \geq 1$, $\exists$ some function φ*

$$\varphi : \{1, 2 \dots n\} \to V$$

with $\varphi(i) = v_i \quad \forall i$

Thus, $X = \text{im}\varphi$.

X ordered, φ need not be injective.

Definition 22 (k-linear combination). *k-linear combination of list $v_1 \dots v_n$ in V, $V \equiv$ vector space over field k, is vector v of form*

$$v = a_1 v_1 + \dots + a_n v_n = \sum_{i=1} a_i v_i \quad \forall a_i \in k, \quad \forall i$$

Definition 23 (list). *If list $X = v_1 \dots v_m$ in vector space V, then* ***subspace spanned by*** *X, $\langle v_1 \dots v_m \rangle :=$ set of all k-linear combinations of $v_1 \dots v_m$. Also, say $v_1 \dots v_m$ spans $\langle v_1 \dots v_m \rangle$.*

Lemma 1 ($\langle v_1 \dots v_m \rangle$ is smallest subspace of V containing $v_1 \dots v_m$). (i) *Every intersection of subspaces of V is itself a subspace.*

(ii) *If $X = v_1 \dots v_m$ list in V, then intersection of all subspaces of V containing X is $\langle v_1 \dots v_m \rangle$, subspace spanned by $v_1 \dots v_m$, so $\langle v_1 \dots v_m \rangle$ is smallest subspace of V containing X.*

cf. (Lemma 3.71 Rotman (2010) [29])

Proof. (i) Consider $\bigcap_{\alpha \in I} V_\alpha$, $\forall \alpha \in I$, V_α subspace of V

(i) $0 \in V_\alpha$, $\forall \alpha \in I$, so $0 \in \bigcap_{\alpha \in I} V_\alpha$,

(ii) Let $u, u' \in \bigcap_{\alpha \in I} V_\alpha$. Then $u, u' \in V_\alpha$, $\forall \alpha \in I$. Consider $\beta \in I$. $u, u' \in V_\beta$, so $u + u' \in V_\beta$. Without loss of generality, $u + u' \in V_\alpha$, $\forall \alpha \in I$. Then $u + u' \in \bigcap_{\alpha \in I} V_\alpha$

(iii) Let $u \in \bigcap_{\alpha \in I} V_\alpha$. Consider $\alpha \in k$. Since $u \in V_\alpha$, $\forall \alpha \in I$, $au \in V_\alpha$, $\forall \alpha \in I$. Then $au \in \bigcap_{\alpha \in I} V_\alpha$

(ii) Let $X = \{v_1 \dots v_m\}$, let $\mathcal{S} \equiv$ family of all subspaces of V containing X.

$\bigcap_{S \in \mathcal{S}} S \subseteq \langle v_1 \dots v_m \rangle$ because $\langle v_1 \dots v_m \rangle \in \mathcal{S}$, since,

$\langle v_1 \dots v_m \rangle$ is *a* subspace of V containing X.

If $S \in \mathcal{S}$, then $S \ni v_1 \dots v_m$. As shown above, $\forall v \in \langle v_1 \dots v_m \rangle$, $v \in S$, and thus $v \in \bigcap_{S \in \mathcal{S}} S$. $\langle v_1 \dots v_m \rangle \subseteq \bigcap_{S \in \mathcal{S}} S$. □

Were all terminology in algebra consistent,
$\langle v_1 \dots v_m \rangle \equiv$ subspace *generated* by X.

Reason for different terms is that group theory, rings, vector spaces developed independently of each other.

Example 3.72 of Rotman (2010) [29]

(i)
(ii)
(iii) **polynomial vector space; polynomials as a vector space**.
Vector space need not be spanned by finite list.
e.g. $V = k[x]$,
Suppose $X = f_1(x) \dots f_m(x)$ finite list in V.
If $d =$ largest degree of any of $f_i(x)$,
then every (nonzero) k-linear combination of $f_1(x), \dots f_m(x)$ has degree at most d.
Thus $x^{d+1} \notin \langle f_1(x) \dots f_m(x) \rangle$, so X doesn't span $k[x]$

Definition 24 (finite-dimensional vector space; infinite-dimensional vector space). *Vector space V is* ***finite-dimensional*** *if it's spanned by a finite list; otherwise V is* ***infinite-dimensional***.

Proposition 8 (linear dependent span properties). *If vector space V, list $X = v_! \dots v_m$ spanning V, following are equivalent:*

(i) *X isn't shortest spanning list*
(ii) *some v_i is in subspace spanned by others, i.e. $v_i \in \langle v_i \dots \widehat{v}_i \dots v_m \rangle$,*
(iii) *$\exists a_1 \dots a_m$ not all 0 s.t. $\sum_{l=1}^m a_l v_l = 0$*

Proof. (i) $\Longrightarrow$ (ii). If X isn't hosrtest spanning list, then 1 of vectors in X can be thrown out, and shorter list still spans, i.e. cf. Lemma 1(Lemma 3.71, Rotman (2010) [29]); let $\mathcal{S} \equiv$ family of all subspaces of V containing X.

EY: 20180610 Let $\bigcap_{S \in \mathcal{S}} S$. $\bigcap_{S \in \mathcal{S}} S \neq \langle v_1 \dots v_m \rangle$, $\bigcap_{S \in \mathcal{S}} S \subset \langle v_1 \dots v_m \rangle$
$\exists v \in \langle v_1 \dots v_m \rangle$, say$v = \sum_{i=1}^m a_i v_i$ s.t. $\exists S \in \mathcal{S}$, s.t. $v \notin S$.

(ii) $\Longrightarrow$ (iii) If $v_i = \sum_{j \neq i} c_j v_j$, define $a_i = -1 \neq 0$, $a_j = c_j$, $\forall j \neq i$. Then $\sum_{l=1}^m a_l v_l = -v_i + \sum_{j \neq i} c_j v_j = 0$

(iii) $\Longrightarrow$ (i) Suppose for $i \in 1 \dots m$, $a_i \neq 0$. $v_i = -\sum_{j \neq i} \frac{a_j}{a_i} v_j$. $\langle v_1 \dots \widehat{v}_i \dots v_m \rangle$ still spans V (i.e. deleting v_i gives a shorter list, which still spans).

For instance, if $v \in \langle v_1 \dots v_m \rangle$, $v = \sum_{l=1}$ □

Exercise 3.67. Suppose $\dim V > 1$. Then $\exists$ at least 2 elements in a basis of V, say e_1, e_2. (Thm. 3.78 of Rotman (2010) [29], "Every finite-dim. vector space V has a basis; Def. of dim, "number of elements in a basis of V").

Consider subspaces $\langle e_1 \rangle$, $\langle e_2 \rangle$, subspaces spanned by e_1, e_2, respectively. Whether $V = \langle e_1, e_2 \rangle$ or $V = \langle e_1, e_2 \rangle$, $\langle e_1 \rangle, \langle e_2 \rangle \neq \{0\}$ nor V. Contradiction of hypothesis.

Thus, "If only subspaces of a vector space V are $\{0\}$ and V itself, $\dim(V) \leq 1$."

Proposition 9 (Matrix representation of linear transformation; 3.94 of Rotman (2010) [29]). *If linear transformation $T : k^n \to k^m$, then $\exists A \in Mat_k(m, n)$ s.t.*

$$T(y) = Ay, \qquad \forall y \in k^n$$

Proof. Let $\begin{matrix} (e_1 \dots e_n) & \text{standard basis of } k^n \\ (e_1' \dots e_m') & \text{standard basis of } k^m \end{matrix}$

Define $A = [a_{ij}]$, s.t. $T(e_j) = A_{*j} = A_{ij} e_i'$ (jth column),

If $\begin{matrix} S : k^n \to k^m \\ S(y) = A(y) \end{matrix}$, then

$$T(e_j) = a_{ij} e_i' = A e_j$$

and so $\forall y = y_j e_j \in k^n$,

$$T(y) = T(y_j e_j) = y_j T(e_j) = y_j A_{ij} e_i' = Ay$$

□

6. Modules

6.1. R-modules.

cf. Sec. 7.1 Modules of Rotman (2010) [29]

Definition 25 (R-module). *R-module is (additive) abelian group M,*

equipped with scalar multiplication $\begin{matrix} R \times M \to M \\ (r, m) \mapsto rm \end{matrix}$

s.t. $\forall m, m' \in M$, $\forall r, r', 1 \in R$

(i) $r(m + m') = rm + rm'$
(ii) $(r + r')m = rm + r'm$
(iii) $(rr')m = r(r'm)$
(iv) $1m = m$

Example 7.1

(i) $\forall$ *vector space* over field k is a k-module. (by inspection of the axioms for a vector space, associativity, distributivity!)

(ii) $\forall$ abelian group is a $\mathbb{Z}$-module, by laws of exponents (Prop. 2.23)
Indeed, for

$$\begin{aligned} \mathbb{Z} \times M &\to M \\ (r,m) &\mapsto rm \equiv m^r \end{aligned}$$

and so

$$r(m \cdot m') \equiv (m \cdot m')^r = m^r (m')^r = rm + rm'$$

(since M abelian)

(iii) For commutative ring, scalar multiplication, defined to be given multiplication of elements of R

$$\begin{aligned} R \times R &\to R \\ (a,b) &\mapsto ab \end{aligned}$$

For reference, recall some of the properties of a commutative ring:

$$\begin{aligned} ab &= ba \\ a(bc) &= (ab)c \\ 1a &= a \\ a(b+c) &= ab + ac \end{aligned}$$

$\forall$ ideal I in R is an R-module,

for if $i \in I$, $r \in R$, then $ri \in I$.

$0 \in I$

$\forall a, b \in I$, $a + b \in I$

If $a \in I$, $r \in R$, then $ra \in I$.

(iv)

(v) Let linear $T : V \to V$, V finite-dim. vector space over field k.
Recall $k[x] \equiv$ set of polynomials with coefficients in k.

Define
$$\begin{aligned} k[x] \times V &\to V \\ f(x)v &= \left(\sum_{i=0}^m c_i x^i \right) v = \sum_{i=0}^m c_i T^i(v) \end{aligned} \qquad \forall f(x) = \sum_{i=0}^m c_i x^i \in k[x]$$

$\Longrightarrow$ denote $k[x]$-module V^T.

Special case: Let $A \in \text{Mat}_k(n,n)$, let linear $T : k^n \to k^n$.

$$T(w) = Aw$$

vector space k^n is $k[x]$-module if we define scalar multiplication

$$\begin{aligned} k[x] \times k^n &\to k^n \\ f(x)w &= \left(\sum_{i=0}^m c_i x^i \right) w = \sum_{i=0}^m c_i A^i w \end{aligned}$$

$\forall f(x) = \sum_{i=0}^m c_i x^i \in k[x]$

In $(k^n)^T$, $xw = T(w)$

In $(k^n)^A$, $xw = Ax$

$T(w) = Ax$ and so $(k^n)^T = (k^n)^A$ (EY : 20151015 because of induction?)

Definition 26 (R-homomorphism (or R-map))**.** *If ring R, R-modules M, N, then function $f : M \to N$, if $\forall m, m' \in M$, $\forall r \in R$,*

$$\begin{aligned} f(m + m') &= f(m) + f(m') \\ f(rm) &= rf(m) \end{aligned}$$

Example 7.2. of Rotman (2011) on pp. 425 [29]]

(i) If R field, then R-modules are vector spaces and R-maps are linear transformations. Isomorphisms are then nonsingular linear transformations.

(ii)

(iii)

(iv)

(v) Let linear $T : V \to V$, let $v_1 \dots v_n$ be basis of V, let A be matrix of T relative to this basis.
Let $e_1 \dots e_n$ be standard basis of k^n.

Define $\varphi : V \to k^n$

$\varphi(v_i) = e_i$

$$\begin{aligned} \varphi(xv_i) &= \varphi(T(v_i)) = \varphi(v_j a_{ji}) = a_{ji}\varphi(v_j) = a_{ji} e_j \\ x\varphi(v_i) &= A\varphi(v_i) = Ae_i \end{aligned}$$

$\Longrightarrow \varphi(xv) = x\varphi(v) \quad \forall v \in V$

By induction on $\deg f$, $\varphi(f(x)v) = f(x)\varphi(v) \qquad \forall f(x) \in k[x] \qquad \forall v \in V$

$\Longrightarrow \varphi$ is $k[x]$-map

$\Longrightarrow \varphi$ is $k[x]$-isomorphism of V^T and $(k^n)^A$.

Proposition 10 (7.3 of Rotman (2011) [29])**.** *Let vector space over field k, V, let linear $T, S : V \to V$
Then $k[x]$-modules V^T, V^S are $k[x]$-isomorphic iff $\exists$ vector space isomorphism $\varphi : V \to V$ s.t. $S = \varphi T \varphi^{-1}$.*

Proof. If $\varphi : V^T \to V^S$is a $k[x]$-isomorphism,

$$\varphi(f(x)v) = f(x)\varphi(v) \quad \forall v \in V, \forall f(x) \in k[x]$$

if $f(x) = x$, then $\varphi(xv) = x\varphi(v)$

$$\begin{aligned} xv &= T(v) \\ x\varphi(v) &= S(\varphi(v)) \\ \Longrightarrow & \varphi \circ T(v) = S \circ \varphi(v) \Longrightarrow \varphi \circ T = S \circ \varphi \end{aligned}$$

φ isomorphism, so $S = \varphi \circ T \circ \varphi^{-1}$

Conversely, if given isomorphism $\varphi : V \to V$ s.t. $S = \varphi T \varphi^{-1}$, then $S\varphi = \varphi T$.

$$S\varphi(v) = \varphi T(v) = \varphi(xv) = x\varphi(v)$$

Then by induction, $\varphi(x^n v) = x^n \varphi(v)$ (for $S^n \varphi(v) = x^n \varphi(v) = (\varphi T \varphi^{-1})^n \varphi(v) = \varphi T^n v = \varphi(x^n v)$).
By induction on $\deg(f)$, $\varphi(f(x)v) = f(x)\varphi(v)$. □

Corollary 2 (7.4 of Rotman (2011) [29])**.** *Let k be a field,
Let $A, B \in Mat_k(n,n)$.
Then $k[x]$-modules $(k^n)^A$, $(k^n)^B$ are $k[x]$-isomorphic.
(recall, $k[x] \equiv$ set of polynomials with coefficients in $k = \{\sum_{i=0}^m c_i x^i | c_i \in k\}$, and define scalar multiplication*

$$\begin{aligned} k[x] \times k^n &\to k^n \\ f(x)w &= \left(\sum_{i=0}^m c_i x^i \right) w = \sum_{i=0}^m c_i A^i w, \qquad \forall f(x) = \sum_{i=0}^m c_i x^i \in k[x] \end{aligned}$$

)

iff $\exists$ nonsingular P with

$$B = PAP^{-1}$$

Proof. Define

$$T : k^n \to k^n$$
$$T(y) = A(y) \quad \text{where } y \in k^n \text{ is a column.}$$

By Example 7.1 (v) of Rotman (2011) [29], recall,
and so for $k[x]$-module, $(k^n)^T = (k^n)^A$.
Similarly, define

$$S : k^n \to k^n$$
$$S(y) = B(y)$$

Denote corresponding $k[x]$-module by $(k^n)^B$.
Given $(k^n)^A \cong (k^n)^B$ (isomorphic), by Prop. 10,
$\exists$ isomorphism $\varphi : k^n \to k^n$ s.t. $B = \varphi A \varphi^{-1}$.
By Prop. 9, i.e. Prop. 3.94 of Rotman (2011) [29], in that every linear transformation has a matrix representation (even in the standard "Euclidean" basis), $\exists P \in \text{Mat}_k(n, n)$, s.t.

$$\varphi(y) = Py \qquad y \in k^n$$

(P nonsingular because φ isomorphism)
Thus,

$$B\varphi(y) = \varphi A(y)$$
$$BPy = P(Ay) \qquad \forall y \in k^n$$
$$\Longrightarrow PA = BP \text{or } B = PAP^{-1}$$

Conversely, given $B = PAP^{-1}$, P nonsingular matrix,
define isomorphism

$$\varphi : k^n \to k^n$$
$$\varphi(y) = Py \qquad \forall y \in k^n$$

By Prop. 10,
$(k^n)^B$, $(k^n)^A$ are $k[x]$-isomorphic.
i.e. $\varphi : (k^n)^A \to (k^n)^B$ is a $k[x]$-module isomorphism.

□

Definition 27 ($\text{Hom}_R(M, N)$)**.**
(19)

$$Hom_R(M,N) = \{ \text{ all } R\text{-homomorphisms } M \to N \} = \{ f | f : M \to N, \text{ s.t. } \forall m, m' \in M, \forall r \in R, \begin{matrix} f(m+m') = f(m) + f(m') \\ f(rm) = rf(m) \end{matrix} \}$$

If $f, g \in Hom_R(M,N)$,
define

$$(20) \qquad \begin{aligned} f + g &: M \to N \\ f + g &: m \mapsto f(m) + g(m) \end{aligned}$$

Proposition 11 ($\text{Hom}_R(M,N)$ R-module, 7.5 of Rotman (2011) [29])**.** *If M, N R-modules, where R commutative ring, then $Hom_R(M,N)$ R-module,*
with addition

$$f + g : M \to N \qquad \forall f, g \in Hom_R(M,N)$$
$$f + g : m \mapsto f(m) + g(m)$$

and scalar multiplication

$$rf : m \mapsto f(rm)$$

Moreover, distributive laws:
If $p : M' \to M$, $q : N \to N'$, then

$$(f+g)p = fp + gp \text{ and } q(f+g) = qf + qg$$

$\forall f, g \in Hom_R(M,N)$

Proof. $\forall f, g \in \text{Hom}_R(M,N)$, $\forall r, r', 1 \in R$,

(i)
$$r(f+g)(m) = (f+g)(rm) = f(rm) + g(rm) = rf(m) + rg(m) = (rf + rg)(m)$$

(ii)
$$(r + r')f(m) = f((r+r')m) = f(rm + r'm) = f(rm) + f(r'm) = (rf + r'f)(m)$$

(iii)
$$(rr')f(m) = f(rr'm) = rf(r'm) = f(r'rm) = f(rr'm) \Longrightarrow (rr')f = r(r'f)$$

(iv)
$$1f(m) = f(1m) = f(m) \Longrightarrow 1f = f$$

□

Definition 28. *if R-module M, the submodule N of M, denoted $N \subseteq M$, is additive subgroup N of M, closed under scalar multiplication $rn \in N$ whenever $n \in N$, $r \in R$*

Definition 29 (quotient module M/N)**.**
***quotient module** M/N -*
For submodule N of R-module M, then,
remember M abelian group, N subgroup,
quotient group M/N equipped with scalar multiplication

$$r(m+N) = rm + N$$
$$M/N = \{m + N | m \in M\}$$

natural map

$$(21) \qquad \begin{aligned} \pi &: M \to M/N \\ m &\mapsto m + N \end{aligned}$$

easily seen to be R-map.
Scalar multiplication in quotient module well-defined:
If $m + N = m' + N$, $m - m' \in N$, so $r(m - m') \in N$ (because N submodule), so

$$rm - rm' \in N \text{ and } rm + N = rm' + N$$

Proposition 12 (7.15 of Rotman (2010) [29])**.** (i) $S \bigsqcup T \simeq M$

(ii) $\exists$ *injective R-maps* $\begin{matrix} i : S \to M \\ j : T \to M \end{matrix}$, *s.t.*

$$(22) \qquad \begin{aligned} M &= im(i) + im(j) \text{ and} \\ im(i) &\bigcap im(j) = \{0\} \end{aligned}$$

(iii) $\exists$ *R-maps*

$$\begin{aligned} i &: S \to M \\ j &: T \to M \end{aligned}$$

s.t. $\forall\, m \in M$, $\exists!$

$$\begin{aligned} s &\in S \\ t &\in T \end{aligned}$$

with $m = is + jt$.

(iv) $\exists$ *R-maps*

$$\begin{aligned} i &: S \to M & p &: M \to S \\ j &: T \to M & q &: M \to T \end{aligned}$$

s.t.

$$\begin{aligned} pi &= 1_S & pj &= 0 \\ qj &= 1_T & , qi &= 0 \end{aligned} \quad ip + jq = 1_M$$

Proof.

- (i)$\to$ (ii) Given $S \bigsqcup T \simeq M$,

let $\varphi : S \bigsqcup T \to M$ be this isomorphism.

Define

$$\begin{aligned} i &:= \varphi \lambda_S & (\lambda_S &: s \mapsto (s,0)) & i &: S \to M \\ j &:= \varphi \lambda_T & (\lambda_T &: t \mapsto (0,t)) & j &: T \to M \end{aligned}$$

i,j are injections, being composites of injections.

If $m \in M$, $\exists!\,(s,t) \in S \bigsqcup T$, s.t. $\varphi(s,t) = m$.

Then

$$m = \varphi(s,t) = \varphi((s,0) + (0,t)) = \varphi\lambda_S(s)\varphi\lambda_T(t) = is + jt \in \text{im}(i) + \text{im}(j)$$

Let $c \in \text{im}(i) + \text{im}(j)$. Since $\begin{aligned} i &: S \to M \\ j &: T \to M \end{aligned}$, $c \in M$.

$\Longrightarrow M = \text{im}(i) + \text{im}(j)$.

If $x \in \text{im}(i) \bigcap \text{im}(j)$,

$$\begin{aligned} x &= i(s) \text{ for some } s \in S \\ x &= j(t) \text{ for some } t \in T \end{aligned}$$

$$is = jt = \varphi\lambda_S(s) = \varphi\lambda_T(t) = \varphi(s,0) = \varphi(0,t)$$

φ isomorphism, so $\exists\, \varphi^{-1} \Longrightarrow (s,0) = (0,t)$, so $s = t = 0$. $x = 0$

- (ii)$\to$ (iii) Given $\begin{aligned} i &: S \to M \\ j &: T \to M \end{aligned}$, s.t. $M = \text{im}(i) + \text{im}(j)$, so

$\forall\, m \in M$, $m = i(s) + j(t)$ for some $s \in S, t \in T$.

Suppose $\begin{aligned} s' &\in S \\ t' &\in T \end{aligned}$, s.t. $m = i(s'_+ j(t')$.

$$i(s - s') = j(t - t') \in \text{im}(i) \bigcap \text{im}(j) = \{0\}$$

So $s = s', t = t'$, since i,j injective.

- (iii)$\to$ (iv)

Given $\forall\, m \in M$, $\exists!\, s \in S, t \in T$ s.t.

$$m = i(s) + j(t)$$

Define

$$\begin{aligned} p &: M \to S & q &: M \to T \\ p(m) &:= s & q(m) &:= t \end{aligned}$$

$$\begin{aligned} pi(s) &= s & pj(t) &= 0 \\ qj(t) &= t & qi(s) &= 0 \end{aligned} \quad (ip + jq)(m) = ip(m) + jq(m) = i(s) + j(t) = m$$

$\square$

6.2. **Vector Spaces as a Module.** Lang made the key insight on vector spaces as a whole in Sec 5. "Vector Spaces" in pp. 139-140 of Lang (2005) [30]:

Theorem 13 (Existence of a basis for vector spaces, Thm. 5.1 Lang (2005) [30])**.** *Let V be a vector space over a field K, assume $V \neq \{0\}$.*

Let Γ be a set of generators of V over K and let S be a subset of Γ which is linearly independent.

Then $\exists$ basis $\mathcal{B}$ of V s.t. $S \subset \mathcal{B} \subset \Gamma$.

Indeed, while this wikipedia article [1] on Vector space does a good job generalizing the properties defining a vector in a vector space, a vector's properties is separate from what *characterizes* a vector space. Here, we can *specify* a vector space by its generators, and furthermore, from Thm. 13, it has a basis that characterizes a vector space. This can be useful for implementation in C++.

7. Categories; Category Theory

7.1. **Categories.** cf. 7.2 Categories of Rotman (2010) [29]

7.1.1. *Russell paradox, Russell set.*

Definition 30 (Russell set)**.** *Russell set - set S that's not a member of itself, i.e. $S \notin R$*

If R is family of all Russell sets,
Let $X \in R$. Then $X \notin X$. But $X \in R$. $X \notin R$.
Let $R \notin R$. Then R in family of Russell Sets. $R \in R$. Contradiction.
Then consider *class* as primitive term, instead of set.

Definition 31 (Category)**.** *Category $\mathcal{C}$ (Rotman's notation) $\equiv$ **C** (my notation), consists of class obj($\mathcal{C}$) (Rotman's notation) $\equiv$ Obj(**C**) $\equiv$ Obj**C** (my notation) of objects, set of morphisms Hom(A,B) $\forall\,(A,B)$ of ordered tuples of objects, composition*

$$\begin{aligned} Hom(A,B) \times Hom(B,C) &\to Hom(A,C) \\ (f,g) &\mapsto gf \end{aligned}$$

, s.t.

(1) $\exists\, \mathbf{1}$, $\forall\, f : A \to B$, $\exists\, \begin{aligned} 1_A &: A \to A \\ 1_B &: B \to B \end{aligned}$ *, s.t.* $1_B \cdot f = f = f \cdot 1_A$*, and*

(2) *associativity,* $\forall \begin{aligned} f &: A \to B \\ g &: B \to C \\ h &: C \to D \end{aligned}$*, then* $h \circ (g \circ f) = (h \circ g) \circ f$

[1] https://en.wikipedia.org/wiki/Vector_space

In summary,

$$\mathbf{C} := (Obj(\mathbf{C}), Mor\mathbf{C}, \circ, \mathbf{1}) \equiv (Obj\mathbf{C}, Mor\mathbf{C}, \circ_{\mathbf{C}}, \mathbf{1}_{\mathbf{C}}) \tag{23}$$

s.t.

$$Mor\mathbf{C} = \bigcup_{A,B \in Obj\mathbf{C}} Hom(A,B)$$

Examples (7.25 of Rotman (2010)[29]):

(i) $\mathbf{C} = \text{Sets}$
(ii) $\mathbf{C} = \text{Groups} = \text{Grps}$
(iii) $\mathbf{C} = \text{CommRings}$
(iv) $\mathbf{C} = {}_R\mathbf{Mod}$, if $R = \mathbb{Z}$, ${}_{\mathbb{Z}}\mathbf{Mod} = \mathbf{Ab}$, i.e. $\mathbb{Z}-$modules are just abelian groups.
(v) $\mathbf{C} = \mathbf{PO}(X)$, If partially ordered set X, regard X as category, s.t. $\mathbf{Obj}, \mathbf{PO}(X) = \{x | x \in X\}$, $\forall \text{Hom}(x,y) \in \mathbf{Mor}_{\mathbf{PO}}(X)$, $\text{Hom}(x,y) = \begin{cases} \emptyset & \text{if } x \not\preceq y \\ \kappa^x_y & \text{if } x \preceq y \end{cases}$ where $\kappa^x_y \equiv$ unique element in Hom set when $x \preceq y$) s.t.

$$\kappa^y_z \kappa^x_y = \kappa^x_z$$

Also, notice that

$$1_x = \kappa^x_x$$

Definition 32 (isormorphisms or equivalences)**.** *$f : A \to B$, $f \in Hom(A,B)$, if $\exists$ **inverse** $g : B \to A$, $g \in Hom(B,A)$, s.t.*

$$gf = 1_A$$
$$fg = 1_B$$

*and if $\mathbf{C} =$ **Top**, equivalences (isomorphisms) are homeomorphisms.*

Feature of category ${}_R\mathbf{Mod}$ not shared by more general categories: *Homomorphisms can be added.*

Definition 33 (pre-additive Category)**.** *category* $\mathbf{C}$

We can force 2 overlapping subsets A, B to be disjoint by "disjointifying" them: e.g. consider $(A \cup B) \times \{1,2\}$, consider

$A' = A \times \{1\}$.
$B' = B \times \{2\}$

$$\Longrightarrow A' \cap B' = \emptyset$$

since $(a,1) \neq (b,2) \quad \forall a \in A, \forall b \in B$.

Let bijections $\alpha : A \to A'$, $\quad \alpha : a \mapsto (a,1)$, denote $A' \bigcup B' \equiv A \coprod B$.
$\beta : B \to B' \quad \beta : b \mapsto (b,2)$

From Rotman (2010) [29], pp. 447,

Definition 34. ***coproduct*** *$A \coprod B \equiv C \in Obj(\mathcal{C})$*

In my notation,
coproduct

$$\begin{aligned} &(\mu_1, A_1 \coprod A_2) \\ &(\mu_2, A_1 \coprod A_2) \end{aligned} \tag{24}$$

where injection (morphisms)

$$\begin{aligned} \mu_1 &: A_1 \to A_1 \coprod A_2 \\ \mu_2 &: A_1 \to A_1 \coprod A_2 \end{aligned} \tag{25}$$

s.t.

$$\forall A \in \text{Obj}\mathbf{A}, \forall f_1, f_2 \in \text{Mor}\mathbf{A} \text{ s.t. } \begin{aligned} f_1 &: A_1 \to A \\ f_2 &: A_2 \to A \end{aligned}$$

then

$$\begin{aligned} &\exists ! [f_i] \equiv [f_1, f_2] \in \text{Mor}\mathbf{A}, [f_1,f_2] : A_1 \coprod A_2 \to A \text{ s.t.} \\ &[f_1,f_2]\mu_1 = f_1 \\ &[f_1,f_2]\mu_2 = f_2 \end{aligned} \tag{26}$$

i.e.

$$\begin{tikzcd} & A \\ A_1 \arrow[ur, "f_1"] \arrow[r, "\mu_1"] & A_1 \coprod A_2 \arrow[u, dashed, "{[f_1, f_2]}"'] \\ A_2 \arrow[uur, "f_2"] \arrow[ur, "\mu_2"'] & \end{tikzcd} \tag{27}$$

So to generalized, for $i \in I$, (finite set I?)
coproduct $(\mu_j, \coprod_{i \in I} A_i)_{j \in I}$, where
(family of) injection (morphisms) $\mu_j : A_j \to \coprod_{i \in I} A_i$
s.t.

$$\forall A \in \text{Obj}\mathbf{A}, \forall f_i \in \text{Mor}\mathbf{A}, i \in I, f_i : A_i \to A$$

then

$$\begin{aligned} &\exists ! [f_i] \equiv [f_i]_{i \in I} \in \text{Mor}\mathbf{A}, [f_i] : \coprod_{i \in I} A_i \to A \text{ s.t.} \\ &[f_i]\mu_j = f_j \qquad \forall j \in I \end{aligned} \tag{28}$$

i.e.

$$\begin{tikzcd} & A \\ A_j \arrow[ur, "f_j"] \arrow[r, "\mu_j"] & \coprod_{i \in I} A_i \arrow[u, dashed, "{[f_i]}"'] \end{tikzcd} \tag{29}$$

For notation purposes only, recall that it's denoted the sets $\text{Hom}(A,B)$ in ${}_R\mathbf{Mod}$ by

$$\text{Hom}_R(A,B)$$

i.e., in my notation, for $A,B \in \text{Obj}_R\mathbf{Mod}$, $\text{Hom}(A,B) \subset \text{Mor}({}_R\mathbf{Mod})$, $\text{Hom}(A,B) \equiv \text{Hom}_R(A,B)$

Definition 35 (pre-additive category)**.** *category* $\mathbf{C}$ *is **pre-additive** if* $\forall$ *Hom*(A,B), *Hom*(A,B) *equipped with binary operation* $+$ *s.t.* $\forall\, f,g \in$ *Hom*(A,B),

(1) *if* $p: B \to B'$, *then*

$$p(f+g) = pf + pg \in Hom(A,B')$$

(2) *if* $q: A' \to A$, *then*

$$(f+g)q = fq + gq \in Hom(A',B)$$

and

$$f+g = g+f \qquad \textit{(additive abelian)}$$

7.1.2. *Examples of extra assumptions on sets,* ${}_R$***Mod*** *we take for granted.* In Prop. 7.15(iii) Rotman (2010) [29],

direct sum $M = A \oplus B$ if $\exists$ homomorphisms
$$\begin{aligned} p &: M \to A & pi &= 1_A \\ q &: M \to B & qj &= 1_B, \\ i &: A \to M & pj &= 0 \\ j &: B \to M & qi &= 0 \end{aligned} \quad \text{s.t.}$$
$$ip + jq = 1_M$$

direct sum $M = A \oplus B$ uses property that morphisms can be added ${}_R\mathbf{Mod}$ has this property. **Sets** don't.

In Corollary 7.17,
direct sum in terms of arrows,
$\exists$ map $\rho : M \to S$ s.t. $\rho(s) = s$. Moreover $\ker\rho = \text{im} j$, $\text{im}\rho = \text{im} i$ and $\rho(s) = s$, $\forall\, s \in \text{im}\rho$.

$$S \xrightarrow{\ i\ } M \xleftarrow{\ j\ } T \qquad \text{and } M \simeq S \textstyle\coprod T,$$

where $i : s \mapsto s$ (i.e. inclusions)
$j : t \mapsto t$

This makes sense in **Sets**, but doesn't make sense in arbitrary categories because image of morphism may fail, e.g. $\text{Mor}(\mathcal{C}(G))$ are elements in $\text{Hom}(*,*) = G$, not functions.

Categorically, object S is (equivalent to) retract of object M, $S, M \in \text{Obj}\mathbf{C}$, if $\exists$ morphisms $i, p \in \text{Mor}(\mathbf{C})$, s.t.

$$i : S \to M$$
$$p : M \to S$$

s.t. $pi = 1_S$, $(ip)^2 = ip$ (for modules, define $\rho = ip$)

Definition 36 (free products)**.** ***free products*** *are coproducts in groups*

Prop. 7.26, Rotman (2010) [29]

Proposition 13 (7.26, Rotman)**.** *If A,B are R-modules,*
then their coproducts in ${}_R$***Mod*** *exists, and it's the direct sum* $C = A \coprod B$.

Proof. Define

$$\begin{aligned} \mu &: A \to C & \nu &: B \to C \\ \mu &: a \mapsto (a,c) & \nu &: b \mapsto (0,b) \end{aligned} \qquad \text{(Rotman's notation)} \begin{aligned} \alpha &: A \to C \\ \beta &: B \to C \end{aligned}$$

Let X be a module, $f : A \to X$, $g : B \to X$ homomorphisms ☐

Define

$$\theta : C \to X$$
$$\theta : (a,b) \mapsto f(a) + g(b)$$

$$\theta\mu(a) = \theta(a,0) = f(a)$$
$$\theta\nu(b) = \theta(0,b) = g(b)$$

so diagram commutes, i.e.

$$\begin{tikzcd} & X & \\ A \arrow[ur, "f"] \arrow[r, "\mu"] & C \arrow[u, dashed, "\theta"] & B \arrow[ul, "g"'] \arrow[l, "\nu"'] \end{tikzcd}$$

If $\psi : C \to X$ makes diagram commute,

$$\psi((a,0)) = f(a) \qquad \forall\, a \in A$$
$$\psi((0,b)) = g(b) \qquad \forall\, b \in B$$

and since ψ is a homomorphism, $\psi((a,b)) = \psi((a,0)) + \psi((0,b)) = f(a) + g(b) = \theta((a,b))$. $\psi = \theta$.

Prop. 7.27, Rotman (2010) [29]

Proposition 14 (7.27, Rotman)**.** *If category $\mathcal{C} = \mathbf{C}$, and if $A,B \in Obj\mathbf{C}$, then $\forall$ 2 coproducts of A,B, if they $\exists$, are equivalent.*

Proof. Suppose C,D coproducts of A,B. Suppose coproducts
$$\begin{aligned} \mu_A &: A \to C, & \nu_A &: A \to D \\ \mu_B &: B \to C, & \nu_B &: B \to D \end{aligned}$$

$$\begin{tikzcd} & D & \\ A \arrow[ur, "\nu_A"] \arrow[r, "\mu_A"] & C \arrow[u, dashed, "\theta"] & B \arrow[ul, "\nu_B"'] \arrow[l, "\mu_B"'] \end{tikzcd}$$

Just substitute $X = D$ in diagram above.

Then substitute again:

$$\begin{tikzcd} & C & \\ A \arrow[ur, "\mu_A"] \arrow[r, "\nu_A"] & D \arrow[u, dashed, "\psi"] & B \arrow[ul, "\mu_B"'] \arrow[l, "\nu_B"'] \end{tikzcd}$$

Then combine the 2 diagrams: $\psi\theta = 1_C$. Likewise by label symmetry of C,D, $\theta\psi = 1_D$.

Then C,D are equivalent. ☐

Exer. 7.29 on pp. 459 of Rotman (2010) [29]

Definition 37. *If $A, B \in Obj\mathbf{C}$, then their* ***product****; $A\prod B = P \in Obj\mathbf{C}$, and morphisms* $\begin{matrix} p : P \to A \\ q : P \to B \end{matrix}$ *s.t.* $\forall X \in Obj\mathbf{C}$,

$\forall\, f : X \to A \in Mor\mathbf{C}$,
$\quad g : X \to B \in Mor\mathbf{C}$
$\exists !\, \theta : X \to P$, *s.t.*

$$\begin{tikzcd} & X \arrow[dl, "f"'] \arrow[d, dashed, "\theta"] \arrow[dr, "g"] & \\ A & A\prod B \arrow[l, "p"'] \arrow[r, "q"] & B \end{tikzcd}$$

In the notation of Kashiwara and Schapira (2006) [1],

$$\begin{tikzcd} & Y \arrow[dl, "f_0"'] \arrow[d, dashed, "{[f_0 f_1]}"] \arrow[dr, "f_1"] & \\ X_0 & X_0\prod X_1 \arrow[l, "p_0"'] \arrow[r, "p_1"] & X_1 \end{tikzcd}$$

In general

$$\begin{tikzcd} & Y \arrow[dl, "f_j"'] \arrow[d, dashed, "{[f_i]_{i\in I}}"] \\ X_j & \prod_{i\in I} X_i \arrow[l, "p_i"] \end{tikzcd}$$

product *of X_i's,*

$$\prod_i X_i \equiv \prod_{i\in I} X_i$$

given by

(30) $$\prod_i X_i := \varprojlim \alpha$$

When $X_i = X$, $\forall\, i \in I$, denote product by $X^{\prod I} \equiv X^I$.

e.g. Cartesian product $P = A \times B$ of 2 sets A, B, $A, B \in \mathrm{Obj}\mathbf{Sets}$.
Define

$$\begin{matrix} p : A\times B \to A & q : A \times B \to B \\ p(a,b) \mapsto a & q(a,b) \mapsto b \end{matrix}$$

If $X \in \mathrm{Obj}\mathbf{Sets}$,

if $\begin{matrix} f : X \to A, \\ g : X \to B \end{matrix}$ then $\begin{matrix} \theta : X \to A \times B \\ \theta : x \mapsto (f(x), g(x)) \in A \times B \end{matrix}$

Proposition 15 (7.28 Rotman (2010); equivalence of products, if it exists)**.** *If $A, B \in Obj\mathbf{C}$, then $\forall$ 2 products of A and B, should they exist, are equivalent.*

Proof. Suppose C, D products of A, B. Suppose products $\begin{matrix} p_A : C \to A, & q_A : D \to A \\ p_B : C \to B, & q_B : D \to B \end{matrix}$

$$\begin{tikzcd} & D \arrow[dl, "q_A"'] \arrow[d, dashed, "\theta"] \arrow[dr, "q_B"] & \\ A & C \arrow[l, "p_A"'] \arrow[r, "p_B"] & B \end{tikzcd}$$

Just substitute $X = D$ in diagram above.
Then substitute again:

$$\begin{tikzcd} & C \arrow[dl, "p_A"'] \arrow[d, dashed, "\psi"] \arrow[dr, "p_B"] & \\ A & D \arrow[l, "q_A"'] \arrow[r, "q_B"] & B \end{tikzcd}$$

Then combine the 2 diagrams: $\psi\theta = 1_C$. Likewise by label symmetry of C, D, $\theta\psi = 1_D$.
Then C, D are equivalent. □

7.1.3. *Products of Modules and Sets.*

Proposition 16 (7.29 Rotman (2010); products of R-modules are equivalent)**.** *If commutative ring R,*
R-modules A, B,
then $\exists$ their (categorical) product $A \sqcup B$, in fact

(31) $$A \sqcap B \cong A \sqcup B$$

Proof. If $A \sqcup B \cong M$, then $\exists$ R-maps, $\begin{matrix} i : S \to M, & p : M \to S & \text{s.t. } pi = 1_A & \text{and } pj = 0, \text{ and } ip + jq = 1_M, \text{ i.e.} \\ j : T \to M & q : M \to T & qj = 1_B & qi = 0 \end{matrix}$

$$\begin{tikzcd} A \arrow[r, shift left, "i"] & M \arrow[l, shift left, "p"] \arrow[r, shift right, "q"'] & B \arrow[l, shift right, "j"'] \end{tikzcd}$$

If module X, since $\begin{matrix} f : X \to A \\ g : X \to B \end{matrix}$ are homomorphisms, define $\begin{matrix} \theta : X \to A \sqcup B \\ \theta(x) = if(x) + jg(x) \end{matrix}$ so that

$$\begin{tikzcd} & X \arrow[dl, "f"'] \arrow[d, dashed, "\theta"] \arrow[dr, "g"] & \\ A & A\sqcup \arrow[l, "p"'] \arrow[r, "q"] & B \end{tikzcd}$$

since, $\forall x \in X$,

$$p\theta(x) = pif(x) + pjg(x) = pif(x) + 0 = f(x)$$

since $ip + jq = 1_{A \sqcup B}$

$$\psi = ip\psi + jq\psi = if + jf = \theta$$

so product is unique. □

Definition 38. *Let R be commutative ring,*
let $\{A_i : i \in I\}$ be indexed family of R-modules.
***direct product** $\prod_{i\in I} A_i$ is cartesian product (i.e. set of all I-tuples (a_i) whose ith coordinate a_i lies in $A_i \quad \forall i$) with coordinate wise addition and scalar multiplication:*

$$(a_i) + (b_i) = (a_i + b_i)$$
$$r(a_i) = (ra_i)$$

where $r \in R$, $a_i, b_i \in A_i$, $\forall i$

cf. Thm. 7.32 of Rotman (2010) [29]

Theorem 14 (7.32, Rotman)**.** *Let commutative ring R.*
$\forall$ R-module A, $\forall$ family $\{B_i | i \in I\}$ of R-modules,

$$Hom_R(A, \coprod_{i\in I} B_i) \simeq \coprod_{i \in I} Hom_R(A, B_i) \tag{32}$$

via R-isomorphism

$$\varphi : f \mapsto (p_i f)$$

where p_i are projections of product $\coprod_{i\in I} B_i$

Proof. Let $a \in A$, $f, g \in \mathrm{Hom}_R(A, \coprod_{i\in I} B_i)$.

$$\varphi(f+g)(a) = (p_i(f+g))(a) = (p_i(f(a) + g(a))) = (p_i f + p_i g)(a)$$

φ additive.

$\forall i, \forall r \in R$, $p_i rf = rp_i f$ (since product of R-modules, $\coprod_{i\in I} B_i$ is also an R-module of $\mathrm{Obj}_R\mathbf{Mod}$, by def. of product).

$$\varphi rf \mapsto (p_i rf) = (rp_i f) = r(p_i f) = r\varphi(f)$$

So φ is R-map.

If $(f_i) \in \coprod_i \mathrm{Hom}_R(A, B_i)$, then $f_i : A \to B_i \quad \forall i$

By Rotman's Prop. 7.31 (If family of R-modules $\{A_i | i \in I\}$, then direct product $C = \coprod_{i\in I} A_i$ is their product in ${}_R\mathbf{Mod}$),

By def. or product, $\exists!$ R-map , $\theta : A \to \coprod_{i\in I} B_i$ s.t. $p_i\theta = f_i \quad \forall i$

$$\begin{tikzcd} & A \arrow[dl, "f_i"'] \arrow[d, dashed, "\theta"] \\ B_i & \prod_{i\in I} B_i \arrow[l, "p_i"] \end{tikzcd}$$

Then

$$f_i = (p_i\theta) = \varphi(\theta)$$

, and so φ *surjective.*

Suppose $f \in \ker\varphi$, so $\theta = \varphi(f) = (p_i f)$. Thus $p_i f = 0 \quad \forall i$

$$\begin{tikzcd} & A \arrow[dl, "0"'] \arrow[d, "f"] \\ B_i & \prod_{i\in I} B_i \arrow[l, "p_i"] \end{tikzcd}$$

But 0-homomorphism also makes this diagram commute, so uniqueness of homomorphism $A \to \prod B_i$ gives $f = 0$. □

8. Applications of Category Theory: Finite State Machines (FSM)

Definition 39 (Finite State Machines ≡ Finite State Automaton)**.** *A deterministic finite state machine or acceptor deterministic finite state machine is a quintuple $(\Sigma, S, s_0, \delta F)$ where*
$\Sigma \equiv$ input alphabet (finite, non-empty set of symbols)
$S \equiv$ finite, non-empty set of states
$s_0 \equiv$ initial state, $s_0 \in S$
$\delta \equiv$ state-transition function; $\delta : S \times \Sigma \to S$ (in a nondeterministic finite automaton, it would be $\delta : S \times \Sigma \to \mathcal{P}(S)$), i.e. δ would return a set of states; $\mathcal{P}(S) \equiv$ set of all subsets of S, including $\emptyset$ and $S \equiv$ power set.
$F \equiv$ set of final states, (possibly empty subset of S; $F \subseteq S$ or $F \subseteq S \cup \{\emptyset\}$)
Finite State Machine (FSM) is also known as a Finite State Automaton.

cf. Black, Paul E (12 May 2008). "Finite State Machine". *Dictionary of Algorithms and Data Structres*. U.S. National Institute of Standards and Technology (NIST).

For both deterministic and non-deterministic FSMs, it's conventional to allow δ to be a partial function, i.e. $\delta(q, x)$ doesn't have to be defined for every combination of $q \in S$, $x \in \Sigma$

If FSM M is in state q; the next symbol (input?) is x and $\delta(q, x)$ not defined; then M can announce an error (i.e. reject the input (???)).

Definition 40 (Alphabet)**.** *alphabet := nonempty set of symbols $\equiv \Sigma$*
*string := finite sequence of members (i.e. symbols) of an underlying base set (i.e. **alphabet**)*
$\Sigma^n \equiv$ set of all strings of length n.

In Curino and Spivak (2011)[18], on pp. 6, Sec. 4.2 "A new concept of state transformations", they say that "Legend has it that Eilenberg and MacLane spent the effort to invent category theory because they needed to formalize the concept now known as natural transformation."

It goes and says "The definition of natural transformations between states naturally captures much of the semantics for what remains unchanged when performing an update."

Part 2. Category Theory

9. Note on notation

From the section on "Terminology" of the Preface of Barr and Wells (1998) [3]:

> "In most scientific disciplines, notation and terminology are standardized, of- ten by an international nomenclature committee. (Would you recognize Einstein's equation if it said $p = HU^2$?) We must warn the nonmathematician reader that such is not the case in mathematics. There is no standardization body and terminology and notation are individual and often idiosyncratic."

To try to bridge the difference choice of notation and through comparison, suggest the "best" notation that's easy to remember and easy to use, I'll present all the different types of notation that I come across as much as I can. My plan of attack is the following:

(1) I'll try to present different types of notation and reference the authors of the text when I can.
(2) I'll try to defer to the notation used in Wikipedia, first.
(3) I'll make a final decision of what notation works best (for me).

10. Category $\mathbf{A}$, (definition)

Definition 41 (Category $\mathbf{A}$). ***category*** $\mathbf{A}$ *is quadruple* $\mathbf{A} = (Obj(\mathbf{A}), Mor\mathbf{A}, 1, \circ)$

$$\mathbf{A} = (Obj(\mathbf{A}), Mor\mathbf{A}, 1, \circ) \tag{33}$$

s.t.

(1) $Obj(\mathbf{A})$ *is a* class, *whose elements*, $A \in Obj(\mathbf{A})$, *are called* objects
(2) $Mor\mathbf{A}$ *is a* class.
   (a) *From Adámek, Herrlich, and Strecker (2004)* [4], *Kashiwara and Schapira (2006)* [1], $\forall A, B \in Obj(\mathbf{A})$, $\exists\, Hom(A,B) \subseteq Mor(\mathbf{A})$. *Therefore,*

$$Mor\mathbf{A} = \bigcup_{A,B\in Obj(\mathbf{A})} Hom(A,B) \tag{34}$$

   (b) $\forall f \in Hom(A,B)$, $f : A \to B \in Hom(A,B)$ *is a* morphism. *Leinster (2014)* [8] *also calls them on pp. 10* ***maps*** *or* ***arrow*** *from* A *to* B.

$$A \xrightarrow{f} B$$

(3) $\forall A \in Obj(\mathbf{A})$, $\exists 1_A : A \to A$, *i.e.* $\exists \mathbf{1}_A \in Hom_{\mathbf{A}}(A,A) \equiv Hom(A,A)$,

$$A \xrightarrow{1_A} A \quad \text{or} \quad \begin{tikzcd} A \arrow[loop left, "\mathbf{1}_A"] \end{tikzcd}$$

(4) ***composition:*** $\forall A, B, C \in Obj\mathbf{A}$, *define* ***composition*** *to be a map*

$$\begin{aligned} Hom_{\mathbf{A}}(A,B) \times Hom_{\mathbf{A}}(B,C) &\to Hom_{\mathbf{A}}(A,C) \\ (f,g) &\mapsto g \circ f \end{aligned} \tag{35}$$

, *i.e.*

$\forall f : A \to B \in Hom(A,B)$, *i.e.* $f, g \in Mor\mathbf{A}$,
$g : B \to C \in Hom(B,C)$
then $g \circ f : A \to C \in Hom(A,C)$, $g \circ f \in Mor\mathbf{A}$ *i.e.*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & B \arrow[r, "g"] & C \end{tikzcd}$$

s.t.

(a) associativity $\forall \begin{matrix} f : A \to B \\ g : B \to C \\ h : C \to D \end{matrix}$, $h \circ (g \circ f) = (h \circ g) \circ f$ *i.e.*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rrr, bend left=50, "(h \circ g) \circ f"] \arrow[rr, bend right, "g \circ f"'] \arrow[rrr, bend right=50, "h \circ (g \circ f)"'] & B \arrow[r, "g"] \arrow[rr, bend left, "h \circ g"] & C \arrow[r, "h"] & D \end{tikzcd}$$

(b) $\forall f : A \to B \in Hom(A,B)$, $1_B \circ f = f$ *and* $f \circ 1_A = f$ *i.e.* $\forall f \in Hom_{\mathbf{A}}(A,B)$,

$$\begin{tikzcd} A \arrow[loop left, "1_A"] \arrow[r, "f"] & B \arrow[loop right, "1_B"] \end{tikzcd}$$

(c) *Adámek, Herrlich, and Strecker (2004)* [4] *posited further that* $Hom(A,B) \in Mor\mathbf{A}$ *pairwise disjoint (i.e.* $Hom(A,B) \bigcap Hom(C,D) \neq \emptyset$ *if* $C \neq A$ *or* $D \neq B$*)*

10.1. Examples of categories.

- $\mathbf{Set} = (\mathrm{Obj}(\mathbf{Set}), \mathrm{Hom}\mathbf{Set}, \mathbf{1}, \circ)$ where
  $\mathrm{Obj}(\mathbf{Set})$ is the class of all sets
  $\mathrm{Hom}\mathbf{Set}$ is the class of all functions on a set to another set
- **Vec**

$$\begin{aligned} \mathrm{Obj}\mathbf{Vec} &\equiv \text{all real vector spaces} \\ \mathrm{Mor}\mathbf{Vec} &\equiv \text{all linear transformations between them (between real vector spaces)} \end{aligned}$$

- *preorder* cf. Ch. 1, Example 7, pp. 8, Awodey (2010) [19].
  preorder P, set P equipped with binary relation $p \leq q$, that's both
  reflexive: $a \leq a$, and
  transitive: if $a \leq b$ and $b \leq c$, then $a \leq c$.

  $\forall$ preorder P can be regarded as a category:
  $\mathrm{Obj}(P) = P$
  $a \to b$ iff $a \leq b \in \mathrm{Mor}P$.

identity: $1 : a \to a$ since $a \leq a$
composition: $f : a \to b$, $g : b \to c$, $(g \circ f)a = g(b) = c$, $a \leq b$ and $b \leq c$, so $a \leq c$.

Going in the other direction, a category with at most 1 arrow between 2 objects determines a preorder.

- **Monoid**. Consider a monoid as a triple $(M, \cdot, e)$.
Every semigroup $(M, \cdot)$ (recall that a *semigroup* is a set S with binary operation $\cdot$, i.e. s.t.

$$S \times S \xrightarrow{\cdot} S$$
$$\forall a, b, c \in S, \ (a \cdot b) \cdot c = a \cdot (b \cdot c) \qquad \text{(associativity)}$$

(but no inverse, necessarily!)) that also has a unit e can be made into a category $\mathbf{C}$
$\Longrightarrow \mathbf{C}(M, \cdot, e) = (\text{Obj}(\mathbf{C}), \text{Hom}(\mathbf{C}), \mathbf{1}, \circ)$, a category $\mathbf{C}$ with only 1 object, i.e. $\text{Obj}(\mathbf{C}) = \{M\}$, so that
$\text{Obj}(\mathbf{C}) = \{M\}$
$\text{Hom}(M, M) = M$
$\mathbf{1}_M = e$
$y \circ x = y \cdot x$

cf. pp. 10, Example 12 of Awodey (2010) [19].

