# MATH 1208, Honors Math B Notes

## Contents



## 1 Vector Spaces

**Definition.** A field is a set $F$ together with two binary operations $+, \cdot$ and two special elements $0, 1$ such that the operations are distributive, commutative associative, and invertible (0 does not have a multiplicative inverse).

*Example.* Given a field $F$, define $2 = 1 + 1, 3 = 2 + 1, \ldots$. It may happen that $p = 0$ for $p$ a prime. So we can't always divide by natural numbers.

**Definition.** Let $F$ be a field. A vector space over $F$ is a set $V$ together with two operations:

$$+ : V \times V \to V \text{ (addition)}$$
$$\cdot : F \times V \to V \text{ (scalar multiplication)}$$

and an element $0 \in V$, such that they obey the following:

1. $\forall x, y \in V, x + y = y + x$
2. $\forall x, y, z \in V, x + (y + z) = (x + y) + z$ and $\forall a, b \in F, x \in V, a \cdot (b \cdot x) = (a \cdot b) \cdot x$
3. $\forall c \in F, x, y \in V, c \cdot (x + y) = c \cdot x + c \cdot y$ and $\forall c, d \in F, x \in V, (c + d) \cdot x = c \cdot x + d \cdot x$
4. $\forall x \in F, x + 0 = x, 1 \cdot x = x$

*Remark.* Call elements of $V$ vectors and call elements of $F$ scalars. Apostol will also call vector spaces linear spaces.

**Prop.** *Easy facts about vector spaces:*
*Let $V$ be a vector space over $F$.*

1. *$\forall x \in V$, $(-1) \cdot x$ is the unique $y$ such that $x + y = 0$.*
2. *$\forall x \in V$, $0 \cdot x = 0$ and $\forall c \in F, c \cdot 0 = 0$.*
3. *$\forall c \in F, \forall x \in V, (-c) \cdot x = c \cdot (-x) = -c \cdot x$.*
4. *If $c \cdot x = 0$ then either $c = 0$ or $x = 0$ and if $c \cdot x = c \cdot y$, then $c = 0$ or $x = y$.*

*Proof.* (2) Given $x \in V$, we have that $0 + 0 = 0 \implies (0 + 0) \cdot x = 0 \cdot x \implies 0 \cdot x = 0$. ☐

**Definition.** Given $F$ a field, let $F^n$ be the $n$-times Cartesian product of $F$. Formally, this is the set of functions from an $n$-element set to $F$. Write such a function $(f_1, f_2, \ldots, f_n)$.
Let $0 = (0, 0, \ldots, 0)$, $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$, and $c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$.
It is easy to check that this satisfies all necessary axioms.
We call $x_i$ the $i^{th}$ component of $x = (x_1, x_2, \ldots, x_n) \in F^n$. Two vectors are equal if and only if their components are all equal.

*Example.* $F^0 = 0$. This is a vector space! Trivial, but still.
*Example.* $F^1$ "is" $F$ as a set. But the data of these objects are not quite the same.

*Example.* Let $F$ be a field and $S$ a set. Denote the set of functions from $S \to F$ as $\mathcal{F}(S, F)$. This is also a vector space. 0 is the zero function. $(f+g)(s) = f(s)+g(s)$ and $(cf)(s) = cf(s)$.

*Example.* More generally, take a set $S$ and a vector space $V$ (over $F$), then $\mathcal{F}(S, V)$ is also a vector space.

**Definition.** Let $V$ be a vector space over $F$. A nonempty subset $W \subseteq V$ is a subspace if $x, y \in W \implies x + y \in W$ and $x \in W, c \in F \implies cx \in W$.

If $W \subseteq V$ is a subspace, then $W$ is a vector space itself (with induced multiplication and addition from the larger vector space).

*Example.*

- $W = \{0\}$
- $W = V$
- Any line through the origin in $\mathbb{R}^2$
- Any line not through the origin is not a subspace of $\mathbb{R}^n$
- For vector space $V$ over $F$, take $v \in V$ with $v \neq 0$. Then $W = \{cv \mid c \in F\}$ is a subspace (this is a general line through the origin)
- Take $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$. Then there are lots of subspaces here, such as those that are bounded, those that are smooth, analytic, etc.
- Take $V = \mathcal{F}([a, b], \mathbb{R})$. Then there are still lots of subspaces, such as the bounded functions, which themselves have subspaces such as integrable and continuous functions.

**Prop.** *Let $V$ be a vector space. Let $W_1 \subseteq V, W_2 \subseteq V$ be subspaces. Then $W_1 \cap W_2$ is a subspace of $V$ as well.*

*Proof.* If $x, y \in W_1 \cap W_2$, then we have that $x, y \in W_1, x, y \in W_2 \implies x + y \in W_1, x + y \in W_2 \implies x + y \in W_1 \cap W_2$. Similarly, we have that $cx \in W_1 \cap W_2$. □

# 2 Linear Maps

**Definition.** If $V, W$ are two vector spaces over $F$, then a linear map $V \to W$ is a function $T : V \to W$ such that $T(v_1 + v_2) = T(v_1) + T(v_2), T(cv_1) = cT(v_1)$.

*Example.*

- Limits, derivatives, integrals, etc.
- $id_V : V \to V$

- If $T : U \to V$ is a linear map and $W$ is a subspace of $U$, then the restriction of $T\Big|_W : W \to V$ is still linear
- If $W \subseteq U$ is a subspace, $id_U\Big|_W : W \to U$ is linear (inclusion function)

**Definition.** For $x_1, x_2, \dots, x_n \in U$, where $U$ is a vector space over $F$ and also $c_1, c_2, \dots, c_n \in F$, the linear combination of $x_i$ with coefficients $c_i$ is the vector $\sum_{i=1}^n c_i x_i$.

**Prop.** *$T : U \to V$ is linear $\iff$ $\forall x_1, \dots, x_n \in U, \forall c_1, \dots, c_n \in F, T(\sum_{i=1}^n c_i x_i) = \sum_{i=1}^n c_i T(x_i)$.*
*Linear means that a transformation preserves linear combinations.*

**Definition.** Suppose $T$ is a linear map. Denote $T^{-1}(\{0\})$ by $\ker(T)$ and $T(U)$ by $\operatorname{im}(T)$. These are both vector spaces.

*Proof.* Just have to check closure under $+, \cdot$.
Let $u_1, u_2 \in \ker(T)$. Then $T(u_1) + T(u_2) = T(u_1 + u_2) = 0$ and $cT(u_1) = T(cu_1) = 0$. Thus, we have that it's closed. The claim about the image follows similarly. □

*Example.* Let $F = \mathbb{R}$ and $V$ be the set of differentiable functions. Then the linear map of the differentiation $D : V \to \mathcal{F}(\mathbb{R}, \mathbb{R})$. The kernel is just constant functions.

**Prop.** $\ker(T) = \{0\} \iff$ *$T$ is injective, and* $\operatorname{im}(T) = V \iff$ *$T$ is surjective.*

**Definition.** An isomorphism is a bijective linear map.

**Prop.** *A linear map $f : V \to W$ is bijective $\iff$ it has a linear inverse.*

*Proof.* The only thing that we really have to check is that if we are given a linear function that its inverse is linear (the opposite direction is obvious).
Let that function be $f : V \to W$ and the inverse $g : W \to V$. Given $w_1, w_2 \in W$, we have that $g(w_1 + w_2) = g(f(g(w_1)) + f(g(w_2))) = g(f(g(w_1) + g(w_2))) = g(w_1) + g(w_2)$. Similarly, we have that $g(cw_1) = g(cf(g(w_1))) = g(f(cg(w_1))) = cg(w_1)$. □

**Definition.** If $V, W$ are vector spaces over $F$, then $\mathcal{L}(V, W) := \{\text{Linear Maps } V \to W\} \subseteq \{\text{functions } V \to W\}$.

**Prop.** *If $f : U \to V$ and $g : V \to W$ are linear maps, then $g \circ f$ is also linear.*

*Proof.*

$$\begin{aligned}(g \circ f)(cu_1) &= g(f(cu_1)) \\ &= g(cf(u_1)) \\ &= cg(f(u_1)) \\ &= c(g \circ f)(u_1)\end{aligned}$$

It goes similarly for addition. □

**Definition.** Let $S \subseteq V$, where $V$ is a vector space over $F$.

$S$ spans $V$ if for every $v \in V$, $\exists s_1, \dots, s_n \in S$ and $c_1, \dots, c_n \in F$ such that $v = \sum_{i=1}^n c_i s_i$.

$S$ is linearly independent if whenever $\sum_{i=1}^n c_i s_i = 0$ with $c_i \in F, s_i \in S$, we have $for alli, c_i = 0$.

**Prop.** *In terms of a linearly independent set, there is exactly one way to express any vector in the span of that set.*

*Example.* If $0 \in S$, then $S$ is dependent ($c0 = 0$ for any $c$).

*Example.* $\mathbb{R}^3$ is spanned by $(1,0,0), (1,1,0), (1,0,1)$.

This can be seen from $c_1(1,0,0) + c_2(1,1,0) + c_3(1,0,1) = (0,0,0)$, which clearly needs $c_1 = c_2 = c_3 = 0$.

*Example.* The set with vectors $(1,0), (0,1), (1,-1)$ is dependent. We have that $(1,0)-(0,1)-(1,-1) = 0$.

*Example.* The set of standard basis elements is independent.

*Example.* We can develop easy rules for low dimension.

1. $S = \emptyset$ is independent.
2. $S = \{v\}$ is independent $\iff v \neq 0$.
3. $S = \{v_1, v_2\}$ is independent $\iff \neg \exists c \in F \mid v_1 = cv_2$ or $v_2 = cv_1$.

**Definition.** $S \subseteq V$ is a basis if $S$ is linearly independent and $S$ spans $V$.

*Example.* Let $V = \mathbb{R}[x]$.

We will show that $S = \{x^n \mid n \in \mathbb{Z}_{\geq 0}\}$ is a linearly independent basis for $V$. $S$ spans $V$ by definition; a polynomial is exactly just $\sum_{i=0}^n c_i x^i$.

From an earlier problem in series, we have that $\sum_{i=0}^n c_i x^i = 0 \implies c_i = 0$.

*Example.* The vectors $f(x) = x(x-1), g(x) = (x-1)(x-2), h(x) = x(x-2)$ are independent.

For $c_1 f(x) + c_2 g(x) + c_3 h(x) = 0$, if we plug in $x = 1$, then we will have that $c_3 = 0$; similarly, we can arrive at $c_1 = c_2 = 0$ by taking $x = 2, 1$.

**Lemma.** *Let $V$ be a vector space over $F$, and $v_1, v_2, \dots, v_n$ a sequence of elements of $V$. Suppose that $\{v_1, \dots, v_n\}$ is dependent. Then for some $k$, $v_k$ may be written as a linear combination of $v_1, \dots, v_{k-1}$.*

*Proof.* The set $\{v_1, \dots, v_n\}$ being dependent means that $\exists c_i \in F$ not all 0 such that $\sum_{i=1}^n c_i v_i = 0$. Let $c_k$ be the last nonzero coefficient. Then $\sum_{i=1}^k c_i v_i = 0 \implies c_k v_k = -\sum_{i=1}^{k-1} c_i v_i \implies v_k = -\frac{1}{c_k} \sum_{i=1}^{k-1} c_i v_i$. □

**Definition.** A vector space is finite-dimensional if it has a finite basis. If it does not, it is infinite-dimensional.

**Theorem 1.** *Let* $\{v_1, \dots, v_n\}$ *be a basis of* $V$. *Suppose* $\{u_1, \dots, u_k\}$ *is independent. Then* $k \leq n$.

*Proof.* We will prove that if $k > n$, then $\{u_1, \dots, u_k\}$ are independent. We have that $u_1 \in \operatorname{span}\{v_1, \dots, v_n\}$, and so $\{v_1, \dots, v_n, u_1\}$ is linearly dependent. The lemme has that some element is a linear combination of previous elements.

If it's $u_1$, then $u_1 = 0$ and we are done.

If not, then $v_i \in \operatorname{span}\{u_1, v_1, v_2, \dots, v_{i-1}\}$ and $\{u_1, v_1, v_2, \dots, \hat{v}_i, \dots, v_n\}$ spans $V$ and thus the set $\{u_1, v_1, \dots, \hat{v}_i, \dots, v_n, u_2\}$ is dependent. We can induct with this argument until we have found a basis $\{u_1, \dots, u_n\}$ spans $V$. Thus, we have that $u_{n+1}$ is a linear combination of the previous elements and we are done. □

**Corollary.** *Any two finite bases have the same number of elements.*

**Definition.** For finite-dimensional vector space $V$, we define the dimension of $V$ as the number of elements in any basis of $V$, and we call it $\dim V$.

*Example.* $\dim \mathbb{R}^n = n$.