10.1.1. *Monoid as a category.* Recall the definition of a *category* $\mathbf{A}$: category $\mathbf{A} = (\text{Obj}(\mathbf{A}), \text{Mor}\mathbf{A}, 1, \circ)$ s.t.
$A \in \text{Obj}\mathbf{A}$,
$\forall A, B \in \text{Obj}\mathbf{A}$, $\exists \text{Hom}(A, B) \subseteq \text{Mor}\mathbf{A}$,
$\forall f \in \text{Hom}(A, B)$, $f : A \to B$ is a morphism.
$\forall A \in \text{Obj}\mathbf{A}$, $\exists 1_A : A \to A$, i.e. $\exists 1_A \in \text{Hom}_{\mathbf{A}}(A, A)$.
$\forall A, B, C \in \text{Obj}\mathbf{A}$,

$$\text{Hom}_{\mathbf{A}}(A, B) \times \text{Hom}_{\mathbf{A}}(B, C) \to \text{Hom}_{\mathbf{A}}(A, C)$$
$$(f, g) \mapsto g \circ f$$

Recall the definition of a *monoid*: monoid $(M, \cdot, e)$, set M,
$\cdot : M \times M \to M$,
$e \in M$ s.t. $\forall x \in M$, $e \cdot x = x \cdot e = x$

$\Longrightarrow \mathbf{C}(M)$ (monoid as a category)
$\text{Obj}\mathbf{C}(M) = \{M\}$
$\text{Hom}(M, M) = \text{Mor}\mathbf{C}(M)$ and $\forall f \in \text{Hom}(M, M)$,

$$f = m \in M$$

i.e. the morphism *is* an element in monoid M (there's a 1-to-1 correspondence).
$\forall n \in M$, $f : n \mapsto m \cdot n \in M$.
$1_M : M \to M$
$1_M : x \mapsto e \cdot x = x$
$m \cdot n \cdot x = m \cdot (n \cdot x) = (m \cdot n) \cdot x$ (composition)

10.1.2. *Examples of monoids.* Examples of monoids: $\mathbb{N}, \mathbb{Q}, \mathbb{R}$ under either addition and $e = 0$, or multiplication and $e = 1$.
$\forall$ set X, set of functions from X to X, $\text{Hom}_{\mathbf{Sets}}(X, X)$ is a monoid under composition.

In general,
$\forall$ object $C \in \text{Obj}(\mathbf{C})$, $\forall$ category $\mathbf{C}$, $\text{Hom}_{\mathbf{C}}(C, C)$ is a monoid under composition of $\mathbf{C}$.

Since monoids are structured sets, $\exists$ category $\mathbf{Mon}$, s.t. $\text{Obj}\mathbf{Mon} \ni$ monoids, $\text{Mor}\mathbf{Mon} \ni$ functions that preserve monoid structure.

In detail, homomorphism from monoid M to monoid N is function $h : M \to N$ s.t. $\forall m, n \in M$.

$$h(m \cdot_M n) = h(m) \cdot_N h(n)$$

and

$$h(e_M) = e_N$$

Check: Consider functor $F : M \to N$.
$F : \text{Hom}_M(a, b) \to \text{Hom}_N(F(a), F(b))$

$$m \cdot a = b \qquad n \cdot_N F(a) = F(b)$$
$$F(m) \cdot_N F(a) = F(b) = F(m \cdot_M a)$$

so functor F is a monoid homomorphism, and so
a monoid homomorphism from M to N is the same thing as a functor from M regarded as a category to N regarded as a category.

- **Poset** cf. Ch. 1, pp. 6, Awodey (2010) [19]. partially ordered set or poset is a set A, equipped with binary relation $a \leq_A b$ s.t. $\forall a, b, c \in A$,
reflexivity: $a \leq_A a$
transitivity: if $a \leq_A b$, $b \leq_A c$, then $a \leq_A c$
antisymmetry: if $a \leq_A b$, $b \leq_A a$, then $a = b$.

An arrow from a poset A to a poset B is a function,

$$m : A \to B$$

that is monotone, i.e. $\forall a, a' \in A$,

$$\text{if } a \leq_A a', \text{ then } m(a) \leq_B m(a')$$

category $\mathbf{Pos}$ of posets, and monotone functions,

$$\forall A \in \text{Obj}(\mathbf{Pos}), \exists 1_A : A \to A, \text{ since}$$

since if $a \leq_A a'$, then $a \leq_A a'$, so 1_A monotone.

composition: $\forall f : A \to B$, $g : B \to C$, $\forall A, B, C \in \text{Obj}(\mathbf{Pos})$, f, g monotone,
$\forall a \leq_A a'$, $f(a) \leq_B f(a')$, and since
$\forall b \leq_B b'$, $g(b) \leq_C g(b')$, then $g(f(a)) \leq_C g(f(a'))$, or $(g \circ f)(a) \leq_C (g \circ f)(a') \forall a \leq_A a'$, so $g \circ f : A \to C$ is monotone.

$\text{Mor}\mathbf{Pos} = \bigcup_{A, B \in \text{Obj}(\mathbf{Pos})}$ monotone functions $A \to B$

A *poset* is a *preorder*, satisfying additional condition of antisymmetry: if $a \leq b$, and $b \leq a$, $a = b$.
e.g. $\forall X$, power set $P(X)$ is a poset under inclusion $U \subseteq V$, $\forall$ subsets $U, V \subseteq X$.
  - cf. Ch. 1, Example 8, pp. 9, Awodey (2010) [19]. **Pos** categories and functors: functor $F : \mathbf{D} \to \mathbf{Q}$ between poset categories $\mathbf{P}, \mathbf{Q}$.
  They are the monotone functions.

  Proof. Consider
  $$F : \text{Hom}_{\mathbf{P}}(A, B) \to \text{Hom}_{\mathbf{Q}}(F(A), F(B))$$
  $f \in \text{Hom}_{\mathbf{P}}(A, B)$ so if $a \leq_A a'$, $f(a) \leq_B f(a')$
  Let $g \in \text{Hom}_{\mathbf{Q}}(F(A), F(B))$, so if $c \leq_{F(A)} c'$, $g(c) \leq_{F(B)} g(c')$.

  $F(f) \in \text{Hom}_{\mathbf{Q}}(F(A), F(B))$ (by definition of a functor), so if $c \leq_{F(A)} c'$, $F(f)(c) \leq_{F(B)} F(f)(c')$ or $F(f(c)) \leq_{F(B)} F(f(c'))$.

  So F itself is monotone. $\square$

- cf. *An example from computer science*, Example 10, pp. 9 of Awodey (2010) [19]. Given a functional programming language L, $\exists$ associated category,
  Obj$\mathbf{L} \ni$ data types of L
  Mor$\mathbf{L} \ni$ computable functions of L ("processes", "procedures", "programs")
  composition (of 2 such programs) $X \xrightarrow{f} Y \xrightarrow{g} Z$ given by applying g to output of f
  (notation note: $g \circ f = f; g$)
  identity is the "do nothing" program.

  if $\mathbf{C}(L)$ is the category just defined, then
  denotational semantics of language L in category $\mathbf{D}$ of say Scott demands is simply functor
  $$S : \mathbf{C}(L) \to \mathbf{D}$$
  since S assigns domains to the types of L,
  continuous functions to programs.
  This example and Example 9, pp. 9, of Awodey (2010) [19] are related to the notion of "Cartesian closed category."
  From Barr and Wells (2012) [20], a *functional programming language* L has
  (1) primitive data types (built into the language)
  (2) constants of each type
  (3) operations, which are functions between types
  (4) constructors, which are applied to data types and operations, to produce derived data types and operations of the language.

  For a $\mathbf{C}(L)$ category corresponding to functional programming language L,

  $\mathbf{C}(L)$ category corresponding to functional programming language L,
  Obj$\mathbf{C}(L) = \{$ types of L $\}$
  Mor$\mathbf{C}(L) = \{$ operations (functions; primitive and derived) of L $\}$

  identity: $\forall$ type $T \in \text{Obj}\mathbf{C}(L)$, $\exists 1_T : T \to T$ i.e. $1_T \in \text{Hom}(T,T)$.

  composition: $\forall$ types $T, U, V \in \text{Obj}\mathbf{C}(L)$, defines composition
  $$\text{Hom}(T,U) \times \text{Hom}(U,V) \to \text{Hom}(T,V)$$
  $$(f,g) \mapsto g \circ f$$
  Assume L has a do nothing operation 1_T, and composition constructor, $\text{Hom}(T,U) \times \text{Hom}(U,V) \to \text{Hom}(T,V)$.

  Add additional type 1 s.t. $\forall$ type $T \in \text{Obj}\mathbf{C}(L)$, $\exists!$ operation (function) to $1, T \to 1$
  Interpret each constant c of type A as arrow (morphism) $c : 1 \to A$.

  As a concrete example (cf. Example 2.2.5 of Barr and Wells (2012) [20]), suppose a simple language with 3 data types, `int`, `bool`, `char`.

  For `int`, $\exists$ constant $0 : 1 \to \texttt{int}$, consider $\texttt{succ} : \texttt{int} \to \texttt{int} \in \text{Hom}(\texttt{int}, \texttt{int})$,
  For `bool`, $\exists$ constants $\texttt{true} : 1 \to \texttt{bool}$, $\texttt{false} : 1 \to \texttt{bool}$; consider $\neg \in \text{Hom}(\texttt{bool}, \texttt{bool})$, s.t.
  $$\neg \circ \texttt{true} = \texttt{false}$$
  $$\neg \circ \texttt{false} = \texttt{true}$$
  For `char`, $\exists$ constants $c : 1 \to \texttt{char}$, $\forall \texttt{char} c$.
  Since $\forall T, U \in \text{Obj}(\mathbf{C}(L))$, $\exists \text{Hom}(T,U) \subseteq \text{Mor}\mathbf{C}(L)$. So consider $\texttt{ord} : \texttt{char} \to \texttt{int}$, $\texttt{chr} : \texttt{int} \to \texttt{char}$, s.t.
  $$\texttt{char} \circ \texttt{ord} = 1_{\texttt{char}}$$
  So Obj$\mathbf{C}(L) = \{\texttt{int}, \texttt{bool}, \texttt{char}, 1\}$,
  Mor$\mathbf{C}(L)$ consists of all programs.
  See 5.3.14 Record Types and 5.7.6, 14.2 for flow of control in Barr and Wells (2012) [20] to complete the programming language. For here, see section **??**
- **Groups**, cf. pp. 11, Def. 1.4, Sec. 1.5 **Isomorphisms**, Awodey (2010) [19].
  A *group* G is a monoid with inverse g^{-1}, $\forall g \in G$.
  Thus, G is a category with 1 object, $\forall f \in \text{Mor}G$ is an isomorphism.
  $\forall$ set X, group $\text{Aut}(X)$ of automoprhisms (or "permutations") of X i.e. $f : X \to X$
  $\text{Aut}(X)$ is closed under composition $\circ$ because a permutation of a permutation is another permutation.
  Homomorphisms of groups $h : G \to H$ is just a homomorphism of monoids, preserving inverses.
  Given h, then $h(gg') = h(g)h(g')$.
  G, H are monoids. Then h monoid homomorphism.

Concrete categories: informally, categories which objects are sets, possibly equipped with some structure, and arrows are certain, possibly structure-preserving functions.
From pp. 10 of Mac Lane (1978) [5], Leinster (2014) [8], Examples 1.1.8 (Categories as mathematical structures), pp. 13, and Sec. 1.4 Examples of Categories, 5. Finite Categories, pp.7 of Awodey (2010) [19].
Category **1**. Obj$\mathbf{1} = \{A\}$
Mor$\mathbf{1} = \{1_A\}$; $1_A : A \to A$.

The only composition to consider is $1_A \circ 1_A : A \to A$, and $1_A \circ 1_A \circ 1_A$, so on and so for.
Category **2**. Obj$\mathbf{2} = \{A, B\}$.
$\text{Hom}(A,B) = \{f\}$, $f : A \to B$.

The only "non-trivial" composition to consider is $1_B \circ f \circ 1_A : A \to B$.

Category **3**. pp. 11, Sec. 2 Categories of Mac Lane (1978) [5], Example 5. "Finite Categories", Sec. 1.4, pp. 7 of Awodey (2010) [19].
Obj$\mathbf{3} = \{A, B, C\}$
$\text{Hom}(A,B) = \{f\}$
$\text{Hom}(A,C) = \{g\}$
$\text{Hom}(B,C) = \{h\}$
$\text{Hom}(C, \cdot) = \emptyset$
$\text{Hom}(B,A) = \emptyset$

The only "non-trivial" compositions to consider are the following: $1_C \circ h \circ f : A \to C$,
$1_C \circ g$, $1_C \circ h$ (and likewise)

10.2. **Duality, opposite category.** Given a category $\mathbf{A} = (\text{Ob}, \text{hom}_{\mathbf{A}}, 1, \circ)$,

Definition 42 (dual opposite category)**.** ***dual*** *or* ***opposite*** *category of* $\mathbf{A} = (Obj(\mathbf{A}), Mor\mathbf{A}, \mathbf{1}, \circ)$, *denoted* $\mathbf{A}^{op}$, *is*

$$\mathbf{A}^{op} = (Obj(\mathbf{A}), Mor\mathbf{A}^{op}, \mathbf{1}, \circ^{op}) \tag{36}$$

s.t.

- $$Obj(\mathbf{A}^{op}) = Obj(\mathbf{A}) \tag{37}$$
- $\forall A, B \in Obj(\mathbf{A}^{op})$, $Hom_{\mathbf{A}^{op}}(A,B) \subseteq Mor\mathbf{A}^{op}$,

$$Hom_{\mathbf{A}^{op}}(A,B) = Hom_{\mathbf{A}}(B,A) \subseteq Mor\mathbf{A} \tag{38}$$

- Define *the new composition*

(39)
$$\begin{aligned} f \circ^{op} g \text{ of } g &\in Hom_{\mathbf{A}^{op}}(C,B) \\ f &\in Hom_{\mathbf{A}^{op}}(B,A) \\ &\text{then} \\ f \circ^{op} g &= g \circ f \end{aligned}$$

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & B \arrow[r, "g"] & C \end{tikzcd} \qquad \text{dual : } \begin{tikzcd} A & B \arrow[l, "f"'] & C \arrow[l, "g"'] \arrow[ll, bend left, "f \circ^{op} g"] \end{tikzcd}$$

or, equivalently (notation-wise)

$$\begin{tikzcd} A & B \arrow[l, "f"'] & C \arrow[l, "g"'] \arrow[ll, bend left, "f \circ g"] \end{tikzcd} \qquad \text{dual : } \begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g \circ^{op} f"'] & B \arrow[r, "g"] & C \end{tikzcd}$$

in that

$$\begin{aligned} g \circ^{op} f \text{ of } f &\in Hom_{\mathbf{A}^{op}}(A,B) \\ g &\in Hom_{\mathbf{A}^{op}}(B,C) \\ &\text{then} \\ g \circ^{op} f &= f \circ g \end{aligned}$$

i.e. (in summary) $\mathbf{C}^{\text{op}}$ s.t. $\text{Obj}\mathbf{C}^{\text{op}} = \text{Obj}\mathbf{C}$,
Denote $\forall \overline{C} \in \text{Obj}\mathbf{C}^{\text{op}}$, $\overline{C} = C \in \text{Obj}\mathbf{C}$.
$\text{Hom}_{\mathbf{C}^{\text{op}}}(\overline{A}, \overline{B}) \ni \overline{f} : \overline{C} \to \overline{D}$, in $\mathbf{C}^{\text{op}}$, for $f : D \to C$ in $\mathbf{C}$.

$$1_{\overline{C}} = 1_C$$
$$\overline{f} \circ \overline{g} \equiv \overline{f} \circ_{\text{op}} \overline{g} = g \circ f$$

Diagram in $\mathbf{C}$:

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "g \circ f"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

Diagram in $\mathbf{C}^{\text{op}}$:

$$\begin{tikzcd} \overline{A} & \overline{B} \arrow[l, "\overline{f}"'] \\ & \overline{C} \arrow[u, "\overline{g}"'] \arrow[ul, "\overline{f} \circ \overline{g}"] \end{tikzcd}$$

e.g. **Sets** is dual to category of complete, atomic Boolean algebras.
e.g. if $\mathbf{A} = (M, \cdot, e)$ monoid, then $\mathbf{A}^{\text{op}} = (M, \widehat{\cdot}, e)$ where $a \widehat{\cdot} b = b \cdot a$

10.2.1. *Example.*

- Vec^{op}

$$\mathbf{Vec}^{\text{op}} = (\text{Obj}(\mathbf{Vec}), \text{Hom}_{\mathbf{Vec}^{\text{op}}}, 1, \circ^{\text{op}})$$

s.t.

$$\text{Hom}_{\mathbf{Vec}^{\text{op}}}(W,V) = \text{Hom}_{\mathbf{Vec}}(V,W)$$

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & V \arrow[r, "g"] & W \end{tikzcd} \qquad \begin{tikzcd} U & V \arrow[l, "f"'] & W \arrow[l, "g"'] \arrow[ll, bend left, "f \circ^{\text{op}} g"] \end{tikzcd}$$

10.3. **Kinds of morphisms.** cf. pp. 11, Definition 1.3 of Awodey (2010) [19]

Definition 43 (isomorphism)**.** ***isomorphism*** *- $\forall$ category $\mathbf{C}$, Mor$\mathbf{C} \ni$ morphism $f : A \to B$ is an isomorphism if $\exists\, g : B \to A$ s.t. $f \circ g = 1_B$, $g \circ f = 1_A$, g unique. g called inverse of f, f^{-1}*

$$\begin{tikzcd} A \arrow[loop left, "1_A"] \arrow[r, "f"] & B \arrow[l, bend left, "g"] \end{tikzcd} \qquad \begin{tikzcd} B \arrow[loop left, "1_B"] \arrow[r, "g"] & A \arrow[l, bend left, "f"] \end{tikzcd}$$

Theorem 15 (Cayley)**.** *cf. pp. 11, Sec. 1.5.* ***Isomorphisms****, Awodey (2010)* [19]
$\forall$ group G,

$$G \cong \text{ group of permutations}$$

(isomorphic $\equiv \cong$)

Proof. Define Cayley representation $\overline{G}$ of G to be following group of permutations:
underlying set of $\overline{G}$ is just G,
$\forall\, g \in G$, we have permutation $\overline{g}$, defined $\forall\, h \in G$,

$$\overline{g}(h) = g \cdot h$$

Let $\overline{g} = \overline{h}$.
$\forall\, g' \in G$,
$\overline{g}(g') = gg'$
$\overline{h}(g') = hg'$
$g' \in G$, so $\exists\, (g')^{-1}$ s.t. $g'(g')^{-1} = e$, so $gg'(g')^{-1} = g = h$.

Define homomorphisms $i : G \to \overline{G}$, $\quad j : \overline{G} \to G$
$$i(g) = \overline{g} \qquad j(\overline{g}) = g$$

$$
\begin{aligned}
& i(gg') = \overline{(gg')} \\
& \overline{(gg')} = h = gg'h = g(g'h) = g(\overline{g}'(h)) = \overline{g}(\overline{g}'(h)) = i(g) = i(g')h \\
& \Longrightarrow i(gg') = i(g) \circ i(g') \\
& j(\overline{g}) = j(\overline{g}')h = j(\overline{g})(g'h) = gg'h = j(\overline{gg'})h = j(\overline{g} \cdot \overline{g}')h \\
& j(\overline{g}) \cdot j(\overline{g}') = j(\overline{g} \cdot \overline{g}')
\end{aligned}
$$

$$
\begin{aligned}
& i \circ j = 1_{\overline{G}},\ j \circ i = 1_G \text{ since} \\
& i \circ j(\overline{g}) = i \circ g = \overline{g} \\
& j \circ i(g) = j(\overline{g}) = g
\end{aligned}
$$

i, j are isomorphisms between $G, \overline{G}$, which is in category **Groups** of groups and group homomorphisms.
permutations $\overline{g} \in \overline{G}$ are themselves isomorphisms in **Sets**.
Cayley's Thm. says that any abstract group can be represented as a "concrete" one, i.e. group of permutations of a set. □

Example 1.1.5 (Leinster (2014) [8]) isomorphisms in **Set** are exactly bijections.
$\Longrightarrow$ function has 2-sided inverse iff function is injective and surjective (this is not trivial)
Example 1.1.7 Leinster (2014) [8] isomorphisms in **Top** are exactly homeomorphisms.
bijective map in **Top** is not necessarily an isomorphism:

$$
\begin{aligned}
& [0,1) \to \{ z \in \mathbb{C} | |z| = 1 \} \\
& t \mapsto e^{2\pi i t}
\end{aligned}
$$

is a continuous bijection but not a homeomorphism (cont. bijection with cont. f^{-1}).

$$\ln \left(\frac{w}{2\pi i} \right) = f^{-1}(w)$$

Example 1.1.8 Leinster (2014) [8] (Categories as mathematical structures) (b) **Discrete** categories contain no maps at all apart from identities. Just a class of objects.

Exercise 1.1.13 pp. 16, Leinster (2014) [8].
Suppose for morphism $f : A \to B$, $\exists$ inverse g, s.t. $f \circ g = 1_B$, $g \circ f = 1_A$
Now $f^{-1} \circ f = 1_A$, $f \circ f^{-1} = 1_B$
$f^{-1} \circ f \circ g = f^{-1} \circ 1_B = 1_A \circ g = f^{-1} \circ 1_B \Longrightarrow g = f^1$

Definition 44 (endomorphism)**.** ***endomorphism*** *- morphism with same source and target, that is, morphism* $f : A \to A$

Definition 45 (automorphism)**.** ***automorphism*** *- endomorphism which is an isomorphism*

Definition 46 (parallel)**.** ***parallel*** *- 2 morphisms* f, g *are parallel if they have same source and same target:*

$$
\begin{aligned}
& f : A \to B \\
& g : A \to B
\end{aligned}
$$

Definition 47 (monomorphism)**.** *monomorphism - morphism* $f : A \to B$ *is a monomorphism if* $\forall$ *pair of parallel* $g_1 : C \to A$, $g_2 : C \to A$

(40) $$f \circ g_1 = f \circ g_2 \text{ implies } g_1 = g_2$$

i.e.

$$
\begin{tikzcd}
C \arrow[r, "f \circ g_1"] & B \arrow[l, bend left, "f \circ g_2"]
\end{tikzcd}
\quad \text{implies } C \xrightarrow{g_1 = g_2} A
$$

Definition 48 (epimorphism)**.** ***epimorphism*** *- morphism* $f : A \to B$ *is an epimorphism if* $f^{op} : B^{op} \to A^{op}$ *is a monomorphism in* $\mathbf{A}^{op}$.

Hence f *epimorphism iff* $\forall$ *parallel morphisms* $g_1 : B \to C$, $g_2 : B \to C$, $g_1 \circ f = g_2 \circ f$
implies $g_1 = g_2$

Proposition 17 (monomorphism, epimorphism iff injective)**.** f *monomorphism iff* $f\circ : Hom_{\mathbf{A}}(C, A) \to Hom_{\mathbf{A}}(C, B)$ *injective* $\forall C \in Obj(\mathbf{A})$, *i.e.*

(41) $$
\begin{gathered}
Hom_{\mathbf{A}}(C, A) \xrightarrow{f\circ} Hom_{\mathbf{A}}(C, B) \\
g_1, g_2 \xmapsto{f\circ} f \circ g_1, f \circ g_2 \\
\textit{then} \\
f\circ \textit{ is injective if} \\
f \circ g_1 = f \circ g_2 \Longrightarrow g_1 = g_2
\end{gathered},
$$

f *epimorphism iff map* $\circ f : Hom_{\mathbf{A}}(B, C) \to Hom_{\mathbf{A}}(A, C)$ *injective* $\forall C \in Obj(\mathbf{A})$

(42) $$
\begin{gathered}
Hom_{\mathbf{A}}(B, C) \xrightarrow{\circ f} Hom_{\mathbf{A}}(A, C) \\
g_1, g_2 \xmapsto{\circ f} g_1 \circ f, g_2 \circ f \\
\textit{then} \\
\circ f \textit{ is injective if} \\
g_1 \circ f = g_2 \circ f \Longrightarrow g_1 = g_2
\end{gathered},
$$

Definition 49 (inverses)**.** $\forall$ *2 morphisms,* $f : X \to Y$, $g : Y \to X$ *s.t.* $f \circ g = 1_Y$,
f *is called left inverse of* g, g *is called right inverse of* f.
We also say, g *is a section of* f, *or* f *is a cosection of* g.
f *is an epimorphism,* g *is a monomorphism.*

10.4. More definitions with categories.

Definition 50 (subcategory)**.** *category* $\mathbf{A}'$, $\mathbf{A}' \subset \mathbf{A}$, *if* $Obj(\mathbf{A}') \subset Obj(\mathbf{A})$, $Hom_{\mathbf{A}'}(A, B) \subset Hom_{\mathbf{A}}(A, B)$, $\forall A, B \in \mathbf{A}'$.
Composition in $\mathbf{A}'$ *is induced by composition in* $\mathbf{A}$.
identity morphisms in $\mathbf{A}'$ *are identity morphisms in* $\mathbf{A}$

Definition 51 (full subcategory)**.** *subcategory* $\mathbf{A}'$ *of* $\mathbf{A}$ *is full if* $Hom_{\mathbf{A}'}(A, B) = Hom_{\mathbf{A}}(A, B)$, $\forall A, B \in \mathbf{A}'$

Definition 52 (saturated subcategory)**.** *full subcategory* $\mathbf{A}'$ *of* $\mathbf{A}$ *saturated if* $A \in \mathbf{A}$ *belongs to* $\mathbf{A}'$ *whenever* A *is isomorphic to object of* $\mathbf{A}'$

Definition 53 (discrete category)**.** ***discrete*** *- discrete category if all morphisms are identity morphisms.*

Definition 54 (nonempty category)**.** ***nonempty*** *- nonempty category if* $Obj(\mathbf{A})$ *is nonempty*

Definition 55 (groupoid)**.** ***groupoid*** *- category* $\mathbf{A}$ *is a* ***groupoid*** *if all morphisms are isomorphisms.*

Definition 56 (finite category)**.** *finite - finite category if set of all morphisms in* $\mathbf{A}$ *(hence, in particular, set of objects) is a finite set*

Definition 57 (connected)**.** *connected category* $\mathbf{A}$ *if it's nonempty, and* $\forall A, B \in Obj\mathbf{A}$, $\exists$ *finite sequence of objects* $(A_0 \dots A_n)$, $A_0 = A$, $A_n = B$, *s.t. at least 1 of the sets* $Hom_{\mathbf{A}}(A_j, A_{j+1})$ *or* $Hom_{\mathbf{A}}(A_{j+1}, A_j)$ *is nonempty* $\forall j \in \mathbb{N}$, *with* $0 \leq j \leq n - 1$

Definition 58 (monoid M)**.** ***monoid*** M *(set endowed with internal product with associative and unital law) is nothing but a category with only 1 object (to* M, *associate category* $\mathbf{M}$, *with single object* A, *and morphisms* $Hom_{\mathbf{M}}(A, A) = M$*)*

cf. Def. 1.2.5 of Kashiwara and Schapira (2006) [1].

Definition 59 (Morphisms as a category)**.** *Let category $\mathcal{C} \equiv \mathbf{A}$.*
$Mor(\mathbf{A})$ is a category.
$Obj(Mor(\mathbf{A})) = Mor\mathbf{A}$ (objects of category $Mor(\mathbf{A})$ are morphisms in $\mathbf{A}$).

Let $\begin{aligned} f &: X \to Y \\ g &: X' \to Y' \end{aligned}$ *,* $f, g \in Mor(\mathbf{A})$ *(i.e.* $\begin{aligned} f &\in Hom(X,Y) \\ g &\in Hom(X',Y') \end{aligned}$ *, for* $X, Y, X', Y' \in Obj(\mathbf{A})$*)*
Then

$$Hom_{Mor(\mathbf{A})}(f,g) = \{u : X \to X', v : Y \to Y'; g \circ u = v \circ f\}$$

Composition and identity in $Mor(\mathbf{A})$ are the obvious ones.
So

$$Obj(Mor(\mathbf{A})) = Mor(\mathbf{A})$$

$$Mor(Mor(\mathbf{A})) = \bigcup_{f,g \in Mor(\mathbf{A})} Hom(f,g) = \bigcup_{f,g \in Mor(\mathbf{A})} \{u : X \to X', v : Y \to Y'; g \circ u = v \circ f\}$$

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "u"] & Y \arrow[d, "v"] \\ X' \arrow[r, "g"] & Y' \end{tikzcd}$$

cf. Def. 1.2.6 of Kashiwara and Schapira (2006) [1].

Definition 60. (1) *object $P \in \mathcal{C} \equiv \mathbf{A}$ is called* initial *if $\forall X \in \mathbf{A}$, ($\equiv \forall x \in Obj(\mathbf{A})$), $Hom_{\mathbf{A}}(P, X) \simeq \{pt\}$.*
(Denote by $\emptyset_{\mathbf{A}}$ an initial object *in $\mathbf{A}$).*
(Note that if P_1 and P_2 are initial, then $\exists!$ isomorphism $P_1 \simeq P_2$)
(2) *P is terminal in $\mathbf{A}$ if P is initial in $\mathbf{A}^{op}$, i.e.*
$\forall X \in \mathbf{A}$, $Hom_{\mathbf{A}}(X, P) \simeq \{pt\}$.
Denote $pt_{\mathbf{A}}$ a terminal object in $\mathbf{A}$.
(3) *P is zero (0) object if it's both initial and terminal.*
Such a P is denoted by 0.
If $\mathbf{A}$ has a zero object, $\forall$ object $X, Y \in \mathbf{A} \equiv Obj\mathbf{A}$, the morphism obtained as composition $X \to 0 \to Y$ is still denoted by $0 : X \to Y$.
(Note that composition of $0 : X \to Y$, and any morphism $f : Y \to Z$ is $0 : X \to Z$)

cf. Example 1.2.7 of Kashiwara and Schapira (2006) [1]. Example
(i) In category **Set**, $\emptyset$ initial, $\{\mathrm{pt}\}$ terminal.
(ii) Zero module 0 is zero object in $\mathrm{Mod}(R)$.

Notation 1.2.8 of Kashiwara and Schapira (2006) [1]:
(1) $\mathbf{Pt} \equiv$ category with a single object and a single morphism (the identity of this object)
(2) $\emptyset \equiv$ empty category with no objects (hence, no morphisms)
(3) $\bullet \to \bullet \equiv$ category which consists of 2 objects, say a, b, and 1 morphism, $a \to b$, other than $\mathrm{id}_a, \mathrm{id}_b \equiv 1_a, 1_b$. Denote this category by $\mathbf{Arr}$.

cf. Example 1.2.9 of Kashiwara and Schapira (2006) [1].
Let R be a ring. Let $N \in \mathrm{Mod}(R^{\mathrm{op}})$, $M \in \mathrm{Mod}(R)$.
Category $\mathbf{C}$:
$\mathrm{Obj}\mathbf{C} \ni (f, L)$, where $L \in \mathrm{Mod}(\mathbb{Z})$, f bilinear map $f : N \times M \to L$ (i.e. it's $\mathbb{Z}$-bilinear and satisfies

$$f(na, m) = f(n, am), \quad \forall a \in R$$

)
Morphism from $f : N \times M \to L$ to $g : N \times M$ is a linear map $h : L \to K$ s.t. $h \circ f = g$.
Since any bilinear map $f : N \times M \to L$ (i.e. any object of $\mathbf{C}$) factorizes uniquely through

$$u : N \times M \to N \otimes_R M$$

object $(u, N \otimes_R M)$ is initial in $\mathbf{C}$

11. Functors

cf. Def. 1.2.10 of Kashiwara and Schapira (2006) [1], with terminology from pp. 13, Sec. 3. "Functors" from Mac Lane (1978) [5]

Definition 61 ((covariant) Functor)**.** (1) *(covariant functor) Let categories $\mathbf{C}$, $\mathbf{D}$.*
(covariant) ***functor*** *$F : \mathbf{C} \to \mathbf{D}$ consists of (Mac Lane (1978) [5] says 2 suitably related functions)*
- *map $F : Obj(\mathbf{C}) \to Obj(\mathbf{D})$ (i.e. $\forall C \in Obj(\mathbf{C})$, $F(C) \in Obj(\mathbf{D})$),*
(Mac Lane (1978) [5] calls this the object function *T or F in our notation; we'll call it the* object map*) and*
- *maps $F : Hom_{\mathbf{C}}(X,Y) \to Hom_{\mathbf{D}}(F(X), F(Y))$, so that*

$$F(f) : F(X) \to F(Y) \text{ or } F(f)(F(X)) = F(f(X))$$

$\forall X, Y \in Obj(\mathbf{C})$ s.t.

$$F(1_X) = 1_{F(X)} \qquad \forall X \in \mathbf{C}$$

$$F(g \circ f) = F(g) \circ F(f) \qquad \begin{aligned} &\forall f : X \to Y, \\ &g : Y \to Z \end{aligned} \qquad X, Y, Z \in Obj(\mathbf{C}) \tag{43}$$

Mac Lane (1978) [5] calls this the arrow function *(also written as T for Mac Lane's notation; F for our notation); we'll call F to be the* morphism map.
(2) *(composition law for functors)*
For categories $\mathbf{A}, \mathbf{B}, \mathbf{C}$, functors $F : \mathbf{A} \to \mathbf{B}$, $G : \mathbf{B} \to \mathbf{C}$,
Composition $G \circ F : \mathbf{A} \to \mathbf{C}$, is a functor defined by

$$\begin{aligned} (G \circ F)(X) &= G(F(X)) \qquad \forall X \in Obj\mathbf{A}, \text{ and} \\ (G \circ F)(f) &= G(F(f)), \qquad \forall \text{ morphism } f \in Mor(\mathbf{C}) \end{aligned} \tag{44}$$

Diagrammatically,

$$X \xrightarrow{f} Y \quad \xrightarrow{F} \quad F(X) \xrightarrow{F(f)} F(Y)$$

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & Y \arrow[r, "g"] & Z \end{tikzcd} \quad \xrightarrow{F} \quad \begin{tikzcd} F(X) \arrow[r, "F(f)"] \arrow[rr, bend right, "F(g \circ f)"'] & F(Y) \arrow[r, "F(g)"] & F(Z) \end{tikzcd}$$

i.e.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "F"] & Y \arrow[d, "F"] \\ F(X) \arrow[r, "F(f)"] & F(Y) \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} X \arrow[r, "f"] \arrow[d, "F"] \arrow[rr, bend left, "g \circ f"] & Y \arrow[r, "g"] \arrow[d, "F"] & Z \arrow[d, "F"] \\ F(X) \arrow[r, "F(f)"] \arrow[rr, bend right, "F(g\circ f) = F(g)\circ F(f)"'] & F(Y) \arrow[r, "F(g)"] & F(Z) \end{tikzcd}$$

11.1. **Examples of Functors.** cf. pp. 13, 3. Functors, Maclane (1978) [5]
Power Set Functor.

$$\begin{gathered} P : \mathbf{Set} \to \mathbf{Set} \\ P : \mathrm{Obj}\mathbf{Set} \to \mathrm{Obj}(\mathbf{Set}) \\ P : X \in \mathrm{Obj}\mathbf{Set} \mapsto 2^X \in \mathrm{Obj}\mathbf{Set}, \text{ i.e.} \\ P : X \mapsto 2^X = \{S | S \subseteq X\} \\ P : \mathrm{Hom}_{\mathbf{Set}}(X,Y) \to \mathrm{Hom}_{\mathbf{Set}}(P(X), P(Y)) \\ P(f) : P(X) \to P(Y) \text{ or } P(f)(P(X)) = P(f(X)) \\ P(f) : S \in 2^X \mapsto f(S) \in 2^Y \text{ or } f(S) \subseteq Y \end{gathered}$$

$$P(1_X) = 1_{PX} \text{ and } P(g \circ f) = P(g)P(f), \text{ so } P \text{ defines a functor}$$

Algebraic topology; singular homology in a given dim. n ($n \in \mathbb{N}$) Maclane (1978) [5]
$\forall$ topological space X, assign $X \mapsto$ abelian group $H_n(X) \equiv n$th homology group of X, also $\forall$ cont. $f : X \to Y$ of spaces corresponding group homomorphism.

$$\begin{gathered} H_n(f) : H_n(X) \to H_n(Y) \\ H_n : \mathrm{Obj}(\mathbf{Top}) \to \mathrm{Obj}(\mathbf{Ab}) \\ H_n : X \mapsto H_n(X) \\ H_n(f)(H_n(X)) = H_n(f(X)) \end{gathered}$$

$\Longrightarrow H_n : \mathbf{Top} \to \mathbf{Ab}$.
e.g. if $X = Y = S^1$, $H_1(S^1) = \mathbb{Z}$
So group homomorphism $H_1(f) : \mathbb{Z} \to \mathbb{Z}$ determined by integer $d \in \mathbb{Z}$ (image of 1)
d is usual "degree" of cont. map $f : S^1 \to S^1$.
In this case, and in general, homotopic maps $f, g : X \to Y$
Then $\exists$ homotopy $H : X \times [0,1] \to Y$ s.t. $H(x,0) = f(x)$; $H(x,1) = g(x) \forall x \in X$
Example 1.2.3 (forgetful functors) Leinster (2014) [8], pp. 18

(a) $U : \mathbf{Grp} \to \mathbf{Set}$
if group G then $U(G)$ is underlying set of G.
If group homomorphism $f : G \to H$,
$U(f)$ is function f itself.
So U forgets group structure of groups and forgets group homomorphisms are homomorphisms.

(c) $\mathbf{Ab} \equiv$ category of abelian groups
$\mathbf{Ring} \to \mathbf{Ab}$ forgets multiplicative structure.
$U : \mathbf{Ring} \to \mathbf{Mon}$ forgets additive structure; $\mathbf{Mon} \equiv$ category of monoids

- (monoid) homomorphism $h : M \to N$, $\forall$ monoid M, N s.t. $\forall m, n \in M$,
$h(m \cdot n) \equiv h(m \cdot_M n) = h(m) \cdot h(n) \equiv h(m) \cdot_N h(n)$ and
$h(e_M) = e_N$.

  $h : m \mapsto h(m)$ so $h : \mathrm{Obj}\mathbf{C}(M) \to \mathrm{Obj}\mathbf{C}(N)$.
Let $m \cdot n \equiv f_m(n)$, $f_m \in \mathrm{Hom}\mathbf{C}(M)$

$$\begin{aligned} &h(m) \cdot h(n) \equiv f_{h(m)}(h(n)),\ f_{h(m)} \in \mathrm{Hom}\mathbf{C}(N) \\ &h(f_m) = h(m) \\ &h(f_m)(h(n)) = h(m) \cdot h(n) = h(m \cdot n) = h(f_m \cdot n) \end{aligned}$$

Compare the last statement with $F(f)(F(X)) = F(f(X))$.
- poset P, Q (set P, equipped with $\leq$), poset as category $\mathbf{P}, \mathbf{Q}$, $\mathrm{Obj}\mathbf{P} = P$, $\mathrm{Obj}\mathbf{Q} = Q$
$\mathrm{Hom}\mathbf{P} \ni m$ s.t. $m : x \mapsto y$ iff $x \leq y$.

  Let functor $F : \mathbf{P} \to \mathbf{Q}$ s.t.

$$\begin{gathered} F : \mathrm{Obj}\mathbf{P} \to \mathrm{Obj}\mathbf{Q} \\ F : x \mapsto F(x) \\ F : \mathrm{Hom}\mathbf{P} \to \mathrm{Hom}\mathbf{Q} \\ F : m \mapsto F(m) \end{gathered}$$

s.t. for $m(x) = y$ and $x \leq y$,
$F(m) : F(x) \mapsto F(y)$ and $F(x) \leq F(y)$ (so that $F(m)$ is also monotone), i.e.

$$F(m)(F(X)) = F(m(X)) = F(y)$$

Kleene closure itself is a functor from $\mathbf{Set}$ to $\mathbf{Set}$, from A to A^*, from f to f^*.
It's the composition of $U \circ F$ of underlying functor $U : \mathbf{Mon} \to \mathbf{Set}$ and free functor $F : \mathbf{Set} \to \mathbf{Mon}$.

- *arrow category* $\mathbf{C}^{\to}$ of category $\mathbf{C}$:

$$\mathrm{Obj}\mathbf{C}^{\to} = \mathrm{Mor}\mathbf{C}$$

Given $f : A \to B$, $f' : A' \to B'$, morphism (arrow) $g : f \mapsto f'$ is a "commutative square", i.e. $g : \mathrm{Hom}_{\mathbf{C}}(A \to B) \to \mathrm{Hom}_{\mathbf{C}}(A' \to B')$

$$\begin{tikzcd} A \arrow[r, "g_1"] \arrow[d, "f"] & A' \arrow[d, "f'"] \\ B \arrow[r, "g_2"] & B' \end{tikzcd}$$

where $g_1, g_2 \in \mathrm{Mor}\mathbf{C}$,
i.e. such a morphism (arrow) is a pair of morphisms (arrows) $g = (g_1, g_2)$ in $\mathbf{C}$ s.t.

$$g_2 \circ f = f' \circ g$$

identity morphism (arrow) 1_f on object $f : A \to B$ is pair $(1_A, 1_B)$.
Composition of arrows is done componentwise:

$$(h_1, h_2) \circ (g_1, g_2) = (h_1 \circ g_1, h_2 \circ g_2)$$

$$\begin{tikzcd}
A \arrow[r, "g_1"] \arrow[d, "f"] \arrow[rr, bend left, "h_1 \circ g_1"] & A' \arrow[r, "h_1"] \arrow[d, "f'"] & A'' \arrow[d, "f''"] \\
B \arrow[r, "g_2"] \arrow[rr, bend right, "h_2 \circ g_2"'] & B' \arrow[r, "h_2"] & B''
\end{tikzcd}$$

Observe that there are 2 functors:

$$\mathbf{C} \xleftarrow{\mathbf{dom}} \mathbf{C}^{\rightarrow} \xrightarrow{\mathbf{cod}} \mathbf{C}$$

- *algebra* cf. pp. 14, Sec. 3, "Functors" of Mac Lane (1978) [5].
  $\forall$ commutative ring K, set of all non-singular $n \times n$ matrices with entries in K = general linear group $GL_n(K)$.
  $\forall$ homomorphism $f : K \to K'$ of rings produces
  homomorphism $GL_n f : GL_n(K) \to GL_n(K)$ of groups.

  $\forall\, n \in \mathbb{N}$ (natural numbers), defined functor $GL_n : \mathbf{CRng} \to \mathbf{Grp} \equiv GL_n \mathbf{CommRing} \to \mathbf{Grp}$.

  $\forall$ group G, set of all products of commutators $xyx^{-1}y^{-1}$ $(x, y \in G)$ is a normal subgroup $[G, G]$ of G called commutator subgroup.

  Since $\forall$ homomorphism $G \to H$ of groups carries commutators to commutators.
  $G \mapsto [G, G]$ defines evident functor $\mathbf{Grp} \to \mathbf{Grp}$,
  while $G \mapsto G/[G, G]$ define functor $\mathbf{Grp} \to \mathbf{Ab}$, factor-commutator functor.
  TODO - understand the previous commutators.

11.1.1. *Endofunctor.*

Definition 62 (Endofunctor)**.** *Consider* ***endofunctor*** $T : \mathbf{C} \to \mathbf{C}$*. This implies that, for the*

object map $T : Obj\mathbf{C} \to Obj\mathbf{C}$,
$T : X \to Y$

object map T *acts like a morphism, in Mor*$\mathbf{C}$ *(!!!). In this case* T *acts like* $T \in Hom_{\mathbf{C}}(X, Y)$.

morphism map $T : Hom_{\mathbf{C}}(X, Y) \to Hom_{\mathbf{C}}(F(X), F(Y))$ *s.t.*

$$\begin{aligned}
& T(f) \equiv Tf : T(X) \to T(Y) \text{ so that } Tf \in Hom_{\mathbf{C}}(T(X), T(Y)) \\
& T(1_X) = 1_{T(X)} \in Hom_{\mathbf{C}}(T(X), T(Y)) (\textit{implied from Kashiwara and Schapira (2006) [1] in that} 1_X \in Hom(X, X)) \\
& T(g \circ f) = T(g) \circ T(f), \forall f : X \to Y,\ g : Y \to Z,\ X, Y, Z \in Obj\mathbf{C}
\end{aligned} \tag{45}$$

So $T(f) \equiv Tf$ *is indeed a morphism map.*

11.1.2. *Hom functors.* Let $\mathbf{C}$ locally-small category (i.e. category s.t. hom-classes are actually sets and not proper classes). $\forall\, A, B \in \mathrm{Obj}\mathbf{C}$,

Definition 63 (covariant Hom functor)**.** ***covariant Hom functor*** $Hom(A, -) : \mathbf{C} \to \mathbf{Set}$

$$\begin{aligned}
& Hom(A, -) : Obj\mathbf{C} \to 2^{Mor(\mathbf{C})} \\
& Hom(A, -) : X \mapsto Hom(A, X) \\
& Hom(A, -) : Mor\mathbf{C} \to (Mor(\mathbf{C}) \to Mor(\mathbf{C})) \\
& Hom(A, -) : Hom(X, Y) \to (Hom(A, X) \to Hom(A, Y)) \\
& Hom(A, -) : f \mapsto Hom(A, f) \text{ where} \\
& Hom(A, f) : Hom(A, X) \to Hom(A, Y) \\
& \qquad g \mapsto f \circ g \quad \forall g \in Hom(A, X)
\end{aligned} \tag{46}$$

<u>Claim:</u> $\mathrm{Hom}(A, -)$ is a functor.

Proof.
- *identity*: $\mathrm{Hom}(A, -) : 1_X \mapsto \mathrm{Hom}(A, 1_X)$ where

$$\begin{aligned}
\mathrm{Hom}(A, 1_X) : & \mathrm{Hom}(A, X) \to \mathrm{Hom}(A, X) \\
& g \mapsto 1_X \cdot g = g, \quad \forall g \in \mathrm{Hom}(A, X)
\end{aligned}$$

  Hence

$$\mathrm{Hom}(A, -)(1_X) = 1_{\mathrm{Hom}(A,X)} \in (\mathrm{Hom}(A, X) \to \mathrm{Hom}(A, X))$$

- composition: Let

$$\begin{aligned}
& f \in \mathrm{Hom}(X, Y) \\
& g \in \mathrm{Hom}(Y, Z) \\
& k \in \mathrm{Hom}(A, X)
\end{aligned}$$

$$\mathrm{Hom}(A, g) \circ \mathrm{Hom}(A, f)(k) = \mathrm{Hom}(A, g)(f \circ k) = g \circ f \circ k = (g \circ f)(k) = \mathrm{Hom}(A, g \circ f)(k)$$

□

Definition 64 (contravariant Hom functor)**.** ***contravariant Hom functor*** $Hom(-, B) : \mathbf{C} \to \mathbf{Set}$

$$\begin{aligned}
& Hom(-, B) : Obj\mathbf{C} \to 2^{Mor(\mathbf{C})} \\
& Hom(-, B) : X \mapsto Hom(X, A) \\
& Hom(-, B) : Mor\mathbf{C} \to (Mor(\mathbf{C}) \to Mor(\mathbf{C})) \\
& Hom(-, B) : Hom(X, Y) \to (Hom(Y, B) \to Hom(X, B)) \\
& Hom(-, B) : h \mapsto Hom(h, B) \text{ where} \\
& Hom(h, B) : Hom(Y, B) \to Hom(X, B) \\
& \qquad g \mapsto g \circ h \quad \forall g \in Hom(Y, B)
\end{aligned} \tag{47}$$

$\mathrm{Hom}(A, -)$, $\mathrm{Hom}(-, B)$ are related in a natural manner: for

$$\begin{aligned}
f : B \to B',\ & f \in \mathrm{Hom}(B, B') \\
g : A \to A',\ & g \in \mathrm{Hom}(A, A') \\
\mathrm{Hom}(A, f) : & \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, B') \\
\mathrm{Hom}(g, B) : & \mathrm{Hom}(A, B) \to \mathrm{Hom}(A', B)
\end{aligned}$$

$$\begin{tikzcd}
\mathrm{Hom}(A,B) \arrow[r, "\mathrm{Hom}(g,B)"] \arrow[d, "\mathrm{Hom}(A,f)"'] & \mathrm{Hom}(A',B) \arrow[d, "\mathrm{Hom}(A',f)"] \\
\mathrm{Hom}(A,B') \arrow[r, "\mathrm{Hom}(g,B')"] & \mathrm{Hom}(A',B')
\end{tikzcd}$$

Definition 65 (2-variable Hom functor)**.** *2-variable Hom functor $Hom(-,-)$ is a bifunctor.*

$$Hom(-,-) : \mathbf{C}^{op} \times \mathbf{C} \to \boldsymbol{Set}$$

$$\begin{tikzcd}
 & D \\
C \arrow[ur, "f\circ g\circ h"] \arrow[r, "g\circ h"] \arrow[d, "h"'] & B \arrow[u, "f"'] \\
A \arrow[ur, "g"'] &
\end{tikzcd}$$

$Hom(-,-)$ is a functor such that:

$$\begin{gathered}
Hom(h,f)(g) = f \circ g \circ h \\
Hom(h,f) : Hom(A,B) \to Hom(C,D) \\
h : C \to A,\ h \in Hom(C,A) \\
f : B\dot{D},\ f \in Hom(B,D) \\
(h,f) \in Hom(C,A) \times Hom(B,D) \mapsto Hom(h,f) \in (Hom(A,B) \to Hom(C,D)) \\
Hom(-,-) : Hom(C,A) \times Hom(B,D) \to (Hom(A,B) \to Hom(C,D))
\end{gathered}$$

and

$$\begin{gathered}
Hom(-,-) : Obj\mathbf{C}^{op} \times \mathbf{C} \\
Hom(-,-) : (C,D) \mapsto Hom(C,D)
\end{gathered}$$

12. Construction of Categories

12.1. **Slice category.** cf. Construction 4 of Awodey (2010) [19], pp. 15.

Slice category $\mathbf{C}/C$, of category $\mathbf{C}$ over object $C \in \mathrm{Obj}\mathbf{C}$,
$\mathrm{Obj}\mathbf{C}/C \ni f \in \mathrm{Mor}\mathbf{C}$ s.t. $\mathrm{cod}(f) = C$
$\mathrm{Mor}\mathbf{C}/C \ni g$ from $f : X \to C$ to $f' : X' \to C$ is a morphism (arrow) $g : X \to X'$ in $\mathbf{C}$ s.t. $f' \circ g = f$

$$\begin{tikzcd}
X \arrow[rr, "g"] \arrow[dr, "f"'] & & X' \arrow[dl, "f'"] \\
 & C &
\end{tikzcd}
\qquad
\begin{tikzcd}
X \arrow[loop right, "1_{\mathbf{C}/C}"] \arrow[d, "f"'] \\
C
\end{tikzcd}
\qquad
\begin{tikzcd}
X \arrow[r, "g"] \arrow[dr, "f"'] & X' \arrow[r, "h"] \arrow[d, "f'"] & X'' \arrow[dl, "f''"] \\
 & C &
\end{tikzcd} \tag{48}$$

If $\mathbf{C} = \mathbf{P}$ poset category, $p \in \mathbf{P}$ (i.e. $p \in \mathrm{Obj}(\mathbf{P})$), then

$$\mathbf{P}/p \simeq \downarrow (p)$$

Slice category $\mathbf{P}/p$ is just the "principal ideal" $\downarrow (p)$ of elements $q \in \mathbf{P}$ with $q \leq p$.

12.1.1. *S-indexed set X and indexed functions as morphisms of slice category* $\mathbf{Set}/S$.

Definition 66 (typed set)**.** *S-indexed set is set X together with function $\tau : X \to S$.*

If $x \in X$, $\tau(x) = s$, then x is of ***type*** *s, so X is a* ***typed set****.*

$\{\tau^{-1}(s) | s \in S\}$, $\tau^{-1}(s) \subset X \equiv$ family of sets indexed by S.

cf. 2.6.11 of Barr and Wells (2012) [20]

Definition 67. *set X typed by $S \to$ set X' typed by S that preserves typing (element of type $s \mapsto$ element of type s), is exactly an arrow (morphism) of slice category* $\mathbf{Set}/S$*, called* ***indexed function*** *or* ***typed function****.*

cf. 2.6.13 Indexed functions of Barr and Wells (2012) [20]

12.1.2. *Underlying functor of the slice F.* Let $f : A \to C$, $f \in \mathrm{Hom}(A,C) \subset \mathrm{Mor}\mathbf{C}$ and $f \in \mathrm{Obj}\mathbf{C}/C$.

Let $f, g \in \mathrm{Obj}\mathbf{C}/C$, s.t. $f \in \mathrm{Hom}(A,C)$, $g \in \mathrm{Hom}(B,C)$
$h \in \mathrm{Hom}(f,g) \subseteq \mathrm{Mor}\mathbf{C}/C$ in that
$h : f \to g$ s.t. $h : A \to B$, so that $g \circ h = f$.

underlying functor of the slice F (notation U in Barr and Wells (2012) [20]):

$$\begin{aligned}
&F : \mathbf{C}/C \to \mathbf{C} \\
&F : \mathrm{Obj}\mathbf{C}/C \to \mathrm{Obj}\mathbf{C} \\
&F : f \mapsto F(f) = A \\
&F : \mathrm{Mor}\mathbf{C}/C \to \mathrm{Mor}\mathbf{C} \\
&F : h \mapsto h
\end{aligned} \tag{49}$$

Special case: $\mathbf{C} = \mathbf{Set}$, $C = S \ni \mathrm{Obj}\mathbf{Set}$, i.e. S is a set.
For $f \in \mathrm{Obj}\mathbf{Set}/S$, $f : T \to S$, where T is a set, object f is an S-indexed set. (e.g. $x(i) \in T \subset \mathbb{R}$, where for some $y \in T \subset \mathbb{R}$, $\exists i \in \mathbb{Z}$ s.t. $y \mapsto i$).

$$\begin{aligned}
&F : \mathbf{Set}/S \to \mathbf{Set} \\
&F : f \mapsto T
\end{aligned}$$

The underlying functor F forgets the S-indexing.