**Prop.** *If* $V$ *is a finite dimensional vector space, then any linearly independent set* $\{v_1, \dots, v_n\}$ *may be extended to a basis.*

*Proof.* If $\{v_1, \dots, v_n\}$ spans $V$, then we are finished. Otherwise pick any $v_{k+1} \in V - \operatorname{span}\{v_1, \dots, v_k\}$. By the lemma earlier, this new set is still linearly independent! Continue inductively up to $k = \dim V$ and that is a basis. □

**Prop.** *If* $V$ *is finite dimensional and* $\{v_1, \dots, v_n\}$ *spans* $V$, *then there is a subset that is a basis for* $V$.

*Proof.* If no $v_i$ is a linear combination of the preceding elements, then we're done by the earlier lemma.

Otherwise, consider the first $v_j$ such that $v_j \in \operatorname{span}\{v_1, \dots, v_{j-1}\}$. We have that $\{v_1, \dots, \hat{v}_j, \dots, v_n\}$ still spans $V$. Continuing inductively, we have that at some point the set will be linearly independent while still spanning $V$. □

**Prop.** *Let* $\{v_1, \dots, v_n\} \subseteq V$ *which is finite dimensional; any two of the following implies the third:*

1. $\{v_i\}$ *are independent*
2. $\{v_i\}$ *span* $V$
3. $\dim V = n$

*Proof.* $(1,2 \implies 3)$ This follows from the definition of a basis and dimension.

$(1,3 \implies 2)$ Follows from the earlier proposition which lets us expand $\{v_i\}$ into a basis; however, this must have $n$ elements, so we add nothing.

$(2,3 \implies 1)$ Follows from the previous proposition as we can shrink to a basis, but we are already the correct size. □

**Theorem 2.** *(Construction principle) Take vector spaces $V, W$ with $v_1, \ldots, v_n$ an ordered basis for $V$ and $w_1, \ldots, w_n$ are $n$ elements of $W$. Then there exists a unique linear map taking $\forall i, v_i \mapsto w_i$.*

*Proof.* Define $F$ as follows:

$$F(v) = F(\sum_{i=1}^{n} c_i v_i) = \sum_{i=1}^{n} c_i w_i$$

The rest follows fairly easily. □

*Example.* Take $V = \operatorname{span}\{1, x, x^2, x^3, \ldots, x^n\} \subset \mathcal{F}(\mathbb{R}, \mathbb{R})$ and $W = \mathbb{R}^{n+1}$, and with $w_i = e_i$. Then there are two linear maps sending $x^i$ to $e_{i+1}$ and vice versa.

These are inverse maps! Therefore, $V \cong W$.

**Corollary.** *Any finite dimensional vector space over $F$ with dimension $n$ is isomorphic to $F^n$.*

**Theorem 3.** *(rank-nullity) Let $T : V \to W$, with $V$ finite-dimensional. Then, $\operatorname{im} T$ is finite-dimensional and*

$$\dim(\ker T) + \dim(\operatorname{im} T) = \dim V$$

*Proof.* Let $\{v_1, \ldots, v_k\}$ be a basis of $\ker T$. Extend to $\{v_1, \ldots, v_n\}$, a basis of $V$.

We will show that $T(v_{k+1}), \ldots, T(v_n)$ is a basis of $\operatorname{im} T$. Let $y \in \operatorname{im} T$, so $y = T(x)$ for some $x \in V$. Put $x = \sum_{i=1}^{n} c_i v_i$.

$$y = T(x) = T(\sum_{i=1}^{n} c_i v_i) = \sum_{i=1}^{n} c_i T(v_i) = \sum_{i=k+1}^{n} c_i T(v_i)$$

We will show that the rest are linearly independent as well: if $\sum_{i=k+1}^{n} c_i T(v_i) = 0$, then $T(\sum_{i=k+1}^{n} c_i v_i) = 0 \implies \sum_{i=k+1}^{n} c_i v_i \in \ker T \implies \sum_{i=k+1}^{n} c_i v_i = \sum_{i=1}^{k} c_i v_i$.

$$0 = \sum_{i=k+1}^{n} c_i v_i - \sum_{i=1}^{k} c_i v_i \implies c_i = 0$$

Counting the number of elements yields the theorem. □

**Corollary.** *If* $\dim V = \dim W$*, where both are finite dimensional over the same field, and* $T : V \to W$ *is linear, then* $T$ *is injective* $\iff$ *it is surjective.*

*Proof.* $T$ is injective $\iff$ $\ker T = \{0\}$ $\iff$ $\dim(\ker T) = 0$ $\implies$ $\dim(\operatorname{im} T) = \dim V = \dim W$ $\implies$ $V = W$ $\implies$ $\operatorname{im} T = W$. □

*Remark.* If $S, T$ are finite sets and $|S| = |T|$ and $f : S \to T$, then $f$ is injective $\iff$ $f$ is surjective.

*Remark.* This absolutely is not true in the case of infinite dimension; consider $V = W = \mathbb{Z}_{>\prime}, \mathbb{R}$ and the map $T : \{a_1, a_2, a_3, \dots\} \mapsto \{0, a_1, a_2, a_3, \dots\}$. This is injective but not surjective. Similarly, $T' : \{a_1, a_2, \dots\} \mapsto \{a_2, \dots\}$ is surjective but not injective. Both $T, T'$ are linear.

# 3 Matrices

**Definition.** The standard basis vectors of $\mathbb{R}^n$ are called $e_1, e_2, \dots, e_n$ where $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ where the 1 is in the $i^{th}$ place.

Specifically, per the Kronecker Delta,

$$(e_i)_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Where $x_i$ is the $i^{th}$ component of $x = (x_1, x_2, \dots, x_n)$.

**Prop.** $\forall x \in \mathbb{R}^n$, $\forall i, x_i = a_i \iff x = \sum_{i=1}^{n} a_i e_i$*. In other words, we have that* $x = \sum_{i=1}^{n} x_i e_i$*. Any vector in* $\mathbb{R}^n$ *can be uniquely written as a linear combination of the standard basis vectors.*

**Definition.** For $m, n \in \mathbb{Z}_{>0}$, an $m \times n$ matrix over a field $F$ is a $m \times n$ box of elements of $F$. The convention is that the order goes rows $\times$ columns.

More mathematically, an $m \times n$ matrix over $F$ is a function $[m] \times [n] \to F$. We put $A_{ij}$ for $A((i, j))$.

Let the set of $m \times n$ matrices over $F$ be called $M_{m \times n}(F)$.

*Example.*

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$$

**Prop.** $M_{m \times n}(F)$ *is a vector space, where* $(A + B)_{ij} = A_{ij} + B_{ij}$ *for matrices* $A, B$*. Similarly, we have that* $(cA)_{ij} = c(A_{ij})$ *for* $f \in F$*.*

**Theorem 4.** $\mathcal{L}(V, W) \cong M_{m \times n}(F)$

**Prop.** *Let $V, W$ be finite dimensional vector spaces over $F$ and choose ordered bases $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$. Then $m : \mathcal{L}(V, W) \to M_{m \times n}(F)$ defined by $m(T) = A$ where $T(v_j) = \sum_{i=1}^m A_{ij} w_i$ is an isomorphism of vector spaces.*

*Proof.* If $m(T) = A$, $m(S) = B$, then

$$\begin{aligned}(T + S)(v_j) &= T(v_j) + S(v_j) \\ &= \sum_{i=1}^m A_{ij} w_i + \sum_{i=1}^m B_{ij} w_i \\ &= \sum_{i=1}^m (A + B)_{ij} w_i\end{aligned}$$

Proof for scalar multiplication is similar.

To check injectivity, if $m(T) = m(S)$, then $T(v_j) = S(v_j) forallj$; this shows $T = S$ by the uniqueness part of the construction theorem.

For surjectivity, we have that given $A$, let $T$ be the linear map taking $v_j$ to $\sum_{i=1}^m A_{ij} w_i$. Then $m(T) = A$. □

*Example.* Take $V = \mathbb{R}^2, W = \mathbb{R}^3$, and let $T((x, y)) = (x + 2y, -y, -x)$. Taking the standard bases,

$$\begin{aligned}e_1 &\mapsto (1, 0, -1) \\ e_2 &\mapsto (2, -1, 0) \\ m(T) &= \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ -1 & 0 \end{bmatrix}\end{aligned}$$

*Example.* $V = W = \{$real polynomials with degree less than 4$\}$. Pick basis $\{1, x, x^2, x^3\}$, and consider $D : V \to V$, or the linear map of differentiation.

We have that $D(1) = 0, D(x) = 1, D(x^2) = 2x, D(3x^2)$. Then,

$$m(D) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**Definition.** To do the reverse, we can compute $T(v)$ in terms of $A$.

1. Write $v = \sum_{j=1}^n c_j v_j$.

2.

$$\begin{bmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_1 A_{11} + c_2 A_{12} + \cdots + c_n A_{1n} \\ c_1 A_{21} + c_2 A_{22} + \cdots + c_n A_{2n} \\ \vdots \\ c_1 A_{m1} + c_2 A_{m2} + \cdots + c_n A_{mn} \end{bmatrix}$$

Or,

$$T(v) = \sum_{i=1}^{m} (\sum_{j=1}^{n} c_j A_{ij}) w_i$$

Now, take $U, V, W$ as finite dimensional vector spaces over the same field $F$ and $T : U \to V$, $S : V \to W$, and let $A = m(S), B = m(T), C = m(S \circ T)$. Then, if we have that

$$\begin{aligned} S(v_k) &= \sum_{i=1}^{p} A_{ik} w_i \\ T(u_j) &= \sum_{k=1}^{n} B_{kj} v_k \\ (S \circ T)(u_j) = S(T(u_j)) &= \sum_{k=1}^{n} B_{kj} S(v_k) \\ &= \sum_{k=1}^{n} B_{kj} \sum_{i=1}^{p} A_{ik} w_i \\ &= \sum_{i=1}^{p} \left( \sum_{k=1}^{n} A_{ik} B_{kj} \right) w_i \\ c_{ij} &= \sum_{k=1}^{n} A_{ik} B_{kj} \end{aligned}$$

This defines matrix multiplication.

**Theorem 5.** *Matrix multiplication represents composition of linear maps:*

$$m(S) m(T) = m(S \circ T)$$

*Remark.* The fact that function composition is not commutative is reflected in the fact that matrix multiplication is also not commutative.

**Prop.** *Matrix multiplication is associative and distributive; this can be shown by the fact that they are isomorphic to linear maps.*

# 4 Linear Equations

**Definition.** A system of linear equations in $n$ unknowns is

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = c_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = c_4 \end{cases}$$

Let $A = (a_{ij}) \in M_{m\times n}(F)$ and $\vec{c} = (c_1, \ldots, c_m) \in F^m$ and $\vec{x} = (x_1, \ldots, x_n) \in F^n$.
If we consider vectors in $F^n$ to be matrices with a single column, then the whole system is $A\vec{x} = \vec{c}$.

*Remark.* This corresponds to $T_A(\vec{x}) = \vec{c}$.

**Definition.** A system $A\vec{x} = \vec{c}$ is homogeneous if $\vec{c} = \vec{0}$.
In this case, the solution set is just the kernel of the map and is therefore a subspace.

**Definition.** If $W$ is a vector space with a subspace $V$, a set $S$ is a translate of $V$ if $\exists w \in W \mid S = \{v + w \mid v \in V\}$..

**Theorem 6.** *If $A\vec{x} = \vec{c}$ is any linear system, then its solution set is either empty or a translate of a subspace of* $\ker(T_a)$*, where $T_a$ is the linear transformation corresponding to $A$.*

*Proof.* Suppose the solution set is not empty, so $\exists\vec{k}$ such that $A\vec{k} = \vec{c}$. Then for any given vector $v \in F$,

$$A\vec{v} = \vec{c} \iff A(\vec{v} - \vec{k}) = 0 \iff \vec{v} - \vec{k} = 0$$

Thus, $\vec{v} - \vec{k} \in \ker(T_A)$, and $\vec{v}$ must be in the translate of $\ker(T_A)$ by $\vec{k}$. □

**Theorem 7.** $\ker(T_A)$ *is unchanged by performing any of the following operations on $A$ (elementary row operations):*

1. *Interchange any two rows.*
2. *Multiply any row by a nonzero scalar.*
3. *Adding a scalar multiplied by any given row to a different row.*

*Proof.* (1) follows immediately, as it just corresponds to a different order, which follows from the fact that $p \wedge q = q \wedge p$.

(2) If $\lambda \neq 0$ and $A\vec{v} = \vec{0}$, then $\vec{a_i} \cdot \vec{v} = \vec{0}$, where $\vec{a_i}$ is the $i^{th}$ row of $A$. Then, we have that $\vec{a_i} \cdot \vec{v} = \vec{0} \iff (\lambda \vec{a_i}) \cdot \vec{v} = \vec{0}$.

(3) For any $t \in F$, two equations being satisfied corresponds to $\vec{a_i} \cdot \vec{v} = \vec{0}$ and $\vec{a_j} \cdot \vec{v} = \vec{0}$. This holds if and only if $(\vec{a_j} + t\vec{a_i}) \cdot \vec{v} = \vec{0}$, as the dot product is linear in each variable separately. □

**Definition.** A matrix is in reduced row echelon form (rref) if

- In each row, the first nonzero entry (if any) is 1. Call these 1's "leading" or "pivot" ones.
- Each leading one is to the right of all preceding leading ones.
- Each leading one is the only nonzero entry in its column.
- Rows which are all zero are at the bottom of the matrix.

**Definition.** A variable/column of $A$ is bound if it has a leading one in it, and is free otherwise.

**Prop.** *For any choice of the free coordinates, $\exists$ a unique choice of bound coordinates such that $\vec{x} \in \ker(T_A)$.*

**Corollary.** *A basis of $\ker(T_A)$ is given by one at a time setting each of the free variables to 1, the rest to 0, and solving for the rest (unique by above proposition) which are bound.*

**Prop.** *For a matrix $A$ in rref, $\ker(T_A)$ is easily written down.*

**Theorem 8.** *(Gauss-Jordan Elimination) Any matrix can be put in rref via elementary row operations.*

*Proof.* (This is a sketch, since the actual proof is annoying!) Given $A$, if the first column $= \vec{0}$, move on. Otherwise, swap rows until the top left corner has a non-zero entry, say $a_{11}$. Multiply the row by $a_{11}^{-1}$. The next step is to multiply the first column by $a_{i1}$ and add to the $i^{th}$ column.

If the second column $= \vec{0}$ after the leading ones, move on. Otherwise, swap rows to arrange that the first available slot $\neq 0$. Then, multiply by the inverse of the leading entry to normalize to one, and eliminate the rest of the column as before.

Repeat until no longer possible; swap all $\vec{0}$ rows to the bottom, and we are done. □

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -6 & -12 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

*Remark.* Solving $A\vec{x} = \vec{c}$ is equivalent to solving

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - c_1x_{n+1} = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - c_2x_{n+1} = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - c_mx_{n+1} = 0 \end{gathered}$$

with $x_{n+1} = -1$. In some sense, this is equivalent to making the augmented matrix $[A \mid \vec{c}] \in M_{m\times(n+1)}(F)$.

**Definition.** Let $A \in M_{m\times n}$.

- The row space of $A$ is $\text{span}\{\text{rows}\} \subseteq F^n$.
- The column space of $A$ is $\text{span}\{\text{columns}\} \subseteq F^n$.
- The row rank is the rank of the row space.
- The column rank is the rank of the column space.

**Corollary.** *Row (respectively columns) rank is the max number of linearly independent rows (respectively columns).*

**Corollary.** *The column space is the image of the corresponding linear transformation.*

**Theorem 9.** *row rank = column rank =* $\text{rank}(A)$*.*

*Proof.* Elementary row operations don't change the row space. Hence, we have that the row rank of $A$ is the row rank of $A'$ where $A' = \text{rref of } A$.

The row rank of $A'$ is the number of leaning ones, which follows immediately from the properties of rref matrices. This is also the column rank of $A'$, which is easily seen to be the amount of bound variables.

$$\begin{aligned} \text{row rank}(A') &= \text{number of leading ones} \\ &= \text{number of bound variables} \\ &= n - \text{number of free variables} \\ &= n - \dim(\ker(T_A)) \\ &= \dim(\text{im}(T_A)) \\ &= \text{column rank of } A \end{aligned}$$

□

*Remark.* In particular, we have that $\text{rank}(A) \le \min(m, n)$, and that $T_A$ is an isomorphism $\iff m = n$ and $\text{rank}(A) = m$.

**Definition.** The $n \times n$ identity matrix $I_n$ has that $T_{I_n} = \text{id} : V \to V$. Equivalently, it is a box of numbers that has $(I_n)_{ij} = \delta_{ij}$.

**Definition.** $A \in M_{m\times n}(F)$ is invertible if $\exists B \in M_{n\times m}(F)$ such that $AB = I_m$ and $BA = I_m$. We say that $B$ is an inverse to $A$, and write $A^{-1} = B$.

**Corollary.** *A homework problem easily leads us to the fact that only square matrices are invertible. Further, $A$ is invertible $\iff T_A$ is an isomorphism.*

**Prop.** *If an inverse exists, it is unique.*

*Proof.* Suppose that $B, B'$ are both inverses to $A$. Then,

$$B = BI = B(AB') = (BA)B' = IB' = B'$$

□

**Prop.** *For $A, B \in M_{n\times n}$,*

$$AB = I_n \iff BA = I_n$$

*Proof.* We have that $AB = I_n \iff T_A \circ T_B = id_F^n \implies T_A$ is surjective $\implies A$ has a two sided inverse $B'$. However, the above manipulation,

$$B = BI = B(AB') = (BA)B' = IB' = B'$$

still holds. □

**Prop.** *$A \in M_{n\times n}$ is invertible $\iff$ its rref is $I_n$.*

**Theorem 10.** *(Computing inverses) Suppose that $A \in M_{n\times n}$ is invertible and becomes $I_n$ after some sequence of elementary row operations.*

*Then $A^{-1}$ is obtained from $I_n$ by the same operations.*

*Proof.* Each elementary row operation is equivalent to left multiplication by a certain matrix $B$.

Swapping $i, j$ rows is the same as left multiplication by the identity matrix with the $i, j$ rows swapped.

Multiplying a row by $\lambda$ is the same as left multiplication by the identity matrix with that row's leading one replaced by $\lambda$.

Adding a multiple $\lambda$ of another row (say row $i$) to row $j$ is the same as left multiplication by the identity matrix with the element at $j, i$ set to $\lambda$.

Hence,

$$(B_k B_{k-1} dots B_1)A = I_n \implies A^{-1} = (B_k B_{k-1} dots B_1) I_n$$

□

*Remark.* We can augment a matrix by the identity, row reduce, and the right side will be the inverse.

*Example.*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

# 5 Determinants

*Remark.* $n$ vectors in $\mathbb{R}^n$ are only not a basis if the corresponding parallelepiped is 0. This is the overall motivation of the determinant!

**Definition.** If $\det : M_{n\times n}(F) \to F$ is regarded as a function of the row vectors $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$, then we want:

1. $\det(\vec{a_1}, \vec{a_2}, \ldots, c\vec{a_i}, \ldots, \vec{a_n}) = c \det(\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n})$.
2. $\det(\vec{a_1}, \vec{a_2}, \ldots, veca_i + \vec{b}, \ldots, \vec{a_n}) = \det(\vec{a_1}, \vec{a_2}, \ldots, \vec{a_i}, \ldots, \vec{a_n}) + \det(\vec{a_1}, \vec{a_2}, \ldots, \vec{b}, \ldots, \vec{a_n})$.
3. $\vec{a_i} = \vec{a_j}$ for $i \neq j \implies \det(\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}) = 0$.
4. $\det(I_n) = 1$.

**Lemma.** *We will have the following:*

1. *If any* $\vec{a_i} = \vec{0}$, *then* $\det = 0$.
2. $\det(\vec{a_1}, \ldots, \vec{a_i}, \ldots, \vec{a_j}, \ldots, \vec{a_n}) = \det(\vec{a_1}, \ldots, \vec{a_j}, \ldots, \vec{a_i}, \ldots, \vec{a_n})$
3. *If a matrix $A$ is not of full rank, then* $\det(A) = 0$.

*Proof.* (1) $\det(\vec{a_1}, \vec{a_2}, \ldots, 0, \ldots, \vec{a_n}) = 0 \det(\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}) = 0$

(2)

$$\begin{aligned} 0 &= \det(\vec{a_1}, \ldots, \vec{a_i} + \vec{a_j}, \ldots, \vec{a_i} + \vec{a_j}, \ldots, \vec{a_n}) \\ &= \det(\vec{a_1}, \ldots, \vec{a_i}, \ldots, \vec{a_i}, \ldots, \vec{a_n}) \\ &\quad + \det(\vec{a_1}, \ldots, \vec{a_i}, \ldots, \vec{a_j}, \ldots, \vec{a_n}) \\ &\quad + \det(\vec{a_1}, \ldots, \vec{a_j}, \ldots, \vec{a_i}, \ldots, \vec{a_n}) \\ &\quad + \det(\vec{a_1}, \ldots, \vec{a_j}, \ldots, \vec{a_j}, \ldots, \vec{a_n}) \end{aligned}$$

$$\begin{aligned}&= \det(\vec{a_1}, \dots, \vec{a_i}, \dots, \vec{a_j}, \dots, \vec{a_n}) \\ &\quad + \det(\vec{a_1}, \dots, \vec{a_j}, \dots, \vec{a_i}, \dots, \vec{a_n})\end{aligned}$$

(3) The rows being dependent mean that some $\vec{a_k} = \sum_{i=1}^{k-1} c_o \vec{a_i}$. Then,

$$\begin{aligned}\det(\vec{a_1}, \dots, \vec{a_k}, \dots, \vec{a_n}) &= \det(\vec{a_1}, \dots, \sum c_i \vec{a_i}, \dots, \vec{a_n}) \\ &= \sum_{i=1}^{k-1} c_i \det(\vec{a_1}, \dots, \vec{a_i}, \dots, \vec{a_n}) \\ &= 0\end{aligned}$$

□

**Lemma.** *For $A \in M_{n\times n}(F)$, choose a sequence of elementary row operations taking it to rref. Let $R$ be the number of times that we swap rows, and let $\lambda_1, \dots, \lambda_m$ be the nonzero scalars we multiply rows by.*

*Then, we have that*

$$\det(A) = \begin{cases} 0 & \text{rank}(A) < n \\ \frac{(-1)^R}{\prod_i^m \lambda} & \text{rank}(A) = n \end{cases}$$

*Proof.* We already know that swapping rows multiplies the determinant by $-1$.

Similarly, we have that multiplying by $\lambda_i$ also multiplies the determinant.

Lastly,

$$\begin{aligned}&\det(\vec{a_1}, \dots, \vec{a_i} + t\vec{a_j}, \dots, \vec{a_n}) \\ =& \det(\vec{a_1}, \dots, \vec{a_i}, \dots, \vec{a_n}) \\ &+ t\det(\vec{a_1}, \dots, \vec{a_j}, \dots, \vec{a_n}) \\ =& 0\end{aligned}$$

If $\text{rank}(A) < n$, we already know that $\det(A) = 0$. Otherwise, row reduce to $I_n$. Therefore, we have that $\det(A)$ is the scalar such that it is $\det(I_n)$ when multiplied by $(-1)^R \prod \lambda$. Thus, we have that

$$\det(A) \cdot (-1)^R \prod \lambda = \det(I_n) = 1$$

□

**Theorem 11.** *There exists a unique function satisfying the above constraints.*

*Proof.* Suppose that there are two functions det and $\widetilde{\det}$ both satisfy the above axioms. Then, given $A \in M_{n\times n}(F)$, row reduce to $rref$. Then, we have that

$$\det(A) = \begin{cases} 0 & \text{rank}(A) < n \\ \frac{(-1)^R}{\prod_i^m \lambda} & \text{rank}(A) = n \end{cases} = \widetilde{\det}(A)$$

□

**Definition.** The $(i, j)$ cofactor matrix of $A$ is the $(n-1) \times (n-1)$ matrix $A^{ij}$ obtained by deleting the $i^{th}$ row and $j^{th}$ column.

**Lemma.** *Fix some index $i$ and write $\vec{a_i} = \sum_{j=1}^n A_{ij}\vec{e_j}$. Then,*

$$\begin{aligned} &\det(\vec{a_1}, \dots, \vec{a_i}, \dots, \vec{a_n}) \\ =&\det(\vec{a_1}, \dots, \sum_{j=1}^n A_{ij}\vec{e_j}, \dots, \vec{a_n}) \\ =&\sum_{j=1}^n A_{ij} \det(\vec{a_1}, \dots, \vec{e_j}, \dots, \vec{a_n}) \end{aligned}$$

*We can zero out, in the $j^{th}$ term, the $j^{th}$ column via elementary row operations. Call the $j^{th}$ such matrix $B$. Then, we have that*

$$(-1)^{i+j} \det(B) = \det(A_{ij})$$

*Proof.* We simply check that the left side satisfies the axioms such that the transformations applied to $A_{ij}$ also induce the same change in the left side. In that case, we have by uniqueness that they are the same.

The first three come easily. In the last case, if $A^{ij} = I_{n-1}$, then rows of $B$ are

$$\vec{e_1}, \vec{e_2}, \dots, \vec{e_{i-1}}, \vec{e_j}, \vec{e_{i+1}}, \dots, \hat{\vec{e_j}}, \dots, \vec{e_n}$$

It takes now $|i - j|$ row flips to get from $B$ to $I_n$, which is equal to $i + j$ in parity. □

**Theorem 12.** *The determinant exists!*

*Proof.* If det exists, it has that $\det A = \sum_{j=1}^n A_{ij}(-1)^{i+j} \det(A^{ij})$.