12.1.3. *Coslice category.*

$$\begin{tikzcd} X & & X' \arrow[ll, "\overline{g}"'] \\ & C \arrow[ul, "\overline{f}"] \arrow[ur, "\overline{f'}"'] & \end{tikzcd}$$

Coslice category $C/\mathbf{C}$ of category $\mathbf{C}$ under object $C \in \text{Obj}\mathbf{C}$:
$\text{Obj}(C/\mathbf{C}) = \overline{f} \in \text{Mor}\mathbf{C}$ s.t. $\text{dom}\overline{f} = \overline{C} = C$,

$\text{Mor}(C/\mathbf{C}) \ni \overline{g} : \overline{X}' \to \overline{X}$s.t.
$$\overline{g} \circ \overline{f'} = \overline{f}$$
cf. Example 1.8 of Awodey (2010) [19], pp. 15. *pointed sets*
$\mathbf{Sets}_*$ of pointed sets consisting of sets A and distinguished element $a \in A$, and, i.e.
$\text{Obj}\mathbf{Set}_* \ni$ set A with distinguished element $a \in A$,
$\text{Mor}\mathbf{Set}_* \ni f : (A, a) \to (B, b)$, function $f : A \to B$ preserves the "points" $f(a) = b$.

$$\mathbf{Sets}_* = \mathbf{Set}_* \simeq 1/\mathbf{Sets}$$

i.e. $\mathbf{Sets}_*$ isomorphic to coslice category of sets "under" any singleton $1 = \{*\}$

functions $a : 1 \to A$ correspond uniquely to elements
$$a(*) = a \in A$$
morphisms (arrows) $f : (A, a) \to (B, b)$ correspond exactly to

$$\begin{tikzcd} 1 \arrow[r, "a"] \arrow[dr, "b"'] & A \arrow[d, "f"] \\ & B \end{tikzcd}$$

12.2. **Free monoid.** Given set S, e.g. $S =$ "alphabet" A,
free monoid S^* is set S^*.

$S^* =$ set of all lists (finite sequences) of elements of S,
e.g. set of all *words* (finite sequence of letters) *over* $A = A^*$.

Concatenation: $* : \forall\, w, w^* \in A^*$, $w * w' = ww'$
empty list (e.g. empty word "-"), (), $*$, or ϵ is a unit. Or we'll use this notation: `""`

$$\Longrightarrow (S^*, *, \text{""}) \text{is a monoid.}$$

monoid M is **freely generated** by subset A of M, if

(1) "no junk": $\forall\, m \in M$, m can be written as product of elements of A,
$$m = a_1 \dots a_n,\ a_i \in A$$
(2) "no noise": no "nontrivial" relations hold in M, i.e. if $a_1 \dots a_j = a_1' \dots a_k'$, then this is required by the axioms for monoid.

$A^* =$ Kleene closure.

12.2.1. *Free monoid functor: underlying functor.*
$$F : \mathbf{Set} \to \mathbf{Mon}$$
$$F : A \in \text{Obj}\mathbf{Set} \mapsto \text{ free monoid } F(A) = \text{ Kleene closure } A^*$$
For $f : A \to B$, $f \in \text{Hom}_{\mathbf{Set}}(A, B)$,
$$F : f \mapsto F(f) \text{ s.t.}$$
$$F(f)(w) = F(f)(a_1 \dots a_m) = (f(a_1), \dots f(a_m)) = w' \in B^*, \text{ i.e.}$$
$$F(f) = f^* : F(A) \to F(B)$$

13. Universal mapping property

13.1. **Examples of Universal mapping property.** cf. Ch. 0 "Introduction" of Leinster (2014) [8].
Example 0.1. Let $\{1\}$ denote set with 1 element.
Then $\forall$ sets $X \in \mathbf{Set}$, $\exists!$ map $X \to \{1\}$
$$X \xrightarrow{f} \{1\}$$

Proof. $\exists\, X \to \{1\}$, $\forall\, X$ because define $f : X \to \{1\}$ s.t. $f(x) = 1$, $\forall\, x \in X$.
(!) if for $X \to \{1\}$, then $x \mapsto 1$, so map is equal to f. □

Example 0.2. $\forall$ rings R with multiplicative identity 1, $\exists!$ homomorphism $\mathbb{Z} \to R$.
Define $\phi : \mathbb{Z} \to R$ by
$$\phi(n) = \begin{cases} \sum_{i=1}^n 1 & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ -\phi(-n) & \text{if } n < 0 \end{cases}$$
Check that ϕ homomorphism. (TODO)

(!) $\psi(1) = 1$ (homomorphisms preserve multiplicative identities)
homomorphisms preserve addition:
$$\psi(n) = \psi(\sum_{i=1}^n 1) = \sum_{i=1}^n \psi(1) = \sum_{i=1}^n 1 = \phi(n)$$
homomorphisms preserve zero: $\psi(0) = 0 = \phi(0)$
homomorphisms preserve negatives: $\psi(n) = -\psi(-n) = -\sum_{i=1}^{-n} 1 = -\phi(-n) = \phi(n)$.

Lemma 2 (0.3 Leinster (2014) [8])**.** *Let ring A s.t. $\forall$ rings $R, \exists!$ homomorphism $A \to R$ (i.e. A is "initial"). Then $A \cong \mathbb{Z}$ (isomorphic).*

Proof. A initial, so $\exists!$ homomorphism $\phi : A \to \mathbb{Z}$.
$\mathbb{Z}$ initial from Ex. 13.1. $\exists!$ homomorphism $\phi' : \mathbb{Z} \to A$.

$\phi' \circ \phi : A \to A$, $1_A : A \to A$ are homomorphisms.
Since A initial, $\phi' \circ \phi = 1_A$.
$\phi \circ \phi' : \mathbb{Z} \to \mathbb{Z}$. Since $\mathbb{Z}$ initial, $1_{\mathbb{Z}} : \mathbb{Z} \to \mathbb{Z} \Longrightarrow \phi \circ \phi' = 1_{\mathbb{Z}}$.
$$\Longrightarrow A \cong \mathbb{Z} \text{ (isomorphic)}$$
□

Example 0.4. pp. 3 Leinster (2014) [8] Let vector space V with basis $(v_s)_{s\in S}$ (.e.g if V finite-dim., e.g. $S=\{1,2\dots,n\}$). Thus $\forall$ vector space W, $\exists$ natural 1-to-1 correspondence between linear maps $\{V\to W\}$ and functions $S\to W$.

i.e. define $i:S\to V$

$$i(s)=v_s \quad (s\in S)$$

Then universal mapping property (given $i,f,\exists !\overline{f}$)

$$\begin{tikzcd} S \arrow[r, "i"] \arrow[dr, "f"'] & V \arrow[d, dashed, "\overline{f}"] \\ & W \end{tikzcd} \quad \text{or} \quad \begin{tikzcd} V \arrow[r, dashed, "\overline{f}"] & W \\ S \arrow[u, "i"] \arrow[ur, "f"'] & \end{tikzcd}$$

$$\begin{aligned} \{S\to W\} &\to \{\text{ linear } V\to W\} \\ \overline{f}\circ i &\mapsto \overline{f} \end{aligned} \qquad \text{bijective}$$

Example 0.5. pp. 4 Leinster (2014) [8]. Given set S, topological space $D(S)$; equip $D(S)$ with discrete topology (all subsets open).
$\Longrightarrow \forall$ map $S\to$ space X is cont.
i.e.

$$\begin{tikzcd} D(S) \arrow[r, dashed, "\overline{f}"] & X \\ S \arrow[u, "i"] \arrow[ur, "f"'] & \end{tikzcd}$$

$\forall$ topological space X, $\forall f:S\to X$, $\exists !$ cont. $\overline{f}:D(S)\to X$ s.t. $\overline{f}\circ i=f$.
If $D(S)$ equipped with indiscrete topology (open $S=\{\phi,S\}$), then property is false.

Example 0.6. pp. 4 Leinster (2014) [8]. Given vector space U,V,W, bilinear map $f:U\times V\to W$ is linear.

$$\begin{aligned} f(u,v_1+\lambda v_2) &= f(u,v_1)+\lambda f(u,v_2) \\ f(u_1+\lambda u_2,v) &= f(u_1,v)+\lambda f(u_2,v) \end{aligned}$$

$\forall u,u_1,u_2\in U$, $v,v_1,v_2\in V$, scalars λ.
e.g. scalar product (dot product) $\mathbf{u}\cdot\mathbf{v}$ bilniear, cross product bilinear.

$\exists$ "universal bilinear map out of $U\times V$", i.e.
$\exists$ vector space T, $\exists$ bilinear map $b:U\times V\to T$

$$\begin{tikzcd} T \arrow[r, dashed, "\overline{f}"] & W \\ U\times V \arrow[u, "b"] \arrow[ur, "f"'] & \end{tikzcd}$$

i.e. $\forall$ bilinear f, given b, $\exists !$ linear $\overline{f}$, s.t. $\overline{f}\circ b=f$

Lemma 3 (0.7, pp. 5 Leinster (2014) [8]). *Let vector spaces U,V. Suppose* $\begin{aligned} b&:U\times V\to T \\ b'&:U\times V\to T' \end{aligned}$ *, both universal bilinear maps.*

Then $T\cong T'$, i.e. $\exists !$ isomorphism $j:T\to T'$ s.t. $j\circ b=b'$

Proof. Take $b':U\times V\to T' \Longrightarrow \exists b$, $\exists !$ linear map $j:T\to T'$ s.t. $j\circ b=b'$.
Take $b:U\times V\to T$. $\exists b'$, $\exists !$ linear map $j':T'\to T$ s.t. $j'\circ b'=b$

$$\begin{tikzcd} T \arrow[r, dashed, "j"] & T' \arrow[d, dashed, "j'"] \\ U\times V \arrow[u, "b"] \arrow[ur, "b'"'] \arrow[r, "b"'] & T \end{tikzcd}$$

Now linear $j'\circ j:T\to T$, s.t. $(j'\circ j)\circ b=b$.
But also $1_T:T\to T$ linear and $1_T\circ b=b$

$$\Longrightarrow j'\circ j=1_T$$

i.e.

$$\begin{tikzcd} T \arrow[r, dashed, "j'\circ j"] & T \\ U\times V \arrow[u, "b"] \arrow[ur, "b"'] & \end{tikzcd} \quad \text{but so is} \quad \begin{tikzcd} T \arrow[r, dashed, "1_T"] & T \\ U\times V \arrow[u, "b"] \arrow[ur, "b"'] & \end{tikzcd}$$

Similarly for $j\circ j'=1_{T'}$.
So j is an isomorphism. $\square$

By Lemma 3, $\exists !$ tensor product $U\otimes V$, not "a" tensor product.
Example 0.8 of Leinster (2014) [8]. Let $\theta:G\to H$ be homomorphism of groups.

$$\ker(\theta) \overset{i}{\hookrightarrow} G \underset{\epsilon}{\overset{\theta}{\rightrightarrows}} H$$

where i inclusion, ϵ trivial homomorphism, i.e.$i(x) = x \forall x \in \ker(\theta)$, $\epsilon(g) = 1 \quad \forall g \in G$.
map i into G s.t. $\theta \circ i = \epsilon \circ i$ and is universal (Ex. 0.11)
Example 0.9 of Leinster (2014) [8]. Let topological space $X = U \cup V$ be covered by 2 open subsets U, V.

$$\begin{tikzcd}
U \cap V \arrow[r, hook, "i"] \arrow[d, hook, "j"] & U \arrow[d, hook, "j'"] \\
V \arrow[r, hook, "i'"] & X
\end{tikzcd}$$

has universal property

$$\begin{tikzcd}
U \cap V \arrow[r, hook, "i"] \arrow[d, hook, "j"] & U \arrow[d, hook, "j'"] \arrow[ddr, "f"] & \\
V \arrow[r, hook, "i'"] \arrow[drr, "g"] & X \arrow[dr, dotted, "h"] & \\
& & Y
\end{tikzcd}$$

$\forall g, Y, f$ s.t. $f \circ i = g \circ j$, $\exists !$ cont. $h : X \to Y$ s.t. $h \circ j' = f$, $h \circ i' = g$.
Under favorable conditions, induced diagram of fundamental groups.

$$\begin{tikzcd}
\pi_1(U \cup V) \arrow[r, "i_*"] \arrow[d, "j_*"] & \pi_1(U) \arrow[d, "j'_*"] \\
\pi_1(V) \arrow[r, "i'_*"] & \pi_1(X)
\end{tikzcd}$$

has same universal property: Van Kampen's thm.

$$\theta : G \to H$$
$$\epsilon : G \to H$$

θ, ϵ have common domain and codomain.
$i : \ker(\theta) \to G$ is s.t. $\theta \circ i = \epsilon \circ i$

$$\begin{tikzcd}
\ker(\theta) \arrow[r, hook, "i"] & G \arrow[r, shift left, "\theta"] \arrow[r, shift right, "\epsilon"'] & H \\
J \arrow[u, dotted, "u"] \arrow[ur, "a"] & &
\end{tikzcd}$$

$\forall a : J \to G$ s.t. $\theta \circ a = \epsilon \circ a$, then $\exists !\, u : J \to \ker\theta$ s.t.

$$a = i \circ u$$

13.1.1. *Free monoid A^* on set A.* $\exists\, i : A \to A^*$, given any monoid N, any $f : A \to |N|$ (where $|N|$ is the underlying set of monoid N),
$\exists !$ monoid homomorphism $\overline{f} : A^* \to N$ s.t. $|\overline{f}| \circ i = f$.

$$\mathbf{Mon} : \quad \begin{tikzcd} A^* \arrow[r, dotted, "\overline{f}"] & N \end{tikzcd}$$

$$\mathbf{Sets} : \quad \begin{tikzcd}
{|A^*|} \arrow[r, dotted, "|\overline{f}|"] & {|N|} \\
A \arrow[u, "i"] \arrow[ur, "f"'] &
\end{tikzcd}$$

Proposition 18 (Proposition 1.9, Awoday (2010) [19])**.** *] A^* has universal mapping property of free monoid on A.*

Proof. Givne $f : A \to |N|$, define $\overline{f} : A^* \to N$, by

$$\overline{f}(\text{""}) = e_N$$
$$\overline{f}(a_1 \dots a_n) = f(a_1) \dots f(a_n)$$

$\overline{f}$ is thus a homomorphism with $\overline{f}(a) = f(a) \qquad \forall a \in A$.

If $g : A^* \to N$ s.t. $g(a) = f(a) \qquad \forall a \in A$, then $\forall (a_1, \dots a_n) \in A^*$,

$$\overline{f}(a_1 \dots a_n) = \overline{f}(a_1) \dots \overline{f}(a_n) = f(a_1) \dots f(a_n) = g(a_1) \dots g(a_n) = g(a_1 \dots a_n)$$

since $\overline{f} = g$, $\overline{f}$ unique. □

Existence part of UMP is "no noise."
Unique part of UMP is "no junk."
TODO Prop. 1.10, Awodey.

Exercise 1. cf. 3.1.23 Exercises of Barr and Wells (2012) [20]
Given semigroup S, construct monoid $M = S \cup \{e\}$, $e \notin S$.
Let $e = \{S\}$.
Multiplication in M defined as

(1) $xy \in S$ if $x, y \in S$
(2) $xe = ex = x \quad \forall x \in M$

$M \equiv S^1$ in semigroup literature.

(1) S^1 is a monoid (Show).
If $a, b, c \in S \subset M$, then $(ab)c = a(bc)$ (associative holds by definition of semigroup).
$\forall a \in S$, $ae = ea = a$, by how we defined multiplication.
$ee = ee = e$ by either definition, or that multiplication in semigroup S is closed.

$$\{S\}\{S\} = \{S\}\{S\} = \{S\}$$

If for $a, b, c \in M$, any of a, b or $c = e$, use multiplication to show equality for $(ab)c = a(bc)$; e.g.

$$(eb)c = (b) \cdot c = bc = e(bc) = bc$$
$$(ae)c = a \cdot c = a \cdot (ec) = a(ec)$$

S^1 is a monoid.
(2)

14. Actions, Finite State Machines

Definition 68. *Let monoid M with identity 1, and a set S.*
action *of M on S, α,*

$$\begin{aligned} &\alpha : M \times S \to S \text{ s.t.} \\ &\alpha(1, s) = s \quad \forall s \in S \\ &\alpha(mn, s) = \alpha(m, \alpha(n, s)) \quad \forall m, n \in M, \, s \in S \end{aligned} \tag{50}$$

Write $ms \equiv \alpha(m, s)$ then

$$\begin{gathered} 1s = s \\ (mn)s = m(ns) \quad \forall m, n \in M, \, s \in S \end{gathered}$$

Definition 69 (equivariant map)**.** *Let monoid M with actions on sets S, T.*
equivariant map *$\phi : S \to T$ s.t. $m\phi(s) = \phi(ms)$*

Suppose 2 equivariant maps ϕ_{TS}, ϕ_{UT}, sets S, T, U s.t. monoid M has actions on S, T, U:

$$\begin{aligned} \phi_{TS} &: S \to T \\ \phi_{UT} &: T \to U \end{aligned}$$

$$m\phi_{UT} \circ \phi_{TS}(s) = m\phi_{UT}(\phi_{TS}(s)) = \phi_{UT}(m\phi_{TS}(s)) = \phi_{UT}\phi_{TS}(ms)$$

Let $\phi_{UT} \circ \phi_{TS} \equiv \phi_{US}$. ϕ_{US} equivariant.
Associativity: given equivariant maps

$$\begin{aligned} \phi_{TS} &: S \to T \\ \phi_{UT} &: T \to U \\ \phi_{VU} &: U \to V \end{aligned}$$

$$m\phi_{VU}\phi_{UT}\phi_{TS}(s) = m\phi_{VU}(\phi_{UT}\phi_{TS}(s)) = \phi_{VU}(m\phi_{UT}\phi_{TS}(s)) = (\phi_{VU}\phi_{UT})(\phi_{TS}(ms))$$

Thus, equivariant maps obey associativity on $\forall S \in \text{Obj}(M - \mathbf{Act})$, (set S that M has action on)
category $M - \mathbf{Act}$, $\text{Obj}(M - \mathbf{Act}) = \{$ sets that monoid M has action on $\}$
$\text{Mor}(M - \mathbf{Act}) = \{$ equivariant maps ϕ s.t. $\forall m \in M, m\phi(s) = \phi(ms)\}$.

14.1. Actions as functors.

Let α action of monoid M on set S.

Consider $\mathbf{C}(M)$ (monoid as a category; i.e. $\text{Obj}\mathbf{C}(M) = \{M\}$, $\text{Mor}\mathbf{C}(M) = M$).
action α determines function $F_\alpha : \mathbf{C}(M) \to \mathbf{Set}$,

(1) $F_\alpha(*) = S \equiv F_\alpha(\{M\}) = S$
(2) $F_\alpha(m) = s \mapsto \alpha(m, s) \, \forall m \in M, s \in S$

i.e.

$$\begin{aligned} F_\alpha &: \text{Obj}\mathbf{C}(M) \to \text{Obj}\mathbf{Set} \\ F_\alpha &: \{M\} \mapsto S \\ F_\alpha &: \text{Mor}\mathbf{C}(M) \to \text{Mor}\mathbf{Set} \\ F_\alpha &: m \mapsto (s \mapsto \alpha(m, s)) \in \text{Hom}(S, S) \end{aligned}$$

Barr and Wells (1998) [3] uses this notation: $\mathcal{M} = (A, S, s_0, \phi)$ for a "machine." Compare this to the notation used in these notes:

$$\text{FSM} = (\Sigma, S, s_0, \delta)$$

Take note that the finite set or input alphabet $A \equiv \Sigma$ is **not** a monoid.

$\forall$ string $\equiv$ finite sequence $\equiv$ "word" induces sequence of transitions in FSM starting at s_0, and ending on some final state; precisely, define

$$\begin{aligned} &\delta^* : \Sigma^* \times S \to S \\ &\delta^*(\text{""}, s) = s \quad \forall \in s \in S \\ &\delta^*((a)w, s) = \delta(a, \delta^*(w, s)) \quad \forall s \in S, \, w \in \Sigma^*, \, a \in A \end{aligned}$$

Recall free monoid Σ^*, $(\Sigma^*, *, \text{""})$ where $\Sigma^* \in \text{Obj}\mathbf{Set}$ is a set.

Proposition 19. *δ^* is an action of Σ^* on S.*

Proof. Assume

$$\delta^*(wv, s) = \delta^*(w, \delta^*(v, s))$$

then

$$\delta^*((a)wv, s) = \delta(a, \delta^*(wv, s)) = \delta(a, \delta^*(w, \delta^*(v, s))) = \delta^*(aw, \delta^*(v, s))$$

1st., 3rd. equalities from $\delta^*((a)w, s) = \delta(a, \delta^*(w, s))$ (definition of δ^*).
2nd. equality is from inductive hypothesis. □

recognizer - subset $L \subseteq A^*$ of strings which drive FSM from s_0 (start state) to acceptor state, is then set of strings, or language, which is **recognized** by machine FSM, this is the machine as **recognizer**.
transducer - FSM outputs string of symbols (not necessarily in same alphabet) $\forall$ state it enters, or each transition it undergoes.

14.2. Set-valued functors as actions.

cf. 3.2.6 of Barr and Wells (2012) [20].

15. Products, Coproducts

15.1. Sources.

It appears Adámek, Herrlich, and Strecker (2004) [4] defines *sources* to simply give a name and formalize a tuple.

Definition 70 (source)**.** ***source*** *is a tuple: $(a, (f_i)_{i \in I})$, $f_i : A \to A_i$*

15.2. Products.

Definition 71 (Products)**.** *(in Turi's notation [7])*

Given objects C_1, C_2 of category $\mathbb{C}$, ***product*** *(if exists) consists of object $C_1 \times C_2$ of $\mathbb{C}$ and $\pi_1 : C_1 \times C_2 \to C_1$ s.t.*
$\pi_2 : C_1 \times C_2 \to C_2$

$\forall$ object A of $\mathbb{C}$, $\forall f : A \to C_1$, $g : A \to C_2$ $\exists!$ $\langle f, g \rangle : A \to C_1 \times C_2$ s.t. $f = \pi_1 \circ \langle f, g \rangle$, i.e. $g = \pi_2 \circ \langle f, g \rangle$

$$\begin{tikzcd} & A \arrow[dl, "f"'] \arrow[d, dashed, "\langle f, g \rangle"] \arrow[dr, "g"] & \\ C_1 & C_1 \times C_2 \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & C_2 \end{tikzcd}$$

(compare with Leinster (2014) [8])
Let category $\mathcal{A}$, $X, Y \in \mathcal{A}$, ***product*** *of X, Y consists of object P and maps*
(compare this definition with Adámek, Herrlich, and Strecker (2004) [4] and their notation)

product *consisting of*

$$\begin{aligned} C_1 \times C_2 \times \dots \times C_{\mathcal{N}} &\in Obj\mathbf{C} \\ \pi_1 : C_1 \times C_2 \times \dots \times C_{\mathcal{N}} &\to C_1 \\ \pi_2 : C_1 \times C_2 \times \dots \times C_{\mathcal{N}} &\to C_2 \\ &\vdots \\ \pi_{\mathcal{N}} : C_1 \times C_2 \times \dots \times C_{\mathcal{N}} &\to C_{\mathcal{N}} \end{aligned}$$

is s.t.

$$\forall \begin{aligned} A &\in Obj\mathbf{C} \\ f_1 &: A \to C_1 \\ f_2 &: A \to C_2, \\ &\vdots \\ f_{\mathcal{N}} &: A \to C_{\mathcal{N}} \end{aligned}$$

$\exists ! \langle f_1, f_2, \dots, f_{\mathcal{N}} \rangle : A \to C_1 \times C_2 \times \dots \times C_{\mathcal{N}}$ *s.t.*

$$\begin{aligned} f_1 &= \pi_1 \circ \langle f_1, f_2, \dots f_{\mathcal{N}} \rangle \\ f_2 &= \pi_2 \circ \langle f_1, f_2, \dots f_{\mathcal{N}} \rangle \\ &\vdots \\ f_{\mathcal{N}} &= \pi_{\mathcal{N}} \circ \langle f_1, f_2, \dots f_{\mathcal{N}} \rangle \end{aligned}$$

Let's use the notation of Adámek, Herrlich, and Strecker (2004) [4]. Also note that these references *assume* the *universal mapping property* in the definition of a **product**.

Definition 72 (Product (assuming univeral mapping property))**.**

$$\begin{tikzcd} & & A \arrow[dll, dashed, "f"'] \arrow[dl, "f_i"] \arrow[ddl, "f_{i'}"] \\ P \arrow[r, "p_i"] \arrow[dr, "p_{i'}"'] & A_i & \\ & A_{i'} & \end{tikzcd} \tag{51}$$

For the case of only 2 objects associated with I*,*

$$\begin{tikzcd} & A \arrow[dl, "f_1"'] \arrow[d, dashed, "f"] \arrow[dr, "f_2"] & \\ A_1 & P \arrow[l, "p_1"'] \arrow[r, "p_2"] & A_2 \end{tikzcd} \tag{52}$$

product $P, \{p_i, A_i\}_{i \in I}$ *s.t.* $\forall A, \{f_i, A_i\}_{i \in I}$, $\exists !$ *morphism* $f : A \to P$ *s.t.* $f_i = p_i \circ f$, $i \in I$.

This definition is the same one used in pp. 35, Def. 2.16, of Awodey (2010) [19].

Consider the definition that doesn't use the universal mapping property from Pareigis (2004) [9].

Definition 73 (Product without universal property)**.** *Given objects* $A, B \in Obj\mathbf{C}$*, category* $\mathbf{C}$*.*

Object $A \times B$*, morphisms* p_A, p_B*,* $\begin{aligned} p_A &: A \times B \to A \\ p_B &: A \times B \to B \end{aligned}$ *is called a (not the) (categorical) product of* A, B*, projections* p_A, p_B*.*

If $\forall$ *object* $T \in Obj\mathbf{C}$*,* $\exists$ *isomorphism*

$$Mor_{\mathbf{C}}(T, A \times B) \cong Mor_{\mathbf{C}}(T, A) \times Mor_{\mathbf{C}}(T, B) \tag{53}$$

(where $Mor_{\mathbf{C}}$ *is the Cartesian product) s.t.*

$$\begin{tikzcd} & Mor_{\mathbf{C}}(T, A \times B) \arrow[dl, "Mor_{\mathbf{C}}(T, p_A)"'] \arrow[d, dashed, "h"] \arrow[dr, "Mor_{\mathbf{C}}(T, p_B)"] & \\ Mor_{\mathbf{C}}(T, A) & Mor_{\mathbf{C}}(T, A) \times Mor_{\mathbf{C}}(T, B) \arrow[l, "p_{Mor_{\mathbf{C}}(T,A)}"'] \arrow[r, "p_{Mor_{\mathbf{C}}(T,B)}"] & Mor_{\mathbf{C}}(T, B) \end{tikzcd}$$

(54)

$$\begin{tikzcd} & u \arrow[dl, mapsto, "Mor_{\mathbf{C}}(T, p_A)"'] \arrow[d, mapsto, dashed, "h"] \arrow[dr, mapsto, "Mor_{\mathbf{C}}(T, p_B)"] & \\ f & h(u) = (f, g) \arrow[l, mapsto, "p_{Mor_{\mathbf{C}}(T,A)}"'] \arrow[r, mapsto, "p_{Mor_{\mathbf{C}}(T,B)}"] & g \end{tikzcd}$$

where $\begin{aligned} Mor_{\mathbf{C}}(T, p_A)(u) &= f = p_{Mor_{\mathbf{C}}(T,A)} h(u) \\ Mor_{\mathbf{C}}(T, p_B)(u) &= g = p_{Mor_{\mathbf{C}}(T,B)} h(u) \end{aligned}$

Proposition 20 (Equivalent definition of product by universal mapping property)**.** *Given objects* $A, B \in Obj\mathbf{C}$*, category* $\mathbf{C}$*,*

object $A \times B$*, morphisms* $\begin{aligned} p_A &: A \times B \to A \\ p_B &: A \times B \to B \end{aligned}$ *is a (categorical) product,*

iff $\forall$ *object* $T \in Obj\mathbf{C}$*,* $\forall \begin{aligned} f &: T \to A, \\ g &: T \to B \end{aligned}$ $\exists !$ *morphism* $(f, g) : T \to A \times B$*, s.t.*

$$\begin{tikzcd} & T \arrow[dl, "f"'] \arrow[d, dashed, "(f,g)"'] \arrow[dr, "g"] & \\ A & A \times B \arrow[l, "p_A"'] \arrow[r, "p_B"] & B \end{tikzcd}$$

(55)

cf. Proposition "Characterization of products by universal mapping property", 2.7.4 in Pareigis (2004) [9].

Proof. $\Longleftarrow$:

Let $(A \times B, p_A, p_B)$ be a "product" of $A \times B$, obeying the universal (mapping) property. Let $T \in \text{Obj}\mathbf{C}$. Let $u \in \text{Mor}_{\mathbf{C}}(T, A \times B)$.

Define $f := p_A \circ u : T \to A$
$g := p_b \circ u : T \to B$
Then $(f,g) \in \text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B)$.

Define $h(u) := (f,g) \in \text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B)$.
Then,

$$\begin{aligned} \text{Mor}_{\mathbf{C}}(T,p_A)(u) &= p_A \circ u = f = p_{\text{Mor}_{\mathbf{C}}(T,A)} \circ h(u) \\ \text{Mor}_{\mathbf{C}}(T,p_B)(u) &= p_B \circ u = g = p_{\text{Mor}_{\mathbf{C}}(T,B)} \circ h(u) \end{aligned}$$

(so Diagram 54 commutes).

Show: h bijective.
Construct inverse map k:

$$\begin{aligned} & k : \text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B) \to \text{Mor}_{\mathbf{C}}(T, A \times B) \\ & k((f,g)) := u \end{aligned}$$

where $(f,g) \in \text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B)$

By universal (mapping) property, $\exists !\, u : T \to A \times B$ s.t. $p_A \circ u = f$.
$p_B \circ u = g$

Let $(f,g) \in \text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B)$.
Then $(h \circ k)((f \circ g)) = h(u) = (p_A \circ u, p_B \circ u) = (f,g)$. Hence $h \circ k = 1$ or i.e. $h \circ k 1_{\text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B)}$.

Let $u \in \text{Mor}_{\mathbf{C}}(T, A \times B)$.
Then $(k \circ h)(u) = k((f,g)) = k((p_A \circ u, p_B \circ u)) = u'$ where $u' : T \to A \times B$ and

$$\begin{aligned} p_A \circ u' &= p_A \circ k((p_A \circ u, p_B \circ u)) = p_A \circ k((f,g)) = p_A \circ u \quad (\text{since } k((f,g)) := u) \\ p_B \circ u' &= p_B \circ k((p_A \circ u, p_B \circ u)) = p_B \circ k((f,g)) = p_B \circ u \quad (\text{since } k((f,g)) := u) \end{aligned}$$

2nd. equality is by universal (mapping) property.
$\Longrightarrow u = u'$ (by component-wise equality), and
$k \circ h = 1_{\text{Mor}_{\mathbf{C}}(T, A \times B)}$.
$\Longrightarrow h$ isomorphic.

$\Longrightarrow$:

Given morphisms

$$\begin{aligned} f &: T \to A \\ g &: T \to B \end{aligned}$$

then $(f,g) \in \text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B)$.
$\forall$ object $T \in \text{Obj}\mathbf{C}$, $\exists$ isomorphism h s.t. $\text{Mor}_{\mathbf{C}}(T, A \times B) \cong \text{Mor}_{\mathbf{C}}(T,A) \times \text{Mor}_{\mathbf{C}}(T,B)$,
Then $u := h^{-1}((f,g))$ is a unique morphism, $u \in \text{Mor}_{\mathbf{C}}(T, A \times B)$.
Also, by Diagram 54,

$$\begin{aligned} \text{Mor}_{\mathbf{C}}(T,p_A)(u) &= f = p_{\text{Mor}_{\mathbf{C}}(T,A)} \circ h(u) \\ \text{Mor}_{\mathbf{C}}(T,p_B)(u) &= f = p_{\text{Mor}_{\mathbf{C}}(T,B)} \circ h(u) \end{aligned}$$

Now

$$\begin{aligned} \text{Mor}_{\mathbf{C}}(T,p_A)(u) &= p_A \circ u \\ \text{Mor}_{\mathbf{C}}(T,p_B)(u) &= p_B \circ u \end{aligned}$$

(by definition).
So

$$\begin{aligned} f &= p_A \circ u \\ g &= p_B \circ u \end{aligned}$$

So universal mapping property is satisfied. □

Thus,

Definition 74 (Product, generalized, assuming universal property)**.** *Give $A_i \in Obj\mathbf{C}$, $i \in I$, category $\mathbf{C}$, **product** is object $\prod_{i \in I} A_i \in Obj\mathbf{C}$, morphisms $p_j : \prod_{i \in I} A_i \to A_j$, $j \in I$,*
s.t. $\forall$ object $A \in\in Obj\mathbf{C}$,
$\exists !$ morphism $\langle f_i \rangle : A \to \prod_{i \in I} A_i$,
s.t.

$$\begin{tikzcd} & A \arrow[dl, "f_j"'] \arrow[d, dashed, "\langle f_i \rangle"] \\ A_j & \prod_{i \in I} A_i \arrow[l, "p_j"] \end{tikzcd}$$

(56)

i.e. $\pi_j \circ \langle f_i \rangle = f_j \, \forall i,j \in I$

15.2.1. *Example: Set always has products.* $\forall$ sets $X, Y \in \text{Obj(Set)}$, $\exists$ product $X \times Y \in \text{Obj(Set)}$.

Let $A \in \text{Obj(Set)}$, $f_1 : A \to X$, $f_2 : A \to Y$. Define $\langle f_1, f_2 \rangle : A \to X \times Y$, $\langle f_1, f_2 \rangle (a) = (f_1(a), f_2(a))$

Then $\pi_1 \circ \langle f_1, f_2 \rangle (a) = f_1(a)$, $\pi_2 \circ \langle f_1, f_2 \rangle (a) = f_2(a)$ $\Longrightarrow \pi_1 \circ \langle f_1, f_2 \rangle = f_1$, $\pi_2 \circ \langle f_1, f_2 \rangle = f_2$

Suppose $f' : A \to X \times Y$ s.t. $\pi_1 \circ f' = f_1$
$\pi_2 \circ f' = f_2$

Write $f'(a) = (x,y)$

$$\begin{aligned} f_1(a) &= \pi_1 \circ f'(a) = \pi_1(x,y) = x \\ f_2(a) &= \pi_2 \circ f'(a) = \pi_2(x,y) = y \end{aligned} \qquad \Longrightarrow f'(a) = (f_1(a), f_2(a)) = \langle f_1, f_2 \rangle (a)$$

$\langle f_1, f_2 \rangle$ unique.

15.2.2. *Example: "Record Types", classes (in computer science, C++, Python).* cf. pp. 174, 5.3.14 "Record Types" Barr and Wells (1998) [3].

Look at Diagram 56 again.

To allow operations depending on several variables in a functional programming language L,
e.g. assume $\forall$ types A_j, $\forall\, j \in I$, language L has record type $\prod_{i\in I} A_i$, and field selectors p_j

$$p_j : \prod_{i \in I} A_i \to A_j$$

Insist that any data in $\prod_{i\in I} A_i$ be determined completely by those fields A_j, $j \in I$,
then $\forall$ operations f_j, $f_j : A \to A_j$, $\quad \forall\, j \in I$,
there ought to be an unique operation $\{f_i\rangle : A \to \prod_{i \in I} A_i$ s.t.

$$p_j \langle f_i \rangle = f_j, \quad \forall\, i,j \in I$$

This would make $\prod_{i\in I} A_i$ the product of A_i's, with selectors as product projections.

Thus, to say that one can always construct record types in a functional programming language L to say that the corresponding category $C(L)$ has finite products.

Proposition 21. *If product $(A_1 \times \dots \times A_{\mathcal{N}} \xrightarrow{\pi_i} A_i)_{i\in I}$, if $\exists\, i_0 \in I$ s.t. $Hom(A_{i_0}, A_i) \neq \emptyset$, $\quad \forall\, i \in I$, then π_{i_0}* retraction

Proof. $\forall\, i \in I$, choose $f_i \in \mathrm{Hom}(A_{i_0}, A_i)$ with $f_{i_0} = 1_{A_{i_0}}$.

Then $\langle f_i \rangle : A_{i_0} \to A_1 \times \dots \times A_{\mathcal{N}}$ is a morphism s.t.

$$\pi_{i_0} \circ \langle f_i \rangle = f_{i_0} = 1_{A_{i_0}}$$

□

Adámek, Herrlich, and Strecker (2004) [4] and their notation) calls a **sink** what Leinster (2014) [8] calls a **cocone**.

Definition 75. ***sink*** *$((f_i)_{i\in I}, A) \equiv (f_i, A)_I \equiv (A_i \xrightarrow{f_i} A)_I$, object A, family of morphisms $f_i : A_i \to A$*

For the *coproduct*, consider this enlightening comparision:

product $(\prod_{i\in I} A_i, \pi_j)_{j\in I}$	coproduct $(\mu_j, \coprod_{i\in I} A_i)_{j\in I}$
projection $\pi_j : \prod_{i \in I} A_i \to A_j$	injection $\mu_j : A_j \to \coprod_{i\in I} A_i$

$$\begin{tikzcd} & A \arrow[dl, "f_j"'] \arrow[d, "\langle f_i \rangle"] \\ A_j & \prod_{i\in I} A_i \arrow[l, "\pi_j"] \end{tikzcd} \qquad \begin{tikzcd} & A \\ A_j \arrow[ur, "f_j"] \arrow[r, "\mu_j"'] & \coprod_{i\in I} A_i \arrow[u, "{[f_i]}"'] \end{tikzcd}$$

$$C \xrightarrow{\langle f,g \rangle} A \times B \qquad\qquad C \xleftarrow{[f,g]} A + B$$

$$\prod_{i\in I} f_i\,, \text{ or if } i = \{1,2\},\ f \times g \qquad\qquad \coprod_{i \in I} f_i\,, \text{ or if } i = \{1,2\},\ f + g$$

15.2.3. *Examples (of coproducts).*

- if $(A_i)_I$ pairwise-disjoint family of sets, then $(\mu_j, \bigcup_{i\in I} A_i)_{j\in I}$ is coproduct in Set.
  If $(A_i)_I$ arbitrary set-indexed family of sets, then it can be "made disjoint" by pairing each A_i with index i, i.e. by working with $A_i \times \{i\}$ rather than A_i.
  So $\bigcup_{i\in I}(A_i \times \{i\})$ disjoint. Consider
  $$\begin{aligned} \mu_j &: A_j \to \bigcup_{i \in I} A_i \times \{i\} \\ \mu_i(a) &= (a, j) \end{aligned}$$
  $(\mu_j, \bigcup_{i\in I} A_i \times \{i\})_{j\in I}$ is a coproduct in Set.

  Indeed, given $f_j : A_j \to A$,
  $f_j(a) \in A$
  $$\begin{aligned} &[f_i] : \coprod_{i \in I} A_i \times \{i\} \to A \\ &[f_i] \circ \mu_j = f_j \end{aligned}$$
  where
  $$f_j(a) = [f_i] \circ \mu_j(a) = [f_i](a,j) = f_j(a)$$
- Top coproducts are "topological sums"; they're "concrete" coproducts (Adámek, Herrlich, and Strecker (2004) [4])
- Vec (nonconcrete) coproducts called *direct sums*
  direct sum $\bigoplus_{i\in I} A_i$ of vector spaces A_i is subspace of direct product $\prod_{i\in I} A_i$
  consisting of all elements $(a_i)_{i\in I}$ with finite carrier (i.e. $\{i \in I | a_i \neq 0\}$ is finite),
  injections
  $$\begin{aligned} \mu_j &: A_j \to \bigoplus_{i \in I} A_i \\ \mu_j(a) &= (a_i)_{i\in I} \text{ with } a_i = \begin{cases} a & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \end{aligned}$$
- Grp has nonconcrete coproducts, "free products"

16. Naturality; Natural transformations

Definition 76 (Natural Transformation)**.** *If functors F, G, categories $\mathbf{C}, \mathbf{D}$,*

$$\begin{aligned} F &: \mathbf{C} \to \mathbf{D} \\ G &: \mathbf{C} \to \mathbf{D} \end{aligned}$$

*the **natural transformation** η from F to G is a family of morphisms s.t.*

(1) *$\forall\, X \in Obj\mathbf{C}$, $\eta_X : F(X) \to G(X)$ s.t. $F(X), G(X) \in Obj\mathbf{D}$. morphisms $\eta_X \in Hom_{\mathbf{D}}(F(X), G(X)) \subset Mor\mathbf{D}$ is called **component** of η at X.*

(2) *components must be s.t.* $\forall$ *morphism* $f : X \to Y$ *in* $\mathbf{C}$*, i.e.* $f \in Hom_{\mathbf{C}}(X,Y) \in Mor\mathbf{C}$*,*

$$\eta_Y \circ F(f) = G(f) \circ \eta_X \tag{57}$$

i.e.

$$\begin{tikzcd} F(X) \arrow[r, "F(f)"] \arrow[d, "\eta_X"] & F(Y) \arrow[d, "\eta_Y"] \\ G(X) \arrow[r, "G(f)"] & G(Y) \end{tikzcd}$$

If both F, G *contravariant,*

$$\begin{tikzcd} F(X) \arrow[d, "\eta_X"] & F(Y) \arrow[l, "F(f)"'] \arrow[d, "\eta_Y"] \\ G(X) & G(Y) \arrow[l, "G(f)"] \end{tikzcd}$$

cf. Wikipedia, "Natural transformation"
Notation: natural transformation from F to G, $\eta : F \to G$, i.e. family of morphisms $\eta_X : F(X) \to G(X)$ is natural in X.
cf. Adámek, Herrlich, and Strecker (2004) [4].

From pp. 16, Sec. 14 "Natural Transformations", Mac Lane (1978) [5],

$\forall$ arrow $f : C \to C'$ in C, i.e. notation: $\forall$ morphism $f : X \to Y$ in $\mathbf{C}$,
Mac Lane's notation:

$$\begin{tikzcd} c \arrow[d, "f"] & Sc \arrow[r, "\tau c"] \arrow[d, "Sf"] & Tc \arrow[d, "Tf"] \\ c' & Sc' \arrow[r, "\tau c'"] & Tc' \end{tikzcd}$$

My notation:

$$\begin{tikzcd} X \arrow[d, "f"] & FX \arrow[r, "\eta_X"] \arrow[d, "Ff"] & GX \arrow[d, "Gf"] \\ Y & FY \arrow[r, "\eta_Y"] & GY \end{tikzcd}$$

We say that $\tau_c : Sc \to Tc$ is natural in C; $\eta_X : FX \to GX$ is natural in X.

If we think of functor S as giving a picture in B of (all the objects, arrows of) C, then natural transformation τ is set of arrow s mapping (or, translating) the picture S to picture T, with all squares (and parallelograms!) like that above commutative:

Mac Lane's notation:

$$\begin{tikzcd} a \arrow[dr, "f"] \arrow[dd, "h"'] & & Sa \arrow[rr, "\tau a"] \arrow[dr, "Sf"] \arrow[dd, "Sh"'] & & Ta \arrow[dr, "Tf"] \arrow[dd] & \\ & b \arrow[dl, "g"] & & Sb \arrow[rr, "\tau b"] \arrow[dl, "Sg"] & & Tb \arrow[dl, "Tg"] \\ c & & Sc \arrow[rr, "\tau c"] & & Tc & \end{tikzcd}$$

My notation:

$$\begin{tikzcd} X \arrow[dr, "f"] \arrow[dd, "g"'] & & FX \arrow[rr, "\eta_X"] \arrow[dr, "Ff"] \arrow[dd, "Fg"'] & & GX \arrow[dr, "Gf"] \arrow[dd] & \\ & Y \arrow[dl, "h"] & & FY \arrow[rr, "\eta_Y"] \arrow[dl, "Fh"] & & GY \arrow[dl, "Gh"] \\ Z & & FZ \arrow[rr, "\eta_Z"] & & GZ & \end{tikzcd}$$

natural equivalence or i.e. natural isomorphism - natural transformation τ with $\forall$ component τc invertible in B, $\tau : S \cong T$.
notation: η with $\forall$ component η_X invertible in $\mathbf{D}$; $\eta : F \cong G$.
$\Longrightarrow$ inverses $(\tau c)^{-1}$ in B are components of natural isomorphism $\tau^{-1} : T \to S$.
notation: inverses $(\eta_X)^{-1}$ in $\mathbf{D}$ are components of natural isomorphism $\eta^{-1} : G \to F$.

16.1. **Examples of Natural transformations.** e.g. Hurewicz homomorphism $\pi_n(X) \to H_n(X)$ $\forall$ topological space X is a natural transformation from nth homology functor.
$\pi_n : \mathbf{Top} \to \mathbf{Grp}$ to nth homology functor $H_n : \mathbf{Top} \to \mathbf{Grp}$.
e.g. **determinant** det is a natural transformation (cf. Mac Lane (1978) [5]).
Let $\det_K M =$ determinant of $n \times n$ matrix M, with entries in commutative ring K, while
$K^* \equiv$ group of units (invertible elements) of K

Thus M non-singular when $\det_K M$ is a unit, and $\det_K$ is a morphism $GL_n K \to K^*$ of groups.
Because det is defined by same formula $\forall$ ring K, each morphism $f : K \to K'$ of commutative rings leads to

$$\begin{tikzcd} GL_n K \arrow[r, "\det_K"] \arrow[d, "GL_n f"] & K^* \arrow[d, "f^*"] \\ GL_n K' \arrow[r, "\det_{K'}"] & (K')^* \end{tikzcd}$$

$\Longrightarrow$ det : $GL_n \to ()^*$ is natural between 2 functors $\mathbf{CRng} \to \mathbf{Grp}$
TODO: Work out examples of natural transformations involving group commutators on pp. 17 of Mac Lane (1978) [5].
e.g. category **Finord** = all finite ordinal numbers n
category $\mathbf{Set}_f$ = all finite sets (in some universe U)
$\forall$ ordinal $n = \{0, 1, \dots n-1\}$ is a finite set, so inclusion S is a functor $S : \mathbf{FinOrd} \to \mathbf{Set}_f$
$\forall$ finite set X determines ordinal number $n = \#X$, number of elements in X
$\forall X$, choose bijection $\theta_X : X \to \#X$
$\forall$ function $f : X \to Y$, X, Y finite sets,
Choose corresponding function $\#f : \#X \to \#Y$ between ordinals by $\#f = \theta_Y f \theta_X^{-1}$, so that

$$\begin{tikzcd} X \arrow[r, "\theta_X"] \arrow[d, "f"] & \#X \arrow[d, "\#f"] \\ Y \arrow[r, "\theta_Y"] & \#Y \end{tikzcd} \tag{58}$$

$\Longrightarrow \#$ a functor $\# : \mathbf{Set}_f \to \mathbf{FinOrd}$
If X itself an ordinal number, take θ_X to be an identity.
$\Longrightarrow \# \circ S$ is identity functor I' of $\mathbf{FinOrd}$
$S \circ \#$ is not identity functor $I : \mathbf{Set}_f \to \mathbf{Set}_f$ because
$S \circ \#$ sends each finite set X to a special finite set - ordinal number n with same number of elements as X.
However $\theta : I \to S\#$ is a natural isomorphism by square diagram above (Eq. 58).

$$I \cong S \circ \#, \quad I' = \# \circ S$$

equivalence between categories $\mathbf{C}, \mathbf{D} =$

pair of functors $S : \mathbf{C} \to \mathbf{D}$ and
$T : \mathbf{C} \to \mathbf{D}$
natural isomorphisms $I_C \cong T \circ S$, $I_D \cong S \circ T$.

17. Limits

17.1. Pullback.

Definition 77. *For some category* $\mathbf{A}$*, and for*

$$\begin{tikzcd} & Y \arrow[d, "t"] \\ X \arrow[r, "s"'] & Z \end{tikzcd}$$

$X, Y, Z \in Obj\mathbf{A}$.

$s : X \to Z$; $\quad s, t \in Mor\mathbf{A}$
$t : Y \to Z$

Then the ***pullback*** *or "pullback square" consists of* $P \in Obj\mathbf{A}$, $\pi_1 : P \to X$ *s.t.*
$\pi_2 : P \to Y$

$$\begin{tikzcd} P \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] & Y \arrow[d, "t"] \\ X \arrow[r, "s"] & Z \end{tikzcd}$$

commutes and s.t. $\forall$ *commutative square in* $\mathbf{A}$

$$\begin{tikzcd} A \arrow[r, "f_2"] \arrow[d, "f_1"] & Y \arrow[d, "t"] \\ X \arrow[r, "s"] & Z \end{tikzcd}$$

then $\exists ! \, f : A \to P$ *s.t.*

$$\begin{tikzcd} A \arrow[dr, "f"] \arrow[drr, "f_2"] \arrow[ddr, "f_1"'] & & \\ & P \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] & Y \arrow[d, "t"] \\ & X \arrow[r, "s"] & Z \end{tikzcd}$$

Simmons (2011) [6]
cf. 3.3 *Some less simple functors*, pp. 79, of Simmons (2011) [6]

18. Monads

cf. Ch. 6, Monads and Algebras, starting on pp. 137 of Mac Lane (1971) [5].

18.1. Monads in a Category.

18.1.1. *Development of a Monad.* Let's follow the development of a Monad according to Mac Lane (1971) [5] starting on pp. 137 in Sec. 1. "Monads in a Category."

$\forall$ endofunctor $T : X \to X$, T has composites $T^2 = T \circ T : X \to X$, $T^3 = T^2 \circ T : X \to X$ in Mac Lane's notation, X is a category. In our notation, $\mathbf{C}$ denotes the category. And so for endofunctor $T : \mathbf{C} \to \mathbf{C}$, $T^2 = T \circ T : \mathbf{C} \to \mathbf{C}$, $T^3 = T^2 \circ T : \mathbf{C} \to \mathbf{C}$.

Recall from Def. 62, that for endofunctor $T : \mathbf{C} \to \mathbf{C}$, the *object map* $T : \text{Obj}\mathbf{C} \to \text{Obj}\mathbf{C}$ acts just like a morphism in Mor$\mathbf{C}$! So $T : X \to T(X)$ for $X, T(X) \in \text{Obj}\mathbf{C}$. Also, recall that the *morphism map* $T : \text{Mor}_{\mathbf{C}} \to \text{Mor}_{\mathbf{C}}$ acts like the usual morphism map, such that $Tf : T(X) \to T(Y)$ if $f : X \to Y$, for $X, Y \in \text{Obj}\mathbf{C}$.

If $\mu : T^2 \to T$ is a *natural transformation*, then recall the definition of the natural transformation, Def. 76, and apply that onto μ.

$\forall X \in \text{Obj}\mathbf{C}$, $\mu_X : T^2(X) \to T(X)$ with $T^2(X), T(X) \in \text{Obj}\mathbf{C}$. μ_X is a morphism $\in \text{Hom}_{\mathbf{C}}(T^2(X), T(X))$ (!!), as well as being a component of μ at X.

$\forall$ morphism $f : X \to Z$ in $\mathbf{C}$, $X, Z \in \text{Obj}\mathbf{C}$ (the notation of using Z to emphasize that f can map X to *any* arbitrary object in Obj$\mathbf{C}$),

$$\mu_Z \circ T^2 f = Tf \circ \mu_X$$

i.e.

$$\begin{tikzcd} T^2(X) \arrow[r, "\mu_X"] \arrow[d, "T^2 f"] & T(X) \arrow[d, "Tf"] \\ T^2(Z) \arrow[r, "\mu_Z"] & T(Z) \end{tikzcd}$$

Let $T\mu : T^3 \to T^2$ *denote* a natural transformation with components $(T\mu)_X = T(\mu_X) : T^3(X) \to T^2(X)$, while $\mu T : T^3 \to T^2$ denotes another natural transformation with components $(\mu T)_X = \mu_{T(X)}$.

$T\mu$, μT are "horizontal" composites in the sense of Sec. II. 5, "The Category of All Categories". (TODO: Understand what Mac Lane (1971) [5] means when he says this on pp. 137).

Let's jump the gun and consider what natural transformation $\eta : 1_{\mathbf{C}} \to T$ means with respect to the definition of a natural transformation, Def. 76. $1_{\mathbf{C}}$ is the identity functor, mapping everything to itself.

$\forall X \in \text{Obj}\mathbf{C}$, $\eta_X : 1_{\mathbf{C}}(X) \to T(X)$ with $1_{\mathbf{C}}(X) = X, T(X) \in \text{Obj}\mathbf{C}$. η_X is a morphism $\in \text{Hom}_{\mathbf{C}}(1_{\mathbf{C}}(X), T(X)) = \text{Hom}_{\mathbf{C}}(X, T(X))$ (!!!), as well as being a component of η at X. Not only that, it applies T onto X!