Take $\det([c]) = c$. Then, assume that we have a determinant in dimension $n - 1$. Then, define $\det(A) = \sum_{j=1}^n (-1)^{i+j} A_{ij} \det(A^{ij})$. This satisfies the axioms in dimension $n$ (this can be checked, but this is annoying). □

**Definition.** Let $V$ be a finite dimensional vector space over $F$. Take two bases $\{v_1, \dots v_n\}$ and $\{v'_1, \dots, v'_n\}$, such that $v'_j = \sum_{i=1}^n B_{ij} v_i, v_k = \sum_{j=1}^n C_{jk} v'_j$.
Now, we can form matrices $B, C \in M_{n \times n}(F)$. Call $B$ the change of basis matrix from $\{v'\}$ to $\{v\}$, and $C$ the change of basis matrix from $v$ to $v'$.

**Prop.** *$B, C$ from above are inverse matrices:*

$$BC = I_n$$

*Proof.*

$$\begin{aligned} v_k &= \sum_{j=1}^n C_{jk} (\sum_{i=1}^n B_{ij} v_i) \\ &= \sum_{j=1}^n (\sum_{j=1}^n B_{ij} C_{jk}) v_i \\ &= \sum_{i=1}^n \delta_{ik} v_i \end{aligned}$$

This means that $\sum_j B_{ij} C_{jk} = \delta_{ik} \implies BC = I_n$. □

**Prop.** *Let $T : V \to U$ is a linear map for vector spaces over $F$. Then, suppose that $\{u_1, \dots, u_n\}$ and $\{v_1, \dots, v_n\}$ are bases for $U, V$, and that $\{v'_1, \dots, v'_n\}$ is also a basis for $V$.*
*Let $A =$ matrix correlated to $\{u_i\}, \{v_i\}$, $A' =$ matrix correlated to $\{u_i\}, \{v'_i\}$.*

$$A' = AB$$

*Proof.*

$$\begin{aligned} T(v_j) &= \sum_{i=1}^n A_{ij} u_i \\ T(v'_k) &= \sum_{i=1}^n A'_{ik} u_i \\ v'(k) &= \sum_{j=1}^n B_{jk} v_k \\ T(v'_k) &= T(\sum_{j=1}^n B_{jk} v_j) \\ &= \sum_{j=1}^n B_{jk} T(v_j) \end{aligned}$$

$$= \sum_{i=1}^{n}(\sum_{j=1}^{n} A_{ij}B_{jk})u_i$$

Since $u_i$ are linearly independent, $A' = \sum_j A_{ij}B_{jk} \implies A' = AB$. □

**Corollary.** *Similarly, if we have $\{u_1, \dots u_n\}, \{u'_1, \dots u'_n\}$ bases of $U$, and $C$ the change of basis matrix from $\{u_i\}$ to $\{u'_i\}$, and the $A'$ are the same as above with $u, v$ switched, then*

$$A' = CA$$

**Prop.** *If $T : V \to V$ with bases $\{v_i\}, \{v'_i\}$, and $B$ is a change of base matrix, and $A$ is the matrix with $\{v_i\}$ and $A'$ the other one:*

$$A' = B^{-1}AB$$

**Definition.** We say that $A, A' \in M_{n\times n}(F)$ are similar (or conjugate) if $\exists B \in M_{n\times n}(F)$ invertible with

$$A' = B^{-1}AB$$

**Corollary.** *Therefore a matrix is diagonalizable if $\exists D$ diagonal with $A$ similar to $D$.*

**Prop.** *If $A$ and $A'$ are similar, then $det(A) = det(A')$.*

*Proof.*

$$\det(A') = \det(B^{-1})\det(A)\det(B) = \det(A)$$

□

*Remark.* Now we can take the determinant of a linear transformation from a domain to itself!

# 6 Eigenvalues

**Prop.** *Let $V$ be any vector space. If $v_1, \dots v_k$ are eigenvectors and $\lambda_1, \dots, \lambda_k$ are the corresponding eigenvalues, and that these eigenvalues are distinct, then the $v_i$ are linearly independent. eigenvalues.*

*Proof.* Induct on $k$. For $k = 1$, it holds trivially.
Assume for $k$. If

$$\sum_{i=1}^{k+1} a_i v_i = 0$$

$$\implies T(\sum_{i=1}^{k+1} a_i v_i) = 0$$

$$
\begin{aligned}
\implies \qquad & 0 = \sum_{i=1}^{k+1} a_i T(v_i) \\
& = \sum_{i=1}^{k+1} a_i \lambda_i v_i \\
\implies \qquad & 0 = \sum_{i=1}^{k+1} (a_i \lambda_i - a_{k+1}\lambda_{k+1}) v_i \\
\implies \qquad & + \sum_{i=1}^{k} (\lambda_i - \lambda_{k+1}) a_i v_i
\end{aligned}
$$

Thus $(\lambda_i - \lambda_{k+1})a_i = 0$, but the eigenvalues are distinct, so $a_i = 0 \implies a_{k+1} = 0$. □

**Corollary.** *If $A \in M_{n\times n}(F)$ has $n$ distinct eigenvalues, then it is diagonalizable.*

*Proof.* The eigenvectors form a basis by the proposition. □

**Prop.** *$\lambda \in F$ is an eigenvalue $\iff \det(A - \lambda I_n) = 0$.*

*Proof.* $\lambda$ is an eigenvalue $\iff \exists v \neq 0 \mid Av = \lambda v \iff Av - \lambda v = 0 \iff (A - \lambda(A))v = 0 \iff \det(A - \lambda I_n) = 0$. □

**Definition.** The set of all eigenvectors corresponding to a given eigenvalue $\lambda$, together with $\{0\}$, is called the eigenspace $E_\lambda$ of $\lambda$.

**Prop.** *The above is a subspace.*

**Prop.** *If $A \in M_{n\times n}$ has eigenvalues $\lambda_1, \dots, \lambda_k$, then*

$$\sum_{i=1}^{k} \dim(E_{\lambda_i}) = n \iff A \text{ is diagonalizable}$$

**Definition.** For $A \in M_{n\times n}(F)$, the characteristic polynomial of $A$, $p_A(\lambda) = \det(\lambda I - A)$.

*Example.* Consider $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. $p_A(\lambda) = \lambda^2 + 1 = 0$ has no real solutions.

**Definition.** The Complex Numbers!

The set $\mathbb{C}$ of complex numbers is $\mathbb{R}^2$ with usual vector addition, and multiplication defined by

$$(x_1, y_2) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1)$$

Take $0 = (0,0), 1 = (1,0)$. Let $i = (0,1)$.

**Definition.** The conjugate of $x + iy$ is:

$$\overline{x + iy} = x - iy$$

The real part of $x + iy$ is:

$$\text{Re}(x + iy) = x$$

The imaginary part is:

$$\text{Im}(x + iy) = x$$

The absolute value of $x + iy$ is:

$$|x + iy| = \sqrt{x^2 + y^2}$$

**Prop.** $\mathbb{C}$ *is a field.*

*Proof.* Check all the axioms. □

*Remark.* Here are some easy formulas:

- $\text{Re}(z) = \frac{z+\overline{z}}{2}$
- $\text{Im}(z) = \frac{z-\overline{z}}{2i}$
- $\overline{z + w} = \overline{z} + \overline{w}$
- $\overline{zw} = \overline{z} \cdot \overline{w}$
- $|z| = z \cdot \overline{z}$

**Theorem 13.** *(Fundamental Theorem of Algebra)* $\mathbb{C}$ *is algebraically closed. Any polynomial with coefficients in* $\mathbb{C}$ *of degree* $n$ *has* $n$ *roots with multiplicity.*

*Remark.* In particular, if $F$ is algebraically closed, and we have some $A \in M_{n\times n}(F)$, then the characteristic polynomial $p_A(\lambda)$ has $n$ roots with multiplicity.

*Example.* Take $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. We can diagonalize over $\mathbb{C}$, as it has eigenvalues $\pm i$.

# 7 Inner Product Spaces

**Definition.** Let $V$ be a real vector space over $F$ (where $F = \mathbb{R}$ or $\mathbb{C}$). An inner product on $V$ is a function $\langle\rangle : V \times V \to F$ such that

1. $\forall c \in \mathbb{R}, \langle cx, y\rangle = c\langle x, y\rangle$

2. $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle$

3. $\langle x, y\rangle = \overline{\langle y, x\rangle}$

4. $\langle x, x\rangle\mathbb{R}, \langle x, x\rangle \geq 0, \langle x, x\rangle = 0 \iff x = \vec{0}$

**Definition.** An $\mathbb{R}$-vector space with an inner product is called an inner product space (or an Euclidean space).

**Definition.** An $\mathbb{C}$-vector space with an inner product is called a complex inner product space (or an Hermitian space).

**Definition.** The length or norm $||x||$ of a vector is $\sqrt{\langle x, x\rangle}$, where we take the positive square root.

**Definition.** The angle between $x$ and $y$ is

$$\arccos\left(\frac{\langle x, y\rangle}{||x|| \cdot ||y||}\right)$$

*Example.* Take $V = \mathbb{R}^n$.

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n)\rangle = \sum_{i=1}^{n} x_i y_i$$

*Example.* Take $V = \mathbb{R}^2$.

$$\langle (x_1, x_2), (y_1, y_2)\rangle = 2x_1y_1 + x_2y_2 - x_1y_2 - x_2y_1$$

*Example.* Take $V = \{f : [0, 1] \to \mathbb{R} \mid f \text{ is continuous}\}$.

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt$$

*Example.* Take $V = \mathbb{C}^n$.

$$\langle z, w\rangle = \sum_{i=1}^{n} z_i \overline{w_i}$$

*Example.* Take $V = \{f : [0, 1] \to \mathbb{R} \mid f \text{ is continuous}\}$.

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt = \int_0^1 \text{Re}(f(t)\overline{g(t)})dt + i\int_0^1 \text{Im}(f(t)\overline{g(t)})dt$$

**Prop.** *The inner product is skew linear in the second variable.*

$$\langle x, cy\rangle = \overline{c}\langle x, y\rangle$$

**Prop.** *(Cauchy-Schwarz inequality) Let $V, \langle\rangle$ be a real or complex inner product space. Then, $\forall x, y \in V$,*

$$|\langle x, y\rangle| \leq ||x|| \cdot ||y||$$

*Proof.* The inequality is easy to check if $x = 0$ or $y = 0$, or if $x, y$ are nonzero multiples of each other.

Assume, then, that the above is not the case. Then, $x - cy \neq 0 \forall c \in \mathbb{R}$ or $\mathbb{C}$.

$$\begin{aligned} 0 &< ||x - cy||^2 \\ &= \langle x - cy, x - cy\rangle \\ &= \langle x, x\rangle + c\bar{c}\langle y, y\rangle - c\langle x, y\rangle - \overline{c\langle y, x\rangle} \end{aligned}$$

Let $\mu$ be a complex number such that $\mu\langle x, y\rangle < |\langle x, y\rangle|$ and $|\mu| = 1$ (in the real case, $\mu = \pm 1$). Write $c = \mu t$ with $t \in \mathbb{R}$, such that $|c| = |t|$. Then,

$$= ||x||^2 + t^2||y||^2 - 2t|\langle x, y\rangle|$$

Pick $t$ to minimize the right hand side, such that

$$\begin{aligned} t &= \frac{|\langle x, y\rangle|}{||y||^2} \\ 0 &< ||x||^2 + \frac{|\langle x, y\rangle|^2}{||y||^2} \end{aligned}$$

Taking the square root of both sides yields the inequality.

□

**Prop.** *(Triangle inequality) With the same hypotheses as before,*

$$||x + y|| \leq ||x|| + ||y||$$

*Proof.* Note that $|\operatorname{Re}(z) \leq |z|$, $\langle x, y\rangle + \langle y, x\rangle = 2\operatorname{Re}\langle x, y\rangle$.

$$\begin{aligned} (||x|| - ||y||)^2 - ||x + y||^2 &= ||x||^2 + 2||x|| \cdot ||y|| + ||y||^2 - ||x||^2 - ||y||^2 - \langle x, y\rangle - \langle y, x\rangle \\ &= 2||x|| \cdot ||y|| - 2\operatorname{Re}(\langle x, y\rangle) \\ &\geq 2||x|| \cdot ||y|| - 2|\langle x, y\rangle| \\ &\geq 0 \end{aligned}$$

□

**Definition.** Two vectors $x, y$ are orthogonal if $\langle x, y\rangle = 0$.

**Definition.** Let $V$ be an inner product space. A sequence $v_1, \dots, v_n$ is orthonormal if $\forall i, j \langle v_i, v_j \rangle = \delta_{ij}$.

**Prop.** *An orthonormal sequence is linearly independent.*

*Proof.* If $\vec{0} = \sum_{j=1}^{n} a_j v_j$ then for each $k$,

$$0 = \langle \vec{0}, v_k \rangle = \langle \sum_{j=1}^{n} a_j v_j, v_k \rangle = \sum_{j=1}^{n} a_j \langle v_j v_k \rangle = a_k$$

□

**Prop.** *Let $v_1, \dots, v_n$ be an orthonormal basis of $V$ and $x \in V$. Then*

$$x = \sum_{j=1}^{n} \text{ where } a_k = \langle x, v_k \rangle$$

*Proof.* Since $v_1, \dots v_n$ is a basis, $\exists$ a unique $a_j \mid x = \sum_{j=1}^{n} a_j v_j$. By uniqueness, we just have to show that taking $a_k = \langle x, v_k \rangle$ works.

$$\langle x, v_k \rangle = \langle \sum_{j=1}^{n} a_j v_j \rangle = \sum_{j=1}^{n} a_j v_j \langle v_j, v_k \rangle = a_k$$

□

**Definition.** Let $V$ be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, and $S \subseteq V$ a subspace.
The orthogonal complement $S^\perp$ of $S$ is

$$S^\perp := \{v \in V \mid \forall s \in S, \langle s, v \rangle = 0\}$$

It is left as an exercise that $S^\perp$ is a subspace of $V$.

**Prop.** *Let $V$ be an inner product space, $S \subseteq V$ a finite-dimensional subspace with an orthonormal basis. Then $\forall x \in V$, $\exists$ a unique $x^S \in S, x^\perp \in S^\perp \mid x = x^S + x^\perp$.*

*Proof.* We show uniqueness first. Suppose $x^S + x^\perp = x = y^S + y^\perp$, and let $z = x^S - y^S = y^\perp - x^\perp = 0 \in S \cap S^\perp$.
However, $z \in S^\perp \implies \forall w \in S, \langle w, z \rangle = 0$. By picking $w = z, \langle z, z \rangle = 0 \implies z = 0$.
For existence, let $S$ have orthonormal basis $v_1, \dots, v_n$ be an orthonormal basis for $S$, $a_j = \langle x, v_j \rangle$, $x^S = \sum_{j=1}^{n} a_j v_j$, and $x^\perp = x - x^S$.
Now, pick any $y \in S$, where $y = \sum_{k=1}^{n} b_k v_k$.

$$\langle x^\perp, y \rangle = \langle x - x^S, y \rangle$$

$$
\begin{aligned}
&= \langle x, y\rangle - \langle x^S, y\rangle \\
&= \langle x, \sum_{k=1}^n b_k v_k\rangle - \langle \sum_{j=1}^n a_j v_j, \sum_{k=1}^n b_k v_k\rangle \\
&= \sum_{k=1}^n \overline{b}_k \langle x, v_k\rangle - \sum_{j=1}^n \sum_{k=1}^n a_j \overline{b_k} \langle v_j, v_k\rangle \\
&= \sum_{k=1}^n \overline{b_k} a_k - \sum_{k=1}^n a_k \overline{b_k}
\end{aligned}
$$

□

**Theorem 14.** *Any finite dimensional inner product space has an orthonormal basis.*

*Proof.* Induct on the dimension $n$ of $V$. For $n = 0$, we immediately have an orthonormal basis. Assume this for spaces of dimension $n = k - 1$, and $V$ has some basis $u_1, \ldots, u_k$. Further, $S = \text{span}\{u_1, \ldots u_k\}$ has an orthonormal basis $v_1, \ldots v_{k-1}$.

Apply the previous proposition to $S$ and $\text{span}\{u_1, \ldots, u_k\}$. Then, there exists a unique $u^S \in S, u^\perp \in S^\perp \mid u^S + u^\perp = u_k$.

Since $u_k \notin S$, $u^\perp \neq 0$. Let $v_k = \frac{u^\perp}{||u^\perp||}$. Then, for $j < k$,

$$
\langle v_j, v_l\rangle = \langle v_j, \frac{u^\perp}{||u^\perp||}\rangle = \frac{1}{||u^\perp||}\langle v_j, u^\perp\rangle = 0
$$

□

*Remark.* The above also gets a type of formula as well. Call this the Gram-Schmidt process.

**Corollary.** *If $S \subseteq V$, $S$ a finite dimensional subspace, $\exists$ a unique decomposition $x = x^S + x^\perp, x^S \in S, x^\perp \in S^\perp$.*

**Corollary.** *(Pythagorean theorem) For $S \subseteq V$, a finite dimensional subspace, $x \in V$,*

$$
||x||^2 = ||x^S||^2 + ||x^\perp||^2
$$

*Proof.*

$$
\begin{aligned}
||x||^2 &= \langle x, x\rangle \\
&= \langle x^S + x^\perp, x^S + x^\perp\rangle \\
&= \langle x^S, x^S\rangle + \langle x^S, x^\perp\rangle + \langle x^\perp, x^S\rangle \langle x^\perp, x^\perp\rangle \\
&= \langle x^S, x^S\rangle + \langle x^\perp, x^\perp\rangle \\
&= ||x^S||^2 + ||x^\perp||^2
\end{aligned}
$$

□

**Exercise.** *Let $v_1, \dots v_n$ be an orthonormal basis of $V$, and $x = \sum_{j=1}^n x_j v_j, y = \sum_{j=1}^n y_j v_j$. Then,*

$$\langle x, y \rangle = \sum_{j=1}^n x_j \overline{y_j} = \sum_{j=1}^n \langle x, v_j \rangle \overline{\langle y, v_j \rangle}$$

**Definition.** Let $V, W$ be inner product spaces and $T : V \to W$ linear.
An adjoint of $T, T^* : W \to V$ linear, satisfies for $v \in V, w \in W$,

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle$$

**Prop.** *Adjoints are unique if they exist.*

*Proof.* Let $T^*, T'$ both be adjoints. Then,

$$\begin{aligned} \langle x, T'(y) \rangle = \langle x, T^*(y) \rangle = \langle T(x), y \rangle \\ \langle x, T'(y) - T^*(y) \rangle = 0 \end{aligned}$$

Take $x = T'(y) - T^*(y)$.

$$\begin{aligned} \langle T'(y) - T^*(y), T'(y) - T^*(y) \rangle = 0 \\ T'(y) = T^*(y) \\ T' = T^* \end{aligned}$$

□

# 8 Multivariable Calculus

*Remark.* We now care about functions $f : \mathbb{R}^n \to \mathbb{R}^m$.

**Definition.** We will call $f : \mathbb{R} \to \mathbb{R}^n$ a parametric curve, and $f : \mathbb{R}^n \to \mathbb{R}^n$ a vector field.

**Definition.** A closed rectangle in $\mathbb{R}^n$ is a set of the form

$$[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

# 9 Limits and Continuity

**Definition.** For $S \subset \mathbb{R}^n, F : S \to \mathbb{R}^n, c \in S, y \in \mathbb{R}^m$, we put

$$\lim_{x \to c} F(x) = y$$

if $\forall \epsilon > 0, \exists \delta > 0$ such that $\forall x \in S, 0 < ||x - \delta|| < \delta \implies ||F(x) - y|| < \epsilon$.

**Definition.** For $F : S \to \mathbb{R}^n$ as above, say $F$ is continuous at $c \in S$ if

$$\lim_{x \to c} F(x) = F(c)$$

and is just continuous if it is continuous at every $c \in S$.

**Theorem 15.** *If* $\lim_{x \to c} F(x) = y$ *and* $\lim_{x \to c} G(x) = z$*, then*

1. $\lim_{x \to c}[aF(x) + bG(x)] = ay + bz$
2. $\lim_{x \to c}[F(x) \cdot G(x)] = y \cdot z$
3. $\lim_{x \to c} ||F(x)|| = ||y||$

*Proof.* 1 is exactly the same as in single-variable calculus.
2 is *morally* the same as in single-variable calculus.
Note that from Cauchy-Schwarz,

$$\begin{aligned}(F(x) \cdot G(x) - y \cdot z) &= (F(x) - y) \cdot (G(x) - z) + y \cdot (G(x) - z) + (F(x) - y) \cdot z \\ &\leq ||F(x) - y|| \cdot ||G(x) - z|| + ||y|| \cdot ||G(x) - z|| + ||F(x) - y|| \cdot ||z||\end{aligned}$$

Given $\epsilon > 0$, take $\delta > 0$ such that

$$\begin{aligned} 0 < ||x - c|| < \delta \implies \\ ||F(x) - y|| < \sqrt{\frac{\epsilon}{3}} \\ ||G(x) - z|| < \sqrt{\frac{\epsilon}{3}} \\ ||F(x) - y|| < \frac{\epsilon}{3||z||} \\ ||G(x) - z|| < \frac{\epsilon}{3||y||} \end{aligned}$$

Then, we have that $||F(x) \cdot G(x) - y \cdot z|| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$.
3 is left as an exercise. □

**Corollary.** *$F, G : S \to \mathbb{R}^n, h : S \to \mathbb{R}$ continuous, then $F + G$, $F \cdot G$, $hF$ are all continuous.*

**Definition.** Let $\pi_i : \mathbb{R}^m \to \mathbb{R}$, for $1 \leq i \leq m$, be defined by

$$\pi_i(x_1, \ldots, x_m) = x_i$$

This is the $i^{th}$ projection.

Similarly, for $F : \mathbb{R}^n \to \mathbb{R}^m$ define the $i^{th}$ component $F_i : \mathbb{R}^n \to \mathbb{R}$ as follows:

$$F_i = \pi_i \circ F$$

**Prop.**

$$\forall x \in \mathbb{R}^n, F(x) = (F_1(x), F_2(x), \ldots, F_m(x))$$

*Proof.* Each side has component $(\pi_i \circ F)(x) = F_i(x)$. □

**Theorem 16.**

$$\lim_{x \to c} F(x) = (y_1, \ldots, y_m) \iff \forall i, \lim_{x \to c} F_i(x) = y_i$$

**Corollary.** *Which specific inner product doesn't matter!*

*Proof.* ( $\implies$ ) Let $E_i : \mathbb{R}^n \to \mathbb{R}^m$, where $E_i(x) = e_i$. Take $\delta = \pi$ (or whatever your favorite real number is) and it is easy to see that it is continuous. Then,

$$F_i(x) = F(x) \cdot E_i(x)$$

The previous theorem implies that $F_i$ is continuous at $x$.

$$\lim_{x \to c} F_i(x) = \lim_{x \to c} F(x) \cdot \lim_{x \to} E_i(x) = y_i$$

( $\impliedby$ ) We claim that

$$F(x) = \sum_{i=1}^{m} F_i(x) E_i(x)$$

This is fairly easy to see, and is continuous as above. Then,

$$\begin{aligned} \lim_{x \to c} F(x) &= \lim_{x \to c} \left( \sum_{i=1}^{m} F_i(x) E_i(x) \right) \\ &= \sum_{i=1}^{m} \lim_{x \to c} y_i e_i \\ &= (y_1, \ldots, y_m) = y \end{aligned}$$

□

**Theorem 17.** *Any $F : \mathbb{R}^n \to \mathbb{R}^m$ linear $\implies$ $F$ continuous.*

*Proof.* Note that the projections are linear; thus, $F_i = \pi_i \circ F$ must also be linear. By the previous theorem, it suffices to show $m = 1$.

A linear algebra fact is that for any linear function $G : \mathbb{R}^n \to \mathbb{R}$, there exists a vector $c \in \mathbb{R}^n$ such that $G(x) = x \cdot c$.

Therefore, $\lim_{x \to c} F(x) = \lim_{x \to c}(x \cdot a) = \lim_{x \to c} \cdot a = c\dot{a} = F(c)$. □

**Theorem 18.** *If $F : \mathbb{R}^n \to^m$, $G : \mathbb{R}^m \to \mathbb{R}^l$ are continuous, then so is $G \circ F$.*

*Proof.* Exactly the same as in the one dimensional case. □

# 10 Derivatives

**Definition.** For $\epsilon > 0$, the open $\epsilon$-ball around $x \in R^n$ is

$$B_\epsilon(x) := \{y \in \mathbb{R}^n \mid ||x - y|| < \epsilon\}$$

.

**Definition.** A set $U \subseteq \mathbb{R}^n$ is open if $\forall x \in U, \exists \epsilon > 0$ such that $B_\epsilon(x) \subseteq U$.

*Example.* The following are open sets:

1. $\emptyset$, vacuously.
2. $\mathbb{R}^n$, taking any $\epsilon$.
3. $(a, b) \in \mathbb{R}$, for any $x \in (a, b)$ take $\epsilon = \min(x - a, b - x) > 0$.
4. $B_\delta(x)$, using the triangle inequality.
5. $(a_1, b_1) \times \cdots \times (a_n, b_n)$ is open.

The following is *not* an open set:

1. $[a, b]$, considering $a, b$.