$\forall$ morphism $f : X \to Z$ in $\mathbf{C}$, $X, Z \in \text{Obj}\mathbf{C}$,

$$\eta_Z \circ 1_{\mathbf{C}} f = \eta_Z \circ f = Tf \circ \eta_X$$

i.e.

$$\begin{array}{ccc} 1_{\mathbf{C}}(X) & \xrightarrow{\eta_X} & T(X) \\ \downarrow{\scriptstyle 1_{\mathbf{C}} f} & & \downarrow{\scriptstyle Tf} \\ 1_{\mathbf{C}}(Z) & \xrightarrow{\eta_Z} & T(Z) \end{array} \Longrightarrow \begin{array}{ccc} X & \xrightarrow{\eta_X} & T(X) \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle Tf} \\ Z & \xrightarrow{\eta_Z} & T(Z) \end{array}$$

18.1.2. *Definitions and Properties of a Monad.*

Definition 78 (Monad)**.** ***monad*** *$T = \langle T, \eta, \mu \rangle$ in category $\mathbf{C}$ consists of functor $T : \mathbf{C} \to \mathbf{C}$ and 2 natural transformations*

$$\begin{aligned} \eta &: 1_{\mathbf{C}} \to T \\ \mu &: T^2 \to T \end{aligned} \tag{59}$$

s.t.

$$\begin{array}{ccc} T^3 & \xrightarrow{T\mu} & T^2 \\ \downarrow{\scriptstyle \mu T} & & \downarrow{\scriptstyle \mu} \\ T^2 & \xrightarrow{\mu} & T \end{array} \tag{60}$$

being like an "associative" diagram, and

$$\begin{array}{ccccc} 1T & \xrightarrow{\eta T} & T^2 & \xleftarrow{T\eta} & T1 \\ \| & & \downarrow{\scriptstyle \mu} & & \| \\ T & = & T & = & T \end{array} \tag{61}$$

being a diagram that could be said to express the left and right unit laws.

Formally, the monad definition is like that of a monoid M in **Sets**:
set M of elements of monad $\to$ replaced by endofunctor $T : X \to X$
Cartesian product $\times$ of 2 sets $\to$ composite of 2 functors
binary operation $\mu : M \times M \to M$ of multiplication $\to$ transformation $\mu : T^2 \to T$ and
unit (identity) element $\eta : 1 \to M$ $\to$ $\eta : 1_{\mathbf{C}} \to T$.

Thus, call
η the **unit** of monad T. The first commutative diagram, Eq. 60, is the associative law for a monad, and
μ the **multiplication** of monad T. The 2nd. and 3rd. diagrams, Eq. 61 expresses the left and right *unit laws.*

Proposition 22 (Coherence conditions)**.**

$$\begin{aligned} &\mu \circ T\mu = \mu \circ \mu T \quad \textit{(as natural transformations } T^3 \to T,\ \mu \circ T\mu \textit{ and } \mu \circ \mu T, \textit{ that is)} \\ &\mu \circ T\eta = \mu \circ \eta T = 1_T \quad \textit{(as natural transformations } T \to T,\ 1_T \equiv \textit{identity functor)} \end{aligned} \tag{62}$$

For the last statement, it is equal to commutative diagram Eq. 61, since it can be redrawn as follows:

$$\begin{array}{ccc} T & \xrightarrow{\eta T} & T^2 \\ \downarrow{\scriptstyle T\eta} & \diagdown\!\!\diagdown & \downarrow{\scriptstyle \mu} \\ T^2 & \xrightarrow{\mu} & T \end{array}$$

18.2. **Monads and Adjoints.** cf. Ch. 6, Monads and Algebras, pp. 138 of Mac Lane (1971) [5].

Proposition 23. *adjunction $\langle F, G, \eta, \epsilon \rangle : X \to A$, $\exists$ monad in category X (notation $\langle F, G, \eta, \epsilon \rangle : \mathbf{C} \to \mathbf{D}$, $\exists$ monad in category $\mathbf{C}$).*
Specifically, 2 functors $F : \mathbf{C} \to \mathbf{D}$, $G : \mathbf{D} \to \mathbf{C}$ have composite $T = GF$, an endofunctor,
unit η of adjunction is natural transformation $\eta : 1 \to T$,
counit $\epsilon : FG \to 1_A$ of adjunction yields by horizontal composition; a natural transformation

$$\mu = G\epsilon F : GFGF \to GF = T$$

TODO: adjunction.

18.3. **Applications of Monads: Correspondence between Monads in Computer Science and Monads in Category Theory.** cf. monad (in computer science), nLab
For monads in computer science,
maps type X to new type $T(X)$,
equipped with rule for composing 2 functions of form $f : X \to T(Y)$ (called Kleisli functions), $g : Y \to T(Z)$ to function

$$g \circ f : X \to T(Z) \quad \text{(their Kleisi composition)}$$

associative in evident sense, unital with respect to given unit function called $\text{pure}_X : X \to T(X)$ to be thought of as taking a value to the pure computation that simply returns that value.

e.g. when monad $T(-)$ forms product types $T(X) = X \times Q$, some fixed type Q that carries structure of monoid.
Then Kleisi function $f : X \to Y \times Q$ is a function $X \to Y$ that produces as a side effect output of type Q.

Kleisi composition of $f : X \to Y \times Q$ not only evaluates the 2 programs in sequence, but also combines their Q-output using
$g : Y \to Z \times Q$
monoid operation of Q, e.g.

if $fx = (y, q)$ then $(g \circ f)(x) = (z, qq')$
$gy = (z, q')$

Let syntactic category be denoted $\mathbf{C}$.

$\text{Obj}\mathbf{C} \ni X$ $\to$ types of the (programming) language
$\text{Hom}_{\mathbf{C}}(X, Y)$ $\to$ "programs" that takes value of type X as input and returns value of type Y

endofunctor $T : \mathbf{C} \to \mathbf{C}$
object map $T : X \to T(X)$ sends each type $X \in \text{Obj}\mathbf{C}$ to another type $T(X) \in \text{Obj}\mathbf{C}$.

unit natural transformation $\epsilon : 1_{\mathbf{C}} \to T$ provides $\forall$ type X a component morphism $\text{pure}_X : X \to T(X)$ (notation, ncat uses the notation ϵ; we will use η for the "unit" natural transformation of a monad):

$$\begin{aligned} &\eta : 1_{\mathbf{C}} \to T \\ &\eta_X : X \to T(X) \equiv \text{pure}_X : X \to T(X) \end{aligned}$$

multiplication natural transformation $\mu : T^2 \to T$ provides $\mu_X : T^2(X) \to T(X)$ which induces Klesli composition by formula

$$(63)\quad \begin{aligned} (g \circ f) &:= (Y \xrightarrow{g} T(Z)) \circ_{\text{Klesli}} (X \xrightarrow{f} T(Y)) \\ &:= X \xrightarrow{f} T(Y) \xrightarrow{T(g)} T(T(Z)) \xrightarrow{\mu(Z)} TZ \end{aligned}$$

$T(g)$ uses fact that $T(-)$ is functor morphism $\mu(Z)$ is 1 that implements "T-computation".

Now nLab says this about the "bind" operation:
"bind" $\leftrightarrow$ multiplication on monad M,

$$(64)\quad \begin{gathered} MA \to (MB)^A \to MB \text{ equivalent to map } MA \times MB^A \to MB \\ MA \times MB^A \xrightarrow{\text{strength}} M(A \times MB^A) \xrightarrow{M_{\text{eval}_{A,MB}}} MMB \xrightarrow{mB} MB \end{gathered}$$

To try to make sense of Eq. 64, consider the following:

Recall for natural transformation η, the component η_X: $\mu_X : T^2(X) \to T(X)$ and for some morphism $f \in \text{Mor}\mathbf{C}$,

$$\begin{gathered} f : X \to Z \\ Tf : T(X) \to T(Z) \end{gathered}$$

So if $f : X \to T(Y)$ instead,

$$\begin{gathered} Tf : T(X) \to T^2(Y) \\ (\mu_Y \circ Tf) : T(X) \to T(Y) \end{gathered}$$

Form a Cartesian product of $T(X)$ and $\text{Hom}_{\mathbf{C}}(X, T(Y)) \ni f$ as inputs: $T(X) \times \text{Hom}_{\mathbf{C}}(X, T(Y))$.

Then I suggest that Eq. 64 is translated as follows:

$$(65)\quad T(X) \times \text{Hom}_{\mathbf{C}}(X, T(Y)) \xrightarrow{\text{apply } T} T(T(X) \times \text{Hom}_{\mathbf{C}}(X, T(Y))) \xrightarrow{Tf} T^2(Y) \xrightarrow{\mu_Y} T(Y)$$

where apply T is on $f \in \text{Hom}_{\mathbf{C}}(X, T(Y))$ in order to form Tf.

18.3.1. *More on Klesli composition.* cf. Kleisli category, nLab.

Proposition 24. *Klesli composition of $g \circ f$ of $f : X \to T(Y)$ is given by*

$$g : Y \to T(Z)$$

$$X \xrightarrow{f} TY \xrightarrow{Tg} T^2Z \xrightarrow{\mu_Z} TZ$$

i.e.

$$(66)\quad \boxed{g \circ_{Klesli} f = \mu_Z \circ Tg \circ f}$$

with $\mu_Z \circ Tg \circ f : X \to T(Z)$

Proof. Fullness of functor T:

$\forall$ morphism $g : T(X) \to T(Y)$ has antecedent the composition

$$X \xrightarrow{\eta_X} T(X) \xrightarrow{g} T(Y)$$

i.e. $g \circ \eta_X : X \to T(Y)$.

Consider $T(g \circ \eta_X) = Tg \circ T\eta_X : T(X) \to T^2(Y)$ for

$$\begin{aligned} T\eta_X &: T(X) \to T^2(X) \\ Tg &: T^2(X) \to T^2(Y) \\ \mu_Y &: T^2(Y) \to T(Y) \end{aligned}$$

So

$$\mu_Y \circ Tg \circ T\eta_X = \mu_Y \circ T(g \circ \eta_X) : T(X) \to T(Y)$$

Because g is a morphism of algebras (TODO: understand what this means, from Kleisli category, nLab),

$$\mu_Y \circ Tg \circ T\eta_X = g \circ \mu_X \circ T\eta_X = g$$

where $\mu_Y \circ Tg = g \circ \mu_X$ uses natural transformation definition 76 for the property $\mu_Z \circ T^2 f = Tf \circ \mu_X$ (TODO: Show $g = Tf$ is of this form),
and $\mu \circ T\eta = \mu \circ \eta T = 1_T$ (coherence condition from Prop. 22).

Faithfulness:

If $\mu_Y \circ Tf = \mu_Y \circ Tg$, then precompose by η_X yields

$$\mu_Y \circ Tf \circ \eta_X = \mu_Y \circ \eta_{f(X)} \circ f = \mu_Y \circ \eta_{T(Y)} \circ f = f$$

where
$Tf \circ \eta_X = \eta_{T(Y)} \circ f$ from natural transformation def. 76 was used, and $\mu \circ \eta T = 1_T$ (coherence condition) was also used, for $f : X \to T(Y)$.

Similarly for g. $\mu_Y \circ Tf = \mu_Y \circ Tg \Longrightarrow f = g$. □

Notice that the multiplication natural transformation $\mu : T^2 \to T$ was needed to induce the Klesli composition.

18.3.2. *Dictionary between Monads in Computer Science and Monads in Category Theory.*

`join` is $\mu_X : T^2(X) \to T(X)$

`return` is $\eta_X : X \to T(X)$

`fmap` is the morphism map (arrow map) for functor F

`bind` is $\mu_Y \circ Tf : T(X) \to T(Y)$ but ...

`bind` is more nuanced in that, from the Haskell and functional programming perspective, it takes in *two* inputs.

One of them is an element x in the object X, $x \in X$. The other is a "transformation of type $X \to T(Y)$." I believe that would mean a morphism $f : X \to T(Y)$.

What Čukić in Čukić (2018) [10] calls `transform` is the *morphism map* (arrow map) for a functor F. However, note that when it comes to concrete implementation, the concrete implementation applies to *instances* of a type $X \in \text{Obj}\mathbf{C}$. For a given type X, e.g. `int` (the integer type in C or C++), there are many values that will have that type. It would be good that we work out how the definitions, theorems, and commutative diagrams apply in this case.

I don't think there's a great way to reconcile what Haskell and functional programming programmers mean by `bind` with the corresponding concept in category theory.

18.4. **Examples of Monads in Computer Science.**

18.4.1. *Reader Monad; Function Monad; Environment Monad.* $[W,-]:\mathbf{C}\to\mathbf{C}$.
Given $W\in\text{Obj}\mathbf{C}$, (W is a type),
$$[W,-]=(W\to(-))$$
where $[-,-]:\mathbf{C}^{\text{op}}\times\mathbf{C}\to\mathbf{C}$.
$W\mapsto$ internal hom out of W.
$$\begin{aligned}&\eta_X: X\to[W,-](X)\\&\eta_W: W\to[W,-](W)\\&\eta_W: w\mapsto[W,-](W)(w)\quad\text{value to constant functions with that value}\\&\Delta: X\in\text{Obj}\mathbf{C}\xrightarrow{(1,1)}X\times X\\&\Delta: x\mapsto(w,w)\end{aligned}$$
"read in " state of type W (an environment)
wikipedia calls this the Environment Nomad.

`reader`: $T\to E\to T=t\mapsto e\mapsto t$ corresponds to $\eta_X: X\to T(X)$ where $T(X)=\text{Hom}_{\mathbf{C}}(E,X)$ s.t. η_X sends values $x\in X$ to constant functions (on E) to value x.

Therefore, rewrite the unit component at X, i.e. `reader`, as

$$\begin{aligned}&X\to E\to X\qquad &&\eta_X: X\to\text{Hom}_{\mathbf{C}}(E,X)\\&x\mapsto e\mapsto x &&\eta_X: x\mapsto(e\mapsto x)\end{aligned}\tag{67}$$

Thus, the *unit* component at X, called `unit` or `return` by programmers, truly maps a value $x\in X$ to a *constant function*:

$$\eta_X: x\mapsto(e\mapsto x)\in\text{Hom}_{\mathbf{C}}(E,X)\tag{68}$$

`bind`: $(E\to T)\to(T\to E\to T')\to E\to T'=r\mapsto f\mapsto e\mapsto f(re)e$ corresponds to $\mu_Y\circ Tf: T(X)\to T(Y)$ where
$T(X)=\text{Hom}_{\mathbf{C}}(E,X)$, $T(Y)=\text{Hom}_{\mathbf{C}}(E,Y)$, and
$f: X\to T(Y)$ or $f: X\to\text{Hom}_{\mathbf{C}}(E,Y)$
$Tf: T(X)\to T^2(Y)$ or $Tf:\text{Hom}_{\mathbf{C}}(E,X)\to\text{Hom}_{\mathbf{C}}(E,T(Y))=\text{Hom}_{\mathbf{C}}(E,\text{Hom}_{\mathbf{C}}(E,Y))$.

So let's rewrite `bind` in our notation:
$$\begin{aligned}&(E\to X)\to\quad (X\to E\to Y)\to E\to\qquad\qquad Y\\&r\mapsto\qquad\qquad\qquad\qquad f\mapsto e\mapsto(f(r))(e)(e)\end{aligned}$$

To clarify these definitions, rewrite the reader monad: the endomorphism T is defined as

$$\boxed{\begin{aligned}&T:=[E,-]\\&T(X)=[E,X]\end{aligned}}\tag{69}$$

For any morphism $f: X\to T(Y)=[E,Y]$, then
recall the Klesli composition: $g\circ_{\text{Klesli}}f=\mu_Z\circ Tg\circ f$ with $g: Y\to T(Z)=[E,Z]$.
But if you take a look at how wikipedia defines "bind" for the reader monad, then bind *is not* the Klesli composition, but rather it's $\mu_Y\circ Tf$ (!!!):

$$\begin{gathered}\mu_Y\circ Tf: T(X)\to T(Y)\text{ or }\mu_Y\circ Tf:[E,X]\to[E,Y]\\ r\in[E,X]\qquad r(e)\in X\text{ for }e\in E\qquad f(r(e))\in T(Y)=[E,Y]\qquad f(r(e))(e)\in Y\end{gathered}\tag{70}$$

bind, as defined by wikipedia for the reader monad *is not* the Klesli composition.
As an example, consider the following example:
$T(Y)\ni(E\to Y)\in\text{Hom}_{\mathbf{C}}(E,Y)$
$e=$ "dag" $\mapsto y=$ "Hello Dag!"

$T(X)\ni(E\to X)\in\text{Hom}_{\mathbf{C}}(E,X)$
$e\equiv$ `name` $\mapsto x=$ "Hi `name`!"

$f: X\to T(Y)\to f: X\to\text{Hom}(E,Y)$
$f(x)=$ replace "Hi" with "Hello" in X.

`ask`: $E\to E=1_E$.
`ask` operation used to retrieve current context.

`local`: $(E\to E)\to(E\to T)\to E\to T=f\mapsto c\mapsto e\mapsto c(fe)$
local executes a computation in a modified subcontext. $(E\to E)$ modifies the environment, $(E\to T)$ uses the modified environment.
Rewrite `local` in our notation:
$$\begin{aligned}&(E\to E)\to\quad(E\to X)\to E\to\qquad X\\&f\mapsto\qquad\qquad\qquad c\mapsto e\mapsto\quad c(f(e))\end{aligned}$$
To make sense of this, let $f\in[E,E]$, $c\in[E,X]=T(X)$, $c\circ f\in[E,X]$ so $(c\circ f)(e)\in X$.

18.4.2. *Writer Monad.* $W\times(-):\mathbf{C}\to\mathbf{C}$ where $W\in\text{Obj}\mathbf{C}$. If W equipped with the structure of monoid, $W\times(-)$ canonically inherits the structure of a *monad*. Hence, the Writer *Monad*.
Recall W monoid if W is a set equipped with binary operation $\mu: M\times M\to M$ s.t. $(xy)z=x(yz)$ and $1\in M$ s.t. $1\cdot x=x=x\cdot 1$.

18.4.3. *State Monad.* cf. state monad, nLab
Input of type X, output of type Y, mutable state type W.
Consider morphism
$$X\times W\to Y\times W$$
Under (Cartesian product $\vdash$ internal hom) adjunction, this is equivalently given by its adjunt, which is a function of type
$$X\to[W,W\times Y]$$
$[W,W\times(-)]$ is the monad, induced by the above adjunction.

$X\to[W,W\times Y]$ regarded as a morphism.
$[W,W\times Y]:\mathbf{H}\to\mathbf{H}$ is called *state monad* for mutable states of type W.
TODO: Understand the above.
Writer comonad $W\times(-):\mathbf{C}\to\mathbf{C}$,

Given by forming Cartesian product with W.

Reader Monad $\circ$ Writer Comonad is state monad.
cf. wikipedia
unit or "return":
$$\begin{aligned}&X\to\quad S\to X\times S\\&x\mapsto\quad(s\mapsto(x,s))\end{aligned}\tag{71}$$

Compare this to

$$\begin{aligned} \eta_X &: X \to T(X) = [S, S\times X] \\ \eta_X &: x \mapsto (S \to S \times X) \end{aligned}$$

"bind" was given by wikipedia as

$$\begin{aligned} (S\to X\times S) \to \quad (X \to S \to Y \times S) \to S \to \quad & Y\times S \\ m \mapsto \qquad\qquad\qquad\qquad k \mapsto s \mapsto \quad & k(x)(s) \end{aligned} \tag{72}$$

where $(x,t) = m(s)$

18.4.4. *Properties of the State Monad.* For the unit or "return", $\eta_X : X \to T(X) = [S, S\times X]$,

$$\begin{aligned} \eta_X : x \mapsto (S\to S\times X) \text{ s.t.} \\ (S\to S\times X) : s \mapsto (s,x) \end{aligned} \tag{73}$$

Therefore, for the unit component at X of a state monad,

$$\boxed{\eta_X(x)(s) = (s,x) \text{ for } s \in S} \tag{74}$$

Compare this to what wikipedia says how it defines the "`return`" operation in Eq. 71.

Consider any morphism $f : X \to T(Y) = [S, S\times Y]$,

$$\begin{aligned} & f : X \to T(Y) = [S, S\times Y] \\ & f : x \mapsto (S\to S\times Y) \text{ s.t. } (S\to S\times Y) : s \mapsto (t = f_s(s,x), y = f_x(s,x)) \end{aligned} \tag{75}$$

Allow morphism f to have "freedom" to set state to arbitrary t.

$$\begin{aligned} & Tf : T(X) \to T^2(Y) \to Tf : [S, S\times X] \to T([S, S\times Y]) = [S, S\times [S, S\times Y]] \\ & Tf : (S\to S\times X) \mapsto (S \to S\times (S\to S\times Y)) \end{aligned} \tag{76}$$

Then

$$\begin{aligned} & \mu_Y \circ Tf : T(X) \to T(Y) \text{ or } \mu_Y\circ Tf : [S,S\times X] \to [S, S\times Y] \\ & \mu_Y \circ Tf : (S\to S\times X) \mapsto (S\to S\times Y) \end{aligned} \tag{77}$$

Note that

$$\begin{aligned} & \mu_X : T^2(X) \to T(X) \\ & \mu_X : (S\to S\times (S\to S\times X)) \mapsto (S\to S\times X) \end{aligned} \tag{78}$$

$$\begin{aligned} & T^2 f : T^2(X) \to T^3(Y) \\ & T^2 f : (S\to S\times (S\to S\times X)) \mapsto (S\to S\times (S\to S\times (S\to S\times Y))) \end{aligned} \tag{79}$$

Now

$$\begin{aligned} \mu_X \circ T^2 f &: (S\to (S\times (S\to S\times X))) \mapsto (S\to S\times (S\to S\times Y)) \\ Tf\circ \mu_X &: (S\to S\times (S\to S\times X)) \mapsto (S\to S\times (S\to S\times Y)) \end{aligned}$$

and so

$$\Longrightarrow \mu_X T^2 = Tf\mu_X$$

Thus, μ, as defined above, *is* a natural transformation.

Show $\eta_{T(Y)}\circ f = Tf\circ \eta_X$.

$$\begin{aligned} Tf\circ \eta_X(x) &= (S\to S\times (S\to S\times Y)) \\ \eta_{T(Y)}\circ f(x) = \eta_{T(Y)}(S\to S\times Y) &= (S\to S\times (S\to S\times Y)) \end{aligned}$$

Prove now the coherence conditions, Prop. 22. For the first one,

$$\mu\circ T\mu = \mu\circ \mu T$$

$$\begin{gathered} \mu\circ T\mu(S\to (S\to S\times (S\to S\times X))) = \mu\circ T(S\times (S\to S\times X)) = \mu(S\to S\times (S\times (S\to S\times Y))) = \\ = (S\times (S\to S\times Y)) \\ \mu\circ\mu T(S\to (S\to S\times (S\to S\times X))) = \mu^2(S\to S\times (S\to (S\to S\times (S\to S\times Y)))) = \\ = S\times (S\to S\times Y) \end{gathered}$$

$$\Longrightarrow \mu\circ T\mu = \mu\circ\mu T$$

So the first coherence condition is proven.

For the second,

$$\begin{gathered} \mu\circ T\eta(S\to S\times X) = \mu\circ T(S\to S\times (S\to S\times X)) = (S\to S\times (S\to S\times X)) \\ \mu\eta T(S\to S\times X) = \mu\circ (S\to S\times T(S\to S\times X)) = T(S\to S\times X) = (S\to S\times (S\to S\times X)) \end{gathered}$$

Thus, it's been shown that the choice of endomorphism for the state monad, $X\to [S, S\times X]$, fulfills the monad coherence conditions being of this form.

Consider what "bind" means in category theory. Given

$$\begin{aligned} & f : X \to T(Y) = [S, S\times Y] \\ & g : Y \to T(Z) = [S, S\times Z] \\ & f(x) = (S\to S\times Y) \\ & f(x)(s) = (t,y) \text{ where } s,t\in W \\ & g(y) = (S\to S\times Z) \\ & g(y)(t) = (u,z) \end{aligned} \tag{80}$$

Therefore, if we define the usual projection operations on a Cartesian product,

$$\begin{aligned} \Pi_W &: W\times X\to W \qquad & \Pi_X &: W\times X \to X \\ \Pi_W(w,x) &= w & \Pi_X(w,x) &= x \end{aligned} \tag{81}$$

$$\begin{aligned} & \mu_Z\circ Tg\circ f : X \to T(Z) = [S, S\times Z] \ni (W\to W\times Z) \\ & \mu_Z\circ Tg\circ f : x\mapsto (S\to S\times Z) \equiv \mu_Z\circ Tg\circ f(x) \end{aligned} \tag{82}$$

such that

$$\boxed{g(y)(t) = (u,z) \text{ where } y = \Pi_Y(f(x)(s)),\ t = \Pi_W(f(x)(s))} \tag{83}$$

Compare this to what wikipedia says in Eq. 72.

18.4.5. *Continuation Monad.* From nLab,
In a category $\mathbf{C}$ with internal homs $[-,-]$, so internal homs $[-,-]$ belong to $\text{Obj}\mathbf{C}$ (they're objects as well), given object $Y \in \text{Obj}\mathbf{C}$, the *continuation monad* is endofunctor

$$\boxed{X \mapsto [[X,Y],Y]} \text{ i.e. } T(X) = [[X,Y],Y] \text{ for object } Y \in \text{Obj}\mathbf{C} \tag{84}$$

so that $T = X \mapsto [[X,Y],Y]$

From wikipedia,

unit or return:

return type Y, maps type X into functions of type $(X \to Y) \to Y$

$$\begin{aligned} X \to \quad & (X \to Y) \to Y \\ x \mapsto \quad & f \mapsto f(x) = y \end{aligned} \tag{85}$$

Compare this to

$$\begin{aligned} & T : X \to [[X,Y],Y] \\ & \eta_X : X \to T(X) = [[X,Y],Y] \end{aligned} \tag{86}$$

And so for the *unit* of the Continuation monad,

$$\boxed{\begin{aligned} & \eta_X : X \to T(X) = [[X,Y],Y] \\ & \eta_X : x \mapsto (f \to y) \\ & \eta_X(x)(f) = y = f(x) \end{aligned}} \tag{87}$$

bind:

$$\begin{aligned} ((X \to Y) \to Y) \to \quad & (X \to (X' \to Y) \to Y) \to (X' \to Y) \to \quad && Y \\ c \mapsto \quad & f \mapsto k \mapsto && c(t \mapsto ftk) \end{aligned} \tag{88}$$

From nLab, if morphism $f : X \to Y$ is in Klesi category of the continuation monad, then it's a morphism in $\mathbf{C}$ of form $X \to T(Y)$ in $\mathbf{C}$. Hence morphism in original category *of the form* $X \to [[Y,S],S]$ is much like a map from $X \to Y$, only instead of "returning" output Y directly, it instead feeds it into given function $Y \to S$, which hence **continues** the computation.

So for $X \to [[Y,S],S]$, nLab gives the corresponding *computation rule* for function types of the internal hom $[X,S]$:

$$(y \mapsto a(y))(x) = a(x)$$

Hence for $X \to [[X,Y],Y]$, we have $x \mapsto f \mapsto f(x)$ because $f \in [X,Y]$ because f is treated as an object in $\mathbf{C}$.

So do the following "rewrite" or change of notation on Eq. 88:

$$\begin{aligned} & ((X \to Z) \to Z) \to (X \to (Y \to Z) \to Z) \to (Y \to Z) \to Z \\ & T(X) \to (X \to T(Y)) \equiv \text{Hom}(X, T(Y)) \to \text{Hom}(Y,Z) \to Z \end{aligned} \tag{89}$$

Let $c \in T(X) = [[X,Z],Z]$.
Let $k \in [Y,Z]$
Let $f(x) \in [[Y,Z],Z]$ and so
$f(x)(k) \in Z$

Form $x \mapsto f(x)(k) \in [X,Z]$, so $c(x \mapsto f(x)(k)) \in Z$.

So "bind" as defined in wikipedia, Haskell/Functional Programming isn't the Klesli composition, but it's

$$\mu_Y \circ Tf : T(X) \to T(Y) = [[Y,Z],Z]$$

Write this as

$$\boxed{\begin{aligned} & (\mu_Y \circ T)(f)(c) \in T(Y) = [[Y,Z],Z] \\ & (\mu_Y \circ T)(f)(c)(k) \in Z \end{aligned}} \tag{90}$$

call with current continuation function defined as follows:
call/cc

$$\begin{aligned} ((X \to (X' \to Y) \to Y) \to (X \to Y) \to Y) \to (X \to Y) \quad && \to Y \\ f \mapsto k \quad && \mapsto (f(t \mapsto x \mapsto kt)k) \end{aligned} \tag{91}$$

Let's look at Eq. 91, step-by-step.

$$\begin{aligned} & (X \to (X' \to Y) \to Y) \\ & f \end{aligned} \text{ corresponds to}$$

$$\begin{aligned} & X \to [[X',Y],Y] = T(X') \\ & f \equiv f_{XX'} : X \to T(X') \end{aligned}$$

18.4.6. ***future*** *as Continuation Monad.* Čukić (2018) [10] asks first if `std::future` in C++ is a functor. Recall the definition of a functor, Def. 61:

$$\begin{aligned} & F : \mathbf{C} \to \mathbf{D} \text{ with "object" map} \\ & F : \text{Obj}\mathbf{C} \to \text{Obj}\mathbf{D} \\ & F : X \to F(X) \end{aligned}$$

and the functor's "morphism map":

$$\begin{aligned} & Ff : F(X) \to F(Y) \\ & f : X \to Y,\ f \in \text{Hom}_{\mathbf{C}}(X,Y) \end{aligned}$$

Suppose $X =$ `T1`, $Y =$ `T2`. Then in the example on pp. 220, Sec. 10.7.1 Futures as monads" example of Čukić (2018) [10],

$$\begin{aligned} & F(X) = \texttt{future<T1>} \\ & F(Y) = \texttt{future<T2>} \end{aligned}$$

Instead, the morphisms to consider is of this form:

$$f : X \to T(Y)$$

where T is a functor but a special one: the endomorphism. And T is `std::future`.

Here are examples of morphisms f, g:

```
future<std::string> user_full_name(const std::string& login);
future<std::string> to_html(const std::string& text);
```

Now $Tf : T(X) \to T^2(Y)$, how the morphism map for T applies to morphism f, now. Recall that we can compose this with the multiplication natural transformation component at X, μ_X:

$$\begin{aligned} & \mu_X \circ Tf : T(X) \to T(Y) \\ & \mu_X : T^2(X) \to T(X) \end{aligned}$$

In order to compose morphisms, say f, g, recall the "Klesli composition", Prop. 24

$$g \circ_{\text{Klesili}} f = \mu_Z \circ Tg \circ f \tag{92}$$

If we call $\mu_Z \circ T$ "bind", $\mu_Z \circ T : \mathrm{Hom}(Y, T(Z)) \to \mathrm{Hom}(T(Y), T(Z))$. The correspondence is such:

$$(\texttt{bind}(\texttt{to\_html}))(\texttt{user\_full\_name}) \Longleftrightarrow \mu_Z \circ Tg \circ f : X \to T(Z) \tag{93}$$

18.4.7. *Resources for implementations of Monads.* Python and C++ code.

https://github.com/Iasi-C-CPP-Developers-Meetup/presentations-code-samples, namely https://github.com/Iasi-C-CPP-Developers-Meetup/presentations-code-samples/tree/master/radugeorge

https://github.com/dbrattli/OSlash, namely https://github.com/dbrattli/OSlash/tree/master/tests

State Monad implementation in C++

19. Applications of Category Theory on Hybrid Systems

cf. Ames (2006) [11].

19.1. **D-Categories.** D stands for discrete.

Recall that a small category $\mathbf{C}$ is called *small* if both $\mathrm{Obj}(\mathbf{C})$ and $\hom(\mathbf{C})$ are sets, not proper classes.

Definition 79 (Axiomatic D-categories)**.** *Let D-category be a* small *category* $\mathbf{D}$ *s.t.*

(1) $\forall D \in Obj(\mathbf{D})$,
$\exists$ *morphism* $f \in Mor(\mathbf{D})$ *s.t.* $f \neq 1$ *s.t.*
$f \in Hom(D, *)$ *or* $f \in Hom(*, D)$, *but never both,*
i.e. $\forall$ *diagram* $a_0 \xrightarrow{\alpha_1} a_1 \xrightarrow{\alpha_2} \dots \xrightarrow{\alpha_n} a_n$ *in* $\mathbf{D}$, *all but 1 morphism must be identity (i.e. longest chain of composite non-identity morphisms is of length 1).*

(2) *If for* $D \in Obj(\mathbf{D})$, D *is the domain of a non-identity morphism, i.e.* $\exists f_1 \in Mor(\mathbf{D})$ *s.t.*
$$f_1 \in Hom(D, *),\ f_1 \neq 1$$
then $\exists f_2 \in Hom(D, *)$, $f_2 \neq f_1$, $f_2 \neq 1$ *and* $\forall f \in Hom(D, *)$ *s.t.* $f \neq f_1, f_2$, $\quad f = 1$

cf. 1.2.1 Important objects in D-categories, Ames (2006) [11].

Let
$$\mathrm{Mor}\mathbf{A} = \bigcup_{A, B \in \mathrm{Obj}(\mathbf{A})} \mathrm{Hom}(A, B) \qquad \text{(my notation)}$$
$$\mathrm{Mor}(\mathcal{D}) = \bigcup_{(a,b) \in \mathrm{Obj}(\mathcal{D}) \times \mathrm{Obj}(\mathcal{D})} \mathrm{Hom}_{\mathcal{D}}(a, b) \qquad \text{(Ames' notation)}$$

Let
$$\mathrm{Mor}_1\mathbf{A} := \{ A \in \mathrm{Mor}(\mathbf{A}) | A \neq 1 \} \tag{94}$$

For a D-category, consider these subset of $\mathrm{Obj}(\mathbf{D})$,

Definition 80 (Edge set)**.** ***edge set*** *of* $\mathbf{D}$, $E(\mathbf{D})$,
$$\begin{gathered} E(\mathbf{D}) := \{ A \in Obj(\mathbf{D}) | \alpha \in Hom(A, *), \beta \in Hom(A, *), \alpha, \beta \in Mor_1(\mathbf{D}),\ \alpha \neq \beta \} \text{ i.e.} \\ E(\mathbf{D}) := \{ A \in Obj(\mathbf{D}) | \alpha, \beta \in Hom(A, *),\ \alpha, \beta \neq 1,\ \alpha \neq \beta \} \end{gathered} \tag{95}$$

i.e. $\forall A \in E(\mathbf{D})$, $\exists \alpha, \beta \in \mathrm{Mor}(\mathbf{D})$, $\alpha, \beta \neq 1$, s.t. $\alpha, \beta \in \mathrm{Hom}(D, *)$; denote these morphisms by s_a, t_a (this specific choice will define an **orientation**).

Conversely, given morphism $\gamma \in \mathrm{Mor}(\mathbf{D})$, $\gamma \neq 1$, $\exists ! A \in E(\mathbf{D})$ s.t. $\gamma = s_a$ or $\gamma = t_a$, i.e. $\gamma \in \mathrm{Hom}(A, *)$.

Definition 81 (Vertex set)**.** ***vertex set*** *of* $\mathbf{D}$*:*
$$V(\mathbf{D}) = (E(\mathbf{D}))^c \tag{96}$$

Definition 82 (Orientation)**.** *Orientation of D-category* $\mathbf{D}$ *is a pair of functions* (s, t) *between sets.*

Part 3. Category Theory and Databases

20. Types

20.1. **Data models for nested arrays, dictionaries, and tabular data.**

Definition 83 (Container types)**.** ***Container types*** *- arrays, key/value pair dictionaries (or: hashes, association lists)*

Definition 84 (atomic types)**.** *basic atomic types (e.g. numbers, strings, Booleans)*

Definition 85 (nesting)**.** ***nesting****: containers may contain atomic values as well as other containers.*

Definition 86 (flat)**.** *Tabular data model (i.e. tables) is* ***flat****: field contains* ***atomic*** *values.*

cf. Number 2, Data Models and Languages, Grust and Duta (2017) [14]

20.2. **Typed Data, Untyped data, in the relational data model.**

Definition 87. *Untyped data models - text, JSON, and tabular data models (e.g. CSV) do* ***not*** *enforce values (container or atomic) to be of specific types.*

These data models are thus referred to as being ***untyped****.*

cf. Grust and Duta (2017) [14]
cf. Number 4, "The Relational Data Model", Grust and Duta (2017) [14]

Definition 88 (Types)**.** *Let* $\mathbb{T} \equiv$ *set of all data types (built-in and user-defined).*
$\forall$ *value* $v \in \mathbb{V}$ *stored in a relation cell must be of type* $t \in \mathbb{T}$.
e.g. When PostgreSQL starts, $\mathbb{T}$ *initialized as*
$$\mathbb{T} = \{ \texttt{boolean}, \texttt{integer}, \texttt{text}, \texttt{bytea}, \dots \} \tag{97}$$
Consider category ***Text*** *s.t.* $Obj(\mathbf{Types}) \equiv DT \in Obj(\mathbf{Set})$. *Then* $\mathbb{T} \equiv DT$, *denoting* data type.

Definition 89 (Values)**.** $\forall v \in \mathbb{V}$ *stored in a relation cell,* v *is an element of the set of all values* $\mathbb{V}$.
in the relational data model, all values $v \in \mathbb{V}$ *are "atomic."*
$$\mathbb{V} = \{ \texttt{true}, \texttt{false}, 0, -1, 1, -2, 2, \dots \}$$
Here, I'll use the notation V *to denote* $\mathbb{V}$, *the set of all values.*

Definition 90 (Domains)**.** $\forall t \in DT$, *its domain* $dom(t) :=$ *set of all values of type* t *(i.e.* $dom(\cdot)$ *is a function with signature*
$$DT \to 2^V$$
) e.g. $dom(\texttt{integer}) = \{0, -1, 1, -2, 2, \dots\}$
$dom(\texttt{boolean}) = \{\texttt{true}, \texttt{false}\}$

Definition 91 (type specification)**.** *type specification* $:=$ *function* $\pi : U \to \mathbf{DT}$ *(Spivak's notation)* $\equiv \pi : U \to DT$, $U, DT \in \mathbf{Set}$.
set $DT \equiv$ *set of data types for* π, *set* $U \equiv$ *domain bundle for* π.
$\forall t \in DT$, *preimage* $\pi^{-1}(t) \subset U$, $\quad \pi^{-1}(t) \equiv$ *domain of* t, $x \in \pi^{-1}(t) \equiv$ *object of type* T.
To reconcile Grust and Duta (2017) [14]'s definition of types above, use this notation:
$$\boxed{\begin{aligned} &\textbf{\textit{type specification}}\ \pi : V \to DT \\ &\qquad \pi(v) \in DT \\ &dom \equiv \pi^{-1} : DT \to V,\ i.e. \\ &\pi^{-1}(t) \subset t \qquad \forall t \in DT \end{aligned}} \tag{98}$$

cf. Spivak (2009) [15]

Corollary 3 (type specification). *If v has type T, $\pi(v) = T$, $\implies v \in \pi^{-1}(T)$*

Proposition 25 (`CREATE DOMAIN`). *Consider new type t' (SQL command `CREATE DOMAIN`) so*

$$
\begin{gathered}
t' \in DT \\
\pi^{-1}(t') \subseteq \pi^{-1}(t)
\end{gathered} \tag{99}
$$

21. Relational Data Model

relational data model maybe understood as a typed variant of the tabular data model.

(1) $\exists$ only 1 container type: **table** (or: multisets) **of rows**
(2) all rows are of same **row type** which is declared when table is created.
(3) row type **consists** of sequence of **atomic types**.

In the relational data mode, data is exclusively organized in **relations**, i.e. sets of tuples of data. Data in each **attribute** (tuple component) is **atomic**, and of declared **type**.

21.1. Schemata and Relations.

In the relational data model, each **attribute** of a table has a **declared type**.

If attribute has declared type t, the RDBMS will exclusively store values v in that attribute s.t.

(1) $v \in \mathrm{dom}(t)$ i.e. $v \in \pi^{-1}(t)$
(2) v can successfully be casted to type t

21.2. Attributes (Columns).

Let $\mathbb{A}$ denote set of attribute names of all relations.

21.3. Attribute types.

$\forall$ attribute $a \in \mathbb{A}$ has declared (attribute) type $\mathrm{type}(a) = t \in DT$ (i.e. $\mathrm{type}(\cdot)$ is a function with signature $\mathbb{A} \to \mathbb{T}$).

Consider Definition 2.2.3 of Spivak (2009) [15],

Definition 92 (simple schema of type π, (C, σ)). *Let type specification $\pi : V \to DT$.*
***simple schema** of type π consists of pair (C, σ) where C is a finite (totally) ordered set and function $\sigma : C \to DT$. $C \equiv$ **column set** or **set of attributes** for σ and π as* type specification for σ.

Compare the notation above. Conclude that

$$
\begin{gathered}
C \equiv \mathbb{A}, \quad c \equiv a \\
\sigma(c) \equiv \mathrm{type}(a)
\end{gathered}
$$

Here, choose the following notation:

$$
\boxed{
\begin{gathered}
\sigma : C \to DT \\
\pi^{-1} : DT \to V \\
\pi^{-1} \circ \sigma \equiv \mathrm{val} \equiv \text{set of (admissable) attribute values for attribute (column) } a \ (c)
\end{gathered}
}
$$

Also, summarize our definitions with these commutative diagrams:

$$
\begin{tikzcd}
V \arrow[d, "\pi"'] \\
DT
\end{tikzcd}
\quad
\begin{tikzcd}
v \arrow[d, mapsto, "\pi"'] \\
t
\end{tikzcd}
\quad
\begin{tikzcd}
V & \pi^{-1}(DT) \arrow[l, hook', "i"'] \\
DT \arrow[u, "i \circ \pi^{-1}"] \arrow[ur, "\pi^{-1}"] &
\end{tikzcd}
\quad
\begin{tikzcd}
V & V_t \subseteq V \arrow[l, hook', "i"'] \\
t \arrow[ur, mapsto, "\pi^{-1}"] &
\end{tikzcd}
$$

$$
\begin{tikzcd}
V & \pi^{-1}(DT) \arrow[l, hook', "i"'] \\
DT \arrow[u, "i \circ \pi^{-1}"] \arrow[ur, "\pi^{-1}"] & C \arrow[l, "\sigma"'] \arrow[u, "\pi^{-1} \circ \sigma \equiv \mathrm{dom} \circ \mathrm{type}"']
\end{tikzcd}
\quad
\begin{tikzcd}
V & \pi^{-1}(\sigma(c_i)) = V_{c_i} \arrow[l, hook', "i"'] \\
\sigma(c_i) \equiv t_i \arrow[ur, mapsto, "\pi^{-1}"] & c_i \arrow[l, mapsto, "\sigma"'] \arrow[u, mapsto, "\pi^{-1} \circ \sigma"']
\end{tikzcd}
\tag{100}
$$

Definition 93 (Relation Schema). *A relation schema associates a relation (table?) name R with its set of declared attributes (a subset of $\mathbb{A}$)*

$$
(R, \{a_1, \dots a_n\}) \tag{101}
$$

Common notation: $R(a_1, \dots a_n)$, so that R is called a n-ary relation.
More notation: $sch(R) = \{a_1, \dots a_n\}$, and $deg(R) = n$ (degree).
Relational database schema:** a non-empty finite set of relation schemata makes a **relational database schema

$$
\{(R_1, \alpha_1), (R_2, \alpha_2), \dots\} \tag{102}
$$

where $\alpha_i \subseteq \mathbb{A}_i$. In a relational database schema, the relation names R_i are unique.

Definition 94 (Tuple). *Given relation (i.e. table) $R(a_1, \dots, a_n)$, a **tuple** t of R maps attributes to values, i.e. t is a function with signature $\{a_1, \dots a_n\} \to V$ with*

$$
\forall a \in \{a_1, \dots, a_n\} : t(a) \in \mathit{val}(a)
$$

Common notation for $t(a)$ is $t.a$.

Recall that $\mathrm{val}(a) := \mathrm{dom}(\mathrm{type}(a))$.
Take note that **tuple**, defined by Grust and Duta (2017) [14], is the *same* as **record** or **row** r, defined by Spivak (2009) [15].

Definition 95 (record or row). ***record** or **row** on (C, σ)*

$$
\begin{gathered}
r : C \to V_\sigma \equiv V_c \\
r(c) \equiv v_{rc} \in \pi^{-1} \circ \sigma(c)
\end{gathered} \tag{103}
$$

SQL `CREATE TABLE` command prescribes an order of the attributes of a relation. This deviates from relational data model's tuple model (name-to-value mapping).

Definition 96 (Row). *Given SQL table $R(a_1, \dots, a_n) \equiv R(\alpha) \equiv \tau$, a row r of τ is an ordered sequence (a_i is called the ith column)*

$$
r = (v_1, \dots, v_n) \in \mathit{val}(a_1) \times \dots \times \mathit{val}(a_n) \equiv \pi^{-1} \circ \sigma(a_1) \times \dots \times \pi^{-1} \circ \sigma(a_n) \tag{104}
$$

Thus, r is a function $\{1, \dots, n\} \to V$ with $\forall\, i \in \{1, \dots, n\}$, $r(i) \in \pi^{-1} \circ \sigma(a_i)$

The set of tuples (rows) stored in a relation (table) is expected to change frequently.

Definition 97 (Relation instance (state)). *The current finite set of tuples $t_i \equiv r_i$ of relation (table) $R(a_1, \dots a_n) \equiv (A, \sigma)$ is called the relation's **instance** (or **state**).*

$$
\mathit{inst}(R) = \{t_1, t_2, \dots t_m\} \implies \equiv \Gamma^\pi(\sigma) \equiv \Gamma(A, \sigma)
$$

Database (instance) state - The database instance comprises instances of all its relations.

$$= \{\Gamma(A^i, \sigma^i)\}_i$$

21.4. **Constraints.** cf. Number 5, "Constraints" of Grust and Duta (2017) [14].

Definition 98 (Constraints)**.** *An **integrity constraints** specifies conditions which table states have to satisfy at all times.*

Current set of constraints, $\mathbb{C}$, is integral part of database schema:

$$(\{(R_1, \alpha_1), (R_2, \alpha_2), \dots\}, \mathbb{C})$$

Set of constraints $\mathbb{C} \equiv$ set of morphisms of table $\tau \equiv \{(R_1, \alpha_1), (R_2, \alpha_2), \dots\}$, $C \subset Mor\tau$.

RDBMS will refuse table state changes that violate any constraint $c \in C$.

21.5. **Key constraints.**

Definition 99 (Key)**.** *Key of a table $R(a_1, \dots, a_n) \equiv (\alpha, \sigma) \equiv (\alpha, type) :=$ set of columns $K \subseteq \{a_1, \dots, a_n\}$ that unique identifies rows of R:*

$$\forall t, u \in inst(R),\, t.K = u.K \Longrightarrow t = u$$

Read: "If 2 rows agree on the columns in K, they are indeed the same row."

I will give the following change of notation a try:

$$\forall r_i, r_j \in \{r_i\}_i^\tau, \text{ if } r_i(K) = r_j(K), \text{ then } r_i = r_j \quad (105)$$

$$\begin{tikzcd} A \arrow[r, "\sigma"] \arrow[d, "r"'] & DT \\ V_A \arrow[ur, "\pi^{-1}"'] & \end{tikzcd} \qquad \begin{tikzcd} a \arrow[r, mapsto, "\sigma"] \arrow[d, mapsto, "r"'] & t \arrow[dl, mapsto, "\pi^{-1}"] \\ V_A & \end{tikzcd} \quad (106)$$

Here's a "dictionary" between the definitions so far for relational databases and familiar terms for tables filled with data: cf. 08-30, Week 1 slides of Yang (2012) [17].

database - collection of relations (or tables)
relation - table
attributes - columns
tuple - row - record
relation schema - heading (heading for a table?, Grust and Duta (2017) [14])
relation contents - body (body of a table?, Grust and Duta (2017) [14])

CSV	Relational Model	SQL
	Domain	Domain
	Type	Type
	Schema	Schema
File	Relation	Table
Line	Tuple	Row
Field	Attribute	Column

Indeed, even Grust and Duta (2017) [14] remarks: "You will find that textbooks, papers, practitioners, academics, these slides, and even PostgreSQL use a mixture of terminology. Deal with it."

cf. Slides Number 7, "Referential Integrity" of Grust and Duta (2017) [14].

Definition 100 (Foreign Keys)**.** *Let (S, α) and (T, β) denote 2 relational schemata (not necessarily distinct), where $K = \{b_{j_1}, \dots, b_{j_k}\} \subseteq \beta$ is the primary key of T.*
Let $F = \{a_{i_1}, \dots, a_{i_k}\} \subseteq \alpha$ with $type(a_{i_h}) = type(b_{j_h})$, $h \in 1, \dots, k$.

*F is a **foreign key** in S **referencing** T, if*

$$\forall s \in inst(S) : \exists t \in inst(T) : s.F = t.K$$

The $\forall$ and $\exists$ condition validates the assumption of K being a key in target T, i.e. that there **exists a row** in table T whose K identifier matches that of table S

K being a key in target T validates the assumption that there is **no more than one row** of in a table T with a matching key K.

In general, **a foreign key** F **is not a key** in source table S. 2 rows $s_1, s_2 \in \text{inst}(S)$ with $s_1.F = s_2.F$ can refer to the same row in target T.

Here is a great example of a foreign key: `https://www.w3schools.com/sql/sql_foreignkey.asp`

So from w3schools,

"A FOREIGN KEY is a key used to link 2 tables together.

A FOREIGN KEY is a field (or collection of fields) in 1 table that refers to the PRIMARY KEY in another table.

The table containing the foreign key is called the child table, and the table containing the candidate key is called the referenced or parent table.

"Persons" table:

PersonID	LastName	FirstName	Age
1	Hansen	Ola	30
2	Svendson	Tove	23
3	Pettersen	Kari	20

"Orders" table:

OrderID	OrderNumber	PersonID
1	77895	3
2	44678	3
3	22456	2
4	24562	1

Notice that the "PersonID" column in the "Orders" table points to the "PersonID" column in the "Persons" table.

The "PersonID" column in the "Persons" table is the PRIMARY KEY in the "Persons" table.

The "PersonId" column in the "Orders" table is a FOREIGN KEY in the "Orders" table.

Thus, in other words,
given 2 tables (i.e. 2 relational schemata, not necessarily distinct), (S, α), (T, β),
given primary key $PK = \{b_{j_1}, \dots, b_{j_k}\} \subseteq \beta$, primary key of of T, and
given foreign key $FK = \{a_{i_1}, \dots, a_{i_k}\} \subseteq \alpha$ s.t.

$$\begin{gathered} \text{type}(a_{i_h}) = \text{type}(b_{j_h}), \quad h \in 1, 2, \dots k \\ (\text{alternate notation})\sigma(a_{i_h}) = \sigma(b_{j_h}) \end{gathered}$$

then
$\forall s \in \text{inst}(S)$, $\exists t \in \text{inst}(T)$ ((alternate notation) $s \in \Gamma(S)$, $t \in \Gamma(T)$), s.t.

$$s(FK) = t(PK)$$

Describe this categorically as

$$S \xrightarrow{FK} T$$

which is from the introductory talk by Spivak (2012)[16], "What does equivalence of paths mean?", where it states, "arrows represent foreign keys", "from table a to table b," and where "we can take any record in table a and return a record in table b.

21.6. Functional Dependency.

Definition 101 (Functional Dependency (FD))**.** *cf. Slides Number 10, "Functional Dependencies" of Grust and Duta (2017)* [14].