**Prop.**

$$U, V \subseteq \mathbb{R}^n \textit{ open } \implies U \cup V, U \cap V \textit{ open}$$

.

*Proof.* $\exists e_1 \mid B_{\epsilon_1} \subseteq U, e_2 \mid B_{\epsilon_2} \subseteq V$. Take $\epsilon = \min(\epsilon_1, \epsilon_2)$. Then, $B_\epsilon \subseteq U \cap V$. □

**Definition.** Let $F : \mathbb{R}^n \to \mathbb{R}$.

Given $x, y \in \mathbb{R}^n$ let

$$L_{x,y}(t) := x + ty, t \in \mathbb{R}$$

**Definition.** Let $U \subseteq \mathbb{R}^n$ be open, $F : U \to \mathbb{R}$, $x \in U$, $y \in \mathbb{R}^n$. The directional derivative of $F$ at $x$ along $y \neq 0$ is

$$F'(x; y) := (F \circ L'_{x,y})(0)$$

if it exists.

In other words,

$$F'(x; y) = \lim_{t \to 0} \frac{F(x + ty) - F(x)}{t}$$

Note that since $U$ is open, we have that $F(x + ty)$ is defined for $t < \frac{\epsilon}{||y||}$.

Further, note also that $F'(x; y)$ depends on $||y||$, not just the direction of $y$.

**Definition.** The $i^{th}$ partial derivative if $F$ at $x$ is $F'(x; e_i)$. We put usually

$$\frac{\partial F(x)}{\partial x_i}$$

.

For our purposes, we will instead use $D_i(F(x))$. Similarly, let $D_{135}(F)$ denote $D_1(D_3(D_5(F)))$, and so on.

Further, note that

$$D_i(F(x_1, x_2, \ldots, x_n)) = \lim_{t \to 0} \frac{F(x_1, \ldots, x_i + t, \ldots, x_n) - F(x_1, \ldots, x_n)}{t}$$

Thus, $D_i(F(x))$ is computed by differentiating with respect to $x_i$ and leaving others constant.

**Theorem 19.** *(equality of mixed partials): For $U$ open, if $f : U \to \mathbb{R}$ and $D_{ij}(f)$ and $D_{ji}(f)$ both exist and are continuous on $U$, then they are equal as functions on $U$.*

*Proof.* By holding all the other variables constant, we can assume $U \in \mathbb{R}^2$ and $i = 1, j = 2$ without loss of generality. Put now $F(x, y)$ as a function of 2 variables.

Given $(x_0, y_0) \in U$ for $h, k > 0$ consider

$$(*) \quad F(x_0 + h, y_0 + k) - F(x_0 + h, y_0) - F(x_0, y + k) + F(x_0, y_0)$$

Let $G(x) := F(x, y_0 + k) - F(x, y_0)$. The above $(*)$ is then equal to $G(x_0 + h) - G(x_0)$. Note that

$$G'(x) = D_1(F(x, y_0 + k)) - D_1(F(x, y_0))$$

The mean value theorem applied to $G$ implies that $\exists x \in (x_0, x_1 + h)$ such that

$$G'(x_1) = \frac{G(x_0 + h) - G(x_0)}{h} \implies h(G'(x_1)) = (*)$$

The mean value theorem applied to $D_1(F(x_1, y))$ as a function of $y$ implies that $\exists y_1 \in (y_0, y_0 + k)$ such that

$$D_{21}F(x_1, y_1) = \frac{D_1(F(x_1, y_0 + k)) - D_1(F(x_1, y_0))}{k} = \frac{G'(x)}{k}$$

Thus, we have that

$$(*) = hkD_{21}(F(x_1, y_1))$$

A similar argument with $\tilde{G}(y) = F(x_0 + h, y) - F(x_0, y)$ implies that

$$(*) = hkD_{12}(F(x_2, y_2))$$

Because $D_{21}, D_{12}$ are continuous, there exists $\forall \epsilon > 0, \exists \delta > 0$ such that

$$\begin{aligned} &(x, y) \in B_\delta(x_0, y_0) \\ \implies &|D_{21}(F(x, y)) - D_{21}(F(x_0, y_0))| < \frac{\epsilon}{2} \\ \implies &|D_{12}(F(x, y)) - D_{21}(F(x_0, y_0))| < \frac{\epsilon}{2} \end{aligned}$$

Choose $h, k$ such that $||(h, k)|| < \delta$, implying that $(x_1, y_1), (x_2, y_2) \in B_\delta(x_0, y_0)$. Then,

$$\begin{aligned} &|D_{21}(F(x_0, y_0)) - D_{12}(F(x_0, y_0))| \\ \leq &|D_{21}(F(x_0, y_0)) - D_{21}(F(x_1, y_1))| \\ + &|D_{12}(F(x_2, y_2)) - D_{12}(F(x_0, y_0))| \\ < &\epsilon \end{aligned}$$

Since $\epsilon$ is arbitrary, we have that $D_{12}(F(x_0, y_0)) = D_{21}(F(x_0, y_0))$. □

*Remark.* In one variable, differentiable implies continuous. However, consider

$$F(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4} & (x, y) \neq (0, 0) \\ 0 & (0, 0) \end{cases}$$

For the above function, all directional derivatives exist, but $F$ is not even continuous at the origin.

Let $\gamma(t) = (t^2, t)$. Then

$$(F \circ \gamma)(t) = \begin{cases} \frac{t^4}{t^4+t^4} & t \neq 0 \\ 0 & t = 0 \end{cases} = \begin{cases} \frac{1}{2} & t \neq 0 \\ 0 & t = 0 \end{cases}$$

This isn't continuous; thus, since $\gamma$ is continuous, we know that $F$ isn't continuous!

**Definition.** Remember that the derivative in the single variable case is the "best linear approximation" of the function at a point. That is,

$$f(x+h) \approx f(x) + f'(x) \cdot h$$

To imitate such a notion in multiple variables, let $U \subseteq \mathbb{R}^n$ be open, and $F : U \to \mathbb{R}$.
A linear $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ is the total derivative of $F$ at $x$, written $F'(x)$ if

$$\lim_{H \to \vec{0}} \frac{F(x+H) - F(x) - T(H)}{||H||} = 0$$

Note that if $n = 1$, we have that $T(H) = f'(x) \cdot h$, and $\lim_{h \to 0} \frac{f(x+h)-f(x)}{h} = f'(x)$.

**Prop.** *If $F'(x)$ exists, then*

$$(F'(x))(y) = F'(x; y)$$

*Proof.* When $y = 0$, both sides vanish. Assume that $y \neq 0$.
Then, $\forall \epsilon > 0, \exists \delta > 0$ such that $0 < ||H|| < \delta \implies \frac{|F(x+H)-F(x)-F'(x)H}{||H||} < \frac{\epsilon}{||y||}$.
Let $H = ty, t \in \mathbb{R}$. Then, $0 < |t| < \frac{\delta}{||y||} \implies 0 < ||H|| = |t|||y|| < \delta$

$$\begin{aligned} &\implies \frac{F(x+H) - F(x) - tF'(x)(y)}{|t|||y||} < \frac{\epsilon}{||y||} \\ &\implies \left| \frac{F(x+ty - F(x))}{t} - F'(x)(y) \right| < \epsilon \end{aligned}$$

Then, we have that

$$\begin{aligned} F'(x)(y) &= \lim_{t \to 0} \frac{F(x+ty) - F(x)}{t} \\ &= F'(x; y) \end{aligned}$$

□

**Corollary.** *If $F'(x; y)$ is undefined for some $y$ or $F'(x; y)$ nonlinear function of $y$, then $F$ is not differentiable at $x$.*

**Corollary.** *If it exists, $F'(x)$ is unique and is given by the dot product with*

$$(D_1(F(x)), D_2(F(x)), \ldots, D_n(F(x)))$$

*Proof.* Both are linear and both take value $F'(x)(e_k) = D_k(F(x))$ on the standard basis vectors. □

**Definition.** We call the above vector $\nabla F(x)$ and it is the gradient of $F$ at $x$. if $F : U \to \mathbb{R}$, then $\nabla F : U \to \mathbb{R}^n$, and $\nabla F(x) = (D_1(F(x)), \ldots, D_n(F(x)))$.

*Remark.* Geometrically, the gradient points in the direction of greatest ascent. Further, the length is the maximum value of the directional derivative on a unit vector.

The above is that this is equivalent to being perpendicular to level curves.

**Theorem 20.** *$F$ is differentiable at $x$, then $F$ is continuous at $x$.*

*Proof.* Let $T$ be the total derivative of $F$ at $x$. Then,

$$\begin{aligned} 0 &\leq |F(x+H) - F(x)| \\ &= \frac{||H|||F(x+H) - F(x)|}{||H||} \\ &= \frac{||H|||F(x+H) - F(x) - T(H) + T(H)|}{||H||} \\ &\leq \frac{||H|||F(x+H) - F(x) - T(H)| + |T(H)|}{||H||} \\ &= \frac{||H|||F(x+H) - F(x) - T(H)|}{||H||} + |T(H)| \end{aligned}$$

As $H \to 0$, everything $\to 0$, and

$$\lim_{H \to 0} |F(x+H) - F(x)| = 0$$

□

**Definition.** For $k \in \mathbb{Z}_{>0}$, $U \subseteq \mathbb{R}^n$ open, $F : U \to \mathbb{R}$ is $C^k$ if all partial derivatives of order $k$ exist and are continuous, where order $k$ means something of the form

$$D_{i_k} \cdots D_{i_1}(F)$$

As a convention, $C^0$ means continuous, and $F$ is $C^\infty$ or smooth if it is $\forall k, C^k$. All of our favorite functions are $C^\infty$!

**Theorem 21.** *$F$ is $C^1$ on $U$ $\implies$ $F$ is differentiable on $U$.*

*Proof.* $F$ is $C^1$ $\iff$ $\nabla F$ exists. Define $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ as follows:

$$T(H) := H \cdot \nabla F(x)$$

We need to show that

$$\lim_{H \to 0} \frac{F(x+H) - F(x) - T(H)}{||H||} = 0$$

We have that

$$\begin{aligned} F(x+H) - F(x) =& F(x_1 + H_1, \ldots, x_n + H_n) \\ & - F(x_1 + H_1, \ldots, x_{n-1} + H_{n-1}, x_n) \\ & + F(x_1 + H_1, \ldots, x_{n-2} + H_{n-2}, x_{n-1}, x_n) \\ & \vdots \\ & + F(x_1 + H_1, x_2, \ldots, x_n) \\ & - F(x_1, x_2, \ldots, x_n) \end{aligned}$$

Let each set of two lines be $\Delta_i$.
For each $i$, let $g_i : [x_i, x_i + H_i] \to \mathbb{R}$ be given by

$$g_i(y) := F(x_1 + H_1, \ldots, y, x_{i+1}, \ldots, x_n)$$

such that

$$\Delta_i = g_i(x_i + H_i) - g_i(x_i)$$

and

$$g_i'(y) = D_i F(x_1 + H_1, \ldots, y, x_{i+1}, \ldots, x_n)$$

The mean value theorem implies that $\exists c_i \in (x_i, x_i + H_i)$ such that

$$\frac{\Delta_i}{H_i} = g_i'(c_i)$$

Hence,

$$\begin{aligned} F(x+H) - F(x) - T(H) &= \sum_{i=1}^{n} (\Delta_i - D_i F(x) H_i) \\ &= \sum_{i=1}^{n} (g_i'(c_i) - D_i F(x)) H_i \end{aligned}$$

$$= \sum_{i=1}^{n} (D_i F(x_1 + H_1, \dots, c_i, x_{i+1}, x_n) - D_i F(x)) H_i$$

Since $\frac{|H_i|}{||H||} = \frac{\sqrt{H_i^2}}{\sqrt{\sum_{i=1}^n H_i^2}} \leq 1$, we have that

$$\frac{|F(x+H) - F(x) - T(H)}{||H||} \leq \sum_{i=1}^{n} |D_i F(x_1 + H_i, \dots, c_i, \dots x_n) - D_i F(x_1, \dots, x_n)$$

Since $F$ is $C^1$ then $D_i F$ are continuous. Given $\epsilon > 0$, take $\delta > 0$ such that $0 < ||y - x|| < \delta \implies |D_i F(y) - D_i F(x)| < \frac{\epsilon}{n}$.

Take $y_i = (x_1 + H_1, \dots, c_i, \dots, x_n)$. $||y_i - x|| \leq ||H||$. Hence, $0 < ||H|| < \delta \implies 0 < ||y_i - x|| < ||H|| < \delta$ for each $i$. Then, we have that

$$\sum_{i=1}^{n} |D_i F(x_1 + H_i, \dots, c_i, \dots x_n) - D_i F(x_1, \dots, x_n) < \epsilon$$

and we are done. □

*Remark.* We need to be able to vary our point; thus, we actually need this on some open interval. Furthermore, the converse isn't true!

**Definition.** We can extend the total derivative to vector valued functions.

Let $U \subseteq \mathbb{R}^n$ open, and $F : U \to \mathbb{R}^m$. Then, at any $x \in U$, $F$ is differentiable at $x$ with derivative $F'(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, if

$$\lim_{H \to 0} \frac{||F(x+H) - F(x) - F'(x)(H)||}{||H||} = 0$$

**Prop.** *$F$ is differentiable at $x \iff F_i$ is differentiable for each $i$. Furthermore,*

$$F'(x) = (F_1'(x), \dots, F_m'(x))$$

*Proof.* The above limit is zero iff it approaches the zero vector; each individual component must then also be 0. □

**Corollary.** *The matrix of $F'(x)$ in the standard basis vectors is*

$$DF(x) = \begin{bmatrix} D_1 F_1(x) & \dots & D_n F_1(x) \\ D_1 F_2(x) & \dots & D_n F_x(x) \\ \vdots & \ddots & \vdots \\ D_m F_2(x) & \dots & D_n F_x(x) \end{bmatrix}$$

**Definition.** The above matrix is the Jacobian matrix at $x$.

*Example.* For $U = \{(r, \theta) \mid r, \theta \in \mathbb{R}\}$, let

$$F(r, \theta) = (r\cos(\theta), r\sin(\theta))$$

Then, the Jacobian is

$$\begin{bmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{bmatrix}$$

**Theorem 22.** *(Multivariable chain rule) Suppose $F : \mathbb{R}^p \to \mathbb{R}^n$ and $G : \mathbb{R}^n \to \mathbb{R}^m$ or some open sets thereof, and that they are differentiable at $x \in \mathbb{R}^p$ and $F(x) \in \mathbb{R}^n$ respectively. Put $S = F'(x), T = G'(F(x))$.*
*Then, $G \circ F$ is differentiable at $x$ and*

$$(G \circ F)'(x) = T \circ S$$

*Equivalently,*

$$(G \circ F)'(x) = G'(F(x)) \circ F'(x)$$

*or in terms of matrices,*

$$D(G \circ F)(x) = DG(F(x)) \cdot DF(x)$$

**Lemma.** *If $S : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then the function $\mathbb{R}^n \to \mathbb{R}$ given by*

$$\begin{cases} \frac{||S(H)||}{||H||} & H \neq 0 \\ 0 & H = 0 \end{cases}$$

*is bounded.*

*Proof.* Let $A$ be the corresponding matrix, such that

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

where each $a_i$ is a row vector in $\mathbb{R}^n$.
Then, we have that

$$S(H) = \begin{bmatrix} a_i \cdot H \\ \vdots \\ a_m \cdot H \end{bmatrix}$$

$$= \sum_{i=1}^{m} (a_i \cdot H) e_i$$

Let $B = \sum_{i=1}^{m} ||a_i||$. Then,

$$\begin{aligned} ||S(H)|| &= ||\sum_{i=1}^{m} (a_i \cdot H) e_i|| \\ &\leq \sum_{i=1}^{m} ||(a_i \cdot H) e_i|| \\ &= \sum_{i=1}^{m} |a_i \cdot H| \\ &\leq \sum_{i=1}^{m} ||a_i|| \cdot ||H|| = ||H|| \cdot B \end{aligned}$$

Thus, we see that the desired function is bounded by $B$. □

**Lemma.** *Suppose $F$ is differentiable at $x$ with $F'(x) = S$. Let $K(H) = F(x + H) - F(x)$. Then, $\exists \delta > 0$ such that the function*

$$\begin{cases} \frac{||K(H)||}{||H||} & H \neq 0 \\ 0 & H = 0 \end{cases}$$

*is bounded on the $B_\delta(0)$.*

*Proof.* Note

$$\frac{K(H)}{||H||} = \frac{F(x + H) - F(x) - S(H)}{||H||} + \frac{S(H)}{||H||}$$

Then,

$$\frac{||K(H)||}{||H||} \leq \frac{||F(x + H) - F(x) - S(H)||}{||H||} + \frac{S(H)}{||H||}$$

Since $\lim_{H \to 0} \frac{||F(x+H) - F(x) - S(H)||}{||H||} = 0$, taking $\epsilon = 1$, $\exists \delta > 0$ such that on $B_\delta(0)$, the desired function is bounded by $B + 1$. □

*Proof.* Let

$$J(H) = \begin{cases} \frac{F(x+H) - F(x) - S(H)}{||H||} & H \neq 0 \\ 0 & H = 0 \end{cases}$$

Note that $F$ being differentiable at $x$ implies that $J$ is continuous at 0.

Let $Y = F(x), K(H) = F(x+H) - F(x)$. Then,

$$\begin{aligned}&(G \circ F)(x+H) - (G \circ F)(x) - (T \circ S)(H)\\ &= G(Y + K(H)) - G(Y) - T(K(H)) + T(K(H) - S(H))\end{aligned}$$

Since we have that

$$\frac{K(H) - S(H)}{||H||} = J(H)$$

for $H \neq 0$. Then, we have that for $H \neq 0$,

$$\begin{aligned}(*) &= \frac{(G \circ F)(x+H) - (G \circ F)(x) - (T \circ S)(H)}{||H||}\\ &= \begin{cases} \frac{G(Y+K(H)) - G(Y) - T(K(H))}{||K(H)||} \cdot \frac{||K(H)||}{||H||} + (T \circ J)(H) & K(H) \neq 0 \\ 0 & K(H) = 0 \end{cases}\end{aligned}$$

Note that lemma 2 yields some $\delta_1$ such that $0 < ||H|| < \delta_1 \implies \frac{||K(H)||}{||H||} < C$ for some constant $C$. We know that

$$\lim_{K \to 0} \frac{G(Y+K) - G(Y) - T(K)}{||K||} = 0$$

as $G'(Y) = T$. Given $\epsilon > 0$, take $\delta_2$ such that $0 < ||K|| < \delta_2 \implies \frac{||G(Y+K) - G(Y) - T(K)}{||K||} < \frac{\epsilon}{2C}$.
Now, since $F$ is differentiable at $x$, then $F$ is continuous at $x$, as is $K$ at 0. Given $\delta_2$, $\exists \delta_3 > 0$ such that $0 < ||H|| < \delta_3 \implies ||K(H)|| < \delta_2$.

Now, since $T$ is linear and therefore continuous, and as $J$ is continuous at 0, we have that $T \circ J$ is continuous at 0. Finally, $\exists \delta_4 > 0$ such that $0 < ||H|| < \delta_4 \implies ||(T \circ J)(H)|| < \frac{\epsilon}{2}$. Taking $\delta = \min(\delta_1, \delta_2, \delta_3, \delta_4)$, we have that

$$(*) < \epsilon$$

and we are done! □

# 11 Line Integrals

**Definition.** A path or curve in $\mathbb{R}^n$ is just a continuous function $\gamma : [a, b] \to \mathbb{R}^n$ for some $a < b \in \mathbb{R}$.

*Remark.* Note that a path is not the same as the image of the path: the path "remembers" where we are at a particular $t$, and thus where we are going and came from.

**Definition.** A path $\gamma : [a, b] \to \mathbb{R}^n$ is smooth if it is $C^1$ and piecewise smooth if there is a partition $0 < a_1 < \cdots < a_n = b$ such that $\gamma|_{(a_i, a_{i+1})}$ is smooth for each $i$.

*Note that this definition is not useful outside of these notes and the associated course. This is just a convenience thing!*

**Definition.** Let $F$ be a continuous function $U \to \mathbb{R}^n$, where $U$ is a set containing $\mathrm{im}(\gamma)$. We define the line integral as follows

$$\int F \cdot d\gamma := \int_a^b F(\gamma(t)) \cdot ([\gamma'(t)](1))dt$$

Since $F$ is continuous and $\gamma$ is piecewise $C^1$, this is piecewise continuous, hence integrable.

Note that we will write $\gamma'(t)$ to mean $[\gamma'(t)](1)$.

*Remark.* There are annoying geometrical issues with noncontinuous functions, and they will therefore *not* be dealt with.

**Prop.** *If $h : [\widetilde{a}, \widetilde{b}] \to [a, b]$ is continuous and piecewise $C^1$ and bijective then $\widetilde{\gamma} = \gamma \circ h$ is another curve, and is a reparameterization of $\gamma$.*

*Example.* Use the intermediate value thereon to show that either

$$h(\widetilde{a} = a, h(\widetilde{b}) = b, \text{ or } h(\widetilde{a}) = b, h(\widetilde{b}) = a$$

**Prop.** *The above definition is invariant with respect to reparameterizations of $\gamma$. That is,*

$$\int F \cdot d\gamma = \pm \int F \cdot d\gamma$$

*Proof.*

$$\begin{aligned}\int_a^b F(\gamma(t)) \cdot \gamma'(t)dt &= \int_{\widetilde{a}}^{\widetilde{b}} F(\widetilde{\gamma}(s)) \cdot \widetilde{\gamma}'(s)ds \\ &= \int_{\widetilde{a}}^{\widetilde{b}} F(\gamma(h(s))) \cdot \gamma'(h(s))h'(s)\end{aligned}$$

From change of variables in one dimension,

$$\begin{aligned}&= \int_{h(\widetilde{a})}^{h(\widetilde{b})} F(\gamma(t)) \cdot \gamma'(t)dt \\ &= \pm \int F \cdot d\gamma\end{aligned}$$

□

**Definition.** An ordering of the endpoints of a path is called an orientation.

*Example.* $\int F \cdot d\gamma$ is a linear function of $F$.

*Example.* Call image of $\gamma$ a curve $C$, and say $C = C_1 \cup C_2$ if $C_1 = \gamma([a,x]), C_2 = \gamma([x,b])$. Then,

$$\int_C F \cdot d\gamma = \int_{C_1} F \cdot d\gamma + \int_{C_2} F \cdot d\gamma$$

Note that this yields a way to integrate on piecewise $C^1$ paths.

**Theorem 23.** *(Fundamental theorem of line integrals) For $U \subseteq \mathbb{R}^n$ open, $g : U \to \mathbb{R}$, and $\gamma : [a,b] \to U$ from $x = \gamma(a)$ to $y = \gamma(b)$ with $C = \gamma([a,b])$. Then,*

$$\int_C \nabla g \cdot d\gamma = g(y) - g(x)$$

*Proof.*

$$\begin{aligned}
\int_C \nabla g \cdot d\gamma &= \int_a^b \nabla g(\gamma(t))dt \cdot \gamma'(t)dt \\
&= \int_a^b Dg(\gamma(t))D\gamma(t)dt \\
&= \int_a^b D(g \circ \gamma)(t)dt \\
&= (g \circ \gamma)(b) - (g \circ \gamma)(a) \\
&= g(y) - g(x)
\end{aligned}$$

□

**Definition.** If $F = \nabla g$ for some $g$ is continuous, we say $g$ is a potential for $F$ and $F$ is conservative.

*Remark.* If $g$ is a potential for $F$, then so is $g + C$, with $C \in \mathbb{R}$. Note that if $F = \nabla g$,

$$\int_C F \cdot d\gamma = g(y) - g(x)$$

the line integral depends not on the choice of the path, just the endpoints.

**Definition.** A path $\gamma : [a,b] \to \mathbb{R}$ is closed if $\gamma(a) = \gamma(b)$.

*Example.* The unit circle path is $\gamma : [0, 2\pi] \to \mathbb{R}^2$, with

$$\gamma(t) = (\cos(t), \sin(t))$$

Furthermore, if $C$ is closed, we write

$$\oint_C F \cdot d\gamma$$

**Definition.** An open set $U \in \mathbb{R}^n$ is path-connected if $\forall x, y \in U$ there exists a path connecting those two points (i.e. $\gamma : [a, b] \to U$ with $\gamma(a) = x, \gamma(b) = y$).

*Remark.* Note that most definitions here are a bit stronger than needed: for example, paths above are $C^1$, but in general only are continuous in other contexts.