Let $(R, \alpha) \equiv$ relational schema. Given $\beta \subseteq \alpha$, $c \in \alpha$,
***functional dependency** $\beta \to c$ holds in R if*

$$\forall t, u \in \mathit{inst}(R), \, t.\beta = u.\beta \Longrightarrow t.c = u.c$$

Notation: the functional dependency $\beta \to \{c_1, \dots, c_n\}$ abbreviates set of FDs $\beta \to c_1, \dots, \beta \to c_n$.
cf. "Functional dependency", wikipedia
A set of attributes $X \subseteq R$ ($\equiv X \subseteq \alpha$ for $R = (R, \alpha)$) is said to functionally determine another set of attributes $Y \subseteq R$ ($\equiv Y \subseteq \alpha$), written $X \to Y$. ($\equiv \beta \to c$)
iff $\forall X$ value in R is associated with precisely 1 Y value in R; R is then said to satisfy the functional dependency $X \to Y$

e.g. Cars: $\beta =$ vehicle identification number (VIN), $c =$ Engine capacity (because assume a car cannot have 2 engines).
e.g Employee department:
$\beta =$ employee ID, $c =$ employee name.
$\beta =$ employee ID, $c =$ department ID.
$\beta =$ department ID, $c =$ department name.
Recall definition of injective function, $f : X \to Y$,

$$\forall a, b \in X, \, f(a) = f(b) \Longrightarrow a = b$$

Instead of writing the functional dependency (FD) with this notation: given $\beta, c \subseteq \alpha$,

$$\forall r_1, r_2 \in \mathrm{inst}(R) \equiv \Gamma(\alpha), \, r_1(\beta) = r_2(\beta), \Longrightarrow r_1(c) = r_2(c)$$

rewrite this as follows:

$$\begin{gathered} \text{Given} \\ \beta : \Gamma(\alpha) \to \mathrm{val}(\beta) \equiv \mathrm{dom}(\mathrm{type}(\beta)) \equiv \pi^{-1} \circ \sigma(\beta), \\ c : \Gamma(\alpha) \to \mathrm{val}(c) \equiv \mathrm{dom}(\mathrm{type}(c)) \equiv \pi^{-1} \circ \sigma(c), \\ \forall r_1, r_2 \in \mathrm{inst}(R) \equiv \Gamma(\alpha), \, \beta(r_1), \beta(r_2) \in \mathrm{val}(\beta) \\ \text{Then } \beta \in \mathrm{Hom}(\mathrm{inst}(R), \mathrm{val}(\beta)) \equiv \mathrm{Hom}(\Gamma(\alpha), \pi^{-1} \circ \sigma(\alpha)) \end{gathered} \tag{107}$$

So define a **functional dependency** FD as

$$\begin{gathered} \mathrm{FD} : 2^\alpha \to 2^\alpha \\ \mathrm{FD} : \beta \mapsto c \text{ i.e. } \boxed{FD(\beta) = c} \end{gathered} \tag{108}$$

where $\beta, c \in 2^\alpha \equiv$ power set of α, i.e. set of all subsets of α, so that $\beta, c \subseteq \alpha$.
Let $FD^{-1} : 2^\alpha \to 2^\alpha$, s.t. $FD^{-1}(c) = \beta$.
So if $\forall r_1, r_2 \in \mathrm{inst}(R) \equiv \Gamma(\alpha)$, for $\beta \subseteq \alpha$, $\beta(r_1), \beta(r_2) \in \mathrm{val}(\beta)$. Then $\beta \in \mathrm{Hom}(\mathrm{inst}(R), \mathrm{val}(\beta)) \equiv \mathrm{Hom}(\Gamma(\alpha), \pi^{-1} \circ \sigma(\alpha))$.
Suppose $\beta(r_1) = \beta(r_2)$. The *key insight* is the following:

$$\beta(r_1) = \beta(r_2) = \mathrm{FD}^{-1}(c)(r_1) = \mathrm{FD}^{-1}(c)(r_2) = (\mathrm{FD}^{-1} \circ c)(r_1) = (\mathrm{FD}^{-1} \circ c)(r_2) \tag{109}$$

Then $c(r_1) = c(r_2)$ implies that FD^{-1} is *injective.*
Then FD^{-1} is a *monomorphism*, and FD^{-1} is injective, since $\forall r_1, r_2 \in \mathrm{inst}(R) \equiv \Gamma(\alpha)$,

$$\beta(r_1) = \beta(r_2) = (\mathrm{FD}^{-1} \circ c)(r_1) = (\mathrm{FD}^{-1} \circ c)(r_2) \Longrightarrow c(r_1) = c(r_2) \tag{110}$$

$$\mathrm{inst}(R) \xrightarrow{c} \mathrm{val}(C) \xrightarrow{\mathrm{FD}^{-1}} \mathrm{val}(\beta)$$

Then conclude that

Theorem 16. *Given a relational schema (R, α), a set of attributes α, instance of the relation (i.e. a table), $\mathit{inst}(R) \equiv \Gamma(\alpha)$, and $\beta, c \subseteq \alpha$,*
for morphisms β, c,

$$\begin{gathered} \beta \in \mathit{Hom}(\Gamma(\alpha), \mathit{val}(\beta)) \\ c \in \mathit{Hom}(\Gamma(\alpha), \mathit{val}(c)) \end{gathered}$$

*Then for $FD : 2^\alpha \to 2^\alpha$, $FD(\beta) = c$, FD^{-1} is a **monomorphism**, i.e. $\forall r_1, r_2 \in \Gamma(\alpha)$,*

$$\beta(r_1) = \beta(r_2) = (FD^{-1} \circ c)(r_1) = (FD^{-1} \circ c)(r_2) \text{ i.e. } (FD^{-1} \circ c) = (FD^{-1} \circ c)$$

then $c(r_1) = c'(r_2)$ i.e. $c = c'$.

Database Queries and Constraints via Lifting Problems. David I. Spivak. `https://arxiv.org/pdf/1202.2591.pdf`

22. Databases and Categories

cf. Spivak (2012) [16].
$A \xrightarrow{\mathrm{FK}} B$, $A, B \in \mathrm{Obj}(\mathbf{DB})$

Part 4. Reading notes on Cox, Little, O'Shea's *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*

23. Geometry, Algebra, and Algorithms

23.1. Polynomials and Affine Space.
fields are important is that linear algebra works over *any* field

Definition 102 (2)**.** *set of all polynomials in $x_1, \dots, x_n$ with coefficients in k, denoted $k[x_1, \dots, x_n]$*

polynomial f *divides* polynomial g provided $g = fh$ for some $h \in k[x_1, \dots, x_n]$
$k[x_1, \dots, x_n]$ satisfies all field axioms except for existence of multiplicative inverses; commutative ring, $k[x_1, \dots, x_n]$ *polynomial ring*

Exercises for 1. **Exercise 1.** $\mathbb{F}_2$ commutative ring since it's an abelian group under addition, commutative in multiplication, and multiplicative identity exists, namely 1. It is a field since for $1 \neq 0$, the multiplicative identity is 1.

Exercise 2.

(a)
(b)
(c)

23.2. Affine Varieties.

23.3. Parametrizations of Affine Varieties.

23.4. Ideals.

23.5. Polynomials of One Variable.

24. Groebner Bases

24.1. Introduction.

24.2. Orderings on the Monomials in $k[x_1, \dots, x_n]$.

24.3. A Division Algorithm in $k[x_1, \dots, x_n]$.

24.4. Monomial Ideals and Dickson's Lemma.

24.5. The Hilbert Basis Theorem and Groebner Bases.

24.6. Properties of Groebner Bases.

24.7. Buchberger's Algorithm.

25. Elimination Theory

25.1. The Elimination and Extension Theorems.

25.2. The Geometry of Elimination.

26. The Algebra-Geometry Dictionary

26.1. Hilbert's Nullstellensatz.

26.2. Radical Ideals and the Ideal-Variety Correspondence.

27. Polynomial and Rational Functions on a Variety

27.1. Polynomial Mappings.

28. Robotics and Automatic Geometric Theorem Proving

28.1. Geometric Description of Robots.

Part 5. Reading notes on Cox, Little, O'Shea's *Using Algebraic Geometry*

Using Algebraic Geometry. David A. Cox. John Little. Donal O'Shea. Second Edition. Springer. 2005. ISBN 0-387-20706-6 QA564.C6883 2004

29. Introduction

29.1. Polynomials and Ideals. *monomial*

(111) $$x_1^{\alpha_1} \dots x_n^{\alpha_n} \quad (1.1)$$

total degree of x^α is $\alpha_1 + \dots + \alpha_n \equiv |\alpha|$

field k, $k[x_1 \dots x_n]$ collection of all polynomials in $x_1 \dots x_n$ with coefficients k.

polynomials in $k[x_1 \dots x_n]$ can be added and multiplied as usual, so $k[x_1 \dots x_n]$ has structure of commutative ring (with identity)
however, only nonzero constant polynomials have multiplicative inverses in $k[x_1 \dots x_n]$, so $k[x_1 \dots x_n]$ not a field
however set of rational functions $\{f/g | f, g \in k[x_1 \dots x_n], g \neq 0\}$ is a field, denoted $k(x_1 \dots x_n)$

so

$$f = \sum_\alpha c_\alpha x^\alpha$$

where $c_\alpha \in k$
so

$$f \in k[x_1 \dots x_n] = \{f | f = \sum_\alpha c_\alpha x^\alpha, x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}, c_\alpha \in k\}$$

f homogeneous if all monomials have same total degrees
polynomial f is homogeneous if all monomials have the *same total degree*

Given a collection of polynomials $f_1 \dots f_s \in k[x_1 \dots x_n]$, we can consider all polynomials which can be built up from these by multiplication by arbitrary polynomials and by taking sums

Definition 103 (1.3)**.** *Let $f_1 \dots f_s \in k[x_1 \dots x_n]$*
Let $\langle f_1 \dots f_s \rangle = \{p_1 f_1 + \dots + p_s f_s | p_i \in k[x_1 \dots x_n]$ for $i = 1 \dots s\}$

Exercise 1.
(a) $x^2 = x \cdot (x - y^2) + y \cdot (xy)$
(b)

$$p \cdot (x - y^2) = px - py^2$$

and for $pxy = (py)x$
(c)

$$p(y)(x - y^2) = p(y)x - p(y)y^2 \notin \langle x^2, xy \rangle$$

Exercise 2.

$$\sum_{i=1}^s p_i f_i + \sum_{j=1}^s q_j f_j = \sum_{i=1}^s (p_i + q_i) f_i, \quad p_i + q_i \in k[x_1 \dots x_n]$$

$\langle f_1 \dots f_s \rangle$ closed under sums in $k[x_1 \dots x_n]$

If $f \in \langle f_1 \dots f_s \rangle$,
$p \in k[x_1 \dots x_n]$

$$p \cdot f = p \sum_{i=1}^s q_j f_j = \sum_{i=1}^s p q_j f_j, \quad pq_j \in k[x_1 \dots x_n] \text{ so}$$
$$p \cdot f \in \langle f_1 \dots f_s \rangle$$

Done.

The 2 properties in Ex. 2 are defining properties of ideals in the ring $k[x_1 \dots x_n]$

Definition 104 (1.5)**.** *Let $I \subset k[x_1 \dots x_n]$, $I \neq \emptyset$*
I ideal if
(a) $f + g \in I, \quad \forall f, g \in I$
(b) $pf \in I, \quad \forall f \in I,$ *arbitrary* $p \in k[x_1 \dots x_n]$

Thus $\langle f_1 \dots f_s \rangle$ is an ideal by Ex. 2.

we call it the ideal generated by $f_1 \dots f_s$.

Exercise 3. Suppose $\exists$ ideal J, $f_1 \dots f_s \in J$ s.t. $J \subset \langle f_1 \dots f_s \rangle$
if $f \in \langle f_1 \dots f_s \rangle$, $f = \sum_{i=1}^s p_i f_i$, $p_i \in k[x_1 \dots x_n]$

$\forall i = 1 \dots s$, $p_i f_i \in J$ and so $\sum_{i=1}^s p_i f_i \in J$, by def. of J as an ideal.

$$\langle f_1 \dots f_s \rangle \subseteq J \quad \implies J = \langle f_1 \dots f_s \rangle$$

$\implies \langle f_1 \dots f_s \rangle$ is smallest ideal in $k[x_1 \dots x_n]$ containing $f_1 \dots f_s$

Exercise 4. For $I = \langle f_1 \dots f_s \rangle$
$J = \langle g_1 \dots g_t \rangle$

$I = J$ iff $s = t$ and $\forall f \in I$, $f = \sum_{i=1}^t q_i g_i$ and if $0 = \sum_{i=1}^t q_i g_i$, $q_i = 0$, $\forall i = 1 \dots t$, and if $0 = \sum_{i=1}^s p_i f_i$, $p_i = 0$, $\forall i = 1 \dots s$

Definition 105 (1.6).
$$\sqrt{I} = \{ g \in k[x_1 \dots x_n] | g^m \in I \textit{ for some } m \geq 1 \}$$

e.g. $x + y \in \sqrt{\langle x^2 + 3xy, 3xy + y^2 \rangle}$
in $\mathbb{Q}[x,y]$ since

$$(x+y)^3 = x(x^2 + 3xy) + y(3xy + y^2) \in \langle x^2 + 3xy, 3xy + y^2 \rangle$$

- (Radical Ideal Property) $\forall$ ideal $I \subset k[x_1 \dots x_n]$, $\sqrt{I}$ ideal, $\sqrt{I} \supset I$
- **(Hilbert basis Thm.)** $\forall$ ideal $I \subset k[x_1 \dots x_n]$
  $\exists$ finite generating set,
  i.e. $\exists \{ f_1 \dots f_2 \} \subset k[x_1 \dots x_n]$ s.t. $I = \langle f_1 \dots f_s \rangle$
- (Division Algorithm in $k[x]$) $\forall f, g \in k[x]$ (EY : in 1 variable)
  $\forall f, g \in k[x]$ (in 1 variable)
  $f = qg + r$, $\exists !$ quotient q, $\exists$ remainder r

29.2.

29.3. **Gröbner Bases.**

Definition 106 (3.1). *Gröbner basis for* $I \equiv G = \{ g_1 \dots g_k \} \subset I$ *s.t.* $\forall f \in I$, *LT*(f) *divisible by LT*(g_i) *for some* i

- (Uniqueness of Remainders) let ideal $I \subset k[x_1 \dots x_n]$
  division of $f \in k[x_1 \dots x_n]$ by Grö bner basis for I, produces $f = g + r$, $g \in I$, and no term in r divisible by any element of LT(I)

29.4. **Affine Varieties.** affine n-dim. space over k $\quad k^n = \{ (a_1 \dots a_n) | a_1 \dots a_n \in k \}$
$\forall$ polynomial $f \in k[x_1 \dots x_n]$, $(a_1 \dots a_n) \in k^n$
$f : k^n \to k$
$f(a_1 \dots a_n)$ s.t. $x_i = a_i$ i.e.

if $f = \sum_\alpha c_\alpha x^\alpha$ for $c_\alpha \in k$, then
$f(a_1 \dots a_n) = \sum_\alpha c_\alpha a^\alpha \in k$, where $a^\alpha = a_1^{\alpha_1} \dots a_n^{\alpha_n}$

Definition 107 (4.1). *affine variety* $\mathbf{V}(f_1 \dots f_s) = \{ (a_1 \dots a_n) | (a_1 \dots a_n) \in k^n, f_1(x_1 \dots x_n) = \dots = f_s(x_1 \dots x_n) = 0 \}$ *subset* $V \subset k^n$ *is affine variety if* $V = V(f_1 \dots f_s)$ *for some* $\{ f_i \}$*, polynomial* $f_i \in k[x_1 \dots x_n]$

- (Equal Ideals Have Equal Varieties) If $\langle f_1 \dots f_s \rangle = \langle g_1 \dots g_t \rangle$ in $k[x_1 \dots x_n]$, then $\mathbf{V}(f_1 \dots f_s) = \mathbf{V}(g_1 \dots g_t)$

so, recap
if $\langle f_1 \dots f_s \rangle = \langle g_1 \dots g_t \rangle$ in $k[x_1 \dots x_n]$,
then $V(f_1 \dots f_s) = V(g_1 \dots g_t)$

Recall Hilbert basis Thm. $\forall$ ideal $I \subset k[x_1 \dots x_n]$
$$I = \langle f_1 \dots f_s \rangle$$

$\Longrightarrow$ if $I = J$, then $V(I) = V(J)$
think of V defined by I, rather than $f_1 = \dots = f_s = 0$

Exercise 3.

Recall Def. 1.5 Let $I \subset k[x_1 \dots x_n]$

I ideal if $f + g \in I \quad \forall f, g \in I$
$pf \in I$, $\quad \forall f \in I$ arbitrary $p \in k[x_1 \dots x_n]$
Let $f, g \in I(V)$
$$(f+g)(a_1 \dots a_n) = f(a_1 \dots a_n) + g(a_1 \dots a_n) = 0 + 0 = 0 \qquad f + g \in I(V)$$
$$pf(a_1 \dots a_n) = p(a_1 \dots a_n) f(a_1 \dots a_n) = 0 \qquad pf \in I(V)$$
Then $I(V)$ an ideal.
$V = V(x^2)$ in $\mathbb{R}^2$
$I = \langle x^2 \rangle$ in $\mathbb{R}[x,y]$, $\quad I = \{ px^2 | p \in k[x,y] \}$
$I \subset I(V)$, since $px^2 = 0$ for $x^2 = 0$, $(0, b)$, $\quad b \in \mathbb{R}$
But $p(x,y) = x \in I(V)$, as
$$I(V) = \{ f \in k[x_1 \dots x_n] | f(a_1 \dots a_n) = 0, \forall (a_1 \dots a_n) \in V \}$$
$p(0, b) = x = 0$
But $x \notin I$

Exercise 4. $I \subset \sqrt{I}$
Recall Def. 1.6 $\sqrt{I} = \{ g \in k[x_1 \dots x_n] | g^m \in I$ for some $m \geq 1 \}$
$\forall f \in I$, $f = f^1$, $m = 1$, so $f \in \sqrt{I}$, $\quad I \subset \sqrt{I}$
Hilbert basis thm., $\forall$ ideal $I \subset k[x_1 \dots x_n]$ s.t. $I = \langle f_1 \dots f_s \rangle$
$\left\{ {}^{V(I) = \{ (a_1 \dots a_n) | (a_1 \dots a_n) \in k^n, f_1(a_1 \dots a_n) = \dots = f_s(a_1 \dots a_n) = 0 } \right\}$
$\mathbf{I}(\mathbf{V}(I)) = \{ f \in k[x_1 \dots x_n] | f(a_1 \dots a_n) = 0 \quad \forall (a_1 \dots a_n) \in V(I) \}$
Let $g \in \sqrt{I}$, $\quad g^m \in I$, $\quad g^m = g^{m-1} g$
$g^m(a_1 \dots a_n) = 0 = g^{m-1}(a_1 \dots a_n) g(a_1 \dots a_n) = 0$. Then $g(a_1 \dots a_n) = 0$ or $g^{m-1}(a_1 \dots a_m) = 0$
as $g^m \in I$, and $V(I)$ is s.t. $f_1(a_1 \dots a_n) = \dots = f_s(a_1 \dots a_n) = 0$ for $I = \langle f_1 \dots f_s \rangle$

- (Strong Nullstellensatz) if k algebraically closed (e.g. $\mathbb{C}$), I ideal in $k[x_1 \dots x_n]$, then
$$\mathbf{I}(\mathbf{V}(I) = \sqrt{I}$$
- (Ideal-variety correspondence) Let k arbitrary field
$$I \subset I(V(I))$$
$$V(I(V)) = V \quad \forall V$$

Additional Exercises for Sec.4. Exercise 6.

30. Solving Polynomial Equations

30.1.

30.2. **Finite-Dimensional Algebras.** Gröbner basis $G = \{ g_1 \dots g_t \}$ of ideal $I \subset k[x_1 \dots x_n]$,
recall def.: Gröbner basis $G = \{ g_1 \dots g_t \} \subset I$ of ideal I, $\quad \forall f \in I$, LT(f) divisible by LT(g_i) for some i
$f \in k[x_1 \dots x_n]$ divide by G produces $f = g + r$, $g \in I$, r not divisible by any LT(I) uniqueness of r
$f \in k[x_1 \dots x_n]$ divide by G,
Recall from Ch. 1, divide $f \in k[x_1 \dots x_n]$ by G, the division algorithm yields

$$f = h_1 g_1 + \dots + h_t g_t + \overline{f}^G \qquad (2.1) \qquad (112)$$

where remainder $\overline{f}^G$ is a linear combination of monomials $x^\alpha \notin \langle \mathrm{LT}(I) \rangle$
since Gröbner basis, $f \in I$ iff $\overline{f}^G = 0$
$\forall f \in k[x_1 \dots x_n]$, we have coset $[f] = f + I = \{ f + h | h \in I \}$ s.t. $[f] = [g]$ iff $f - g \in I$

We have a 1-to-1 correspondence

$$\text{remainders} \leftrightarrow \text{cosets}$$
$$\overline{f}^G \leftrightarrow [f]$$

algebraic

$$\overline{f}^G + \overline{g}^G \leftrightarrow [f] + [g]$$
$$\overline{\overline{f}^G \cdot \overline{g}^G} \leftrightarrow [f]\cdot[g]$$

$B = \{ x^\alpha | x^\alpha \notin \langle \text{LT}(I)\rangle \}$ is a basis of A, basis monomials, standard monomials

20141023 EY's take

$\forall [f] \in A = k[x_1 \dots x_n]/I, \quad [f] = p_i b_i; \quad b_i \in B = \{ x^\alpha | x^\alpha \notin \langle \text{LT}(I)\rangle \}$

For $I = \langle G \rangle$

e.g. $G = \{ x^2 + \frac{3}{2}xy + \frac{1}{2}y^2 - \frac{3}{2}x - \frac{3}{2}y, xy^2 - x, y^3 - y \}$

$\langle \text{LT}(I)\rangle = \langle x^2, xy^2, y^3\rangle$

e.g. $B = \{1, x, y, xy, y^2\}$

$[f]\cdot[g] = [fg]$

e.g. $f = x$, $g = xy$, $[fg] = [x^2y]$

now $f = h_1 g_1 + \dots + h_t g_t + \overline{f}^G$

30.3.

30.4. **Solving Equations via Eigenvalues and Eigenvectors.**

31. Resultants

32. Computation in Local Rings

32.1. **Local Rings.**

Definition 108 (1.1).

$$k[x_1 \dots x_n]_{\langle x_1 \dots x_n\rangle} \equiv \{ \frac{f}{g} | \text{ rational functions } \frac{f}{g} \text{ of } x_1 \dots x_n \text{ with } g(p) \neq 0 \text{ at } p \}$$

main properties of $k[x_1 \dots x_n]_{\langle x_1 \dots x_n\rangle}$

Proposition 26 (1.2). *Let* $R = k[x_1 \dots x_n]_{\langle x_1 \dots x_n \rangle}$. *Then*

(a) *R subring of field of rational functions* $k(x_1 \dots x_n) \supset k[x_1 \dots x_n]$
(b) *Let* $M = \langle x_1 \dots x_n \rangle \subset R$ *(ideal generated by* $x_1 \dots X_n$ *in R)*
Then $\forall \frac{f}{g} \in R\backslash M$, $\frac{f}{g}$ *unit in R (i.e. multiplicative inverse in R)*
(c) *M maximal ideal in R*

Exercise 1. if $p = (a_1 \dots a_n) \in k^n$, $R = \{ \frac{f}{g} | f, g \in k[x_1 \dots x_n], \, g(p) \neq 0 \}$

(a) R subring of field of rational functions $k(x_1 \dots x_n)$
(b) Let M ideal generated by $x_1 - a_1 \dots x_n - a_n$ in R
Then $\forall \frac{f}{g} \in R\backslash M$, $\frac{f}{g}$ unit in R (i.e. multiplicative inverse in R)
(c) M maximal ideal in R

Proof. let $p = (a_1 \dots a_n) \in k^n$

let $g_1(p) \neq 0$, $g_2(p) \neq 0$

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} = \frac{f_1 g_2 + f_2 g_1}{g_1 g_2} \qquad g_1(p)g_2(p) \neq 0 \text{ so } \frac{f_1}{g_1} + \frac{f_2}{g_2} \in R$$
$$\frac{f_1}{g_1} \cdot \frac{f_2}{g_2} = \frac{f_1 f_2}{g_1 g_2} \qquad g_1(p)g_2(p) \neq 0 \text{ so } \frac{f_1}{g_1}\frac{f_2}{g_2} \in R$$

$f = \frac{f}{1} \in R, \quad \forall f \in k[x_1 \dots x_n]$, so $k[x_1 \dots x_n] \subset R$ □

EY : 20141027, to recap,

Let $V = k^n$

Let $p = (a_1 \dots a_n)$

single pt. $\{p\}$ is (an example of) a variety

$I(\{p\}) = \{ x_1 - a_1 \dots x_n - a_n \rangle \subset k[x_1 \dots x_n]$

$R \equiv k[x_1 \dots x_n]_{\langle x_1 - a_1 \dots x_n - a_n \rangle}$

$$R = \{ \frac{f}{g} | \text{ rational function } \frac{f}{g} \text{ of } x_1 \dots x_n, \, g(p) \neq 0, \, p = (a_1 \dots a_n) \, \}$$

Prop. 1.2. properties

(a) R subring of field of rational functions $k(x_1 \dots x_n) \qquad k(x_1 \dots x_n) \subset R$
(b) $M = \langle x_1 \dots a_1 \dots x_n - a_n \rangle \subset R$. ideal generated by $x_1 - a_1 \dots x_n - a_n$
Then $\forall \frac{f}{g} \in R\backslash M$, $\frac{f}{g}$ unit in R ($\exists$ multiplicative inverse in R)
(c) M maximal ideal in R.
in R we allow denominators that are not elements of this ideal $I(\{p\})$

Definition 109 (1.3). *local ring is a ring that has exactly 1 maximal ideal*

Proposition 27 (1.4). *ring R with proper ideal $M \subset R$ is local ring if* $\forall \frac{f}{g} \in R\backslash M$ *is unit in R*

localization Ex. 8, Ex. 9

parametrization

Exercise 2.

$$x = x(t) = \frac{-2t^2}{1+t^2}$$
$$y = y(t) = \frac{2t}{1+t^2}$$

$k[t]_{\langle t \rangle} \qquad \frac{-2t^2}{1+t^2}$ rational function of t. $1 + t^2 \neq 0$

if $k = \mathbb{C}$ or $\mathbb{R}$

Consider set of convergent power series in n variables

(113) $$(1.5) \qquad k\{x_1 \dots x_n\} = \{ \sum_{\alpha \in \mathbb{Z}^n_{\geq 0}} c_\alpha x^\alpha | c_\alpha \in k, \text{ series converges in some open } U \ni 0 \in k^n \, \}$$

Consider set $k[[x_1 \dots x_n]]$ of formal power series

(114) $$(1.6) \qquad k[[x_1 \dots x_n]] = \{ \sum_{\alpha \in \mathbb{Z}^n_{\geq 0}} c_\alpha x^\alpha | c_\alpha \in k \} \text{ series need not converge}$$

variety V

$k[x_1 \dots x_n]/\mathbf{I}(V) \qquad$ variety V

32.2. **Multiplicities and Milnor Numbers.** if I ideal in $k[x_1 \dots x_n]$, then denote $Ik[x_1 \dots x_n]_{\langle x_1 \dots x_n \rangle}$ ideal generated by I in larger ring $k[x_1 \dots x_n]_{\langle x_1 \dots x_n \rangle}$

Definition 110 (2.1). *Let I 0-dim. ideal in $k[x_1 \dots x_n]$, so $V(I)$ consists of finitely many pts. in k^n.*
Assume $(0 \dots 0) \in V(I)$
multiplicity of $(0 \dots 0) \in V(I)$ is

$$dim_k k[x_1 \dots x_n]_{\langle x_1 \dots x_n \rangle} / Ik[x_! \dots x_n]_{\langle x_1 \dots x_n \rangle}$$

generally, if $p = (a_1 \dots a_n) \in V(I)$
multiplicity of p, $m(p) = \dim k[x_1 \dots x_n]_M / Ik[x_1 \dots x_n]_M$

$$\dim k[x_1 \dots x_n]_M / Ik[x_1 \dots x_n]_M$$

localizing $k[x_1 \dots x_n]$ at maximal ideal $M = I(\{p\}) = \langle x_1 - a_1 \dots x_n - a_n \rangle$

33.

34.

35. Polytopes, Resultants, and Equations

36. Polyhedral Regions and Polynomials

36.1. **Integer Programming.** Prop. 1.12.

Suppose 2 customers A, B ship to same location
A: ship 400 kg pallet taking up $2\, m^3$ volume
B: ship 500 kg pallet taking up $3\, m^3$ volume

shipping firm trucks carry up to 3700 kg, up to $20\, m^3$

B's product more perishable, paying \$ 15 per pallet

A pays \$ 11 per pallet
How many pallets from A, B each in truck to maximize revenues?

$$\begin{aligned} 4A + 5B &\leq 37 \\ 2A + 3B &\leq 20 \\ A, B &\in \mathbb{Z}^*_{\geq 0} \end{aligned} \tag{115} \quad (1.1)$$

maximize $11A + 15B$

integer programming.
max. or min. value of some linear function

$$l(A_1 \dots A_n) = \sum_{i=1}^n c_i A_i$$

on set $(A_1 \dots A_n) \in \mathbb{Z}^n_{\geq 0}$ s.t.
3. Finally, by introducing additional variables; rewrite linear constraint inequalities as equalities. The new variables are called "slack variables"

$$a_{ij} A_j = b_i, \quad A_j \in \mathbb{Z}_{\geq 0} \tag{116} \quad (1.4)$$

introduce indeterminate z_i, $\forall$ equation in (1.4)

$$z_i^{a_{ij} A_j} = z_i^{b_i}$$

m constraints

$$\prod_{i=1}^m z_i^{a_{ij} A_j} = \prod_{i=1}^m z_i^{b_i} = \left(\prod_{i=1}^m z_i^{a_{ij}} \right)^{A_j}$$

Proposition 28 (1.6). *Let k field, define $\varphi : k[w_1 \dots w_n] \to k[z_1 \dots z_m]$ by*

$$\varphi(w_j) = \prod_{i=1}^m z_i^{a_{ij}} \qquad \forall\, j = 1 \dots n$$

and

$$\varphi(g(w_1 \dots w_n)) = g(\varphi(w_1) \dots \varphi(w_n))$$

$\forall$ *general polynomial $g \in k[w_1 \dots w_n]$*
Then $(A_1 \dots A_n)$ integer pt. in feasible region iff $\varphi : w_1^{A_1} \dots w_n^{A_n} \mapsto z_1^{b_1} \dots z_m^{b_m}$

Exercise 3.
Now

$$\begin{gathered} \varphi(w_j) = \prod_{i=1}^m z_i^{a_{ij}} \\ z_i^{a_{ij} A_j} = z_i^{b_i} \end{gathered}$$

If $(A_1 \dots A_n)$ an integer pt. in feasible region, $a_{ij} A_j = b_i$

$$z_i^{a_{ij} A_j} = z_i^{b_i} = \prod_{j=1}^n z_i^{a_{ij} A_j} \Longrightarrow \prod_{j=1}^n \prod_{i=1}^m (z_i^{a_{ij}})^{A_j} = \prod_{i=1}^m z_i^{b_i} = \prod_{j=1}^n \varphi(w_j)^{A_j} = \prod_{j=1}^n \varphi(w_j)^{A_j} = \varphi\left(\prod_{j=1}^n w_j^{A_j} \right) = \prod_{i=1}^m z_i^{b_i}$$

since $\varphi(g(w_1 \dots w_n)) = g(\varphi(w_1) \dots \varphi(w_n))$

If $\varphi : \prod_{j=1}^n w_j^{A_j} \mapsto \prod_{i=1}^m z_i^{b_i}$

$$\varphi\left(\prod_{j=1}^n w_j^{A_j} \right) = \prod_{j=1}^n (\varphi(w_j))^{A_j} = \prod_{i=1}^m z_i^{b_i} = \prod_{j=1}^n \left(\prod_{i=1}^m z_i^{a_{ij}} \right)^{A_j} \Longrightarrow \prod_{j=1}^n z_i^{a_{ij} A_j} = z_i^{b_i}$$

or $a_{ij} A_j = b_i$. So $(A_1 \dots A_n)$ integer pt.

Exercise 4.

$$\prod_{i=1}^m z_i^{b_i} = \prod_{i=1}^m \prod_{j=1}^n z_i^{a_{ij} A_j} = \prod_{j=1}^n \left(\prod_{i=1}^m z_i^{a_{ij}} \right)^{A_j} = \prod_{j=1}^n \varphi(w_j)^{A_j} = \varphi\left(\prod_{j=1}^n w_j^{A_j} \right)$$

So if given $(b_1 \dots b_m) \in \mathbb{Z}^m$, and for a given a_{ij}, $a_{ij} A_j = b_i$

For $m \leq n$, then a_{ij} is surjective, so $\exists\, A_j$ s.t. $\prod_{i=1}^m z_i^{b_i} = \varphi\left(\prod_{j=1}^n w_j^{A_j} \right)$

Proposition 29 (1.8). *Suppose $f_1 \dots f_n \in k[z_1 \dots z_m]$ given*
Fix monomial order in $k[z_1 \dots z_n, w_1 \dots w_n]$ with elimination property:
$\forall$ monomial containing 1 of z_i greater than any monomial containing only w_j

Let $\mathcal{G}$ Gröbner basis for ideal

$$I = \langle f_1 - w_1 \dots f_n - w_n \rangle \subset k[z_1 \dots z_m, w_1 \dots w_n]$$

$\forall f \in k[z_1 \dots z_m]$, let $\overline{f}^{\mathcal{G}}$ be remainder on division of f by $\mathcal{G}$
Then
(a) *polynomial f s.t. $f \in k[f_1 \dots f_n]$ iff $g = \overline{f}^{\mathcal{G}} \in k[w_1 \dots w_n]$*

(b) *if $f \in k[f_1 \dots f_n]$ as in part (a),*
$g = \overline{f}^{\mathcal{G}} \in k[w_1 \dots w_n]$
then $f = g(f_1 \dots f_n)$, giving an expression for f as polynomial in f_j
(c) *if $\forall f_i, f$ monomials, $f \in k[f_1 \dots f_n]$,*
then g also a monomial.

36.2. **Integer Programming and Combinatorics.**

37. Algebraic Coding Theory

38. The Berlekamp-Massey-Sakata Decoding Algorithm

Gröbner Bases, Martin R. Albrecht of the DTU Crypto Group

Part 6. Statistical Mechanics: Ising Model

39. Ising Model

39.1. **Definition of Ising Model.** cf. Wikipedia, "Ising model"

Consider set of lattice sites Λ, each with set of adjacent sites (e.g. **graph**) forming d-dim. lattice.
$\forall$ lattice site $k \in \Lambda$, $\exists$ discrete variable σ_k, s.t. $\sigma_k \in \{-1, 1\}$.
spin configuration $\equiv \sigma = (\sigma_k)_{k \in \Lambda}$ is an assignment of spin value to each lattice site.
i.e.
$d = 1$, consider "line" configuration: $i \in \mathbb{Z}$, $i = 0, 1, \dots L-1$. Lattice site $k \in \Lambda = \Lambda_{d=1}$. $\forall k \in \Lambda$, $\exists$ bijection to its index i, $k \mapsto i$, and $\exists \sigma_k$ i.e.

$$\sigma : \Lambda \leftrightarrow \sigma : \mathbb{Z} \to \mathbb{Z}_2$$
$$\sigma(k) \equiv \sigma_k \leftrightarrow \sigma(i) \equiv \sigma_i \mapsto \{-1, 1\}$$

spin configuration $\sigma : \Lambda \mapsto (\sigma_k)_{k \in \Lambda} \in \{-1, 1\}^{|\Lambda|}$, where $|\Lambda| = L$.
$\forall k \in \Lambda$, $\exists !$ only at most 2 edges, given, for $k \mapsto i$, $i+1, i-1$, $\forall i = 1 \dots L-2$.

$d = 2$, "rectangle" configuration. $(i, j) \in \mathbb{Z}^2$. $\begin{aligned} i &\in 0, 1, \dots L_x - 1 \\ j &\in 0, 1, \dots L_y - 1 \end{aligned}$. Lattice site $\mathbf{k} \in \Lambda = \Lambda_{d=2}$.
$\forall \mathbf{k} \in \Lambda$, $\exists$ bijection to its "grid coordinates" (i, j), $\mathbf{k} \mapsto (i, j)$, and $\exists \sigma_{\mathbf{k}}$ i.e. $\sigma_{\mathbf{k}} = \sigma_{ij} \in \{-1, 1\}$.
spin configuration $\sigma : \Lambda \mapsto (\sigma_{\mathbf{k}})_{\mathbf{k} \in \Lambda} \in \{-1, 1\}^{|\Lambda|}$, where $|\Lambda| \equiv |\Lambda_{d=2}| = L_x L_y$.

$\forall \mathbf{k} \in \Lambda$, $\exists !$ only at most 4 edges, given by $\mathbf{k} \mapsto (i, j), (i \pm 1, j), (i, j \pm 1)$, $\begin{aligned} i &= 1 \dots L_x - 2 \\ j &= 1 \dots L_y - 2 \end{aligned}$.

Note that in both cases, I haven't yet defined the boundary conditions, and leave that to be discussed thoroughly in the future (i.e. following sections).

There are $2^{|\Lambda|}$ number of configurations in any dim. d.
cf. Wikipedia, "Ising model"

39.1.1. *Interaction $J_{ij} \equiv J_{\mathbf{kl}}$, Hamiltonian (energy functional)for a configuration $H(\sigma)$.* $\forall$ 2 adjacent (lattice) sites, $i, j \equiv \mathbf{k}, \mathbf{l} \in \Lambda$, let there be an interaction $J_{ij} \equiv J_{\mathbf{kl}}$ i.e. $\begin{aligned} J &: \Lambda^2 \to \mathbb{R} \\ J &: (\mathbf{k}, \mathbf{l}) \mapsto J_{\mathbf{kl}} \end{aligned}$.
Adjacent means $\exists$ edge $\mathbf{k} \mapsto \mathbf{l}$ (the mapping is the edge)
Suppose $\forall$ site $j \equiv \mathbf{l} \in \Lambda$, $\exists$ external magnetic field $h_j \equiv h_{\mathbf{l}}$ interacting with it.
Given (site) configuration $\sigma : \Lambda \mapsto (\sigma_{\mathbf{k}})_{\mathbf{k} \in \Lambda} \in \{-1, 1\}^{|\Lambda|}$.

$$H(\sigma) = -\sum_{\langle ij \rangle} J_{ij} \sigma_i \sigma_j - \mu \sum_j h_j \sigma_j \equiv H(\sigma(\Lambda)) = -\sum_{\langle \mathbf{kl} \rangle} J_{\mathbf{kl}} \sigma_{\mathbf{k}} \sigma_{\mathbf{l}} - \mu \sum_{\mathbf{k} \in \Lambda} h_{\mathbf{k}} \sigma_{\mathbf{k}} \quad (117)$$

where $\sum_{\langle \mathbf{kl} \rangle}$ is overall pairs of adjacent spins (every pair is counted once),
$\langle \mathbf{k}, \mathbf{l} \rangle \equiv$ sites $\mathbf{k}, \mathbf{l}$ are nearest neighbors.
Note sign in 2nd. term, $-\mu \sum_{\mathbf{k}} h_{\mathbf{k}} \sigma_{\mathbf{k}}$ should be positive because of electron's magnetic moment is antiparallel to its spin, but negative term used conventionally.
Nothing was said about boundary conditions, I propose that it can be either fixed in the summation or by setting $J_{\mathbf{kl}} = 0$.

$\forall \mathbf{k} \in \Lambda$, let $\begin{aligned} \mathbf{y} &: \Lambda \to E \\ \mathbf{y} &: \mathbf{k} \mapsto \{\langle \mathbf{k}, \mathbf{l} \rangle_{\mathbf{l}}\} \end{aligned}$, with $\{\langle \mathbf{k}, \mathbf{l} \rangle\}_{\mathbf{l}}$ be set of all edges from $\mathbf{k}$
Then clearly $\sum_{\langle \mathbf{kl} \rangle} = \frac{1}{2} \sum_{\mathbf{k} \in \Lambda} \sum_{\{\langle \mathbf{kl} \rangle\}_{\mathbf{l}}}$.
Taking into account only interaction between adjoining dipoles, on a square lattice:

$$E(\sigma) = -J \sum_{k,l=0}^{L-1} (\sigma_{kl} \sigma_{k,l+1} + \sigma_{kl} \sigma_{k+1,l})$$

cf. Landau and Lifshitz [23]
EY : 20171223 Things to check from Hjorth-Jensen (2015) [24]:
2-dim. Ising model, with $\mathcal{B} \equiv h_j = 0$, undergoes phase transition of 2nd. order: meaning below given critical temperature T_C, there's spontaneous magnetization with $\langle \mathcal{M} \rangle \equiv \langle \mathbf{M} \rangle \neq 0$. $\langle \mathbf{B} \rangle \to 0$ at T_C with *infinite* slope, a behavior called *critical phenomena*. Critical phenomenon normally marked by 1 or more thermodynamical variables which is 0 above a critical point. In this case, $\langle \mathbf{B} \rangle \neq 0$, such a parameter normally called *order parameter*.
Critical phenomena; we still don't have a satisfactory understanding of system's properties close to the critical point, even for simplest 3-dim. systems. Even mean-field models can predict wrong physics; mean-field theory results in a 2nd.-order phase transition for 1-dim. Ising model, wherea 1-dim. Ising model doesn't predict any spontaneous magnetization at any finite temperature T.
e.g. Consider 1-dim. N-spin system. Assume periodic boundary conditions. Consider state of all spins up, with total energy $-NJ$ and magnetization N. Flip half of spins (e.g. all spins of index $i > N/2$) so 1st half of spins point upwards and last half points downwards. Energy is $-NJ + 4J$, net magnetization 0. This is an example of a possible disordered state with net magnetization 0. Change in energy is too small to stabilize disordered state (to $-NJ$).

Definition 111 (configuration probability). ***configuration probability*** *$P_\beta(\sigma)$ given by Boltzmann distribution:*

$$P_\beta(\sigma) = \frac{\exp(-\beta H(\sigma))}{Z_\beta} = \text{ prob. of configuration } \sigma \equiv \sigma(\Lambda) \equiv (\sigma_{\mathbf{k}})_{\mathbf{k} \in \Lambda} \quad (118)$$

with the partition function as normalization constant Z_β:

$$Z_\beta = \sum_\sigma \exp -\beta H(\sigma) \quad (119)$$

cf. pp. 504 Sec. 151 Phase transitions of the second kind in a 2-dim. lattice, Landau and Lifshitz [23]

$$Z = 2^N (1-x^2)^{-N} \prod_{p,q=0}^{L-1} \left[(1+x^2)^2 - 2x(1-x^2)\left(\cos\frac{2\pi p}{L} + \cos\frac{2\pi q}{L} \right) \right]^{1/2} \tag{120}$$

cf. (151.11) of Landau and Lifshitz [23], where $x = \tanh\theta$, $\theta = J/T \equiv J/\tau = \beta J$.

$$\begin{aligned} \Phi \equiv F = -\tau \ln Z = \\ = -\tau N \ln 2 + \tau N \ln (1-x^2) - \frac{\tau}{2} \sum_{p,q=0}^{L} \ln \left[(1+x^2)^2 - 2x(1-x^2)\left(\cos\frac{2\pi p}{L} + \cos\frac{2\pi q}{L} \right) \right] \end{aligned} \tag{121}$$

Let $\omega_1 = \frac{2\pi p}{L}$ with $p\to 0$ as $L\to\infty$ so $\frac{Ld\omega_1}{2\pi} = dp$ and using $L^2 = N$.
$\omega_2 = \frac{2\pi q}{L}$ with $q\to 0$ as $L\to\infty$ $\frac{Ld\omega_2}{2\pi} = dq$

$$\Phi = -\tau N \ln 2 + \tau N \ln (1-x^2) - \frac{N\tau}{2(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} d\omega_1 d\omega_2 \ln\left[(1-x^2) - 2x(1-x^2)(\cos\omega_1 + \cos\omega_2) \right]$$

$F\equiv\Phi$ has singularity when $(1-x^2) - 2x(1-x^2)(\cos\omega_1+\cos\omega_2)$ in $\ln\left[(1-x^2)-2x(1-x^2)(\cos\omega_1+\cos\omega_2)\right]$.
$(1-x^2)-2x(1-x^2)(\cos\omega_1+\cos\omega_2)$ minimized when $\cos\omega_1 = \cos\omega_2 = 1$ (since $-1<x<1$)

$$\Longrightarrow (1+x^2)^2 - 4x(1-x^2) = 1 + 2x^2 + x^4 - 4x + 4x^3 = (x^2+2x-1)^2 = 0 \Longrightarrow x = \frac{-2\pm\sqrt{4-4(-1)}}{2} = -1+\sqrt{2}$$

$$x = \tanh\theta = \frac{e^\theta - e^{-\theta}}{e^\theta + e^{-\theta}} = \sqrt{2}-1 \text{ or } \qquad \begin{aligned} e^\theta - e^{-\theta} &= \sqrt{2}e^\theta + \sqrt{2}e^{-\theta} - e^\theta - e^{-\theta} \text{ so} \\ (2-\sqrt{2})e^\theta &= \sqrt{2}e^{-\theta} \\ e^{2\theta} &= \frac{\sqrt{2}}{2-\sqrt{2}}\left(\frac{2+\sqrt{2}}{2+\sqrt{2}}\right) \text{ or} \\ 2\theta &= \ln(1+\sqrt{2}) \end{aligned}$$

$$\frac{J}{T_c} = \frac{1}{2}\ln(1+\sqrt{2}) \text{ or}$$

$$\boxed{\tau_c = \frac{2J}{\ln(1+\sqrt{2})}} \tag{122}$$

so that $\tau_C \equiv T_C$ is where phase transition occurs.
Let $t := \tau - \tau_c$. $\theta = \frac{J}{\tau} = \frac{J}{t+\tau_C}$
Expand about minimum
EY:20171230 do this explicitly

$$\begin{gathered} \int_0^{2\pi}\int_0^{2\pi} d\omega_1 d\omega_2 \ln[c_1 t^2 + c_2(\omega_1^2+\omega_2^2)] \\ F\equiv\Phi \simeq a + \frac{1}{2}b(\tau-\tau_c)^2\ln|\tau-\tau_c| \\ C = \frac{\partial^2 F}{\partial\tau} \simeq -b\tau_c \ln|\tau-\tau_c| \end{gathered}$$

with C being heat capacity.

Order parameter $\langle M\rangle \equiv \eta = \text{constant}(\tau_c-\tau)^{1/8} = \begin{cases} 0 & \text{if } \tau > \tau_c \\ \text{constant } (\tau_c-\tau)^{1/8} & \text{if } \tau<\tau_c \end{cases}$

cf. pp. 505 Sec. 151 Phase transitions of the second kind in a 2-dim. lattice, Landau and Lifshitz [23], L.Onsager 1947.

39.2. **An actual calculation of a small number of spins with Ising model.** Sec. 3.7 "An actual calculation" on pp. 76 of Newman and Barkema (1999) [25] goes through a simple actual Monte Carlo calculation as a test case check so to compare this exact calculation/solution to the simulation, as a test of whether the simulation/program is correct. This is done in Sec. 1.3 of Newman and Barkema (1999) [25].

However, none of these promised simple calculations were shown explicitly in Newman and Barkema (1999) [25]. I will forego this simple case.

39.3. **Explicit calculation showing stencil operation on each spin on a periodic lattice grid.** Consider

$$\begin{aligned} H(\sigma) = -\sum_{\langle\mathbf{kl}\rangle} J\sigma_{\mathbf{k}}\sigma_{\mathbf{l}} = -J\sum_{i=0}^{L_x-1}\sum_{j=0}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma_{ij+1}) = \\ = \frac{-J}{2}\left(\sum_{i=0}^{L_x-1}\sum_{j=0}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma_{ij+1}) + \sum_{i=1}^{L_x}\sum_{j=0}^{L_y-1}\sigma_{i-1j}(\sigma_{ij}+\sigma_{i-1j+1}) \right) = \\ = \frac{-J}{2}\left(\sum_{i=0}^{L_x-1}\sum_{j=0}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma_{ij+1}) + \sum_{i=1}^{L_x}\sum_{j=0}^{L_y-1}\sigma_{i-1j}\sigma_{ij} + \sum_{i=0}^{L_x-1}\sum_{j=1}^{L_y}\sigma_{ij-1}\sigma_{ij} \right) \end{aligned}$$

Now for each of these terms,

$$\sum_{i=1}^{L_x}\sum_{j=0}^{L_y-1}\sigma_{i-1j}\sigma_{ij} = \sum_{i=1}^{L_x}\left(\sum_{j=1}^{L_y-1}\sigma_{i-1j}\sigma_{ij} + \sigma_{i-10}\sigma_{i0}\right) = \sum_{i=1}^{L_x-1}\left(\sum_{j=1}^{L_y-1}\sigma_{i-1j}\sigma_{ij}+\sigma_{i-10}\sigma_{i0}\right) + \left(\sum_{j=1}^{L_y-1}\sigma_{L_x-1j}\sigma_{L_xj}\right) + \sigma_{L_x-10}\sigma_{L_x0}$$

$$\sum_{i=0}^{L_x-1}\sum_{j=1}^{L_y}\sigma_{ij-1}\sigma_{ij} = \sum_{j=1}^{L_y-1}\left(\sum_{i=1}^{L_x-1}\sigma_{ij-1}\sigma_{ij} + \sigma_{0j-1}\sigma_{0j}\right) + \sum_{i=1}^{L_x-1}\sigma_{iL_y-1}\sigma_{iL_y} + \sigma_{0L_y-1}\sigma_{0L_y}$$

$$\begin{aligned} \sum_{i=0}^{L_x-1}\sum_{j=0}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma_{ij+1}) = \sum_{i=0}^{L_x-1}\left(\sum_{j=1}^{L_y}\sigma_{ij}(\sigma_{i+1j}+\sigma_{ij+1}) + \sigma_{i0}(\sigma_{i+10}+\sigma_{i1})\right) = \\ \sum_{i=1}^{L_x-1}\left(\sum_{j=1}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma_{ij+1}) + \sigma_{i0}(\sigma_{i+10}+\sigma_{i1})\right) + \sum_{j=1}^{L_y-1}\sigma_{0j}(\sigma_{1j}+\sigma_{0j+1}) + \sigma_{00}(\sigma_{10}+\sigma_{01}) \end{aligned}$$

Apply periodic boundary conditions. Adding up all the terms above, clearly we obtain 1 term which shows the stencil operation for spins on the "interior" of the grid:

$$\sum_{i=1}^{L_x-1}\sum_{j=1}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma_{ij+1}+\sigma_{i-1j}+\sigma_{ij-1})$$

and if we apply *periodic* boundary conditions, neatly, we'll see all the lattice sites at the boundary also will have this stencil operation:

$$\sum_{i=1}^{L_x-1}\sigma_{i0}(\sigma_{i+10}+\sigma_{i1})+\sum_{j=1}^{L_y-1}\sigma_{0j}(\sigma_{1j}+\sigma_{0j+1})+\sigma_{00}(\sigma_{10}+\sigma_{01})+\left(\sum_{i=1}^{L_x-1}\sigma_{iL_y-1}\sigma_{i0}\right)+\sigma_{0L_y-1}\sigma_{00}+\sum_{j=1}^{L_y-1}\sigma_{0j-1}\sigma_{0j}+$$
$$+\sum_{j=1}^{L_y-1}\sigma_{L_x-1j}\sigma_{0j}+\sigma_{L_x-10}\sigma_{00}+\sum_{i=1}^{L_x-1}\sigma_{i-10}\sigma_{i0}$$

Now, we can obtain the following for Hamiltonian, given spin configuration σ with a lattice grid obeying periodic conditions:

$$H(\sigma)=-\frac{J}{2}\sum_{i=0}^{L_x-1}\sum_{j=0}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma i-1j+\sigma_{ij+1}+\sigma_{ij-1})=$$

$$=\frac{-J}{2}\left[\sum_{i=0}^{L_x-1}\left(\sum_{\substack{j=0\\ j\neq j'}}^{L_y-1}\sigma_{ij}(\sigma_{i+1j}+\sigma_{i-1j}+\sigma_{ij+1}+\sigma_{ij-1})+\sigma_{ij'}(\sigma_{i+1j'}+\sigma_{i-1j'}+\sigma_{ij'+1}+\sigma_{ij'-1})\right)+\right.$$
$$\left.\sum_{\substack{j=0\\ j\neq j'}}^{L_y-1}\sigma_{i'j}(\sigma_{i'+1j}+\sigma_{i'-1j}+\sigma_{i'j+1}+\sigma_{i'j-1})+\sigma_{i'j'}(\sigma_{i'+1j'}+\sigma_{i'-1j'}+\sigma_{i'j'+1}+\sigma_{i'j'-1})\right] \quad (123)$$

Consider a psin flip of $\sigma_{i'j'}$. Contribution to ΔH at stencil operation on $\sigma_{i'j'}$, at $(i'j')\in\Lambda$, is

$$\frac{-J}{2}(-\sigma_{i'j'}-\sigma_{i'j'})(\sigma_{i'+1j'}+\sigma_{i'-1j'}+\sigma_{i'j'+1}+\sigma_{i'j'-1})=J\sigma_{i'j'}(\sigma_{i'+1j'}+\sigma_{i'-1j'}+\sigma_{i'j'+1}+\sigma_{i'j'-1})$$

Consider $\sigma_{i'j'}\sigma_{i'+1j'}$. Clearly, term $\sigma_{i-1j'}\sigma_{ij'}$ with $i=i'+1$ only occurs once more in the summation. Thus, we can definitely conclude that for $\Delta H\equiv\Delta H(\Delta\sigma_{i'j'})$ due to a single spin-flip is

$$\Delta H(\Delta\sigma_{i'j'})=2J\sigma_{i'j'}(\sigma_{i'+1j'}+\sigma_{i'-1j'}+\sigma_{i'j'+1}+\sigma_{i'j'-1}) \quad (124)$$

https://www.colorado.edu/physics/phys7240/phys7240_fa12/notes/Week3.pdf Victor Gurarie, Advanced Statistical Mechanics, Fall 2012 Exact solution by transfer matrices for 2-dim. Ising model.

Part 7. Conformal Field Theory; Virasoro Algebra

cf. Schottenloher (2008) [22]

40. Conformal Transformations

40.1. Semi-Riemannian manifolds (review and (key) examples). cf. pp. 7, Ch. 1 "Conformal Transformations and Conformal Killing Fields." Schottenloher (2008) [22]

Semi-Riemannian manifold is a pair (M,g) s.t.
smooth manifold M, $\dim M=n$,
smooth tensor field g s.t. $g: a\in M\mapsto\Omega^2(T_aM)$, i.e. $\forall\, a\in M$, g assigns a a nonnegative and symmetric bilinear form on tangent space T_aM.

In local coordinates, $x^1\dots x^n$ of manifold M,
given chart $\phi:U\to V$, open subset $U\subseteq M$, open subset $V\subseteq\mathbb{R}^n$,

$$\phi(a)=(x^1(a)\dots x^n(a)),\, a\in M$$

[2] https://doc.sagemath.org/html/en/reference/manifolds/sage/manifolds/differentiable/pseudo_riemannian.html

Bilinear form g_a on T_aM, written

$$g_a(X,Y)=g_{\mu\nu}(a)X^\mu Y^\nu$$

Tangent vectors $X=X^\mu\partial_\mu$, $Y=Y^\nu\partial_\nu\in T_aM$ basis $\partial_\mu:=\frac{\partial}{\partial x^\mu}$, $\mu=1\dots n$ of tangent space T_aM, induced by chart ϕ.
By assumption, matrix

$$g_{\mu\nu}(a)$$

Nondegenerate and symmetric, $\forall\, a\in U$, i.e.

$$\det(g_{\mu\nu}(a))\neq 0,\qquad (g_{\mu\nu}(a))^T=(g_{\mu\nu}(a))$$

Differentiating of g_a implies matrix $g_{\mu\nu}(a)$ depends differentiably on a.
That means that in its dependence on local coordinates x^j, coefficient $g_{\mu\nu}=g_{\mu\nu}(x)$ are smooth functions.
In general, $g_{\mu\nu}X^\mu X^\nu>0$ doesn't hold $\forall\, X\neq 0$, i.e. $g_{\mu\nu}(a)$ not required to be positive-definite.
2 important subcases: [2]
Riemannian manifold: metric g positive definite, signature $n=\dim M$.
Lorentz manifold specified as semi-Riemannian manifold with $(p,q)=(n-1,1)$ or $(p,q)=(1,n-1)$.
Metric g has signature $n-2$ (positive convention) or $2-n$ (negative convention).

40.1.1. *Examples (of Riemannian manifolds for Conformal Field Theory).* $\mathbb{R}^{p,q}=(\mathbb{R}^{p,q},g^{p,q})$, $p,q\in\mathbb{N}$, where

$$g^{p,q}(X,Y):=\sum_{i=1}^{p}X^iY^i-\sum_{i=p+1}^{p+q}X^iY^i$$

Hence

$$(g_{\mu\nu})=\begin{pmatrix}1_p & \\ & -1_q\end{pmatrix}=\mathrm{diag}(1\dots 1,-1,\dots-1)$$

$\mathbb{R}^{1,3}=\mathbb{R}^{3,1}$, usual Minkowski space.
$\mathbb{R}^{1,1}$, 2 -dim. Minkowski space (Minkowski plane).
$\mathbb{R}^{2,0}$, Euclidean plane.
$\mathbb{S}^2\subset\mathbb{R}^{3,0}$, compactification of $\mathbb{R}^{2,0}$, structure of Riemannian manifold on 2-sphere $\mathbb{S}^2$ induced by inclusion in $\mathbb{R}^{2,0}$
$\mathbb{S}\times\mathbb{S}\subset\mathbb{R}^{2,2}$, compactification of $\mathbb{R}^{1,1}$. More precisely,
$\mathbb{S}\times\mathbb{S}\subset\mathbb{R}^{2,0}\times\mathbb{R}^{0,2}\simeq\mathbb{R}^{2,2}$ where structure of semi-Riemannian manifold on $\mathbb{S}\times\mathbb{S}$ induced by inclusion into $\mathbb{R}^{2,2}$.
$\mathbb{S}^p\times\mathbb{S}^q\subset\mathbb{R}^{p+1,0}\times\mathbb{R}^{0,q+1}\simeq\mathbb{R}^{p+1,q+1}$ with p-sphere $\mathbb{S}^p=\{X\in\mathbb{R}^{p+1}:g^{p+1,0}(X,X)=1\}\subset\mathbb{R}^{p+1,0}$, q-sphere $\mathbb{S}^q\subset\mathbb{R}^{0,q+1}$ yields a compactification of $\mathbb{R}^{p,q}$ for $p,q\geq 1$
Compact semi-Riemannian manifold denoted by $\mathbb{S}^{p,q}$, for $p,q\geq 0$.
Quadrics $N^{p,q}$ (of Sec. 2.1) are locally isomorphic to $\mathbb{S}^{p,q}$ from point of view of conformal geometry.
For the "*negative convention*":

$$g^{p,q}(X,Y)=-\sum_{i=0}^{p-1}X^iY^i+\sum_{i=p}^{p+q}X^iY^i$$

$$(g_{\mu\nu})=\begin{pmatrix}-1_p & \\ & 1_q\end{pmatrix}=\mathrm{diag}(-1,\dots-1,1\dots 1)$$

$\mathbb{R}^{1,3}$, Minkowski space.
$\mathbb{R}^{1,1}$, 2 -dim. Minkowski space.
$\mathbb{R}^{0,2}$, Euclidean plane.
$\mathbb{S}^2\subset\mathbb{R}^{0,3}$, compactification of $\mathbb{R}^{0,2}$
$\mathbb{S}\times\mathbb{S}\subset\mathbb{R}^{0,2}\times\mathbb{R}^{2,0}\simeq\mathbb{R}^{2,2}$
$\mathbb{S}^p\times\mathbb{S}^q\subset\mathbb{R}^{0,p+1}\times\mathbb{R}^{q+1,0}\simeq\mathbb{R}^{p+1,q+1}$ with p-sphere $\mathbb{S}^p=\{X\in\mathbb{R}^{p+1}:g^{0,p+1}(X,X)=1\}\subset\mathbb{R}^{0,p+1}$, q-sphere $\mathbb{S}^q\subset\mathbb{R}^{q+1,0}$ yields a compactification of $\mathbb{R}^{p,q}$

Definition 112 (Conformal transformation or conformal map)**.** *Let 2 semi-Riemannian manifolds(M,g), (M',g'), $\dim M = \dim M'$, let open $U\subset M$, open $V \subset M'$.*

conformal transformation *or* ***conformal map*** *is a smooth $\varphi : U \to V$ of maximal rank, if $\exists$ smooth $\Omega : U \to \mathbb{R}^+$ s.t.*

$$\varphi^* g' = \Omega^2 g \tag{125}$$

where $\varphi * g'(X,Y) := g'(T\varphi(X), T\varphi(Y))$ and $T\varphi : TU \to TV$ denote tangent map (derivative) of φ.

$\Omega \equiv$ conformal factor *of φ.*

Locally, $y^i = \varphi^i(x)$,

$$\frac{\partial \varphi^i}{\partial x^j} = \frac{\partial y^i}{\partial x^j}$$

Then

$$X = X^k \frac{\partial}{\partial x^k} = X^k \frac{\partial y^i}{\partial x^k}\frac{\partial}{\partial y^i} = X^k \frac{\partial \varphi^i}{\partial x^k}\frac{\partial}{\partial y^k} \in TM$$

and so

$$\begin{gathered}\varphi^* g'(X,Y) = g'(T\varphi(X), T\varphi(Y)) = (g')_{ij} X^k \frac{\partial y^i}{\partial x^k} Y^l \frac{\partial y^j}{\partial x^l} = (g')_{ij} X^k \frac{\partial \varphi^i}{\partial x^k} Y^l \frac{\partial y^j}{\partial x^l} \\ \Longrightarrow (\varphi^* g')_{kl} = (g')_{ij} \frac{\partial y^i}{\partial x^k}\frac{\partial y^j}{\partial x^l} \\ \Longrightarrow (\varphi^* g')_{kl} = (g')_{ij} \frac{\partial \varphi^i}{\partial x^k}\frac{\partial \varphi^j}{\partial x^l} = \Omega^2 g_{kl}\end{gathered}$$

Definition 113. ***extension*** *of G by group A is (given by) an* exact sequence *of group homomorphisms.*

$$1 \longrightarrow A \xrightarrow{i} E \xrightarrow{\pi} G \longrightarrow 1$$

cf. Def. 3.1 of Schottenloher (2008) [22].

Recall that an exact sequence, if
$$\begin{aligned} \text{im}(1\to A) &= \ker(i) \\ \text{im}(i) &= \ker(\pi) \\ \text{im}(\pi) &= \ker(G \to 1)\end{aligned}$$

By Thm., $1 \to A \xrightarrow{i} E$ exact so i injective.
$E \xrightarrow{\pi} G \to 1$ exact so π surjective.

Extension is called **central** if A abelian and image im$i$ is in center of E, i.e. $a \in A, b\in E \Longrightarrow i(a)b = bi(a)$.

40.1.2. *Examples of extensions of G, and central extensions of G (which has a particular E).*

- e.g. central extension has form

$$1 \longrightarrow A \xrightarrow{i} A\times G \xrightarrow{\text{pr}_2} G \longrightarrow 1$$

where
$$\begin{aligned} i : A &\to A\times G \\ a &\mapsto (a,1)\end{aligned}$$

$$\begin{aligned} i(a)(a',g) &= (a,1)(a',g) = (aa',g) = \\ &= (a'a, g\cdot 1) = (a',g)(a,1) = (a',g)i(a)\end{aligned}$$

Notice that what the *exactness* property of an exact sequence does:

$$\text{pr}_2 i(a) = \text{pr}_2(a,1) = 1$$

- e.g. of a nontrivial central extension is exact sequence

$$1 \longrightarrow \mathbb{Z}/k\mathbb{Z} \longrightarrow E\times U(1) \xrightarrow{\pi} U(1) \longrightarrow 1 \tag{126}$$

with $\pi(z) = z^k \quad \forall k \in \mathbb{N}, k\geq 2$, since $E = U(1)$ and $\mathbb{Z}/k\mathbb{Z}$ are not isomorphic.

Also, homomorphism $\tau : U(1) \to E$ with $\pi\circ\tau = 1_{U(1)}$, doesn't exist, since there's no global kth root.

EY : 20170926 It's that in integer division of the argument in a complex number $z\in U(1)$, and exponent multiplication by k, you go from 1 to many and many to 1, depending upon the "branch" you're mapping to for complex numbers.

For $[n] \in \mathbb{Z}/k\mathbb{Z}$,

$$[n] \xmapsto{i} \exp\left(\frac{[n]}{k} 2\pi i\right)$$

and so

$$\ker\pi = \{z | \pi(z) = 1\} \text{ so that } \ker\pi = \{ z = \exp\left(\frac{i2\pi n}{k}\right)\}$$

- e.g. *Semidirect products.*

group G acting on another group H, by homomorphism

$$\tau : G \to \text{Aut}(H)$$

Definition 114 (semi-direct product)**.** ***semidirect product*** *group $G\ltimes H$ is set $H\times G$, with multiplication*

$$(x,g)\cdot(x',g') := (x\tau(g)(x'), gg') \qquad \forall (x,g),(x',g') \in H\times G$$

$$1 \longrightarrow H \xrightarrow{i} G\ltimes H \xrightarrow{\pi} G \longrightarrow 1 \tag{127}$$

with

$$\begin{aligned} i &: H \to G\ltimes H \\ i(x) &= (x,1)\end{aligned} \tag{128}$$

i group homomorphism, since

$$i(x_1x_2) = (x_1x_2,1) = (x_1\tau(1)x_2,1) = (x_1,1)\cdot(x_2,1) = i(x_1)i(x_2)$$

$$\begin{aligned} \pi &: G\ltimes H \to G \\ \pi(x,g) &= g\end{aligned} \tag{129}$$

cf. http://sierra.nmsu.edu/morandi/oldwebpages/math683fall2002/GroupExtensions.pdf

Observe that

$$\pi i(x) = \pi(x,1) = 1 \text{ so } \ker\pi = \text{im} i$$

Definition 115 (Semi-direct product (2); with direct product)**.** ***direct product*** *$G = HK$ if H,K subgroups of group G, s.t.*

- *H and K are normal in G ($gkg^{-1} \in K \quad \forall g\in G, \forall k \in K$)*
- *$H\cap K = \{1\}$*
- *$HK = G$.*

semi-direct product*. Relax the 1st condition (of direct products) so H still normal in G, but K need not be.*

- *H normal in G ($ghg^{-1}\in H, \forall g, \forall h\in H$)*
- *$H\cap K = \{1\}$*
- *$HK = G$*

Connection between Def. 114 and Def. 115 for the *semidirect product*: Consider $\tau : G \to \mathrm{Aut}(H)$.

Consider $G \ltimes H$ - what is the identity $1_{G\ltimes H} \equiv (1_H, 1_G)$ of this group?

$$(x,g)\cdot(1_H,1_G) = (x\tau(g)1_H, g1_G) = (x\tau(g)1_H, g) \Longrightarrow 1_H = \tau(g^{-1})1,\ 1_G = 1$$

and so the inverse, $\forall (x,g)\in G\ltimes H$, $(x,g)^{-1} \equiv ((x^{-1}),(g^{-1}))$,

$$(x,g)(x,g)^{-1} = (x\tau(g)(x^{-1}), g(g^{-1})) = (x\tau(g)(x^{-1}),1) \text{ (if } (g^{-1}) = g^{-1})$$

Moving along,

$$\begin{aligned} & x\tau(g)(x^{-1}) = \tau(g^{-1})1 \\ \Longrightarrow & (x^{-1}) = \tau(g^{-1})x^{-1}\tau(g^{-1})1 \end{aligned}$$

Checking out the H being a normal subgroup of $G\ltimes H$ condition, i.e. $H \lhd G$,

$$\begin{aligned} (x,g)(h,1)(\tau(g^{-1})x^{-1}\tau(g^{-1}),g^{-1}) &= (x\tau(g)h,g)(\tau(g^{-1})x^{-1}\tau(g^{-1}),g^{-1}) = \\ &= (x\tau(g)h\tau(g)\tau(g^{-1})x^{-1}\tau(g^{-1}),1) = (x\tau(g)hx^{-1}\tau(g^{-1}),1) \end{aligned}$$

$\Longrightarrow H$ normal subgroup of $G\ltimes H \equiv H \lhd (G\ltimes H)$.

Notes on Semidirect products

- extension

$$1 \longrightarrow SL(n,\mathbb{R}) \xrightarrow{i} GL(n,\mathbb{R}) \xrightarrow{\det} \mathbb{R}^* \longrightarrow 1 \tag{130}$$

with

$GL(n,\mathbb{R}) \equiv Gl_n(\mathbb{R}) = \{A | A \in \mathrm{Mat}_{\mathbb{R}}(n,n); \det A \neq 0\}$
$\det : GL(n,\mathbb{R}) \to \mathbb{R}^* \equiv \mathbb{R}\backslash\{0\}$, det surjective homomorphism
$SL(n,\mathbb{R}) \equiv Sl_n(\mathbb{R}) = \{A | A \in \mathrm{Mat}_{\mathbb{R}}(n,n);\ \det A = 1\}$

Note that $\ker(\det) = SL(n,\mathbb{R})$.

Now

$$\mathbb{R}^* \simeq \{a1_n | a \in \mathbb{R}^*\}$$

and $\det(a1_n) = a^n$.

If n odd, and $\det(a1_n) = a^n = 1$, then $a=1$. If n even, $a = \{-1,1\}$.

By the second definition of a semi-direct product, Def. 115, it's required that $SL(n,\mathbb{R})\cap\mathbb{R}^* = 1$ (i.e. the intersection is only the identity). This will only be the case if n odd.

cf. `http://sierra.nmsu.edu/morandi/oldwebpages/math683fall2002/GroupExtensions.pdf`

Part 8. Quantum Mechanics

41. The Wave function and the Schrödinger Equation, its probability interpretation, some postulates

cf. Ch. 2 "The Wave Function and the Schrödinger Equation" in **Quantum Mechanics** by Franz Schwabl (2007) [21].

From experimental considerations (Sec. 1.2.2, Schwabl (2007) [21]), with electron diffraction, electrons, e^-, have wavelike properties; let this wave be $\psi(\mathbf{x},t)$.

For free e^- of momentum $\mathbf{p}$, energy $E = \frac{\mathbf{p}^2}{2m}$, in accordance with diffraction experiments, consider as free plane waves

$$\psi(\mathbf{x},t) = C\exp\left(i(\mathbf{k}\cdot\mathbf{x}-\omega t)\right), \qquad \omega = E/\hbar = E,\ \mathbf{k} = \mathbf{p}/hbar = \mathbf{p}$$

with $\hbar = 1$

Hypothesis: wave function $\psi(\mathbf{x},t)$ gives probability distribution

$$\rho(\mathbf{x},t) = |\psi(\mathbf{x},t)|^2$$

$\rho(\mathbf{x},t)d^3x$ = probability of finding e^- at location $\mathbf{x}$ in volume element d^3x.

e.g. e^- waves $\psi_1(\mathbf{x},t)$, $\psi_2(\mathbf{x},t)$

If both slits open, superposition of wave functions $\psi_1(\mathbf{x},t)+\psi_2(\mathbf{x},t)$

Note $|\psi_1(\mathbf{x},t)+\psi_2(\mathbf{x},t)|^2 \neq |\psi_1(\mathbf{x},t)|^2 + |\psi_2(\mathbf{x},t)|^2$ if there are no interference terms.

Important remarks:

(i) Single e^- not smeared out. $\rho(\mathbf{x},t)$ is **not** the charge distribution of e^-, but is the probability density for measuring particle at position $\mathbf{x}$ at time t.
(ii) Prob. distribution doesn't occur by interference of many simultaneously incoming e^-, but one obtains same interference pattern if each e^- enters separately, i.e. even for very low intensity source. Thus, wave function applies to every electron and describes state of single e^-.

cf. 2.2 "The Schrödinger Equation for Free Particles" in **Quantum Mechanics** by Franz Schwabl (2007) [21].

(i) 1st. order DE (differential equation); (ii) linear in ψ for linear superposition (iii) "homogeneous" $\int d^3x|\psi(\mathbf{x},t)|^2 = 1$, (iv) plane waves

$$\psi(\mathbf{x},t) = C\exp\left[i(\mathbf{p}\cdot\mathbf{x} - \frac{p^2}{2m}t)/\hbar\right] \qquad \text{plane waves}$$

Should be solutions of the equations.

From postulates (i-iv),

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t) = \frac{-\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t)$$

Time-dependent Schrödinger equation for free particles.

$$\int_{-\infty}^{\infty} d^3k e^{i\mathbf{k}\cdot\mathbf{x}}e^{-k^2\alpha^2} = \prod_{j=x}^{z}\int_{-\infty}^{\infty} dk_j e^{ik_jx_j}e^{-k_j^2\alpha^2} = \prod_{j=x}^{z}\left(\sqrt{\frac{\pi}{\alpha^2}}\exp\left(\frac{-x_j^2}{4\alpha^2}\right)\right) = \left(\frac{\sqrt{\pi}}{\alpha}\right)^3\exp\left(\frac{-x^2}{4\alpha^2}\right)$$

Part 9. Algebraic Topology

cf. Bredon (1997) [26]

42. Simplicial Complexes

cf. pp. 245, from Sec. 21 Simplicial Complexes of Ch. 4 Homology Theory in Bredon (1997) [26]

$\mathbf{v}_0,\dots\mathbf{v}_n \in \mathbb{R}^\infty$, "affinely independent" if they span an affine n-plane, i.e.

$$\text{if } \left(\sum_{i=0}^n \lambda_i\mathbf{v}_i = 0,\ \sum_{i=0}^n \lambda_i = 0\right),\ \text{then } \Longrightarrow \forall\,\lambda_i = 0$$

If not, then, e.g. $\lambda_0\neq 0$, assume $\lambda_0 = -1$, and solve the equations to get

$$\begin{aligned} \mathbf{v}_0 &= \sum_{i=1}^n \lambda_i\mathbf{v}_i \\ \sum_{i=1}^n \lambda_i &= 1 \end{aligned}$$

i.e. $\mathbf{v}_0$ is in affine space spanned by $\mathbf{v}_1\dots\mathbf{v}_n$.

If $\mathbf{v}_0,\dots\mathbf{v}_n$ affinely independent, then

$$\sigma = (\mathbf{v}_0,\dots\mathbf{v}_n) = \{\sum_{i=0}^n\lambda_i\mathbf{v}_i | \sum_{i=0}^n\lambda_i = 1,\ \lambda_i\geq 0\} \tag{131}$$

is "affine simplex" spanned by $\mathbf{v}_i$; also convex hull of $\mathbf{v}_i$.

$\forall\, k\leq n$, k-face of σ is any affine simplex of form $(\mathbf{v}_{i_1},\dots\mathbf{v}_{i_k})$, where vertices all distinct, so are affinely independent.

Definition 116. *(geometric) simplicial complex $K :=$ collection of affine simplices s.t.*

(1) $\sigma\in K \Longrightarrow$ *any face of* $\sigma\in K$*; and*

(2) *$\sigma, \tau \in K \Longrightarrow \sigma \bigcap \tau$ is a face of both σ and τ, or $\sigma \bigcap \tau = \emptyset$*

If K simplicial complex, $|K| = \bigcup \{\sigma | \sigma \in K\} \equiv$ "polyhedron" of K

Definition 117 (Def. 21.2 of Bredon (1997) [26]). *polyhedron := space X if $\exists$ homeomorphism $h : |K| \xrightarrow{\approx} X$ for some simplicial complex K. h, K is triangulation of X; (map h, complex K)*

Let K finite simplicial complex.
Choose ordering of vertices $\mathbf{v}_0, \mathbf{v}_1 \dots$ of K.
If $\sigma = (\mathbf{v}_{\sigma_0}, \dots \mathbf{v}_{\sigma_n})$ is simplex of K, where $\sigma_0 < \dots < \sigma_n$, then
let $f_\sigma : \Delta_n \to |K|$ be

$$f_\sigma = [\mathbf{v}_{\sigma_b}, \dots \mathbf{v}_{\sigma_n}]$$

in notation of Def. 1.2. Bredon (1997) [26].
Then this gives CW-complex structure on $|K|$ with f_σ as characteristic maps.

Part 10. Graphs, Finite Graphs

43. Graphs, Finite Graphs, Trees

Serre (1980) [28]
cf. Chapter I. Trees and Amalgams, Section 1 Amalgams, Subsection 1.1 Direct limits of Serre (1980) [28]
Let $(G_i)_{i \in I}$, family of groups.
$\forall$ pair (i, j), let F_{ij} = set of homomorphisms of G_i into G_j
Want: group $G = \varinjlim G_i$ and

$$\{f_i | f_i : G_i \to G\} \text{s.t. } f_j \circ f = f_i \quad \forall f \in F_{ij}$$

group G and family $\{f_i\}$ universal in that
(*) if H group, if $\{h_i | h_i : G_i \to H; h_j \circ f = h_i \qquad \forall f \in F_{ij}\}$,
then $\exists ! h : G \to H$ s.t. $h_i = h \circ f_i$
i.e. $\text{Hom}(G, H) \simeq \varprojlim \text{Hom}(G_i, H)$, the inverse limit being taken relative to F_{ij}.
i.e. G direct limit of G_i relative to the F_{ij}.
EY : 20170918 this is my rewrite/reinterpretation:
Let $(G_i)_{i \in I}$, $\forall (i, j) \in I^2$, let $F_{ij} = \{f \equiv f_{ij} | f : G_i \to G_j, f$ homomorphism of G_i into G_j $\}$.
Given group $G = \varinjlim G_i$ (for fixed i), $\{f_i | f_i : G_i \to G | f_j \circ f = f_i \quad \forall f \in F_{ij}\}$, i.e.

$$\begin{tikzcd} G_i \arrow[r, "f_i"] \arrow[d, "f_{ij} \equiv f"'] & G \\ G_j \arrow[ur, "f_j"'] & \end{tikzcd}$$

Then G, $\{f_i | f_i : G_i \to G | f_j \circ f = f_i \quad \forall f \in F_{ij}\}$ **universal**
if $\forall$ group H, $\forall \{h_i | h_i : G_i \to H | h_j \circ f = h_i \quad \forall f \in F_{ij}\}$,

$$\begin{tikzcd} H & G_i \arrow[l, "h_i"] \arrow[r, "f_i"] \arrow[d, "f_{ij} \equiv f"] & G \arrow[ll, bend right, "\exists ! h"'] \\ & G_j \arrow[ul, "h_j"] \arrow[ur, "f_j"'] & \end{tikzcd}$$

then $\exists !\ h : G \to H$, s.t. $h_i = h \circ f_i$ i.e.

Proposition 30. *$\exists !$ pair G, family $(f_i)_{i \in I}$, i.e. (pair consisting of $G, (f_i)_{i \in I}$, unique up to unique isomorphism.*

Proof. Define G by generators and relations.
Take generating family to be disjoint union of those for G_i.
relations - xyz^{-1} where $x, y, z \in G_i$, $z = xy \in G_i$
xy^{-1} where $x \in G_i$, $y \in G_j$, $y = f(x)$ for at least $f \in F_{ij}$.
Thus, existence of $G, \{f_i\}$.
G represents functor $H \mapsto \varprojlim \text{Hom}(G_i, H)$.
Thus, uniqueness (also from universal property). ☐

e.g. groups A, G_1, G_2, homomorphisms $f_1 : A \to G_1$.
$f_2 : A \to G_2$
G obtained by amalgamating A in G_1, G_2 by $f_1, f_2 \equiv G_1 *_A G_2$.
1 can have $G = \{1\}$, even though f_1, f_2 non-trivial.
Application: (Van Kampen Thm.)
Let topological space X be covered by open U_1, U_2.
Suppose $U_1, U_2, U_{12} = U_1 \bigcap U_2$ arcwise connected.
Let basept. $x \in U_{12}$.
Then $\pi_1(X; x)$ obtained by taking 3 groups

$$\pi_1(U_1; x), \pi_1(U_2; x), \pi_1(U_{12}; x)$$

and amalagamating them according to homomorphism

$$\pi_1(U_{12}; x) \to \pi_1(U_1; x)$$
$$\pi_1(U_{12}; x) \to \pi_1(U_2; x)$$

Exercise 1. Let homomorphisms $f_1 : A \to G_1$ amalgam $G = G_1 *_A G_2$.
$f_2 : A \to G_2$
Define subgroups A^n, G_1^n, G_2^n, of A, G_1, G_2 recursively by

$$A^1 = \{1\}$$
$$G_1^1 = \{1\}$$
$$G_2^1 = \{1\}$$

A^n = subgroup of A generated by $f_1^{-1}(G_1^{n-1})$ and $f_2^{-1}(G_2^{n-1})$

$$G_1^n = \text{subgroup of } G_i \text{ generated by } f_i(A^n)$$

Let A^∞, G_i^∞ be unions of A^n, G_i^n resp.
Show that f_i defines injection $A/A^\infty \to G_i/G_i^\infty$.
So the amalgamation is $G \simeq G_1/G_1^\infty *_{A/A^\infty} G_2/G_2^\infty$.
Take the first induction case (for intuition about the solution).

$$A^2 = \langle f_1^{-1}(G_1^1), f_2^{-1}(G_2^1) \rangle = \langle f_1^{-1}(\{1\}), f_2^{-1}(\{1\}) \rangle$$
$$G_i^2 = f_i(A^2)$$

Let $f_i(a) = f_i(b) \in G_i/G_i^\infty$; $a, b \in A/A^\infty$.
Then since $f_i(a), f_i(b) \in G_i/G_i^\infty$, $f_i(a), f_i(b) \in \{gG_i^\infty | g \in G_i\}$ (quotient is defined to be the set of all left cosets of G_i^∞, which has to be a normal subgroup for G_i/G_i^∞ to be a quotient group).
Since $a, b \in A/A^\infty$, suppose we take $a, b \in A$.

And suppose we take

$$f_i(a) = f_i(a)G_i^\infty = f_i(a)f_i(A^{n_a}) = f_i(aA^{n_a})$$
$$f_i(b) = f_i(b)G_i^\infty = f_i(b)f_i(A^{n_b}) = f_i(bA^{n_b})$$

Taking f_i^{-1} (recall for group homomorphisms, they map inverse of element of 1st. group to inverse of image of this element).

$aA^{n_a} = bA^{n_b} \in A/A^\infty$ (This is okay as we've "quotiented out A^∞; so indeed, they're equal)

cf. Subsection 1.2 Structure of amalgams of Serre (1980) [28]

Suppose given group A, family of groups $(G_i)_{i\in I}$, and, $\forall\, i \in I$, injective homomorphism $A \to G_i$.

$*_A G_i \equiv$ direct limit (cf. no. 1.1) of family (A, G_i) with respect to these homomorphisms, call it *sum* (in category theory sense, i.e. product) of G_i with A amalgamated.

e.g. $A = \{1\}$,
$*G_i \equiv$ free product of G_i.

43.0.1. *reduced word.* $\forall\, i \in I$, choose set S_i of right coset representations of G_i modulo A,
assume $1 \in S_i$,
$(a, s) \mapsto as$ is bijection of $A \times S_i$ onto G_i,
$A \times (S_i - \{1\}) \to G_i - A$ (onto)
Let $\mathbf{i} = (i_1 \dots i_n)$, $n \geq 0$, $i_j \in I$, s.t.

$$i_m \neq i_{m+1} \text{ for } 1 \leq m \leq n-1 \tag{132}$$

cf. (T) of Serre (1980) [28].
So *reduced word* m is defined as

$$m = (a; s_1 \dots s_n)$$

where $a \in A, s_1 \in S_{i_1} \dots s_n \in S_{i_n}$, and $s - j \neq 1 \,\forall\, j$.
$f \equiv$ canonical homomorphism of A into group $G = *_A G_i$
$f_i \equiv$ canonical homomorphism of G_i into group $G = *_A G_i$
EY : 20170611 (Further explanations, basic examples, from me):
Given $A, \{G_i\}_{i\in I}$, injective (group) homomorphisms $\{f_i : A \to G_i\}_i$.
$G_i \backslash f_i(A) = \{f_i(A)g | g \in G_i\}$.
Right coset representation of $f_i(A)g \mapsto g$.

e.g. A, G_1, G_2, $f_1 : A \to G_1$.
$f_2 : A \to G_2$

$$G_1 \backslash f_1(A) = \{f_1(A)g | g \in G_1\}$$
$$G_2 \backslash f_2(A) = \{f_2(A)g | g \in G_2\}$$

$\mathbf{i} = (i_1 \dots i_n)$, $i_j \in I$, $i_m \neq i_{m+1}$ for $1 \leq m \leq n-1$.
Consider $(1212 \dots 12)$
$m = (a; f_1 g_2 f_3 g_4 \dots f_{2n-1}, g_{2n})$ where f's $\in S_1 \subset G_1$, g's $\in S_2 \subset G_2$.
and so

Definition 118 (reduced word)**.** ***reduced word*** *of type* $\mathbf{i}$, m,

$$m = (a; s_1 \dots s_n) \tag{133}$$

where $a \in A, s_1 \in S_{i_1}, \dots s_n \in S_{i_n}$, $s_j \neq 1 \quad \forall\, j$,
$\mathbf{i} = (i_1 \dots i_n)$, $i_j \in I$, *s.t.* $i_m \neq i_{m+1}$ *for* $1 \leq m \leq n-1$,
with $S_i = \{g | g \in f_i(A)g \in f_i(A)G_i\}$

Theorem 17 (1 of Serre (1980) [28])**.** $\forall\, g \in G$, $\exists$ *sequence* $\mathbf{i}$ *s.t.* $i_m \neq i_{m+1}$ *for* $1 \leq m \leq n-1$ *and reduced word*

$$m = (a; s_1 \dots s_n)$$

of type $\mathbf{i}$ *s.t.*

$$g = f(a) f_{i_1}(s_1) \dots f_{i_n}(s_n)$$

Furthermore, $\mathbf{i}$ and m unique.
Remark. Thm. 1 implies $f; f_i$ injective.
Then identify A and G_i with images $f(A), f_i(G_i)$ in G, and reduced decomposition (*) of $g \in G$

$$g = as_1 \dots s_n, \quad a \in A,\ s_1 \in S_{i_1} - \{1\} \dots s_n \in S_{i_n} - \{1\}$$

Likewise, $G_i \bigcap G_j = A$ if $i \neq j$.
In particular, $S_i - \{1\}$ pairwise disjoint in G.

Proof. Let $X_i \equiv$set of reduced words of type $\mathbf{i}$, $X = \coprod X_i$.
Make G act on X.
In view of universal property of G, sufficient to make $\forall\, i, G_i$ act,
check action induced on A doesn't depend on i
Suppose then that $i \in I$, and let Y_i = set of reduced words of form $(1; s_1 \dots s_n)$, with $i_1 \neq i$.
EY : 20170611
Recall that

$$S_i = \{g | g \in f_i(A)g \in f_i(A)G_i\}$$
$$A \times S_i \to G_i \text{ onto}$$
$$A \times (S_i - \{1\}) \to G_i - A \text{ onto}$$
$$(a, s) \mapsto as \text{ bijection}$$

Let Y_i = set of reduced words of form $(1; s_1 \dots s_n) = \{(1; s_1 \dots s_n) | 1 \in A; s_1 \in S_{i_1} \dots s_n \in S_{i_n}; \mathbf{i} = (i_1 \dots i_n),\ i_j \in I \text{ s.t. } i_m \neq i_{m+1} \text{ for } 1 \leq m \leq n-1\}$.

$$A \times Y_i \to X = \coprod_i X_i$$
$$(a, (1; s_1 \dots s_n)) \mapsto (a; s_1 \dots s_n)$$
$$A \times \{S_i - \{1\}) \times Y_i \to X$$
$$((a, s), (1; s_1 \dots s_n)) \mapsto (a; s, s_1 \dots s_n)$$

and remember that X_i = set of reduced words of type $\mathbf{i}$.
It's clear that this yields a bijection $A \times Y_i \bigcup A \times (S_i - \{1\}) \times Y_i \to X$.
Let $x \in X$. Then $x \in X_{\mathbf{i}}$ for some $\mathbf{i}$. So x is a reduced word of type $\mathbf{i}$: $x = (a; s_1 \dots s_n)$. Then clearly $x = (a; s_1 \dots s_n) \mapsto (a, (1; s_1 \dots s_n)) \in A \times Y_i$. □

cf. pp. 13, Sec. 2. Trees, 2.1 Graphs of Serre (1980) [28]

Definition 119 (1. of Serre (1980) [28])**.** ***graph*** $\Gamma = (X, Y, Y \to X \times X, Y \to Y)$, *where* *set* $X =$ *vert* Γ
set $Y =$ *edge* Γ

$$Y \to X \times X$$
$$y \mapsto (o(y), t(y))$$
$$Y \to Y$$
$$y \mapsto \overline{y}$$

s.t. $\forall\, y \in Y$, $\overline{\overline{y}} = y$, $\overline{y} \neq y$, $o(y) = t(\overline{y})$.
vertex $P \in X$ *of* Γ.
(oriented) edge $y \in Y$, $\overline{y} \equiv$ *inverse edge.*

origin of $y :=$ *vertex* $o(y) = t(\overline{y})$.
terminus of $y :=$ *vertex* $t(y) = o(\overline{y})$
extremities of $y := \{o(y), t(y)\}$
If 2 vertices ***adjacent****, they're extremities of some edge.*
orientation of graph $\Gamma = Y_+ \subset Y =$ *edge* Γ *s.t.* $Y = Y_+ \coprod \overline{Y}_+$*. It always exists.*
oriented graph defined, up to isomorphism, by giving 2 sets X, Y_+ *and* $Y+ \to X \times X$*.*
corresponding set of edges is $Y = Y_+ \coprod \overline{Y}_+$ *where* $\overline{Y}_+ \equiv$ *copy of* Y_+

43.0.2. *Realization of a Graph.* cf. Realization of a Graph in Serre (1980) [28].
Let graph Γ, $X = \text{vert}\Gamma$, $Y = \text{edge}\Gamma$.
topological space $T = X \coprod Y \times [0,1]$, where X, Y provided with discrete topology.
Let R be finest equivalence relation on T for which

$$\begin{aligned} &(y,t) \equiv (\overline{y}, 1-t) \\ &(y,0) \equiv o(y) \qquad \forall\, y \in Y, \forall\, t \in [0,1] \\ &(y,1) \equiv t(y) \end{aligned} \tag{134}$$

quotient space $\text{real}(\Gamma) = T/R$ is *realization* of graph Γ. (realization is a functor which commutes with direct limits).
Let $n \in \mathbb{Z}^+$. Consider oriented graph of $n+1$ vertices $0, 1, \dots n$,

Definition 120. *path (of length n) in graph* Γ *is morphism* c *of* $Path_n$ *into* Γ

orientation given by n edges $[i, i+1]$, $0 \leq i < n$, $o([i,i+1]) = i$
$t([i,i+1]) = i+1$

For $n \geq 1$,
$(y_1 \dots y_n)$ sequence of edges $y_i = c([i-1,i])$ s.t.

$$t(y_i) = o(y_{i+1}), \qquad 1 \leq i < n \text{ determine } c$$

If $P_i = c(i)$,
c is a path from P_0 to P_n, and P_0 and P_n are *extremities of the path c.*
pair of form $(y_i, y_{i+1}) = (y_i, \overline{y}_i)$ in path is **backtracking**.
path (of length $n-2$), from P_0 to P_n given (for $n > 2$) by $(y_1 \dots y_{i-1}, y_{i+2} \dots y_n)$
If $\exists$ path from P to Q in Γ, $\exists$ one without backtracking (by induction)
direct limit $\text{Path}_\infty = \varinjlim \text{Path}_n$ provides notion of infinite path.
$\text{Path}_\infty \ni$ infinite sequence $(y_1, y_2, \dots)$ of edges s.t. $t(y_i) = o(y_{i+1}) \quad \forall\, i \geq 1$.

Definition 121 (connected graph; Def. 3 of Serre (1980) [28])**.** *graph connected if* $\forall$ *2 vertices, 2 vertices are extremities of at least 1 path.*
maximal connected subgraphs (under relation of inclusion) are connected components *of graph.*

43.0.3. *Circuits.* Let $n \in \mathbb{Z}^+$, $n \geq 1$.
Consider

set of vertices $\mathbb{Z}/n\mathbb{Z}$, orientation given by n edges $[i, i+1]$, $(i \in \mathbb{Z}/n\mathbb{Z})$ with $o([i,i+1]) = i$
$t([i,i+1]) = i+1$

Definition 122 (circuit; Def. 4 of Serre (1980) [28])**.** *circuit (length n) in graph is subgraph isormorphic to* $Circ_n$*.*

i.e. subgraph = path $(y_1 \dots y_n)$, without backtracking, s.t. $P_i = t(y_i)$, $(1 \leq i \leq n)$ distinct, s.t. $P_n = o(y_1)$

$n = 1$ case: Circ_1, $\mathbb{Z}/\mathbb{Z} = \{0\}$, 1 edge, $[0,1]$, $0 \in \mathbb{Z}/1\mathbb{Z}$, $o([0,1]) = 0$
$t([0,1]) = 1$

Note Circ_1 has automorphism of order 2, which changes its orientation, i.e.
$\exists$ automorphism $\sigma \in \text{Aut}(\text{Circ}_1)$ s.t. $|\sigma| = 2$, i.e. $\sigma^2 = 1$.
loop := circuit of length 1; so loop $\in \text{Circ}_1$.
path (y_1), $P_1 = t(y_1) = o(y_1)$.
$n = 2$ case: Circ_2, $\mathbb{Z}/2\mathbb{Z} = \{0,1\}$, 2 edges $[0,1], [1,2]$,

path (y_1, y_2), $(1 \leq i \leq 2)$, $P_1 = t(y_1)$
$P_2 = t(y_2) = o(y_1)$

43.1. **Combinatorial graphs.** Let $(X, S) \equiv$ simplicial complex of dim. ≤ 1, with
$X \equiv$ set
$S \equiv$ set of subsets of X with 1 or 2 elements, containing all the 1-element subsets.
associates with it a graph $\Gamma = (X, \{(P,Q)\})$.
X is its set of vertices.
edges $= \{(P,Q) \in X \times X\}$ s.t. $P \neq Q$, $\{P,Q\} \in S$, with $\overline{(P,Q)} = (Q,P)$
$o(P,Q) = P$
$t(P,Q) = Q$

In this graph, 2 edges with same origin and same terminus are equal. This is equivalent to (see following Def.)

Definition 123 (combinatorial; Def. 5 of Serre (1980) [28])**.** *graph is combinatorial if it has no circuit of length* ≤ 2

Conversely, it's easy to see that
every combinatorial graph Γ derived (up to isomorphism) by construction above from simplicial complex (X, S), where
$X = \text{vert}\Gamma$
$S =$ set of subset $\{P, Q\}$ of X s.t. P and Q either adjacent or equal.

Part 11. Tensors, Tensor networks; Singular Value Decomposition, QR decomposition, Density Matrix Renormalization Group (DMRG), Matrix Product states (MPS)

44. Introductions to Tensor Networks

José Barbon (IFT-CSIC, Univ. Autonoma de Madrid) gave the https://youtu.be/nsxgAOAEgbg for the workshop "Black Holes, Quantum Information, Entanglement, and all that," (29 May-1 June, 2017, with the organizing committee of Thibault Damour (IHES), Vasily Pestun (IHES), Eliezer Rabinovici (IHES & Hebrew Univ. of Jerusalem).
In the talk,
cf. 43:13

The church of the doubled Hilbert space. Any thermal box can be obtained by tracing over a second identical copy, if appropriately entangled into a global pure state.

$$\rho_R = \text{Tr}_L \sum_n C_n \Psi_n^L \otimes \Psi_n^R$$

$$(C_n)_{\text{thermal}} = \left[\frac{e^{-\beta E_n}}{\sum_m e^{-\beta E_M}} \right]^{1/2}$$

But!!
If the entanglement basis is taken to be the high-energy band of two "entangled" CFTs ...

$$|TFD\rangle \sim \sum_{E_n} e^{-\beta E_n/2} |E_n\rangle_L \otimes |E_n\rangle_R$$

neglecting the tiny e^{-S} spacings. we can approximate by continuous spectrum of fields in the background of an AdS black hole, to get ...

$$\int_E e^{-\beta E/2} |E\rangle_L \otimes |E\rangle_R$$

The HH state of the bulk fields!

cf. 46:16

SLOGAN: EPR = ER Maldacena-Susskind

Accumulating a density of entanglement of $S \gg 1$ well-separated Bell pairs within a transversal size of order $(GS)^{1/2}$ seems to generate a geometrical bridge of area GS.

cf. 49:26

Parametrizing complexity of entanglement. Pick a tensor decomposition of Hilbert space of dimension $\exp(S)$ into S factors of $O(1)$ dimension.

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_S$$

A tensor of S indices gives a generic state.

cf. 50:27

The decomposition of the big tensor in small building blocks gives a notion of "complexity of entanglement"

rather simple entanglement pattern

somewhat more complex entanglement pattern

picture from M von Raamsdonk

cf. 55:10

A list of open questions & problems.

- Need exactly calculable toy models of AdS/CFT along the lines of SYK model
- Give a "renormalized" definition of quantum complexity for continuum CFTs
- Can tensor networks describe bulk gravitons?
- What is the space-time meaning of quantum complexity saturation?
- Can we define approximate local observables for black hole inferiors?
- Are there obstructions related to firewalls and/or fuzzballs?

Workshop introductory overview by José Barbon for the Institut des Hautes Études Scientifiques (IHÉS) gave me the first impetus to understand tensor networks as I sought to also understand the condensates of entanglement pairs within the black hole.

A Google search for introductions to tensor networks that are on arxiv ("Introduction Tensor Network arxiv") yielded Bridgeman and Chubb's course notes (bf. Bridgeman and Chubb (2017) [33]).

44.1. **List of stuff I want to look at/do/study.** I would like to compare/contrast the following:

- Rotman (2010) [29], Ch. 8, but starting from 8.4 Tensor Products, pp. 574
- Jeffrey Lee (2009) [32], Ch. 7 Tensors
- http://www.irisa.fr/sage/bernard/publis/SVD-Chapter06.pdf, https://math.stackexchange.com/questions/694339/parallel-algorithms-for-svd

Maldacena and Susskind (2013) [38]

Lectures on Gravity and Entanglement. Mark Van Raamsdonk [41]

- Consider as physical system AdS-Schwarzchild black hole
- CFT
  - PFL Lectures on Conformal Field Theory in $D \geq 3$ Dimensions, Rychkov (2016) [39].

Evenbly and Vidal (2011) [40], Tensor network states and geometry

Loose ends (might not be useful links)

- https://arxiv.org/pdf/1506.06958.pdf
- https://arxiv.org/pdf/1512.02532.pdf One-point Functions in AdS/dCFT from Matrix Product States

Numerical implementation strategy: 1st, CUDA cuSolver, 2nd, Numerical Recepes version, 3rd, parallel algorithm review.

44.2. **Tensor operations; Tensor properties.**

44.2.1. *rank.* r = rank tensor of dim. $d_1 \times \cdots \times d_r$ is element of $\mathbb{C}^{d_1 \times \cdots \times d_r}$

Tensor product

$$[A \otimes B]_{i_1 \dots i_r, j_1 \dots j_s} := A_{i_1 \dots i_r} \cdot B_{j_1 \dots j_s} \tag{135}$$

44.2.2. *Trace.* Given tensor A, xth, yth indices have identical dims. ($d_x = d_y$), partial trace over these 2 dims. is simply joint summation over that index

$$[\text{Tr}_{x,y} A]_{i_1 \dots i_{x-1} i_{x+1} \dots i_{y-1} i_{y+1} \dots i_r} = \sum_{\alpha=1}^{d_x} A_{i_1 \dots i_{x-1} \alpha i_{x+1} \dots i_{y-1} \alpha i_{y+1} \dots i_r} \tag{136}$$

44.2.3. *Contraction.*

44.2.4. *Group and splitting, Bridgeman and Chubb (2017)* [33]. "Rank is a rather fluid concept in the study of tensor networks." Bridgeman and Chubb (2017) [33].

$\mathbb{C}^{a_1 \times \cdots \times a_n} \simeq \mathbb{C}^{b_1 \times \cdots \times b_m}$ isomorphic as vector spaces if $\prod_i a_i = \prod_i b_i$.

We can "group" or "split" indices to lower or raise rank of given tensor, resp.

Consider contracting 2 arbitrary tensors.

If we group together indices which are and are not involved in contraction,

"It should be noted that not only is this reduction to matrix multiplication pedagogically handy, but this is precisely the manner in which numerical tensor packages perform contraction, allowing them to leverage highly optimised matrix multiplication code." (cf. Bridgeman and Chubb (2017) [33]; check this)

"Owing to freedom in choice of basis, precise details of grouping and splitting aren't unique." (cf. Bridgeman and Chubb (2017) [33]).

1 specific choice of convention:

tensor product basis, defining basis on product space by product of respective bases.

"The canonical use of tensor product bases in quantum information allows for grouping and splitting described above to be - dealt with implicitly."

$$|0\rangle \otimes |1\rangle \equiv |0\rangle \tag{137}$$

and precisely this grouping,

$$\begin{gathered} |0\rangle \otimes |1\rangle \in \text{Mat}_{\mathbb{C}}(2,2), \text{ whilst} \\ |01\rangle \in \mathbb{C}^4 \end{gathered} \tag{138}$$

Suppose rank $n+m$ tensor T, group its first n indices, last m indices together.

$$T_{I,J} := T_{i_1 \dots i_n, j_1 \dots j_m}$$

where

$$\begin{aligned} I &:= i_1 + d_1^{(i)} i_2 + d_1^{(i)} d_2^{(i)} i_3 + \cdots + d_1^{(i)} \dots d_{n-1}^{(i)} i_n \\ J &:= j_1 + d_1^{(j)} j_2 + d_1^{(j)} d_2^{(j)} j_3 + \cdots + d_1^{(j)} \dots d_{m-1}^{j)} j_m \end{aligned}$$

EY : 20170627 to elaborate, consider a functor `flatten` that does what's described above, in the context of category theory (and so this is the generalization):

$$\mathbb{K}^{d_1^{(i)}} \times \mathbb{K}^{d_2^{(i)}} \times \dots \times \mathbb{K}^{d_n^{(i)}} \times \mathbb{K}^{d_1^{(j)}} \times \mathbb{K}^{d_2^{(j)}} \times \dots \times \mathbb{K}^{d_m^{(j)}} \xrightarrow{\text{flatten}} \mathbb{K}^{\prod_{p=1}^n d_p^{(i)}} \times \mathbb{K}^{\prod_{q=1}^m d_q^{(j)}}$$

$$T_{i_1 \dots i_n, j_1 \dots j_m} \xmapsto{\text{flatten}} T_{I,J}$$

(139)
$$\{0,1,\dots d_1^{(i)}\} \times \{0,1,\dots d_2^{(i)}\} \times \dots \times \{0,1,\dots d_n^{(i)}\} \times \{0,1,\dots d_1^{(j)}\} \times \{0,1,\dots d_2^{(j)}\} \times \dots \times \{0,1,\dots d_m^{(j)}\} \xrightarrow{\text{flatten}}$$
$$\xrightarrow{\text{flatten}} \{0,1,\dots \prod_{p=1}^n d_p^{(i)} - 1\} \times \{0,1,\dots \prod_{q=1}^m d_q^{(j)} - 1\}$$

$$(i_1, i_2, \dots i_n, j_1, j_2 \dots j_m) \xmapsto{\text{flatten}} (I,J) := (i_1 + d_1^{(i)} i_2 + \dots + d_1^{(i)} \dots d_{n-1}^{(i)} i_n, j_1 + d_1^{(j)} j_2 + \dots + d_1^{(j)} \dots d_{m-1}^{(j)} j_m)$$

It doesn't make sense to call this "row-major" or "column-major" ordering generalization, because we are not dealing with only 2 indices where we can definitely say the first index indexes the "row" and the second index indexes the "column." At most, possibly, you can alternatively have this:

$$(i_1 \dots i_n, j_1 \dots j_m) \xmapsto{\text{flatten}} (I,J) := (d_2^{(i)} \dots d_n^{(i)} i_1 + d_3^{(i)} \dots d_n^{(i)} i_2 + \dots + i_n, d_2^{(j)} \dots d_m^{(j)} j_1 + \dots + j_m)$$

Note that this is all 0-*based counting* (i.e. we start counting from 0 just like in C,C++,Python, etc.). If you really wanted 1-based counting, you'd have to complicate the above formulas as such:

$$(I,J) := (i_1 + d_1^{(i)}(i_2 - 1) + \dots + d_1^{(i)} \dots d_{n-1}^{(i)}(i_n - 1), j_1 + d_1^{(j)}(j_2 - 1) + \dots + d_1^{(j)} \dots d_{m-1}^{(j)}(j_m - 1))$$

Note that formulas are easily checked by pluggin in the minimum and maximum values for the indices and seeing if they make sense (e.g. plug in $(0,0,\dots,0)$ for all indices for 0-based counting and make sure you get back $I = 0$ or $J = 0$).

44.3. **Singular Value Decomposition.**

$$T_{I,J} = \sum_\alpha U_{I,\alpha} S_{\alpha,\alpha} \overline{V}_{J,\alpha}$$

(140)
$$\begin{gathered} \mathrm{Mat}_{\mathbb{K}}(N,M) \xrightarrow{\text{SVD}} \mathrm{Mat}_{\mathbb{K}}(N,P) \times \mathrm{Mat}_{\mathbb{K}}(P,P) \times \mathrm{Mat}_{\mathbb{K}}(M,P) \\ T_{I,J} \xmapsto{\text{SVD}} U_{I,\alpha}, S_{\alpha,\alpha}, \overline{V}_{I,\alpha} \text{ s.t.} \\ T_{I,J} = \sum_\alpha U_{I,\alpha} S_{\alpha,\alpha} \overline{V}_{J,\alpha} \\ T = USV^\dagger \end{gathered}$$

For the higher-dimensional version of SVD,

(141)
$$\mathbb{K}^{d_1^{(i)}} \otimes \dots \otimes \mathbb{K}^{d_N^{(i)}} \otimes \mathbb{K}^{d_1^{(j)}} \otimes \dots \otimes \mathbb{K}^{d_M^{(j)}} \xrightarrow{\text{flatten}} \mathrm{Mat}_{\mathbb{K}}(N,M) \xrightarrow{\text{SVD}} \mathrm{Mat}_{\mathbb{K}}(N,P) \times \mathrm{Mat}_{\mathbb{K}}(P,P) \times \mathrm{Mat}_{\mathbb{K}}(M,P) \xrightarrow{\text{splitting}}$$
$$\xrightarrow{\text{splitting}} \mathbb{K}^{d_1^{(i)}} \otimes \dots \otimes \mathbb{K}^{d_N^{(i)}} \otimes \mathbb{K}^P \times \mathrm{Mat}_{\mathbb{K}}(P,P) \times \mathbb{K}^{d_1^{(j)}} \otimes \dots \otimes \mathbb{K}^{d_M^{(j)}} \otimes \mathbb{K}^P$$
$$T_{i_1 \dots i_N, j_1 \dots j_M} = \sum_\alpha U_{i_1 \dots i_N, \alpha} S_{\alpha,\alpha} \overline{V}_{j_1 \dots j_M, \alpha}$$

45. Density Matrix Renormalization Group; Matrix Product States (MPS)

45.1. **Introduction; physical system (physical setup).** cf. "Density Matrix Renormalization Group/Matrix Product States" lectures by Schollwöck (2017) [36].

Recall the fundamental Hamiltonian (frequently in solid state physics), for *electrons moving in a Hamiltonian potential.*

(142)
$$H = \sum_{j=1}^{e^-} \frac{\mathbf{p}_j^2}{2m_e} + \frac{1}{2} \frac{1}{4\pi\epsilon_0} \frac{q_e^2}{|\mathbf{r}_i - \mathbf{r}_j|^2} + \sum_{j=1}^{e^-} V_{\text{eff}}(\mathbf{r}_j)$$

where $\frac{\mathbf{p}_j^2}{2m_e}$ is the kinetic energy term, $\sum_{j=1}^{e^-} V_{\text{eff}}(\mathbf{r}_j)$ is the lattice potential. The problem is in the 2nd. term, electron-electron interaction, $\frac{1}{2} \frac{1}{4\pi\epsilon_0} \frac{q_e^2}{|\mathbf{r}_i - \mathbf{r}_j|^2}$

Typical models include the following:

- Hubbard model (tight, binding-like model; basis states are not energy states but *Wannier basis* states):

(143)
$$H = -t \sum_{\langle i,j \rangle, \sigma} c_{i\sigma}^\dagger c_{j\sigma} + h.c. + U \sum_i n_{i\uparrow} n_{i\downarrow}$$

where $\langle i,j \rangle$ denotes nearest neighbors, σ index is for all possible states, $h.c.$ stands for hermitian conjugate, and $d \equiv$ number of states of single spin site.
$-t \sum_{\langle i,j \rangle, \sigma} c_{i\sigma}^\dagger c_{j\sigma} + h.c.$ is the kinetic energy term,
$U \sum_i n_{i\uparrow} n_{i\downarrow}$ is the Coulomb energy.
Hilbert space for the Hubbard model is

(144)
$$\{|\emptyset\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}^{\otimes L}, \qquad d = 4$$

- Heisenberg model (large -U Hubbard at half-filling)

(145)
$$H = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = J \sum_{\langle i,j \rangle} \frac{1}{2}(S_i^+ S_j^- + S_j^+ S_i^-) + S_i^z S_j^z)$$

Hilbert space $\{|\uparrow\rangle, |\downarrow\rangle\}^{\otimes L}$, $d = 2$

45.1.1. *Compression of information viewpoint for solid-state Hamiltonians, quantum many-body systems.* .
"emergent" macroscopic quantities, τ, p (temperature, pressure). For

$$H = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = J \sum_{\langle i,j \rangle} \frac{1}{2}(S_i^+ S_j^- + S_j^+ S_i^-) + S_i^z S_j^z)$$

H as classical spins: thermodynamic limit $N \to \infty$. 2 angles required to describe unit vector on unit sphere $(S^3) \implies 2N$ degrees of freedom (linear)
quantum spins: superposition of states, thermodynamic limit: $N \to \infty$, 2^N degrees of freedom (exponential).

45.1.2. *Definitions; notation and conventions.* Quantum system living on L lattice sites; cf. Schollwöck (2017) [36], lattice can be in any dim., effectively most useful in 1-dim., think of the example of a 1-dim. chain of L sites.

d local states per site $\{\sigma_i\rangle$, $i \in \{1,2,\dots L\}$
e.g. spin $\frac{1}{2}$, $d = 2$, $|\uparrow\rangle, |\downarrow\rangle$.
Hilbert space: $\mathcal{H} = \otimes_{i=1}^L \mathcal{H}_i$, $\mathcal{H}_i = \{|1_i\rangle, \dots |d_i\rangle\}$.
Notice, there are *exponentially many coefficients*, c's. Most general state (not necessarily 1-dim.) is

(146)
$$|\psi\rangle = \sum_{\sigma_1 \dots \sigma_L} c^{\sigma_1 \dots \sigma_L} |\sigma_1 \dots \sigma_L\rangle$$

abbreviations: $\{\sigma\} = \sigma_1 \dots \sigma_L$. And so we can write $c^{\{\sigma\}}$.

45.2. **MPS, matrix product states.**

$$|\psi\rangle = \sum_{\sigma_1\dots\sigma_L} M^{\sigma_1}M^{\sigma_2}\dots M^{\sigma_L}|\sigma_1\sigma_2\dots\sigma_L\rangle \tag{147}$$

The $\sum_{\sigma_1\dots\sigma_L}$ means that all basis states participate; Schollwöck is not kicking out any states arbitrarily.

$$c^{\{\sigma\}} = M^{\sigma_1}M^{\sigma_2}\dots M^{\sigma_L} \in \mathbb{C}$$

so

$M^{\sigma_1} \in \text{Mat}_{\mathbb{C}}(1,n_1)$ so to get a scalar in the product of matrices. Likewise, $M^{\sigma_L} \in \text{Mat}_{\mathbb{C}}(m_L,1)$
(variational) constraint is in expansion coefficients.
$\forall$ d local basis states, $|\sigma_i\rangle \in V_i \equiv V$,dim$V = d$, let there be 1 matrix M, i.e. M^{σ_i}.
Thus, dL matrices altogether (in total).
Assume matrix size has upper limit D (a computer limitation).
Up to dLD^2 coefficients, instead of exponentially many ($c^{\{\sigma\}}$, and sum over $\{\sigma\}$).

45.2.1. *Product States and MPS.* Mean-filed approximation/product state misses essential quantum feature: **entanglement**.
Consider 2 spin $\frac{1}{2}$ systems: $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, $\mathcal{H}_i = \{|\uparrow\rangle, |\downarrow\rangle\}$
General state is

$$|\psi\rangle = c^{\uparrow\uparrow}|\uparrow\uparrow\rangle + c^{\uparrow\downarrow}|\uparrow\downarrow\rangle + c^{\downarrow\uparrow}|\downarrow\uparrow\rangle + c^{\downarrow\downarrow}|\downarrow\downarrow\rangle$$

e.g. singlet state: $|\psi\rangle = \frac{1}{\sqrt{2}}|\uparrow\downarrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\uparrow\rangle$.
As an exercise, show that the singlet state cannot be written as product of local coefficients, i.e.

$$c_{\uparrow\downarrow} \neq c^{\uparrow}c^{\downarrow}$$

Instead of writing products of scalars, write product of matrices, i.e. $e^{\sigma_1}\cdot e^{\sigma_2} \to M^{\sigma_1}M^{\sigma_2}$

$$M^{\uparrow 1} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad M^{\downarrow 1} = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad \begin{matrix} M^{\uparrow 2} = \begin{bmatrix} 0 \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \\ M^{\downarrow 2} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \end{matrix}$$

$$M^{\uparrow 1}M^{\downarrow 2} = \frac{1}{\sqrt{2}}$$
$$M^{\downarrow 1}M^{\uparrow 2} = \frac{-1}{\sqrt{2}}$$

45.2.2. *AKLT model (Affleck-Kennedy-Lieb-Tasaki).* MPS is useful even for matrices of dim. 2.

45.3. **General matrix product state (MPS) and SVD (Singular Value Decomposition).** cf. Schollwöck (2017) [36]
The general matrix product state (MPS) is the following:

$$|\psi\rangle = \sum_{\sigma_1\dots\sigma_L} M^{\sigma_1}M^{\sigma_2}\dots M^{\sigma_L}|\sigma_1\sigma_2\dots\sigma_L\rangle \tag{148}$$

where $\sigma_i \in V_i$, dim$V_i = d_i$ and
$M^{\sigma_1} \in \text{Mat}_{\mathbb{C}}(1,D_1)$
$M^{\sigma_2} \in \text{Mat}_{\mathbb{C}}(D_1,D_2)$
$\vdots$
$M^{\sigma_{L-1}} \in \text{Mat}_{\mathbb{C}}(D_{L-2},D_{L-1})$
$M^{\sigma_L} \in \text{Mat}_{\mathbb{C}}(D_{L-1},1)$

Notice the non-unique **gauge degree of freedom**:

$\forall\, A \in \text{Mat}_{\mathbb{C}}(m,n)$, then for $k = \min(m,n)$,

$$A = USV^\dagger \equiv U\Sigma V^\dagger \text{ where} \tag{149}$$

$U \in \text{Mat}_{\mathbb{C}}(m,k)$, $U^\dagger U = 1$ (i.e. U consists of orthonormal columns, or k number of u's $\in \mathbb{C}^m$); if $m = k$, $UU^\dagger = 1$,
$S \in \text{Mat}_{\mathbb{C}}(k,k)$ s.t. $S \in \text{diag}_{\mathbb{C}}(k)$, $s_1 \geq s_2 \geq s_3 \geq \dots s_i \geq 0$, s_j's non-negative "singular values" (adjacent "singular" in name doesn't imply anything), non-vanishing = rank$r \leq k$.
$V^\dagger \in \text{Mat}_{\mathbb{C}}(k,n)$, $V^\dagger V = 1$, (orthonormal rows, or k number of $v \in \mathbb{C}^n$); if $k = n$, $VV^\dagger = 1$
Recall eigenvalue equation and thus so-called eigenvalue decomposition.
For $A \in \text{Mat}_{\mathbb{C}}(m,m)$,

$$Au_j = \lambda_j u_j; \qquad j = 1\dots r;\ r \equiv \text{ rank }, \quad u_j \in \text{Mat}_{\mathbb{C}}(m,1)$$
$$A_{ik}u_{kj} = \lambda_j u_{ij} = u_{ik}\delta_{kj}\lambda_j \Longrightarrow AU = U\Lambda$$

with $U \in \text{Mat}_{\mathbb{C}}(m,r)$, $\Lambda \in \text{Mat}_{\mathbb{C}}(r,r)$.
And so

$$AA^\dagger = USV^\dagger VSU^\dagger = US^2U^\dagger \Longrightarrow (AA^\dagger)U = US^2$$
$$A^\dagger A = VSU^\dagger USV^\dagger = VS^2V^\dagger \Longrightarrow (A^\dagger A)V = VS^2$$

so if we treat U and V, matrices of left, right singular vectors, then S^2 singular value squared are eigenvalues.
Start with

$$|\psi\rangle = \sum_{\sigma_1\dots\sigma_L} c^{\sigma_1\dots\sigma_L}|\sigma_1\dots\sigma_L\rangle \in V \text{ s.t. dim}V = d^L \tag{150}$$

Note the *abuse of notation*: while $c^{\sigma_1\dots\sigma_L} \in \mathbb{C}$ itself, also denote $c^{\sigma_1\dots\sigma_L} \in \mathbb{C}^{d^L}$ as a shorthand for $\sum_{\sigma_1\dots\sigma_L} c^{\sigma_1\dots\sigma_L}|\sigma_1\dots\sigma_L\rangle$
Reshape coefficient vector into matrix of (size) dimension $(d \times d^{L-1})$.

$$\mathbb{C}^{d^L} \xrightarrow{\text{reshape}} \text{Mat}_{\mathbb{C}}(d, d^{L-1})$$
$$c^{\sigma_1\dots\sigma_L} \xmapsto{\text{reshape}} \Psi_{\sigma_1,(\sigma_2\dots\sigma_L)}$$

Then do SVD:

$$\Psi_{\sigma_1,(\sigma_2\dots\sigma_L)} \overset{\text{SVD}}{=} \sum_{a_1} U_{\sigma_1 a_1}S_{a_1a_1}V^\dagger_{a_1,\sigma_2\dots L} = U_{\sigma_1a_1}S_{a_1a_1}V^\dagger_{a_1,\sigma_2\dots\sigma_L}$$

Let's utilize commutative diagrams to summarize the reshaping and SVD operations that we've done.

$$\mathbb{C}^{d^L} = \text{Mat}_{\mathbb{C}}(1,d^L) \xrightarrow{\text{reshape}} \text{Mat}_{\mathbb{C}}(d,d^{L-1}) \xrightarrow{\text{SVD}} \text{Mat}_{\mathbb{C}}(d,r_1)\times\text{Mat}_{\mathbb{C}}(r_1,r_1)\times\text{Mat}_{\mathbb{C}}(r_1,d^{L-1})$$

$$|\Psi\rangle \equiv c^{\sigma_1\dots\sigma_L} \xmapsto{\text{reshape}} \Psi_{\sigma_1,(\sigma_2\dots\sigma_L)} \xmapsto{\text{SVD}} \Psi_{\sigma_1,(\sigma_2\dots\sigma_L)} \overset{\text{SVD}}{=} U_{\sigma_1a_1}S_{a_1a_1}V^\dagger_{a_1,\sigma_2\dots\sigma_L}$$

where I abuse notation for the SVD operation in that, SVD maps a matrix (in this case, Ψ) into 3 matrices, that obey the equality relationship when they're multiplied together (i.e. $\Psi = USV^\dagger$).
Slice U into d row vectors, i.e. for $U \in \text{Mat}_{\mathbb{C}}(d,r_1)$.

$$\text{Mat}_{\mathbb{C}}(d,r_1) \xrightarrow{\text{slice}} \text{Mat}_{\mathbb{C}}(1,r_1)^d$$
$$U_{\sigma_1a_1} \mapsto \{A^{\sigma_1}\} \equiv \{A^{\sigma_1}_{1,a_1}\}_{\sigma_1} \text{ s.t. } A^{\sigma_1}_{1,a_1} = U_{\sigma_1a_1} \text{ and } |\{A^{\sigma_1}_{1,a_1}\}| = d$$

Collecting all the operations, and doing the following notation rewrite,

$$c^{\sigma_1\sigma_2\dots\sigma_L} \mapsto \Psi_{\sigma_1\sigma_2\dots\sigma_L} = \sum_{a_1} A^{\sigma_1}_{1a_1}S_{a_1a_1}V^\dagger_{a_1,\sigma_2\dots\sigma_L} = \sum_{a_1} A^{\sigma_1}_{1a_1}c^{a_1\sigma_2\sigma_3\dots\sigma_L}$$

where

$$c^{a_1\sigma_2\sigma_3\ldots\sigma_L} = S_{a_1a_1}V^\dagger_{a_1\sigma_2\ldots\sigma_L}$$

Do the same procedure again.

$$\mathrm{Mat}_{\mathbb{C}}(r_1, d^{L-1}) \xrightarrow{\text{reshape}} \mathrm{Mat}_{\mathbb{C}}(r_1 d, d^{L-2}) \xrightarrow{\text{SVD}} \mathrm{Mat}_{\mathbb{C}}(r_1 d, r_2) \times \mathrm{Mat}_{\mathbb{C}}(r_2, r_2) \times \mathrm{Mat}_{\mathbb{C}}(r_2, d^{L-2})$$

$$c^{a_1,\sigma_2\sigma_3\ldots\sigma_L} \xmapsto{\text{reshape}} \Psi_{a_1\sigma_2,(\sigma_3\ldots\sigma_L)} \xmapsto{=} \Psi_{a_1\sigma_2,(\sigma_3\ldots\sigma_L)} \overset{\text{SVD}}{=} U_{a_1\sigma_2,a_2}S_{a_2a_2}V^\dagger_{a_2,\sigma_3\ldots\sigma_L}$$

Then slice U into d matrices, and then matrix multiply the S and $V^\dagger$ matrices together:

$$\mathrm{Mat}_{\mathbb{C}}(r_1 d, r_2) \times \mathrm{Mat}_{\mathbb{C}}(r_2, r_2) \times \mathrm{Mat}_{\mathbb{C}}(r_2, d^{L-2}) \xrightarrow{\text{slice and multiply}} \mathrm{Mat}_{\mathbb{C}}(r_1, r_2)^d \times \mathrm{Mat}_{\mathbb{C}}(r_2, d^{L-2})$$

$$\textstyle\sum_{a_2} U_{a_1\sigma_2,a_2}S_{a_2a_2}V^\dagger_{a_2,\sigma_3\ldots\sigma_L} \xmapsto{\quad=\quad} = \sum_{a_2} A^{\sigma_2}_{a_1a_2}c^{a_2,a_3\ldots\sigma_L} \text{ where } A^{\sigma_2}_{a_1a_2} = U_{a_1\sigma_2,\sigma_3\ldots\sigma_L}$$

Thus, generalize the *ith procedure*: for $i = 1 \dots L$,
Let $r_0 = 1$.

$$\boxed{\begin{array}{ccccccc} \mathrm{Mat}_{\mathbb{C}}(r_{i-1}, d^{L-(i-1)}) & \xrightarrow{\text{reshape}} & \mathrm{Mat}_{\mathbb{C}}(r_{i-1}d, d^{L-i}) & \xrightarrow{\text{SVD}} & \mathrm{Mat}_{\mathbb{C}}(r_{i-1}d, r_i) \times \mathrm{Mat}_{\mathbb{C}}(r_i, r_i) \times \mathrm{Mat}_{\mathbb{C}}(r_i, d^{L-i}) & \xrightarrow{\text{slice and multiply}} & \mathrm{Mat}_{\mathbb{C}}(r_{i-1}, r_i)^d \times \mathrm{Mat}_{\mathbb{C}}(r_i, d^{L-i}) \\ c^{a_{i-1},\sigma_i\sigma_{i+1}\dots\sigma_L} & \xmapsto{\text{reshape}} & \Psi_{a_{i-1}\sigma_i,(\sigma_{i+1}\sigma_{i+2}\dots\sigma_L)} & \xmapsto{=} & U_{a_{i-1}\sigma_i,a_i} S_{a_i a_i} V^\dagger_{a_i,\sigma_{i+1}\dots\sigma_L} & \xmapsto{=} & A^{\sigma_i}_{a_{i-1},a_i} c^{a_i,\sigma_{i+1}\dots\sigma_L} \end{array}} \quad (151)$$

Remember that $r_i \leq \min{(r_{i-1}d, d^{L-i})}$ and for $i = L$, there is no need to do a SVD, but only a reshape, and slice and multiply.
Collecting all the A matrices:

$$\boxed{\begin{aligned} & A^{\sigma_1}_{1,a_1} \in \mathrm{Mat}_{\mathbb{C}}(1, r_1); \qquad r_1 \leq d \\ & A^{\sigma_2}_{a_1,a_2} \in \mathrm{Mat}_{\mathbb{C}}(r_1, r_2); \qquad r_2 \leq r_1 d \\ & \vdots \\ & A^{\sigma_i}_{a_{i-1},a_i} \in \mathrm{Mat}_{\mathbb{C}}(r_{i-1}, r_i); \qquad r_i \leq \min{(r_{i-1}d, d^{L-i})} \\ & \vdots \\ & A^{\sigma_L}_{a_{L-1},a_L} \in \mathrm{Mat}_{\mathbb{C}}(r_{L-1}, 1); \qquad r_{L-1} \leq d \end{aligned}} \quad (152)$$

45.3.1. *Left and Right Normalization, A and B matrices, "special gauge" from normalization.* Choose orthonormal basis states $\forall\, a_l$, $\forall\, l = 1, 2, \dots L$ For

$$\begin{aligned} |a_l\rangle &= \sum_{a_{l-1}\sigma_l} M^{\sigma_l}_{a_{l-1}a_l} |a_{l-1}\sigma_l\rangle \\ \langle a_l'| &= \sum_{a_{l-1}'\sigma_l'} \langle a_{l-1}'\sigma_l'| (M^{\sigma_l'}_{a_{l-1}'a_l'})^* \end{aligned}$$

then,

$$\begin{gathered} \delta_{a_l'a_l} = \langle a_l'|a_l\rangle = \sum_{a_{l-1}'\sigma_l', a_{l-1}\sigma_l} M^{\sigma_l'*}_{a_{l-1}'a_l'} M^{\sigma_l}_{a_{l-1}a_l} \langle a_{l-1}'\sigma_l'|a_{l-1}\sigma_l\rangle = \sum_{a_{l-1}\sigma_l} M^{\sigma_l *}_{a_{l-1}a_l'} M^{\sigma_l}_{a_{l-1}a_l} = \\ = \sum_{\sigma_l} ((M^{\sigma_l})^\dagger M^{\sigma_l})_{a_l'a_l} \end{gathered} \quad (153)$$

Left normalization comes from a property of SVD in that $\forall\, U$ matrices, $U^\dagger U = 1$, and so

$$\begin{gathered} (U^\dagger)_{a_i'k_i} U_{k_i a_i} = \delta_{a_i'a_i} = U^*_{k_i a_i'} U_{k_i a_i} = U^*_{a_{i-1}'\sigma_i, a_i'} U_{a_{i-1}''\sigma_i, a_i} = \\ = A^{\sigma_i *}_{a_{i-1}'', a_i'} A^{\sigma_i}_{a_{i-1}'', a_i} = (A^{\sigma_i})^\dagger A^{\sigma_i} = \boxed{\sum_{\sigma_i} (A^{\sigma_i})^\dagger A^{\sigma_i} = 1} \end{gathered} \quad (154)$$

For right normalization, consider doing the operations of Eq. 151 "on the right":

$$\mathrm{Mat}_{\mathbb{C}}(d^L,1) \xrightarrow{\text{reshape}} \mathrm{Mat}_{\mathbb{C}}(d^{L-1},d) \xrightarrow{\text{SVD}} \mathrm{Mat}_{\mathbb{C}}(d^{L-1},r_1)\times \mathrm{Mat}_{\mathbb{C}}(r_1,r_1)\times \mathrm{Mat}_{\mathbb{C}}(r_1,d) \xrightarrow{\text{slice and multiply}} \mathrm{Mat}_{\mathbb{C}}(d^{L-1},r_1)\times \mathrm{Mat}_{\mathbb{C}}(r_1,1)^d$$

$$c^{\sigma_1\sigma_2\dots\sigma_L} \xmapsto{\text{reshape}} \Psi_{\sigma_1\dots\sigma_{L-1},\sigma_L} \xmapsto{=} U_{\sigma_1\dots\sigma_{L-1}a_1}S_{a_1a_1}V^\dagger_{a_1,\sigma_L} \xmapsto{=} \sum_{\sigma_L} c^{\sigma_1\dots\sigma_{L-1},a_1}B^{\sigma_L}_{a_1,1}$$

$$\mathrm{Mat}_{\mathbb{C}}(d^{L-1},r_1) \xrightarrow{\text{reshape}} \mathrm{Mat}_{\mathbb{C}}(d^{L-2},r_1d) \xrightarrow{\text{SVD}} \mathrm{Mat}_{\mathbb{C}}(d^{L-2},r_2)\times \mathrm{Mat}_{\mathbb{C}}(r_2,r_2)\times \mathrm{Mat}_{\mathbb{C}}(r_2,r_1d) \xrightarrow{\text{slice and multiply}} \mathrm{Mat}_{\mathbb{C}}(d^{L-2},r_2)\times \mathrm{Mat}_{\mathbb{C}}(r_2,r_1)^d$$

$$c^{\sigma_1\dots\sigma_{L-1}a_1} \xmapsto{\text{reshape}} \Psi_{\sigma_1\dots\sigma_{L-2},\sigma_{L-1}a_1} \xmapsto{=} U_{\sigma_1\dots\sigma_{L-2},a_2}S_{a_2a_2}V^\dagger_{a_2,\sigma_{L-1}a_1} \xmapsto{=} \sum_{\sigma_{L-1}} c^{\sigma_1\dots\sigma_{L-2},a_2}B^{\sigma_{L-1}}_{a_2,a_1}$$

$$\vdots$$

$$\boxed{\begin{gathered}\mathrm{Mat}_{\mathbb{C}}(d^{L-(i-1)},r_{i-1}) \xrightarrow{\text{reshape}} \mathrm{Mat}_{\mathbb{C}}(d^{L-i},r_{i-1}d) \xrightarrow{\text{SVD}} \mathrm{Mat}_{\mathbb{C}}(d^{L-i},r_i)\times \mathrm{Mat}_{\mathbb{C}}(r_i,r_i)\times \mathrm{Mat}_{\mathbb{C}}(r_i,r_{i-1}d) \xrightarrow{\text{slice and multiply}} \mathrm{Mat}_{\mathbb{C}}(d^{L-i},r_i)\times \mathrm{Mat}_{\mathbb{C}}(r_i,r_{i-1})^d \\ c^{\sigma_1\dots\sigma_{L-(i-1)}a_{i-1}} \xmapsto{\text{reshape}} \Psi_{\sigma_1\dots\sigma_{L-i},\sigma_{L-(i-1)}a_{i-1}} \xmapsto{=} U_{\sigma_1\dots\sigma_{L-i},a_i}S_{a_ia_i}V^\dagger_{a_i,\sigma_{L-(i-1)}a_{i-1}} \xmapsto{=} \sum_{\sigma_{L-(i-1)}} c^{\sigma_1\dots\sigma_{L-i},a_i}B^{\sigma_{L-(i-1)}}_{a_i,a_{i-1}}\end{gathered}} \tag{155}$$

Remember that $r_i \leq \min(d^{L-i},r_{i-1}d)$ and for $i=L$, just do reshape and slice and multiply operations.

Then, finally, the **right normalization** is derived and is such:

$$\begin{gathered} V^\dagger V = 1 \Longrightarrow \\ (V^\dagger V)_{a_ia_i'} = \delta_{a_ia_i'} = V^\dagger_{a_i,\sigma_{L-(i-1)}a_{i-1}}V_{\sigma_{L-(i-1)}a_{i-1},a_i'} = B^{\sigma_{L-(i-1)}}_{a_i,a_{i-1}}(V^\dagger)^\dagger_{\sigma_{L-(i-1)}a_{i-1},a_i'} = \\ = B^{\sigma_{L-(i-1)}}_{a_ia_{i-1}}(V^\dagger)^*_{a_i',\sigma_{L-(i-1)},a_{i-1}} = B^{\sigma_{L-(i-1)}}_{a_i,a_{i-1}}B^{\sigma_{L-(i-1)}}_{a_i',a_{i-1}} = B^{\sigma_{L-(i-1)}}_{a_ia_{i-1}}(B^\dagger)^{\sigma_{L-(i-1)}}_{a_{i-1}a_i'} \qquad \forall i = 1\dots L \\ \Longrightarrow \sum_{\sigma_{L-(i-1)}} B^{\sigma_{L-(i-1)}}(B^\dagger)^{\sigma_{L-(i-1)}} = 1 \end{gathered} \tag{156}$$

cf. Sec. 4, Matrix Product States (MPS) of Schollwöck [35].

Necessarily, given matrix $M \in \mathrm{Mat}_{\mathbb{K}}(M,N)$ (notation in Bridgeman and Chubb (2017) [33] and CUDA Toolkit Documentation; I will follow the notation in Schollwöck [35] since his A,B denote specific physical meaning).

For

$$U \in \mathrm{Mat}_{\mathbb{K}}(N_A, \min(N_A,N_B)) \text{ s.t. } UU^\dagger = 1$$

$$S \in \mathrm{Mat}_{\mathbb{K}}(\min(N_A,N_B),\min(N_A,N_B))$$

s.t. S diagonal with nonnegative $S_{aa}=s_a$, i.e. $S_{ij}=\delta_{ij}s_i$ s.t. $s_i \geq 0 \quad \forall i = 1,2,\dots\min(N_A,N_B)$.

$r \equiv$ (Schmidt) rank of $M :=$ number of nonzero singular values.

Assume $s_1 \geq s_2 \geq \dots \geq s_r \geq 0$.

$V^\dagger \in \mathrm{Mat}_{\mathbb{K}}(\min(N_A,N_B),N_B)$ s.t. $V^\dagger V = 1$.

$$\mathrm{Mat}_{\mathbb{K}}(N_A,N_B) \xrightarrow{\text{SVD}} U_{\mathbb{K}}(N_A,\min(N_A,N_B))\times \mathrm{diag}_{\mathbb{K}}(\min(N_A,N_B))\times U_{\mathbb{K}}(\min(N_A,N_B),N_B)$$

$$M \xmapsto{\text{SVD}} USV^\dagger$$

Optimal approximation of M (rank r by matrix M' (rank $r' < r$) property.

In Frobenius norm $\|M\|_F^2 := \sum_{i,j}|M_{ij}|^2$, induced by inner product $\langle M|N\rangle = \mathrm{tr}M^\dagger N$. Indeed,

$$\mathrm{tr}M^\dagger N = (M^\dagger)_{ik}N_{ki} = \overline{M}_{ki}N_{ki}$$

and so for

$$M' = US'V^\dagger, \qquad S' = \mathrm{diag}(s_1,s_2\dots s_{r'},0\dots) \tag{157}$$

cf. Eq. (19) of Schollwöck [35], i.e. 1 sets all but 1st r' singualr values to 0.

Use singular value decomposition (SVD) to derive Schmidt decomposition of general quantum state.

$\forall$ pure state $|\psi\rangle$ on AB,

$$|\psi\rangle = \sum_{i,j} \Psi_{ij} |i\rangle_A |j\rangle_B$$

where $\{|i\rangle_A\}, \{|j\rangle_B\}$ orthonormal bases of A, B ((complex) Hilbert spaces), with dim. N_A, N_B, respectively.

Let $\Psi_{i,j} \in \text{Mat}_{\mathbb{K}}(N_A, N_B)$.

Then **reduced density operators** $\widehat{\rho}_A, \widehat{\rho}_B$ are such that

$$\begin{aligned} \widehat{\rho}_A &= \text{tr}_B |\psi\rangle\langle\psi| \\ \widehat{\rho}_B &= \text{tr}_A |\psi\rangle\langle\psi| \end{aligned}$$

In matrix form,

$$\begin{aligned} \rho_A &= \Psi\Psi^\dagger \\ \rho_B &= \Psi^\dagger\Psi \end{aligned}$$

Indeed,

$$\begin{aligned} (\rho_A)_{ij} &= \Psi_{ik}\overline{\Psi}_{jk} \\ (\rho_B)_{ij} &= \overline{\Psi}_{ki}\Psi_{kj} \\ |\psi\rangle\langle\psi| &= \sum_{i,j} \Psi_{ij}|i\rangle_A|j\rangle_B \sum_{l,m} \overline{\Psi}_{lm}\langle l|_A \langle m|_B \\ \text{tr}_B|\psi\rangle\langle\psi| &= \sum_{i,j} \Psi_{ik}\overline{\Psi}_{jk} |i\rangle_A \rangle j|_A \end{aligned}$$

In matrix form,

$$\begin{aligned} \rho_A &= \Psi\Psi^\dagger \\ \rho_B &= \Psi^\dagger\Psi \end{aligned}$$

Carry out SVD on Ψ in Eq. (20) of Schollwöck [35],

$$|\psi\rangle = \sum_{i,j} \Psi_{ij}|i\rangle_A|j\rangle_B$$

$$|\psi\rangle = \sum_{ij} \Psi_{ij}|i\rangle_A|j\rangle_B = \sum_{ij} \sum_{a=1}^{\min(N_A,N_B)} U_{ia}S_{aa}\overline{V}_{ja}|i\rangle_A|j\rangle_B = \sum_{a=1}^{\min(N_A,N_B)} \sum_i U_{ia}|i\rangle_A s_a \sum_j \overline{V}_{ja}|j\rangle_B = \sum_{a=1}^{\min(N_A,N_B)} s_a|a\rangle_A|a\rangle_B$$

Due to orthogonality of $U, V^\dagger$, $\{|a\rangle_A\}, \{|a\rangle_B\}$ orthonormal, and can be extended to be orthonormal bases of A, B.

If we restrict the sum to run only over the $r \leq \min(N_A, N_B)$ positive nonzero singular values (i.e., for $\sum_{a=1}^{\min(N_A,N_B)}$, $_a > 0$ $\forall\, a \leq r$, and so

$$|\psi\rangle = \sum_{a=1}^{r} s_a|a\rangle_A|a\rangle_B$$

$r = 1$ (classical) product states. $|\psi\rangle = s_1|1\rangle_A|1\rangle_B$.
$r > 1$ entangled (quantum) states.

Schmidt decomposition on reduced density operators for A and B:

$$\begin{aligned} \widehat{\rho}_A &= \sum_{a=1}^{r} s_a^2 |a\rangle_A\langle a|_A \\ \widehat{\rho}_B &= \sum_{a=1}^{r} s_a^2 |a\rangle_B\langle a|_B \end{aligned}$$

Respective eigenvectors are left and right singular vectors.

Von Neumann entropy can be read off:

$$S_{A|B}(|\psi\rangle) = -\text{tr}\widehat{\rho}_A \log_2 \widehat{\rho}_A = -\sum_{a=1}^{r} s_a^2 \log_2 s_a^2$$

In view of large size of Hilbert spaces, approximate $|\psi\rangle$ by some $|\widetilde{\psi}\rangle$ spanned over state spaces A, B that have dims. r' only.

Since 2-norm of $|\psi\rangle$,

$$\| |\psi\rangle \|_2^2 = \sum_{ij} |\Psi_{ij}|^2 = \|\Psi\|_F^2$$

since

$$\| |\psi\rangle \|_2^2 = \sum_{a=1}^{r} s_a^2 = \sum_{ij} |\Psi_{ij}|^2$$

iff $\{|i\rangle\}, \{|j\rangle\}$ orthonormal. Optimal approx. of 2-norm given by optimal approx. of Ψ by $\overline{\Psi}$ in Frobenius norm, where $\overline{\Psi}$ is matrix of rank r'.

$\overline{\Psi} = US'V^\dagger$, $S' = \text{diag}(s_1, \dots s_{r'}, 0 \dots)$ from above.

$\Longrightarrow$ Schmidt decomposition of approximate state

(158) $$|\overline{\Psi}\rangle = \sum_{a=1}^{r'} s_a|a\rangle_A|a\rangle_B$$

cf. Eq. (27) of Schollwöck [35], where s_a must be rescaled if normalization desired.

45.4. **QR decomposition.** cf. 4.1.2. of Schollwöck [35].

If actual value of singular values not used explicitly, then use *QR decomposition*.

QR decomposition: $\forall\, M \in \text{Mat}_{\mathbb{K}}(N_A, N_B)$,

(159) $$M = QR,\ Q \in U_{\mathbb{K}}(N_A), \text{ i.e. } Q^\dagger Q = 1 = QQ^\dagger,\ R \in \text{Mat}_{\mathbb{K}}(N_A, N_B) \text{ s.t. upper triangular, i.e. } R_{ij} = 0 \text{ if } i > j$$

thin QR decomposition: assume $N_A > N_B$. Then bottom $N_A - N_B$ rows of R are 0, so

$$\begin{aligned} M = Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix} &= \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}\begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1R_1 \\ Q_1 &\in \text{Mat}_{\mathbb{K}}(N_A, N_B) \\ R_1 &\in \text{Mat}_{\mathbb{K}}(N_B, N_B) \end{aligned}$$

While $Q_1^\dagger Q_1 = 1$ in general $Q_1Q_1^\dagger \neq 1$

46. Matrix Product States (MPS)

cf. Section 4.13 Decomposition of arbitrary quantum states into MPS of Schollwöck [35].

Consider lattice of L sites, d-dim. local state spaces $\{\sigma_i\}_{i=1,\dots L}$.

Most general pure quantum state on lattice (assume normalized)

(160) $$|\psi\rangle = \sum_{\sigma_1\dots\sigma_L} c_{\sigma_1\dots\sigma_L}|\sigma_1\dots\sigma_L\rangle$$

cf. Eq. (30) of Schollwöck [35],

46.1. **Left-canonical matrix product state.** cf. Schollwöck [35],

Consider the process of refactoring or "flattening", which I claim to be a functor *flatten*:

$$|\psi\rangle \in \mathcal{H} \text{ s.t. } \dim \mathcal{H} = d^L \mapsto \Psi \in \text{Mat}_{\mathbb{K}}(d, d^{L-1})$$

$$\Psi_{\sigma_1,(\sigma_2\dots\sigma_L)} = c_{\sigma_1\dots\sigma_L}$$

(161) $$\xrightarrow{\text{SVD}} c_{\sigma_1\dots\sigma_L} = \Psi_{\sigma_1,(\sigma_2\dots\sigma_L)} = \sum_a^{r_1} U_{\sigma_1,a_1} S_{a_1,a_1} (V^\dagger)_{a_1,(\sigma_2\dots\sigma_L)} \equiv \sum_{a_1}^{r_1} U_{\sigma_1,a_1} c_{a_1,\sigma_2\dots\sigma_L}$$

i.e.

$$\begin{aligned} &(\mathbb{K}^d)^L \to \text{Mat}_{\mathbb{K}}(1,r) \times \text{Mat}_{\mathbb{K}}(r_1 d, d^{L-2}) \\ &c_{\sigma_1\dots\sigma_L} \mapsto A^{\sigma_1}_{a_1}, \Psi_{(a_1\sigma_2),(\sigma_3\dots\sigma_L)} \end{aligned}$$

s.t.

$$c_{\sigma_1\dots\sigma_L} = \sum_{a_1}^{r_1} A^{\sigma_1}_{a_1} \Psi_{(a_1\sigma_2),(\sigma_3\dots\sigma_L)}$$

where rank $r_1 \leq d$.

$$U \in \text{Mat}_{\mathbb{K}}(d, \min(d,r)) = \text{Mat}_{\mathbb{K}}(d,r)$$

Consider d row vectors A^{σ_1}, $A^{\sigma_1}_{a_1} = U_{\sigma_1,a_1}$.

$$c_{a_1\sigma_2\dots\sigma_L} = \sum_{a_1}^{r_1} A^{\sigma_1}_{a_1} \Psi_{(a_1,\sigma_2),(\sigma_3\dots\sigma_L)} \text{ with}$$

$$\Psi_{(a_1\sigma_2),(\sigma_3\dots\sigma_L)} \in \text{Mat}_{\mathbb{K}}(r_1 d, d^{L-2})$$

So from Eq. (34) of Schollwöck [35],

(162) $$c_{\sigma_1\dots\sigma_L} = \sum_{a_1}^{r_1}\sum_{a_2}^{r_2} A^{\sigma_1}_{a_1} U_{(a_1\sigma_2),a_2} S_{a_2,a_2} (V^\dagger)_{a_2,(\sigma_3\dots\sigma_L)} = \sum_{a_1}^{r_1}\sum_{a_2}^{r_2} A^{\sigma_1}_{a_1} A^{\sigma_2}_{a_1,a_2} \Psi_{(a_2\sigma_3),(\sigma_4\dots\sigma_L)}$$

So for

$$U \in \text{Mat}_{\mathbb{K}}(d, r_1 \times r_2) \mapsto \{A^{\sigma_2}\}_{\sigma_2}, \qquad |\{A^{\sigma_2}\}_{\sigma_2}| = d, \qquad A^{\sigma_2} \in \text{Mat}_{\mathbb{K}}(r_1, r_2)$$

$A^{\sigma_2}_{a_1,a_2} = U_{(a_1,\sigma_2),a_2}$ and multiplied S and $V^\dagger$,

$$SV^\dagger \mapsto \Psi \in \text{Mat}_{\mathbb{K}}(r_2 d, d^{L-3}); \qquad r_2 \leq r_1 d \leq d^2$$

and so continuing the application of SVD and refactoring (what I call applying the *flatten* functor)

$$\xrightarrow{\text{SVD}} c_{\sigma_1\dots\sigma_L} = \sum_{a_1\dots a_{L-1}} A^{\sigma_1}_{a_1} A^{\sigma_2}_{a_1 a_2} \dots A^{\sigma_{L-1}}_{a_{L-2},a_{L-1}} A^{\sigma_L}_{a_L-1} \equiv A^{\sigma_1} A^{\sigma_2} \dots A^{\sigma_{L-1}} A^{\sigma_L}$$

46.1.1. *Matrix Product State (definition).*

Definition 124 (Matrix Product State)**.**

(163) $$|\psi\rangle = \sum_{\sigma_1\dots\sigma_L} A^{\sigma_1} A^{\sigma_2} \dots A^{\sigma_{L-1}} A^{\sigma_L} |\sigma_1 \dots \sigma_L\rangle$$

Maximally, the dims. are

$$(1 \times d), (d \times d^2) \dots (d^{L/2-1} \times d^{L/2}), (d^{L/2} \times d^{L/2-1}) \dots (d^2 \times d), (d \times 1)$$

Since $\forall$ SVD, $U^\dagger U = 1$,

$$\delta_{a_l,a_l'} = \sum_{a_{l-1}a_l} (U^\dagger)_{a_l,(a_{l-1}\sigma_l)} U_{(a_{l-1}\sigma_l),a_l'} = \sum_{a_{l-1}\sigma_l} (A^{\sigma_l})^\dagger_{a_l,a_{l-1}} A^{\sigma_l}_{a_{l-1},a_l'} = \sum_{\sigma_l} ((A^{\sigma_2})^\dagger A^{\sigma_l})_{a_l,a_l'}$$

or

(164) $$\sum_{\sigma_l} (A^{\sigma_l})^\dagger A^{\sigma_l} = 1$$

cf. Eq. (38) of Schollwöck [35],

If for $\{A^{\sigma_l}\}_{\sigma_l}$, $\sum_{\sigma_l} (A^{\sigma_l})^\dagger A = 1$, $\{A^{\sigma_l}\}_{\sigma_l}$ are **left-normalized**; matrix product states that consist of only left-normalized matrices are **left-canonical**.

View Density Matrix Renormalization Group (DMRG) decomposition of universe into blocks A and B, split lattice into parts A,B, where A compries sites 1 through l and B sites $l+1$ through L.

$$\begin{aligned} |a_l\rangle_A &= \sum_{\sigma_1\dots\sigma_l} (A^{\sigma_1} A^{\sigma_2} \dots A^{\sigma_l})_{a_l,1} |\sigma_1 \dots \sigma_l\rangle \\ |a_l\rangle_B &= \sum_{\sigma_{l+1}\dots\sigma_L} (A^{\sigma_{l+1}} A^{\sigma_{l+2}} \dots A^{\sigma_L})_{a_l,1} |\sigma_{l+1} \dots \sigma_L\rangle \end{aligned}$$

s.t. matrix product state (MPS) is

$$|\psi\rangle = \sum_{a_l} |a_l\rangle_A |a_l\rangle_B$$

46.1.2. *Summarize this procedure of constructing, from a pure state, the matrix product state (version) by successive application Singular Value Decomposition (SVD) from the Category Theory point of view.* Consider all applications of SVD to get to a matrix

$$(\mathbb{K}^d)^L \xrightarrow{\text{SVD}} (\text{Mat}_{\mathbb{K}}(1,r_1))^d \times (\text{Mat}_{\mathbb{K}}(r_1,r_2))^d \times \dots \times (\text{Mat}_{\mathbb{K}}(r_{L-2},r_{L-1}))^d \times (\text{Mat}_{\mathbb{K}}(r_{L-1},1))^d$$

$$c_{\sigma_1\dots\sigma_L} \xmapsto{\text{SVD}} c_{\sigma_1\dots\sigma_L} = \textstyle\sum_{a_1\dots a_{L-1}} A^{\sigma_1}_{a_1} A^{\sigma_2}_{a_1 a_2} \dots A^{\sigma_{L-1}}_{a_{L-2},a_{L-1}} A^{\sigma_L}_{a_{L-1}}$$

product state (MPS):

and remember the maximal values that the r_i's can take:

$$\begin{array}{lll} r_1 \leq d & r_{L/2} \leq d^{L/2} & r_{L-2} \leq d^2 \\ r_2 \leq d^2 & r_{L/2+1} \leq d^{L/2-1} & r_{L-1} \leq d \end{array}$$

Let us explicitly note the functors (that were applied) flatten (and its inverse), and the application of SVD, explicitly:

$$(\mathbb{K}^d)^L \xrightarrow{\text{flatten}^{-1}} \text{Mat}_{\mathbb{K}}(d, d^{L-1}) \xrightarrow{\text{SVD}} U_{\mathbb{K}}(d, r_1) \times \text{diag}_{\mathbb{K}}(r_1) \times U_{\mathbb{K}}(r_1, d^{L-1}) \xrightarrow{\cong} (\text{Mat}_{\mathbb{K}}(1, r_1))^d \times \text{Mat}_{\mathbb{K}}(r_1 d, d^{L-2}) \xrightarrow{\text{flatten}} (\text{Mat}_{\mathbb{K}}(1, r_1))^d \times (\mathbb{K}^{r_1}) \times (\mathbb{K}^d)^{L-1}$$

$$c_{\sigma_1 \dots \sigma_L} \xmapsto{\text{flatten}^{-1}} c_{\sigma_1 \dots \sigma_L} = \Psi_{\sigma_1, (\sigma_2 \dots \sigma_L)} \xmapsto{\text{SVD}} \Psi_{\sigma_1, (\sigma_2 \dots \sigma_L)} = \sum_{a_1}^{r_1} U_{\sigma_1 a_1} S_{a_1, a_1} (V^\dagger)_{a_1, (\sigma_2 \dots \sigma_L)} \xmapsto{\cong} c_{a_1 \sigma_2 \dots \sigma_L} = \sum_{a_1}^{r_1} A_{a_1}^{\sigma_1} \Psi_{(a_1, a_2), (\sigma_3 \dots \sigma_L)} \xmapsto{\text{flatten}} c_{a_1 \sigma_2 \dots \sigma_L} = \sum_{a_1}^{r_1} A_{a_1}^{\sigma_1} c_{a_1 \sigma_2 \dots \sigma_L}$$

with $\cong$ in this case denoting an isomorphism (clearly).

In considering some kind of recursive algorithm, so to repeat some series of steps until a matrix product state is obtained, consider this:

$$\begin{aligned} (\mathbb{K}^d)^L &\longrightarrow (\text{Mat}_{\mathbb{K}}(1, r_1))^d \times \mathbb{K}^{r_1} \times (\mathbb{K}^d)^{L-1} \\ c_{\sigma_1 \dots \sigma_L} &\longmapsto c_{\sigma_1 \dots \sigma_L} = \sum_{a_1}^{r_1} A_{a_1}^{\sigma_1} c_{a_1 \sigma_2 \dots \sigma_L} \end{aligned}$$

So in summary, to obtain matrix product states, starting from a matrix,

$$\boxed{\begin{aligned} &\text{Mat}_{\mathbb{K}}(d, d^{L-1}) \longrightarrow (\text{Mat}_{\mathbb{K}}(1, r_1))^d \times \text{Mat}_{\mathbb{K}}(r_1 d, d^{L-2}) \longrightarrow \dots \longrightarrow (\text{Mat}_{\mathbb{K}}(1, r_1))^d \times (\text{Mat}_{\mathbb{K}}(r_1, r_2))^d \times \dots \times (\text{Mat}_{\mathbb{K}}(r_{n-1}, r_n))^d \times (\text{Mat}_{\mathbb{K}}(r_n d, d^{L-(n+1)}))^d \\ &\Psi_{\sigma_1, (\sigma_2 \dots \sigma_L)} \longmapsto \sum_{a_1}^{r_1} A_{a_1}^{\sigma_1} \Psi_{(a_1, \sigma_2), (\sigma_3 \dots \sigma_L)} \longmapsto \dots \longmapsto \sum_{a_1, a_2, \dots a_n}^{r_1, r_2, \dots r_n} A_{a_1}^{\sigma_1} A_{a_1 a_2}^{\sigma_2} \dots A_{a_{n-1} a_n}^{\sigma_n} \Psi_{(a_n \sigma_{n+1}), (\sigma_{n+2} \dots \sigma_L)} \end{aligned}} \tag{165}$$

46.2. **Right-canonical matrix product state.** cf. Schollwöck [35],

We can start from right in order to obtain

$$c_{\sigma_1 \dots \sigma_L} = \Psi_{(\sigma_1 \dots \sigma_{L-1}), \sigma_L} = \sum_{a_{L-1}} U_{(\sigma_1 \dots \sigma_{L-1}), a_{L-1}} S_{a_{L-1}, a_{L-1}} (V^\dagger)_{a_{L-1}, \sigma_L} = \sum_{a_{L-1}} \Psi_{(\sigma_1 \dots \sigma_{L-2}), (\sigma_{L-1} a_{L-1})} B_{a_{L-1}}^{\sigma_L} =$$
$$= \sum_{a_{L-1}, a_{L-2}} U_{(\sigma_1 \dots \sigma_{L-2}), a_{L-2}} S_{a_{L-2}, a_{L-2}} (V^\dagger)_{a_{L-2}, (\sigma_{L-1} a_{L-1})} B_{a_{L-1}}^{\sigma_L} = \sum_{a_{L-2}, a_{L-1}} \Psi_{(\sigma_1 \dots \sigma_{L-3}), (\sigma_{L-2} a_{L-2})} B_{a_{L-2}, a_{L-1}}^{\sigma_{L-1}} B_{a_{L-1}}^{\sigma_L} = \dots$$

or consider

$$(\mathbb{K}^d)^L \xrightarrow{\text{flatten}^{-1}} \text{Mat}_{\mathbb{K}}(d^{L-1}, d) \xrightarrow{\text{SVD}} U_{\mathbb{K}}(d^{L-1}, r_{L-1}) \times \text{diag}_{\mathbb{K}}(r_{L-1}) \times U_{\mathbb{K}}(r_{L-1}, d) \xrightarrow{\cong} \text{Mat}_{\mathbb{K}}(d^{L-2}, d r_{L-1}) \times (\text{Mat}_{\mathbb{K}}(r_{L-1}, 1))^d \xrightarrow{\text{SVD}}$$

$$c_{\sigma_1 \dots \sigma_L} \xmapsto{\text{flatten}^{-1}} c_{\sigma_1 \dots \sigma_L} = \Psi_{(\sigma_1 \dots \sigma_{L-1}), \sigma_L} \xmapsto{\text{SVD}} c_{\sigma_1 \dots \sigma_L} = \sum_{a_{L-1}}^{r_{L-1}} U_{(\sigma_1 \dots \sigma_{L-1}), a_{L-1}} S_{a_{L-1}, a_{L-1}} (V^\dagger)_{a_{L-1}, \sigma_L} \xmapsto{\cong} \begin{gathered} U_{(\sigma_1 \dots \sigma_{L-1}), a_{L-1}} S_{a_{L-1}, a_{L-1}} = \Psi_{(\sigma_1 \dots \sigma_{L-2}), (\sigma_{L-1} a_{L-1})} \\ (V^\dagger)_{a_{L-1}, \sigma_L} = B_{a_{L-1}}^{\sigma_L} \\ c_{\sigma_1 \dots \sigma_L} = \sum_{a_{L-1}} \Psi_{(\sigma_1 \dots \sigma_{L-2}), (\sigma_{L-1}, a_{L-1})} B_{a_{L-1}}^{\sigma_L} \end{gathered} \xmapsto{\text{SVD}}$$

$$\xrightarrow{\text{SVD}} U_{\mathbb{K}}(d^{L-2}, r_{L-2}) \times \text{diag}_{\mathbb{K}}(r_{L-2}) \times U_{\mathbb{K}}(r_{L-2}, d r_{L-1}) \times (\text{Mat}_{\mathbb{K}}(r_{L-1}, 1))^d \xrightarrow{\cong} \text{Mat}_{\mathbb{K}}(d^{L-3}, d r_{L-2}) \times (\text{Mat}_{\mathbb{K}}(r_{L-2}, r_{L-1}))^d \times (\text{Mat}_{\mathbb{K}}(r_{L-1}, 1))^d$$

$$\xmapsto{\text{SVD}} c_{\sigma_1 \dots \sigma_L} = \sum_{a_{L-1}, a_{L-2}} U_{(\sigma_1 \dots \sigma_{L-2}), a_{L-2}} S_{a_{L-2}, a_{L-2}} (V^\dagger)_{a_{L-2}, (\sigma_{L-1} a_{L-1})} B_{a_{L-1}}^{\sigma_L} \xmapsto{\cong} \begin{gathered} U_{(\sigma_1 \dots \sigma_{L-2}), a_{L-2}} S_{a_{L-2}, a_{L-2}} = \Psi_{(\sigma_1 \dots \sigma_{L-3}), (\sigma_{L-2} a_{L-2})} \\ (V^\dagger)_{a_{L-2}, (\sigma_{L-1} a_{L-1})} = B_{a_{L-2} a_{L-1}}^{\sigma_{L-1}} \\ c_{\sigma_1 \dots \sigma_L} = \sum_{a_{L-1}, a_{L-2}} \Psi_{(\sigma_1 \dots \sigma_{L-3}), (\sigma_{L-2}, a_{L-2})} B_{a_{L-2}, a_{L-1}}^{\sigma_{L-1}} B_{a_{L-1}}^{\sigma_L} \end{gathered}$$

with $\cong$ in this case denoting an isomorphism (clearly).

And so we can explicitly state the recursion step, for the purpose of writing numerical implementations/algorithms: $\forall\, l = 1, 2 \dots L$,

$$\begin{aligned} \text{Mat}_{\mathbb{K}}(d^{L-l}, dr_{L-(l-1)}) &\longrightarrow \text{Mat}_{\mathbb{K}}(d^{L-(l+1)}, dr_{L-l}) \times (\text{Mat}_{\mathbb{K}}(r_{L-l}, r_{L-(l-1)}))^d \\ \Psi_{(\sigma_1\dots\sigma_{L-l}),(\sigma_{L-(l-1)}a_{L-(l-1)})} &\longmapsto \Psi_{(\sigma_1\dots\sigma_{L-l}),(\sigma_{L-(l-1)}a_{L-(l-1)})} = \textstyle\sum_{a_{L-l}} \Psi_{(\sigma_1\dots\sigma_{L-(l+1)}),(\sigma_{L-l}a_{L-l})} B^{\sigma_{L-(l-1)}}_{a_{L-l},a_{L-(l-1)}} \end{aligned}$$

and we finally obtained, after successive applications SVD, the matrix product state:

$$\begin{aligned} (\mathbb{K}^d)^L \longrightarrow \text{Mat}_{\mathbb{K}}(d^{L-1}, d) &\longrightarrow (\text{Mat}_{\mathbb{K}}(1, r_1))^d \times (\text{Mat}_{\mathbb{K}}(r_1, r_2))^d \times \dots \times (\text{Mat}_{\mathbb{K}}(r_{L-2}, r_{L-1}))^d \times (\text{Mat}_{\mathbb{K}}(r_{L-1}, 1))^d \\ c_{\sigma_1\dots\sigma_L} \longmapsto \Psi_{(\sigma_1\dots\sigma_{L-l}),\sigma_L} &\longmapsto c_{\sigma_1\dots\sigma_L} = \textstyle\sum_{a_1\dots a_{L-1}} B^{\sigma_1}_{a_1} B^{\sigma_2}_{a_1 a_2} \dots B^{\sigma_{L-1}}_{a_{L-2}a_{L-1}} B^{\sigma_L}_{a_{L-1}} \end{aligned}$$

Since

$$V^\dagger V = 1 \tag{166}$$

, then

$$\delta_{a_l a_l'} = \sum_{\sigma_m a_m} (V^\dagger)_{a_l(\sigma_m a_m)} V_{(\sigma_m a_m) a_l'} = \sum_{\sigma_m a_m} B^{\sigma_m}_{a_l a_m} \overline{B}^{\sigma_m}_{a_l' a_m} \Longrightarrow \sum_{\sigma_m} \boxed{B^{\sigma_m}(B^{\sigma_m})^\dagger = 1} \tag{167}$$

The B-matrices that obey this condition are referred to as **right-normalized** matrices. A matrix product state (MPS) entirely consisting of a product of these right-normalized matrices is called **right-canonical**.

46.3. **Matrix Product Operators (MPO).** The form of a general operator, $\widehat{O}$ is the following:

$$\widehat{O} = \sum_{\{\sigma\}} \sum_{\{\sigma'\}} c^{\sigma_1\dots\sigma_L,\sigma_1'\dots\sigma_L'} |\sigma_1 \dots \sigma_L\rangle\langle\sigma_1' \dots \sigma_L'| \in \mathcal{H} \otimes \mathcal{H}^* \tag{168}$$

with $\dim\mathcal{H} = \dim\mathcal{H}^* = d^L$.

For MPO, do the same decomposition as done in Eq. 151 or in **??**, but with the double index $\sigma_i\sigma_i'$ taking the role of index σ_i in MPS (i.e. do this substitution and the decomposition will proceed *exactly* as before).

$$
\begin{tikzcd}
\mathrm{Mat}_{\mathbb{C}}(d^L, d^L) \arrow[d, "\text{reorder}"] & & & \\
\mathrm{Mat}_{\mathbb{C}}(1, (d^2)^L) \arrow[r, "\text{reshape}"] & \mathrm{Mat}_{\mathbb{C}}(d^2, (d^2)^{L-1}) \arrow[r, "\text{SVD}"] & \mathrm{Mat}_{\mathbb{C}}(d^2, r_i) \times \mathrm{Mat}_{\mathbb{C}}(r_i, r_i) \times \mathrm{Mat}_{\mathbb{C}}(r_i, (d^2)^{L-1}) \arrow[r, "\text{slice and multiply}"] & \mathrm{Mat}_{\mathbb{C}}(1, r_i)^d \times \mathrm{Mat}_{\mathbb{C}}(r_i, (d^2)^{L-1})
\end{tikzcd}
$$

$$
\begin{tikzcd}
c^{\sigma_1 \dots \sigma_L, \sigma_1' \dots \sigma_L'} \arrow[d, mapsto, "\text{reorder}"] & & & \\
c^{\sigma_1 \sigma_1' \sigma_2 \sigma_2' \dots \sigma_L \sigma_L'} \arrow[r, mapsto, "\text{reshape}"] & \Psi_{\sigma_1 \sigma_1', (\sigma_2 \sigma_2', \sigma_3 \sigma_3' \dots \sigma_L \sigma_L')} \arrow[r, mapsto, "="] & U_{\sigma_1 \sigma_1', a_1} S_{a_1 a_1} V^\dagger_{a_1, (\sigma_2 \sigma_2' \dots \sigma_L \sigma_L'} \arrow[r, mapsto, "="] & \sum_{\sigma_1 \sigma_1'} A^{\sigma_1 \sigma_1'}_{1, a_1} c^{a_1, \sigma_2 \sigma_2' \dots \sigma_L \sigma_L'}
\end{tikzcd}
$$

$$
\boxed{
\begin{tikzcd}
\mathrm{Mat}_{\mathbb{C}}(r_{i-1}, (d^2)^{L-(i-1)}) \arrow[r, "\text{reshape}"] & \mathrm{Mat}_{\mathbb{C}}(r_{i-1} d^2, (d^2)^{L-i}) \arrow[r, "\text{SVD}"] & \mathrm{Mat}_{\mathbb{C}}(r_{i-1} d^2, r_i) \times \mathrm{Mat}_{\mathbb{C}}(r_i, r_i) \times \mathrm{Mat}_{\mathbb{C}}(r_i, (d^2)^{L-i}) \arrow[r, "\text{slice and multiply}"] & \mathrm{Mat}_{\mathbb{C}}(r_{i-1}, r_i)^{d^2} \times \mathrm{Mat}_{\mathbb{C}}(r_i, (d^2)^{L-i}) \\
c^{a_{i-1}, \sigma_i \sigma_i' \sigma_{i+1} \sigma_{i+1}' \dots \sigma_L \sigma_L'} \arrow[r, mapsto, "\text{reshape}"] & \Psi_{a_{i-1} \sigma_i \sigma_i', (\sigma_{i+1} \sigma_{i+1}' \sigma_{i+2} \sigma_{i+2}' \dots \sigma_L \sigma_L')} \arrow[r, mapsto, "="] & U_{a_{i-1} \sigma_i \sigma_i', a_i} S_{a_i a_i} V^\dagger_{a_i, \sigma_{i+1} \sigma_{i+1}' \dots \sigma_L \sigma_L'} \arrow[r, mapsto, "="] & \sum_{\sigma_i \sigma_i'}^{max} A^{\sigma_i \sigma_i'}_{a_{i-1}, a_i} c^{a_i, \sigma_{i+1} \sigma_{i+1}' \dots \sigma_L \sigma_L'}
\end{tikzcd}
}
\tag{169}
$$

46.3.1. *Numerical implementation; both in BLAS and cuBLAS.* As stated in the CUDA Toolkit Documentation v8.0 for cuSOLVER, under section 5.3.6. `cusolverDn<t>gesvd()` and Remark 1, `gesvd` "only supports" `m>=n`, for matrix you want to decompose $A \in \mathrm{Mat}_{\mathbb{K}}(m, n)$. So number of rows must be greater than or equal to number of columns. And so we can only consider right-normalized matrices in a practical implementation.

I suspect it's the same in BLAS.

Consider the very first step, $l = 1$, in a procedure to calculate the matrix product state.

$$
\mathrm{Mat}_{\mathbb{K}}(d^{L-1}, d) \xrightarrow{\text{SVD}} U_{\mathbb{K}}(d^{L-1}, r_{L-1}) \times \mathrm{diag}_{\mathbb{K}}(r_{L-1}) \times U_{\mathbb{K}}(r_{L-1}, d) \xrightarrow{\cong} \mathrm{Mat}_{\mathbb{K}}(d^{L-2}, d r_{L-1}) \times (\mathrm{Mat}_{\mathbb{K}}(r_{L-1}, 1))^d
$$

$$
\Psi_{(\sigma_1 \dots \sigma_{L-1}), \sigma_L} \xmapsto{\text{SVD}} = \sum_{a_{L-1}}^{r_{L-1}} U_{(\sigma_1 \dots \sigma_{L-1}), a_{L-1}} S_{a_{L-1}, a_{L-1}} (V^\dagger)_{a_{L-1}, \sigma_L} \xmapsto{\cong} \begin{aligned} U_{(\sigma_1 \dots \sigma_{L-1}), a_{L-1}} S_{a_{L-1}, a_{L-1}} &= \Psi_{(\sigma_1 \dots \sigma_{L-2}), (\sigma_{L-1} a_{L-1})} \\ (V^\dagger)_{a_{L-1}, \sigma_L} &= B^{\sigma_L}_{a_{L-1}} \end{aligned}
$$

with $\cong$ in this case denoting an isomorphism, the *reshaping* of a matrix into different matrix size dimensions, which should be the inverse of a "flatten" functor, which I'll denote as flatten^{-1} as well (and this is this same isomorphism we're talking about).

Let's deal with the specific procedure of flatten^{-1}, how it reshapes indices in accordance with different matrix size dimensions, and with the so-called "stride" when going from, say, 2-dimensional indices to a "flattened" 1-dimensional index.

Note also as a practical numerical implementation design point, LAPACK's linear algebra BLAS library package and CUBLAS assumes *column*-major ordering.

Consider $i = 1, 2, \dots L-1$ (for site i) (or for 0-based counting, starting to count from 0, $i = 0, 1, \dots L-2$; be aware of this difference as in practical numerical implementation, in C, C++, Python, it assumes 0-based counting).

For a state space of dimension d, we can consider the specific example of $d = 2$, representing say a spin-1/2 system. Then index σ_i can be 0 or 1: $\sigma_i \in \{0, 1\}$. In general, $\sigma_i \in \{0, 1, \dots d-1\}$. I may use d or 2 in the context of the number of states (basis vectors) of the spin system (state vector space).

Consider site i. Suppose the spin system there interacts most with sites $i-1$, $i+1$, and then next sites $i-2$, $i+2$, etc. So the values at $\sigma_{i-1}, \sigma_{i+1}$, etc. are most important in calculating interactions with spin system at site i.

Then we seek this reshaping of the matrix index - assuming 0-based counting/ordering, for $l = 1$:

$$
\{0, 1\}^{L-1} \xrightarrow{(\text{flatten})^{-1}} \{0, 1, \dots 2^{L-1} - 1\}
$$

$$
(\sigma_0, \sigma_1, \dots \sigma_{L-2}) \xmapsto{(\text{flatten})^{-1}} I_{L-1} := \sigma_0 + 2\sigma_1 + \dots + 2^i \sigma_i + \dots + 2^{L-2} \sigma_{L-2} = \sum_{i=0}^{L-2} 2^i \sigma_i
$$

In this way, states of a site i are closest in memory addresses in the allocation of a 1-dim. array, on CPU or GPU memory, so that memory access operations should be efficient.

Assuming SVD doesn't change the striding, and defining the result of matrix multiplication:

$$
U_{(\sigma_0, \sigma_1 \dots \sigma_{L-2}), a_{L-1}} S_{a_{L-1}, a_{L-1}} =: (US)_{(\sigma_0 \dots \sigma_{L-2}), a_{L-1}} \in \mathrm{Mat}_{\mathbb{K}}(d^{L-1}, r_{L-1})
$$

We can reshape (i.e. $(\text{flatten})^{-1}$) in such a manner:

$$\mathrm{Mat}_{\mathbb{K}}(d^{L-1}, r_{L-1}) \xrightarrow{(\text{flatten})^{-1}} \mathrm{Mat}_{\mathbb{K}}(d^{L-2}, dr_{L-1})$$

$$(US)_{(\sigma_0 \dots \sigma_{L-2}),a_{L-1}} \xmapsto{(\text{flatten})^{-1}} \Psi_{(\sigma_0,\sigma_1,\dots\sigma_{L-3}),(\sigma_{L-2}a_{L-1})}$$

$$\{0,1,\dots 2^{L-1}-1\} \times \{0,1,\dots r_{L-1}-1\} \xrightarrow{(\text{flatten})^{-1}} \{0,1,\dots 2^{L-2}-1\} \times \{0,1,\dots dr_{L-1}-1\}$$

$$I_{L-1}, a_{L-1} \xmapsto{(\text{flatten})^{-1}} I_{L-1} \mod 2^{L-2}, \tfrac{I_{L-1}}{2^{L-2}} + da_{L-1}$$

Reshaping $V^\dagger$ at iteration $l=1$ can be done as follows:

$$U_{\mathbb{K}}(r_{L-1}, d) \xrightarrow{(\text{flatten})^{-1}} (\mathrm{Mat}_{\mathbb{K}}(r_{L-1}, 1))^d$$

$$(V^\dagger)_{a_{L-1},\sigma_{L-1}} \xmapsto{(\text{flatten})^{-1}} (V^\dagger)_{a_{L-1},\sigma_{L-1}} = B^{\sigma_{L-1}}_{a_{L-1}}$$

$$\{0,1,\dots r_{L-1}-1\} \times \{0,1,\dots d-1\} \xrightarrow{(\text{flatten})^{-1}} (\{0,1,\dots r_{L-1}-1\})^d$$

$$a_{L-1}, \sigma_{L-1} \xmapsto{(\text{flatten})^{-1}} a_{L-1}$$

Let's do this same procedure, reshaping or $(\text{flatten})^{-1}$, for a general l iteration.

$$\mathrm{Mat}_{\mathbb{K}}(d^{L-l}, r_{L-l}) \xrightarrow{(\text{flatten})^{-1}} \mathrm{Mat}_{\mathbb{K}}(d^{L-(l+1)}, dr_{L-l})$$

$$(US)_{(\sigma_0 \dots \sigma_{L-(l+1)}),a_{L-l}} \xmapsto{(\text{flatten})^{-1}} \Psi_{(\sigma_0,\sigma_1,\dots\sigma_{L-(l+2)}),(\sigma_{L-(l+1)}a_{L-l})}$$

$$\{0,1,\dots d^{L-l}-1\} \times \{0,1,\dots r_{L-l}-1\} \xrightarrow{(\text{flatten})^{-1}} \{0,1,\dots d^{L-(l+1)}-1\} \times \{0,1,\dots dr_{L-l}-1\}$$

$$I_{L-l}, a_{L-l} \xmapsto{(\text{flatten})^{-1}} I_{L-l} \mod d^{L-(l+1)}, \tfrac{I_{L-l}}{d^{L-(l+1)}} + da_{L-l}$$

$$U_{\mathbb{K}}(r_{L-l}, dr_{L-(l-1)}) \xrightarrow{(\text{flatten})^{-1}} (\mathrm{Mat}_{\mathbb{K}}(r_{L-l}, r_{L-(l-1)}))^d$$

$$(V^\dagger)_{a_{L-l},(\sigma_{L-l}a_{L-(l-1)})} \xmapsto{(\text{flatten})^{-1}} (V^\dagger)_{a_{L-l},(\sigma_{L-l}a_{L-(l-1)})} = B^{\sigma_{L-l}}_{a_{L-l},a_{L-(l-1)}}$$

$$\{0,1,\dots r_{L-l}-1\} \times \{0,1,\dots dr_{L-(l-1)}-1\} \xrightarrow{(\text{flatten})^{-1}} (\{0,1,\dots r_L-1\} \times \{0,1,\dots r_{L-(l-1)}-1\})^d$$

$$a_{L-l}, (\sigma_{L-l}a_{L-(l-1)}) := a_{L-l}, \sigma_{L-1} + da_{L-(l-1)} \xmapsto{(\text{flatten})^{-1}} a_{L-l}, \tfrac{(\sigma_{L-1}a_{L-(l-1)})}{d}; \sigma_{L-l} = (\sigma_{L-l}a_{L-(l-1)}) \mod d$$

46.3.2. *Numerical implementations of initial states.* Something else that shouldn't be overlooked is the numerical implementation of *initial states*, the c's of a state $|\psi\rangle = \sum_{\{\sigma\}} c^\sigma |\{\sigma\}\rangle$ for a many-body quantum system. Remember what the postulates of quantum mechanics say and interpret accordingly (and correctly). While we call them "probability amplitudes", one should be careful about what physical interpretation we may (or may not!) assign them. One thing's for certain: $c \in \mathbb{C}$ and normalization of the quantum state: $|\langle\psi|\psi\rangle|^2 = 1$

Here are some setups to try:
$d=2, L=2,\ d^L = 2^2 = 4$.

$$\begin{bmatrix} c_{\uparrow\uparrow} & c_{\uparrow\downarrow} & c_{\downarrow\uparrow} & c_{\downarrow\downarrow} \end{bmatrix} \mapsto \begin{bmatrix} c_{\uparrow\uparrow} & c_{\uparrow\downarrow} \\ c_{\downarrow\uparrow} & c_{\downarrow\downarrow} \end{bmatrix}$$

Singlet state: $|\psi\rangle = \frac{1}{\sqrt{2}}|\uparrow\downarrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\uparrow\rangle$, $\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & 0 \end{bmatrix}$
$d=2, L=3, d^L = 2^8 = 8$
For notational convenience, let $\uparrow \equiv 1$, $\downarrow \equiv 0$

$$\begin{bmatrix} c_{000} & c_{001} & c_{010} & c_{011} & c_{100} & \dots & c_{111} \end{bmatrix} \mapsto \begin{bmatrix} c_{000} & c_{001} & \dots & c_{011} \\ c_{100} & c_{001} & \dots & c_{111} \end{bmatrix}$$

$d=3, L=2,\ d^L = 3^2 = 9$

$$\begin{bmatrix} c_{-1-1} & c_{-10} & c_{-11} & \dots & c_{11} \end{bmatrix} \mapsto \begin{bmatrix} c_{-1-1} & c_{-10} & c_{-11} \\ c_{0-1} & c_{00} & c_{01} \\ c_{1-1} & c_{10} & c_{11} \end{bmatrix}$$

Part 12. **Algebraic Geometry**

47. Affine and Projective Varieties

cf. Harris (1992)[43]
For (algebraically closed) field K,
vector space K^n,
affine space $\mathbb{A}^n_K \equiv \mathbb{A}^n = K^n$, but origin plays no special role in affine space.
Affine variety $X \subset \mathbb{A}^n :=$ common zero locus of collection of polynomials $f_\alpha \in K[z_1 \dots z_n] :=$

$$X = \{Z | f_\alpha(Z) = 0 \quad \forall \alpha, \quad f_\alpha \in K(z_1 \dots z_n), Z = (z_1 \dots z_n)\}$$

47.1. **Projective Space and Projective Varieties.** Projective space over field K = set of 1-dim. subspaces of vector space $K^{n+1} \equiv \mathbb{P}^n_K \equiv \mathbb{P}^n = (K^{n+1} - \{0\})/K^*$,
where $(K^{n+1} - \{0\})/K^*$ is the quotient of $K^{n+1} - \{0\}$ by the action of the group K^n acting by scalar multiplication.

$\mathbb{P}(V) \equiv \mathbb{P}V \equiv$ projective space of 1-dim. subspaces of a vector space V over field K.

$P \in \mathbb{P}^n$ usually written as homogeneous vector $[Z_0 \dots Z_n]$, by which be mean line spaced by $(Z_0 \dots Z_n) \in K^{n+1}$.

For U_n s.t. $\forall P \in U_n \subset \mathbb{P}^n \subset V^{n+1}$, $Z_n \neq 0$. Then $[Z_0 \dots Z_n] \sim \left[\frac{Z_0}{Z_n}, \dots, \frac{Z_{n-1}}{Z_n}, 1 \right] \cong \left[\frac{Z_0}{Z_n}, \dots \frac{Z_{n-1}}{Z_n} \right] \in K^n$.

$\forall v \neq 0$, $v \in V$, $[v]$ = corresponding pt. in $\mathbb{P}V \cong \mathbb{P}^n$

Polynomial $F \in K[Z_0 \dots Z_n]$ on vector space K^{n+1} doesn't define a function on $\mathbb{P}^n$, but
if F is homogeneous of degree d,
then since
$$F(\lambda Z_0, \dots, \lambda Z_n) = \lambda^d F(Z_0 \dots Z_n)$$
it does make sense to talk about 0 locus of polynomial F.

Definition 125 (Projective variety)**.** *projective variety* $X \subset \mathbb{P}^n = \{P | F_\alpha(P) = 0 \ \forall \alpha,\ F_\alpha(\lambda P) = \lambda^d F_\alpha(P)\}$ = *zero locus of a collection of homogeneous polynomials* F_α.

Group $PGL_{n+1}K$ acts on space $\mathbb{P}^n$ (in Lecture 18, $PGL_{n+1}K$ are automorphisms of $\mathbb{P}^n$).

Varieties $X, Y \subset \mathbb{P}^n$ are projectively equivalent, if they're congruent, modulo this group.

Note that if $\mathbb{P}^n = \mathbb{P}V$ is projective space associated with vector space V,

- homogeneous coordinates on $\mathbb{P}V$ correspond to elements of dual space V^*
- similarly, space of homogeneous polynomials of degree d on $\mathbb{P}V$ naturally identified with vector space $\text{Sym}^d(V^*)$

Meaning, set of linear coordinates on vector space V, $\dim V = n$, over field K (so $V = K^n$), $\alpha_i \equiv z_i$, $i = 1 \dots n$, is a *basis* (α_i) of V^*, since
$$\begin{aligned} \alpha : V &\to K^n \\ v &\mapsto (\alpha_1(v), \dots \alpha_n(v)) \end{aligned} \quad \text{i.e.} \quad \equiv \quad \begin{aligned} z : V &\to K^n \\ v &\mapsto (z_1(v) \dots z_n(v)) \end{aligned}$$

Now $\mathbb{P}(V) = (V \backslash \{0\})/K^*$ and homogeneous coordinates on $\mathbb{P}(V)$ are just linear coordinates on V up to action K^*

cf. "Correspondence between the projective space associated to a vector space and the dual space of the vector space?", stackexchange, Can dual vector spaces be thought of as linear coordinate functions? stackexchange

From $Z_i \in V^*$, $i = 0, 1 \dots n$, $Z_i : V \to K$, $Z_i : v \mapsto Z_i(v) = Z_i \in K$,
let f be a homogeneous polynomial of degree d on $\mathbb{P}V$:
$$f = \sum a_{i_0 i_1 \dots i_n} z_0^{i_0} z_1^{i_1} \dots z_n^{i_n}$$
where summation $\sum$ is over $0 \leq i_0, i_1, \dots i_n \leq d$ s.t. $\sum_{j=0}^n i_j = d$.

$\dim \text{Sym}^d(V^*) = \binom{d+n}{n}$

$\{z_0^{i_0} z_1^{i_1} \dots z_n^{i_n}\}_{\substack{0 \leq i_0, i_1 \dots i_n \leq d \\ \sum_{j=0}^n i_j = d}}$ form a basis for $\text{Sym}^d(V^*)$

Let $U_i \subset \mathbb{P}^n$, $U_i = \{[Z_0 \dots Z_n] | Z_i \neq 0\}$. Then $[Z_0 \dots Z_n] \sim \left[\frac{Z_0}{Z_i} \dots \frac{Z_{i-1}}{Z_i}, 1 \dots \frac{Z_n}{Z_i} \right] \equiv [z_0, \dots z_{i-1}, 1, z_i \dots z_{n-1}] \cong (z_0, z_1 \dots z_{n-1}) \in K^n$.

So there's a bijection $U_i \to K^n$.

Geometrically, this map is associating line $L \subset K^{n+1}$ not contained in hyperplane $(Z_i = 0)$, its pt. p of intersection with affine plane $(Z_i = 1) \subset K^{n+1}$.

Coordinates z_j on U_i are called affine or Euclidean coordinates on projective space or open set U_i
- open sets U_i comprise standard cover of $\mathbb{P}^n$ by affine open sets.

If $X \subset \mathbb{P}^n$ is a variety, $X_i = X \cup U_i$ is affine variety:
if X given by polynomials $F_\alpha \in K[Z_0, \dots Z_n]$, then e.g. X_0 will be zero locus of polynomials
$$f_\alpha(z_0 \dots z_n) = F_\alpha(Z_0 \dots Z_n)/Z_0^d = F_\alpha(1, z_1 \dots z_n)$$
where $d = \deg F_\alpha$.

For (projective) variety $X \subset \mathbb{P}^n$, $X = \{P | F_\alpha(P) = 0, \forall \alpha, F_\alpha \text{ homogeneous}, P = [Z_0, Z_1 \dots Z_n] \in \mathbb{P}^n\}$, obtain affine variety $X_i = X \cup U_i$ as follows: for
$$z_j = \begin{cases} \frac{Z_{j-1}}{Z_i} & j \leq i \\ \frac{Z_j}{Z_i}, & j > i \end{cases}$$
$$f_\alpha(z_1 \dots z_n) = f_\alpha \left(\frac{Z_0}{Z_i}, \dots \frac{Z_{i-1}}{Z_i}, \frac{Z_{i+1}}{Z_i} \dots \frac{Z_n}{Z_i} \right) = \frac{1}{Z_i}^{d_\alpha} F_\alpha(Z_0 \dots Z_n) = F_\alpha(z_1 \dots z_i, 1, z_{i+1}, \dots z_n)$$

If $F_\alpha(Z_0 \dots Z_n) = 0$, then $f_\alpha(z_1 \dots z_n) = 0$

$\forall$ projective variety X, X is union of affine varieties.

If affine variety $X_i \subset K^n \cong U_i \subset \mathbb{P}^n$, by def. X_i given by polynomials $\{f_\alpha\}_\alpha$
$$f_\alpha(z_1 \dots z_n) = \sum a_{i_1 \dots i_n} z_1^{i_1} \dots z_n^{i_n} = 0$$
of degree d_α (i.e. $i_1 + \dots i_n = d_\alpha$)

Then
$$F_\alpha(Z_0 \dots Z_n) = Z_i^{D_\alpha} F_\alpha \left(\frac{Z_0}{Z_i} \dots \frac{Z_n}{Z_i} \right) = Z_i^{D_\alpha} f_\alpha(z_1 \dots z_n) = \sum a_{i_1 \dots i_n} Z_i^{D_\alpha - \sum i_l} Z_0^{i_0} \dots Z_n^{i_n} =$$
$$= \sum a_{i_1 \dots i_n} Z_i^{D_\alpha - d_\alpha} Z_0^{i_0} \dots \widehat{Z}_i^{i_i} \dots Z_n^{i_n}$$

47.1.1. *Example: ellipse.*

$$\mathbb{P}^n \to U_Z \cong K^n$$
$$[X, Y, Z] \mapsto (x, y) = \left(\frac{X}{Z}, \frac{Y}{Z} \right) \tag{170}$$

Consider
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \text{ or } f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \tag{171}$$

For affine variety $X_Z \subset K^2$,
$$F(X, Y, Z) = \left(\frac{X^2}{Z^2 a^2} + \frac{Y^2}{Z^2 b^2} - 1 \right) Z^2 = \frac{X^2}{a^2} + \frac{Y^2}{b^2} - Z^2 \tag{172}$$

48. Algebraic Curves; Conic sections

cf. Reid (2013) [42].

cf. Ch. 0 "Woffle" of Reid (2013) [42].

Given field k, $k[x_1 \dots x_n]$ colelction of all polynomials in $x_1 \dots x_n$, with coefficients in k,
$$f \in k[x_1 \dots x_n] = \{f | f = \sum_\alpha c_\alpha x^\alpha, x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}, c_\alpha \in k\}$$

Variety is (roughly) locus defined by polynomial equations
$$V = \{P \in k^n | f_i(P) = 0\} \subset k^n,\ f_i \in k[x_1 \dots x_n]$$
e.g. plane curves C: $(f(x, y) = 0) \subset \mathbb{R}^2$ or $\mathbb{C}^2$

Groups of transformations (i.e. transformation groups) are of central importance throughout geometry; properties of geometric figures must be invariant under appropriate kind of transformations before they're significant.

affine change of coordinates in $\mathbb{R}^2$ is of form
$$T(\mathbf{x}) = A\mathbf{x} + B \quad \text{(affine change of coordinates)} \tag{173}$$
where $\mathbf{x} = (x, y) \in \mathbb{R}^2$, A 2×2 invertible matrix (i.e. $A \in GL(2, \mathbb{R})$), $B \in \mathbb{R}^2$.

If A orthogonal, transformation T is *Euclidean.*

$\forall$ nondegenerate conic can be reduced to "standard form" by Euclidean transformation.

projectivity or projective transformation $\mathbb{P}^2_{\mathbb{R}}$ is map $T(\mathbf{X}) = M\mathbf{X}$, $M \in GL(3,\mathbb{R})$.

Understand T on affine piece $\mathbb{R}^2 \subset \mathbb{P}^2_{\mathbb{R}}$ is partially defined map $\mathbb{R}^2 \to \mathbb{R}^2$; it's a fractional linear transformation.

$$\begin{aligned} (x,y) &\xmapsto{\cong} [x,y,1] \\ (x,y) &\mapsto \begin{pmatrix} A\begin{pmatrix} x \\ y \end{pmatrix} + B \\ cx+dy+e \end{pmatrix} \end{aligned}$$

where

$$M = \left(\begin{array}{c|c} A & B \\ c \quad d & e \end{array} \right)$$

e.g. 2 different photographs of same (plane) object are obviously related by a projectivity.

For inhomogeneous quadratic polynomial q, homogeneous quadratic polynomial Q, then there exists bijection

$$q \in K[x,y] \xmapsto{\cong} Q \in K[X,Y,Z]$$
$$q(x,y) = ax^2 + bxy + cy^2 + dx + ey + f \mapsto Q(X,Y,Z) = aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2$$

so

$$q(x,y) = Q\left(\frac{X}{Z}, \frac{Y}{Z}, 1\right) \text{ with } x = X/Z,\, y = Y/Z$$

inverse:

$$Q = Z^2 q(X/Z, Y/Z)$$

48.0.1. *"Line at infinity" and asymptotic directions.* cf. Ch. 1 of Reid (2013)

Points of $\mathbb{P}^2$ with $Z=0$, $[X,Y,0]$, form *line at infinity*, a copy of $\mathbb{P}^1_{\mathbb{R}} = \mathbb{R}\cup\{\infty\}$ (since $[X,Y] \mapsto X/Y$) define bijection $\mathbb{P}^1_{\mathbb{R}} \to \mathbb{R}\cup\{\infty\}$.

Line in $\mathbb{P}^2$, L, $L := \{[X,Y,Z] | aX + bY + cZ = 0\}$.

L passes through $(X,Y,0) \Longleftrightarrow aX + bY = 0$.

(a) hyperbola $\left(\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\right)$. Recall that the lines of asymptotes (asymptotic lines). They are found in the following manner:

$$\frac{(bx-ay)(bx+ay)}{a^2b^2} = 1 \text{ or } \frac{bx-ay}{a^2b^2} = \frac{1}{bx+ay} \xrightarrow{x,y\to\infty} \frac{bx-ay}{a^2b^2} = 0 \text{ or } y = \frac{b}{a}x$$

Now, $\left(\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\right)$ in $\mathbb{R}^2$ corresponds in $\mathbb{P}^2_{\mathbb{R}}$ to $C : \left(\frac{X^2}{a^2} - \frac{Y^2}{b^2} = Z^2\right)$.

This meets $(Z=0)$ in 2 pts. $(a, \pm b, 0) \in \mathbb{P}^2_{\mathbb{R}}$, corresponding to asymptotic lines of hyperbola, $y = \frac{b}{a}x$, $y = \frac{-b}{a}x$

For affine piece $U_x \subset \mathbb{P}^2_{\mathbb{R}}$, $U_x = \{p \in \mathbb{P}^2_{\mathbb{R}} | p = [X,Y,Z] \text{ s.t. } X \neq 0\}$, then bijection $U_x \to \mathbb{R}^2$,

$$[X,Y,Z \sim [1, \frac{Y}{X}, \frac{Z}{X}] \mapsto (u,v) = \left(\frac{Y}{X}, \frac{Z}{X}\right), \text{ so}$$
$$C : X^2/a^2 - Y^2/b^2 = Z^2 \mapsto u^2 + \frac{v^2}{b^2} = \frac{1}{a^2} \text{ or } \frac{u^2}{1/a^2} + \frac{v^2}{(b/a)^2} = 1 \qquad \text{(an ellipse!)}$$

(b) $y = mx^2$ (parabola) in $\mathbb{R}^2 \mapsto C : YZ = mX^2$ in $\mathbb{P}^2_{\mathbb{R}}$.

For $Z=0$, C meets $Z=0$ at single pt. $[0,1,0] \sim [0,Y,0]$. So in $\mathbb{P}^2$, "2 branches of parabola meet at infinity."

48.0.2. *Classification of conics in $\mathbb{P}^2$.* cf. 1.6. Classification of conics in $\mathbb{P}^2$, Reid (2013) [42]

Let K be any field of characteristic $\neq 2$.

Recall 2 linear algebra results for quadratic forms:

Proposition 31. *$\exists$ bijections*
$\{$ homogeneous quadratic polynomials $\} = \{$ quadratic forms $K^3 \to K$ $\} \cong \{$ symmetric bilinear forms on K^3 $\}$ given by

$$aX^2 + 2bXY + cY^2 + 2dXZ + 2eYZ + fZ^2 \cong \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \text{ since}$$
$$\begin{bmatrix} X & Y & Z \end{bmatrix} \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = aX^2 + 2bXY + cY^2 + 2dXZ + 2eYZ + fZ^2$$

Quadratic form nondegenerate if corresponding bilinear form nondegenerate, i.e. matrix is nonsingular.

Theorem 18. *Let V be vector space over K, quadratic form $Q : V \to K$, then $\exists$ basis of V s.t.*

$$Q = \epsilon_1 x_1^2 + \epsilon_2 x_2^2 + \dots + \epsilon_n x_n^2 \text{ with } \epsilon_i \in K \tag{174}$$

This theorem is proved by Gram-Schmidt orthogonalization.

For $\lambda \in K\backslash\{0\}$, $x_i \mapsto \lambda x_i$ takes $\epsilon_i \mapsto \lambda^{-2}\epsilon_i$.

Corollary 4. *In a suitable coordinate system, any conic in $\mathbb{P}^2$ is one of*

(a) *nondegenerate conic $C : (X^2 + Y^2 - Z^2 = 0)$*

48.0.3. *Parametrization of a conic.* Let C be a nondegenerate, nonempty conic of $\mathbb{P}^2_{\mathbb{R}}$.

Then by Corollary 4 (cf. Corollary 1.6 (cf. Reid (2013) [42]), and taking new coordinates $[X+Z, Y, Z-X]$,

$$X^2 + Y^2 - Z^2 = 0 \mapsto (X+Z)^2 + Y^2 - (Z-X)^2 = X^2 + 2XZ + Z^2 + Y^2 - (Z^2 - 2ZX + X^2) = Y^2 + 4XZ = 0$$

$\Longrightarrow C$ is projectively equivalent to curve $(Y^2 = XZ)$.

This is a curve parametrized by

$$\begin{aligned} \Phi : \mathbb{P}^1_{\mathbb{R}} &\to C \subset \mathbb{P}^2_{\mathbb{R}} \\ [U,V] &\mapsto [U^2, UV, V^2] \end{aligned}$$

This is because

$$[X,Y,Z] \sim [X^2, XY, XZ] = [X^2, XY, Y^2]$$

and so let $U = X, V = Y$. Note that if $X \mapsto X + Z$, then $U = X + Z$.

Inverse map $\Psi = \Phi^{-1}$, $\Psi : C \to \mathbb{P}^1_{\mathbb{R}}$ given by

$$[X,Y,Z] \mapsto [X,Y] = [Y,Z]$$

$[X,Y]$ defined if $X \neq 0$, $[Y,Z]$ defined if $Z \neq 0$.

Φ, Ψ are inverse isomorphisms of varieties.

cf. Ch. 2 "Cubics and the group law" of Reid (2013) [42].

cf. Sec. 2.1 "Examples of parametrized cubics" in Ch. 2 of Reid (2013) [42].

Nodal cubic: $C : (y^2 = x^3 + x^2) \subset \mathbb{R}^2$, is image of map $\varphi : \mathbb{R}^1 \to \mathbb{R}^2$, $t \mapsto (t^2 - 1, t^3 - t)$, since

$$(t^2-1)^3 + (t^2-1)^2 = t^6 - 3t^4 + 3t^2 - 1 + t^4 - 2t^3 + 1 = t^6 - 2t^4 + t^2 = t^2(t^4 - 2t^2 + 1) = t^2(t^2-1)^2 = y^2$$

Cuspidal cubic $C : (y^2 = x^3) \subset \mathbb{R}^2$ is image of $\varphi : \mathbb{R}^1 \to \mathbb{R}^2$, $t \mapsto (t^2, t^3)$

48.0.4. *Curve $y^2 = x(x-1)(x-\lambda)$ has no rational parametrization.* cf. Sec. 2.2 "Curve $y^2 = x(x-1)(x-\lambda)$" in Ch. 2 of Reid (2013) [42].

$f = f(t)$ *rational function* if it's a quotient of 2 polynomials.

Lemma 4. *Let $\overline{K}$ algebraically closed field, $p, q \in \overline{K}[t]$ coprime elements (i.e. if $\exists x$ s.t. p/x, $q/x \in \overline{K}$ (i.e. $x|p$, $x|q$), then $x = 1$),*
assume 4 distinct linear combinations (i.e. $\lambda p + \mu q$ for 4 distinct ratios $(\lambda : \mu) \in \mathbb{P}^1 K$) are squares in $\overline{K}[t]$,
then $p, q \in \overline{K}$

cf. Lemma 2.3 of Reid (2013) [42]

Proof. (Fermat's method of "infinite descent")
Without loss of generality,

$$p' = ap + bq$$
$$q' = cp + dq$$

$a, b, c, d \in K$, $ad - bc \neq 0$.
Hence, assume 4 given squares are

$$p, p-q, p-\lambda q, q$$

i.e. $\lambda p + \mu q$, for $\lambda = 1, \mu = 0$; $\lambda = 1, \mu = -1$; $\lambda = 1, \mu = -\lambda$; $\lambda = 0, \mu = 1$
Since a, b, c, d arbitrary linear transformation.
Then $p = u^2, q = v^2$, $u, v \in \overline{K}[t]$ are coprime, with

$$\max(\deg u, \deg v) < \max(\deg p, \deg q)$$

Suppose $\max(\deg p, \deg q) > 0$ and is minimal among all p, q satisfying lemma condition.
Then

$$p - q = u^2 - v^2 = (u-v)(u+v)$$
$$p - \lambda q = u^2 - \lambda v^2 = (u - \mu v)(u + \mu v)$$

where $\mu = \sqrt{\lambda}$, are squares in $\overline{K}[t]$.
So by u, v being coprime,
Then $u - v, u + v, u - \mu v, u + \mu v$ are squares.
This contradicts minimality of $\max(\deg p, \deg q)$

□

Theorem 19 ($y^2 = x(x-1)(x-\lambda)$ has no rational parametrization)**.** *Let K be field of characteristic $\neq 2$, let $\lambda \in K$, $\lambda \neq 0, 1$; let $f, g \in K(t)$ be rational functions s.t.*

$$f^2 = g(g-1)(g-\lambda)$$

Then $f, g \in K$.

EY (20181229). Recall, *characteristic of ring* R (e.g. field), $\text{char}(K)$, smallest number of times 1 must using ring's multiplicative identity 1 in a sum to get additive identity (0).
$\text{char}(K) = 0$ for case that $\underbrace{1 + \dots + 1}_{n} = \sum_{i=1}^n 1 \neq 0 \quad \forall n \in \mathbb{Z}^+$.
Theorem 19 is equivalent to $\nexists$ nonconstant map $\mathbb{R}^1 \to C : (y^2 = x(x-1)(x-\lambda))$ given by rational functions.

Proof. $K[t]$ UFD; unique factorization domain (given).
EY: 20181229, recall the definitions: integral domain - nonzero commutative ring in which product of any 2 nonzero elements is nonzero.
unique factorization domain is an integral domain R s.t. $\forall x \in R$, $x \neq 0$, x can be written as

$$x = u p_1 p_2 \dots p_n, \quad n \geq 0$$

with irreducible elements p_i of R, unit u.

$$\Longrightarrow \begin{aligned} f &= r/s \qquad r, s \in K[t] \text{ and coprime} \\ g &= p/q \qquad p, q \in K[t] \text{ and coprime} \end{aligned}$$

$$\Longrightarrow f^2 = g(g-1)(g-\lambda) = \frac{r^2}{s^2} = \frac{p}{q}\left(\frac{p-q}{q}\right)\left(\frac{p-\lambda q}{q}\right) \Longrightarrow r^2 q^3 = s^2 p(p-q)(p-\lambda q)$$

r, s are coprime, so RHS s^2 must divide q^3.
p, q are coprime, LHS q^3 must divide s^2
EY (20181229): observe that LHS and RHS are different and equal. How to get them into the same form? Try to divide both sides!

$$\Longrightarrow s^2 | q^3 \text{and } q^3 | s^2, \text{ so } s^2 = aq^3 \text{ with } a \in K$$

Then $aq = (s/q)^2$ is square in $K[t]$
Then $r^2 = ap(p-q)(p-\lambda q)$
Consider factorization into primes $\Longrightarrow$ nonzero constants $b, c, d \in K$, s.t. $bp, c(p-q), d(p-\lambda q)$ are all squares in $K[t]$.
Let algebraic closure $\overline{K}$ (algebraic extension of K s.t. $\overline{K}$ algebraically closed, i.e. $\forall$ nonconstant $f(x) \in K[x]$ has a root in K).
Then $\forall p, q \in \overline{K}(t)$, by lemma, $p, q \in \overline{K}$. Then $r, s \in \overline{K}$. Then $\boxed{f, g \in \overline{K}}$.

□

cf. Sec. 2.4 "Linear systems" in Ch. 2 of Reid (2013) [42].
Let $S_d \equiv \{$ forms of degree d in (X, Y, Z) $\}$; recall form is just a homogeneous polynomial.
$\forall F \in S_d$, $\exists$ unique form for F: $F = \sum a_{ijk} X^i Y^j Z^k$, $a_{ijk} \in K$, and $\sum \equiv \sum_{\substack{i,j,k \geq 0 \\ i+j+k=d}}$.
$\Longrightarrow S_d$ is K-vector space with basis $\{Z^d, XZ^{d-1}, YZ^{d-1}, \dots X^{d-2}Y^2 \dots Y^d\}$, where

$$\boxed{\dim S_d = \binom{d+2}{2}}$$

(to see this, imagine d stars, 2 bars, and the 2 bars distinguish which are X's, Y's, or Z's).
For $P_1 \dots P_n \in \mathbb{P}^2$, let

$$S_d(P_1 \dots P_n) = \{F \in S_d | F(P_i) = 0 \quad \forall i = 1 \dots n\} \subset S_d$$

$\forall$ condition $F(P_i) = 0$ (e.g. $F(X_i, Y_i, Z_i) = 0$, where $P_i = (X_i, Y_i, Z_i)$) is 1 linear condition on F, so $S_d(P_1 \dots P_n)$ is a vector space of dim $\geq \binom{d+2}{2} - n$

Lemma 5 (Special case of Nullstellensatz)**.** (i) *Let $L \subset \mathbb{P}^2_K$ be a line; if $F \equiv 0$ on L, then F divisible in $K[X, Y, Z]$ by equation of L, i.e. $F = H \cdot F'$, where H is equation of L, and $F' \in S_{d-1}$.*
(ii) *Let $C \subset \mathbb{P}^2_K$ be nonempty nondegenerate conic; if $F = 0$ on C, then F divisible in $K[X, Y, Z]$, by equation of C, i.e. $F = QF'$, where Q is equation of C, and $F' \in S_{d-2}$.*

cf. Lemma 2.5 of Reid (2013).

Proof. (i) By change of coordinates, assume $H = X$, Then, $\forall F \in S_d$, $\exists!\, F = X \cdot F'_{d-1} + G(Y, Z)$, since, just gather together all monomials involving X into 1st. summand, and what's left must be a polynomial Y, Z.
Since
$F = 0$ on L, $F(0) = 0 = 0 \cdot F'_{d-1} + G(Y, Z) \Longrightarrow G(Y, Z) = 0 \quad \forall Y, Z$.
Otherwise, if $G(Y, Z) \neq 0$, then it has at most d zeros on $\mathbb{P}^1_K$, whereas if K is infinite, then so is $\mathbb{P}^1_K$.
(ii) By change of coordinates $Q = XZ - Y^2$,
Consider why

$$F = QF'_{d-2} + A(X, Z) + YB(X, Z)$$

where $d-2$ in F'_{d-2} denotes the *degree* of the polynomial (to be $d-2$).

This is because if $Y^2 = XZ - Q$, then $F(Y^2 = XZ - Q)$ has degree ≤ 1 in Y, and so would have the form

$$F(Y^2 = XZ - Q) = A(X,Z) + YB(X,Z)$$

C is a parametrized conic given by

$$X = U^2, Y = UV, Z = V^2$$

so that,
$F = 0$ on $C \iff A(U^2, V^2) + UVB(U^2, V^2) = 0$ on C
$\implies A(U^2, V^2) + UVB(U^2, V^2) = 0 \in K[U, V]$.
$\implies A(X,Z) = B(X,Z) = 0$

Since here the last equality comes by considering separately terms of even and odd degrees in form

$$A(U^2, V^2) + UVB(U^2, V^2)$$

□

cf. Exercises to Ch. 2, Reid (2013)

Exercise 2.2. Let $\varphi : \mathbb{R}^1 \to \mathbb{R}^2$.
$t \mapsto (t^2, t^3)$
$\forall$ polynomial $f \in \mathbb{R}[X,Y]$, s.t. $f = 0$ on image $C = \varphi(\mathbb{R}^1)$, f divisible by $Y^2 - X^3$.

Proof. Given $\varphi(t) = (t^2, t^3) = (x, y)$, then $y^2 = x^3 \quad \forall t \in \mathbb{R}$, or $y^2 - x^3 = 0$.
Let $q = q(x,y) = y^2 - x^3 \in K[x,y]$.
Suppose f of degree d.
Then

$$f = qf'_{d-2} + a(x) + yb(x)$$

This is because, if $y^2 = q - x^3$, $f(y^2 = q - x^3)$ has degree ≤ 1 in y, so would have the previous form.
Now

$$f(y^2 = q - x^3) = 0 = 0 + a(x) + yb(x)$$

$f = 0$ on $C = \varphi(\mathbb{R}^1) \implies a(x) + yb(x) = 0 = a(t^2) + t^3 b(t^2) = 0$.
Suppose for $t_1 > 0$, $t_1^3 b(t_1^2) = -a(t_1^2)$.
Consider $-t_1 < 0$:

$$\implies -t_1^3 b(t_1^3) = -a(t_1^2) \implies a(t_1^2) = 0 \quad \forall t_1 > 0$$

Then $b(t_1^2) = 0 \; \forall t_1 > 0$.
$\implies f = qf'_{d-2}$ where $q = y^2 - x^3$.
K needs to have "negative numbers" (i.e. additive inverses) to exist, for this proof to work. □

References

[1] Masaki Kashiwara and Pierre Schapira. **Categories and Sheaves**. *Grundlehren der mathematischen Wissenschaften.* Volume 332. 2006. Springer-Verlag Berlin Heidelberg. eBook ISBN 978-3-540-27950-1
[2] David S. Dummit, Richard M. Foote. **Abstract Algebra**. 3rd. Ed. Wiley; (July 14, 2003). ISBN-13: 978-0471433347
[3] Michael Barr, Charles Wells. **Category Theory for Computing Science**. http://www.tac.mta.ca/tac/reprints/articles/22/tr22.pdf, http://www.math.mcgill.ca/triples/Barr-Wells-ctcs.pdf
[4] Jiří Adámek, Horst Herrlich, George E. Strecker. **Abstract and Concrete Categories The Joy of Cats**. 2004. http://katmat.math.uni-bremen.de/acc/acc.pdf
[5] Saunders Mac Lane. **Categories for the Working Mathematician**. Graduate Texts in Mathematics. Volume 5. 1978. Springer-Verlag New York. eBook ISBN 978-1-4757-4721-8
[6] Harold Simmons. **An Introduction to Category Theory**. 1st Edition. Cambridge University Press; 1 edition (November 21, 2011). ISBN-13: 978-0521283045
[7] Daniele Turi. **Category Theory Lecture Notes**. September 1996 – December 2001. http://www.dcs.ed.ac.uk/home/dt/CT/categories.pdf
[8] Tom Leinster. **Basic Category Theory** (Cambridge Studies in Advanced Mathematics) 1st Edition. 2014. ISBN-13: 978-1107044241
[9] Prof. Dr. B. Pareigis. **Category Theory**. Summer Semester 2004. June 14, 2004. http://www.mathematik.uni-muenchen.de/~pareigis/Vorlesungen/04SS/Cats1.pdf
[10] Ivan Čukić. **Functional Programming in C++: How to improve your C++ programs using functional techniques**. Manning Publications; 1st edition (November 19, 2018). ISBN-13: 978-1617293818
[11] Aaron David Ames. "A Categorical Theory of Hybrid Systems." PhD dissertation. Dec. 11, 2006. Technical Report No. UCB/EECS-2006-165. http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-165.html
[12] David A. Cox. John Little. Donal O'Shea. **Using Algebraic Geometry**. Second Edition. Springer. 2005. ISBN 0-387-20706-6 QA564.C6883 2004
[13] David Cox, John Little, Donal O'Shea. **Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra**, Fourth Edition, Springer
[14] Torsten Grust, Christian Duta. *Datenbanksysteme I.* Winter 2017-2018. Eberhard Karls Universität Tübingen. https://www-db.informatik.uni-tuebingen.de/teaching/DatenbanksystemeIWS2017-2018.html
[15] David I. Spivak. *Simplicial Databases.* arXiv: 0904.2012v1 [cs.DB] 13 Apr 2009.
[16] David I. Spivak. *Categorical Databases.* Presented on 2012/01/13. http://math.mit.edu/~dspivak/informatics/talks/CTDBIntroductoryTalk
[17] Jun Yang. *CompSci 316: Introduction to Database Systems.* Fall 2012. Duke University. https://www2.cs.duke.edu/courses/fall12/compsci316/index.html
[18] Carlo Curino. David Spivak. *Category Theory as a Unifying Database Formalism.* 2011. http://math.mit.edu/~dspivak/informatics/notes/unorganized/PODS.pdf
[19] Steve Awodey. **Category Theory** (Oxford Logic Guides) 2nd Edition. Oxford Logic Guides (Book 52). Oxford University Press; 2 edition (August 13, 2010). ISBN-13: 978-0199237180
[20] Michael Barr, Charles Wells. **Category Theory for Computing Science**. 2012
[21] Franz Schwabl. **Quantum Mechanics**. Springer-Verlag Berlin Heidelberg. 2007. 4th Edition. ISBN 978-3-540-71933-5
[22] Schottenloher, Martin. **A Mathematical Introduction to Conformal Field Theory**. Springer, 2008.
[23] L.D. Landau and E.M. Lifshitz. **Statistical Physics**, Third Edition, Part 1: Volume 5 (Course of Theoretical Physics, Volume 5) 3rd Edition. Butterworth-Heinemann; 3 edition (January 15, 1980). ISBN-13: 978-0750633727
[24] M. Hjorth-Jensen, **Computational Physics**, University of Oslo (2015) http://www.mn.uio.no/fysikk/english/people/aca/mhjensen/
[25] M.E.J. Newman and G.T. Barkema. **Monte Carlo Methods in Statistical Physics**. Oxford University Press. 1999.
[26] Glen E. Bredon. **Topology and Geometry**. Graduate Texts in Mathematics (Book 139). Springer; Corrected edition (October 17, 1997). ISBN-13: 978-0387979267
[27] Joseph Rotman. **An Introduction to Algebraic Topology**. *Graduate Texts in Mathematics.* Volume 119. 1988. Springer-Verlag New York. eBook ISBN 978-1-4612-4576-6
[28] Jean-Pierre Serre (Author), J. Stilwell (Translator). **Trees** (Springer Monographs in Mathematics) 1st ed. 1980. Corr. 2nd printing 2002 Edition. ISBN-13: 978-3540442370
[29] Joseph J. Rotman, **Advanced Modern Algebra** (Graduate Studies in Mathematics) 2nd Edition, American Mathematical Society; 2 edition (August 10, 2010), ISBN-13: 978-0821847411
[30] Serge Lang. **Algebra** (Graduate Texts in Mathematics) 3rd Edition. *Graduate Texts in Mathematics* (Book 211). Springer; 3rd edition (June 21, 2005). ISBN-13: 978-0387953854
[31] Edward Scheinerman. **C++ for Mathematicians: An Introduction for Students and Professionals**. 1st Edition. CRC Press; 1 edition (June 8, 2006). ISBN-13: 978-1584885849
[32] Jeffrey M. Lee. **Manifolds and Differential Geometry**, *Graduate Studies in Mathematics* Volume: 107, American Mathematical Society, 2009. ISBN-13: 978-0-8218-4815-9
[33] Jacob C. Bridgeman and Christopher T. Chubb. *Hand-waving and Interpretive Dance: An Introductory Course on Tensor Networks:* ***Lecture Notes***. arXiv:1603.03039 **[quant-ph]**
[34] Ulrich Schollwoeck. *The density-matrix renormalization group. Rev. Mod. Phys.* **77**, 259 (2005) arXiv:cond-mat/0409292 **[cond-mat.str-el]**
[35] Ulrich Schollwoeck. *The density-matrix renormalization group in the age of matrix product states. Annals of Physics* **326**, 96 (2011). arXiv:1008.3477 **[cond-mat.str-el]**
[36] Ulrich Schollwöck, et. al. *Numerical methods for correlated many-body systems.* **2017 Arnold Sommerfeld School**.
Matrix product states (MPS), Density matrix renormalization group (DMRG) Lecture 1
Matrix product states (MPS), Density matrix renormalization group (DMRG) Lecture 2
Matrix product states (MPS), Density matrix renormalization group (DMRG) Hands on Session 1
Matrix product states (MPS), Density matrix renormalization group (DMRG) Hands on Session 2
Ulrich Schollwöck (LMU): *Matrix product states (MPS), Density matrix renormalization group (DMRG)*
Ulrich Schollwöck (LMU): *Matrix product states (MPS), Density matrix renormalization group (DMRG)*
[37] José L.F. Barbón and Eliezer Rabinovici. "Holographic Complexity And Spacetime Singularities." arXiv:1509.09291 **[hep-th]**
[38] Juan Maldacena, Leonard Susskind. "Cool horizons for entangled black holes." arXiv:1306.0533 **[hep-th]**
[39] Slava Rychkov. "EPFL Lectures on Conformal Field Theory in $D \geq 3$ Dimensions." arXiv:1601.05000 **[hep-th]**
[40] G. Evenbly, G. Vidal. "Tensor network states and geometry." arXiv:1106.1082 **[quant-ph]**
[41] Mark Van Raamsdonk. "Lectures on Gravity and Entanglement." arXiv:1609.00026 **[hep-th]**
[42] Miles Reid. **Undergraduate Algebraic Geometry**. October 20, 2013. https://homepages.warwick.ac.uk/staff/Miles.Reid/MA4A5/UAG.pdf
[43] Joe Harris. **Algebraic Geometry: A First Course**. Graduate Texts in Mathematics. 1992. Springer-Verlag New York. eBook ISBN 978-1-4757-2189-8