**Theorem 24.** *For $U$ open and path-connected, $F : U \to \mathbb{R}^n$ continuous, the following are equivalent:*

1. *$F$ is conservative.*
2. *$\int_C F \cdot d\gamma$ depend only on endpoints.*
3. *$\oint_C F \cdot d\gamma = 0$ for all closed $C$.*

*Proof.* We have shown that $1 \implies 2$; for $2 \implies 3$, take the path $\widetilde{C}$ given by $\widetilde{\gamma}(t) = \gamma(a) = \gamma(b)$. This has $\widetilde{\gamma}'(t) = 0$, such that

$$\oint_C F \cdot d\gamma = 0$$

Now, for $3 \implies 2$, consider for two paths with the same endpoints $C, \widetilde{C}$, the new path $C\#\widetilde{C} = C \cup \widetilde{C}$, where this is the concatenation of the two paths $C$ and $\widetilde{C}$ backwards. Then,

$$0 = \oint_{C\#\widetilde{C}} F \cdot d\gamma = \int_C F \cdot d\gamma - \int_{\widetilde{C}} F \cdot d\gamma \implies \int_C F \cdot d\gamma = \int_{\widetilde{C}} F \cdot d\gamma$$

It now suffices to show $2 \implies 1$. Pick $x_0 \in U$. Define

$$g(x) = \int_C F \cdot d\gamma$$

where $C$ is any path starting at $x_0$ and ending at $x$. We will show $\nabla g = F$.
For $x \in U$, $U$ open, $\exists \delta > 0$ such that $B_\delta(x) \subseteq U$. For $|s| < \delta$.

$$\gamma : [0, |s|] \to B_\delta(x)$$
$$C_s = \gamma(t) = \begin{cases} x + te_i & s \geq 0 \\ x - te_i & s \leq 0 \end{cases}$$

Now, $C\#C_s$ is a path from $x_0$ to $x + se_i$.

$$g(x + se_i) - g(x) = \int_{C\#C_s} F \cdot d\gamma - \int_C F \cdot d\gamma$$

$$
\begin{aligned}
&= \int_{C_s} F \cdot d\gamma \\
&= \int_0^{|s|} F(x \pm te_i) \cdot (\pm e_i) dt \\
&= \pm \int_0^{|s|} F_i(x \pm te_i) dt \\
&= \int_0^{s} F_i(x + te_i) dt
\end{aligned}
$$

Returning to $g$,

$$
\begin{aligned}
D_i g(x) &= \lim_{s \to 0} \frac{g(x + se_i) - g(x)}{s} \\
&= \lim_{s \to 0} \frac{1}{s} \int_0^s F_i(x + te_i) dt \\
&= F_i(x + 0e_i) = F_i(x)
\end{aligned}
$$

Finally, we have that $\nabla g(x) = F(x)$. □

**Prop.** *If $U \subseteq \mathbb{R}^n$ open, $F : U \to \mathbb{R}^n$ is conservative and $C^1$, then $\forall i, j \in \{1, \ldots, n\}$,*

$$D_i F_j = D_j F_i$$

*Proof.*

$$F = \nabla g \implies F_i = D_i g \implies D_i F_j = D_j D_i g = D_j D_i g = D_j F_i$$

The equality of mixed partials yields the middle equality. □

**Definition.** We say $F$ is closed if the above condition holds.

*Remark.* We show in the next homework that if $U = \mathbb{R}^2 - \{(0,0)\}$, we can construct a closed vector field that is not conservative.

**Definition.** An open set $U \subseteq \mathbb{R}^n$ is star-shaped if $\exists x_0 \in U$ such that $\forall x \in U, \forall t \in [0,1], x_0 + t(x - x_0) \in U$.

**Theorem 25.** *(Poincaré lemma) $U$ star-shaped $\implies$ every closed $C^1$ vector field is conservative.*

*Proof.* Let $F : U \to \mathbb{R}^n$ be $C^1$ and closed. Let $x_0 \in U$ be a base point as in the above definition of star-shaped. Without loss of generality, take $x_0 = 0$, and let

$$g(x) = \int_C F \cdot d\gamma$$

where $C$ is given by $\gamma : [0,1] \to U, \gamma(t) = tx \implies \gamma'(t) = x$. We will show $\nabla g = F$.

$$\begin{aligned}
D_i g(x) &= D_i \int_C F \cdot d\gamma \\
&= D_i \int_0^1 F(tx) \cdot x dt \\
&= \int_0^1 D_i(F(tx) \cdot x) dt \\
D_i(F(tx) \cdot x) &= D_i \sum_{j=1}^n F_j(tx) x_j \\
&= \sum_{j=1}^n D_i(F_j(tx)) x_j + \delta_{ij} F_j(tx) \\
&= F_i(tx) + \sum_{j=1}^n \sum_{k=1}^n D_k F_j(tx) D_i(tx) \\
&= F_i(tx) + \sum_{j=1}^n \sum_{k=1}^n \delta_{ik} t D_k F_j(tx) x_i \\
&= F_i(tx) + \sum_{j=1}^n D_j F_i(tx)(tx_i) \\
&= F_i(tx) + \nabla F_i(tx) \cdot tx \\
&= \frac{dh_i}{dt}
\end{aligned}$$

where $h_i(t) = tF_i(tx)$.

Finally,

$$D_i g(x) = \int_0^1 \frac{dh_i}{dt}(t) dt = h_i(1) - h_i(0) = F_i(x)$$

□

*Remark.* We assumed an unproved result above, namely that one can pull the derivative inside the integral. Some of this material will be skipped, but we need to prove the following:

**Lemma.** *Let $g(h,t)$ be continuous $U \to \mathbb{R}$ for some $U \subseteq \mathbb{R}^2$ open, containing $\{y\} \times [a,b]$. Then, $\forall \epsilon > 0, \exists \delta > 0 \mid \forall t \in [a,b], \forall h \in \mathbb{R}$,*

$$|h - y| < \delta \implies |g(h,t) - g(y,t)| < \epsilon$$

*This creates some type of uniform continuity in $t$.*

*Proof.* The proof is omitted, but can be done in the same scooching method as done for continuity in the single variable case. □

**Definition.** Let $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$ be a rectangle. Then interior of $Q$ is $(a_1, b_1) \times \cdots \times (a_n, b_n)$.

**Theorem 26.** *Let $Q \subset \mathbb{R}^n$ be a rectangle, $R = Q \times [a, b] \subset \mathbb{R}^{n+1}$, with coordinates $(x_1, \ldots, x_n, t)$. Then, let $\psi : R \to \mathbb{R}$ be $C^1$, and $\phi : Q \to \mathbb{R}$ be*

$$\phi(x) = \int_a^b \psi(x, t)dt$$

*Then $\forall x$ in the interior of $Q$,*

$$D_k\phi(x) = \int_a^b D_k\psi(x, t)dt$$

*Proof.* Fix $x$ in the interior of $R$. Then, $\exists \delta_1 > 0$ such that $|h| < \delta \implies x + he_k$ is still in the interior of $R$.

$$\phi(x + he_k) - \phi(x) = \int_a^b (\psi(x + he_k, t) - \psi(x, t))dt$$

Put $I(h, x, t) = \psi(x + he_k, t) - \psi(x, t) - hD_k(x, y)$. Then, we will show that

$$\lim_{h \to 0} \frac{1}{h} \int_a^b I(h, x, t)dt = 0$$

The mean value theorem states that $\forall h, x, t$, $\exists z$ on the line between $x, x + he_k$ such that

$$I(h, x, t) = h(D_k\psi(z, t) - D_k\psi(x, y))$$

Note that the above $z$ is not necessarily a continuous function of $t$. The lemma, however, applied to the function $g(h, t) = D_k\psi(x + he_k, t)$ at $y = 0$ yields $\delta \in (0, \delta_1)$ such that $\forall t \in [a, b], \forall h \in \mathbb{R}$,

$$|h| < \delta \implies |D_k\psi(x + he_k) - D_k\psi(x, t)| < \frac{\epsilon}{2(b - a)}$$

Hence, $\forall t \in [a, b], |h| < \delta \implies ||z - x|| < \delta$

$$|D_k\psi(z, t) - D_k\psi(x, t)| = |I(h, x, t)| < \frac{|h|\epsilon}{2(b - a)}$$

Integrating,

$$\frac{1}{|h|}\left|\int_a^b I(h, x, t)dt\right| \le \frac{1}{|h|}\int_a^b |I(h, x, t)|dt \le \frac{\epsilon}{2} < \epsilon$$

□

*Remark.* Apply this to $x \in Q \subset U$ open and star-shaped, with $F : U \to \mathbb{R}^n$ $C^1$ and the above $[a, b] = [0, 1]$, such that

$$\psi(x, t) = F(tx) \cdot x$$

to justify the earlier result.

## 12 Multivariate Integrals

**Definition.** A partition of a closed rectangle $[a_1, b_1] \times \cdots \times [a_n, b_n]$ is a partition of each of the individual $[a_i, b_i]$. A closed subrectangle of $P$ is a product of closed sub-intervals in $P$.

**Definition.** Let $Q \subseteq \mathbb{R}^n$ be a closed rectangle. Say $f : Q \to \mathbb{R}$ is a step function if $\exists$ a partition $P$ of $Q$ such that $\forall$ subrectangles $R$ of $P$, $f|_{\text{int}\, R}$ is a constant $f_R$.

*Remark.* Note that this specifies nothing about the boundary planes, meaning that step functions can be really terrible!

**Definition.** We define the integral on step functions as follows:

$$\int_Q f := \sum_R f_R \text{vol}(R)$$

where $\text{vol}([a_1, b_1] \times \cdots \times [a_n, b_n]) := \prod_{i=1}^n (b_i - a_i)$.

**Definition.** If $f : Q \to \mathbb{R}$ is bounded,

$$\underline{I}(f) := \sup \left\{ \int_Q s \mid s \leq f, s \text{ a step function} \right\}$$
$$\overline{I}(f) := \inf \left\{ \int_Q s \mid t \geq f, t \text{ a step function} \right\}$$

Then, $f$ is integrable if $\overline{I}(f) = \underline{I}(f)$, and

$$\int_Q f := \overline{I}(f) = \underline{I}(f)$$

*Remark.* It still holds with the same proof that $f$ is integrable $\iff$ $\forall \epsilon > 0$, $\exists s, t$ step functions with $s \leq f \leq t$ such that $\int_Q (t - s) < \epsilon$.

**Theorem 27.** *If $f : Q \to \mathbb{R}$ is continuous, then it is integrable.*

*Proof.* This will use results from Apostol that $f$ is continuous $\implies$ $f$ is bounded, and that we can always find some partition $P$ such that for any $\epsilon > 0$, any subrectangle of $P$, $\text{span } f|_R = \sup f|_R - \inf f|_R < \epsilon$.

Suppose that $|f| \leq K$, and take a partition $P$ such that $\text{span } f|_R < \frac{\epsilon}{2\text{vol}(Q)}$ for any subrectangle of $P$. Let

$$s(x) = \begin{cases} \inf_{x \in R} f(x) & x \in \text{int } R \\ -K & \text{otherwise} \end{cases}$$

$$t(x) = \begin{cases} \sup_{x \in R} f(x) & x \in \text{int } R \\ K & \text{otherwise} \end{cases}$$

Then, $s, t$ are step functions with $s \leq f \leq t$.

$$\int_Q (t - s) = \sum_R (\text{span } f|_R)\text{vol}(R) \leq \sum_R \frac{\epsilon}{2\text{vol}(Q)}\text{vol}(R) = \frac{\epsilon}{2} < \epsilon$$

□

**Prop.** *The multivariate integral shares a lot of the same formal properties as the one variable case. For example, it is still linear.*

**Lemma.** *Suppose we have $g(x_1, \ldots, x_m, t_1, \ldots, t_n)$ is continuous on rectangle $Q$ and $y \in \mathbb{R}^m$. Then, $\forall \epsilon$, $\exists \delta > 0$ such that $\forall t \in Q$, $\forall x \in \mathbb{R}^m$,*

$$|x - y| < \delta \implies |g(x, t) - g(y, t)| < \epsilon$$

**Theorem 28.** *(Continuity of iterated integrals) If $g : [a_1, b_1] \times \cdots \times [a_{m+n}, b_{m+n}] \to \mathbb{R}$ is continuous, then*

$$\int_{a_{m+1}}^{b_{m+1}} \cdots \left( \int_{a_{n+m-1}}^{b_{n+m-1}} \left( \int_{a_{m+n}}^{b_{m+n}} g(x_1, \ldots, x_m, t_1, \ldots, t_n) dt_n \right) dt_{n-1} \right) \ldots dt_1$$

*exists and is continuous.*

*Proof.* By a very easy induction argument, we only need to prove $n = 1$. In this case consider $g(x_1, \ldots, x_m, t)$ continuous on $R \times [a, b]$. We will show

$$f(x_1, \ldots, x_m) = \int_a^b g(x_1, \ldots, x_m, t) dt$$

is continuous. Note that existence is immediate from the continuity of $g$. The uniform continuity lemma tells us that $\forall y \in R, \forall \epsilon > 0, \exists \delta > 0$ such that $\forall x \in R, \forall t \in [a, b]$,

$$|y - x| < \delta \implies |g(y, t) - g(x, y)| < \frac{\epsilon}{b - a}$$

$$\implies \int_a^b -\frac{\epsilon}{b-a} \leq \int_a^b (g(y,t) - g(x,t))dt \leq \int_a^b \frac{\epsilon}{b-a} dt$$
$$\implies -\epsilon \leq f(y) - f(x) \leq \epsilon$$

□

*Remark.* In particular, $g : Q \to \mathbb{R}$ continuous implies that the iterated integral

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} g(x_1, \dots, x_n) dx_n \cdots dx_1$$

exists.

**Theorem 29.** *If $\int_Q f$ and $\int \cdots \int f(x_1, \dots, x_n) dx_n \cdots dx_1$ both exist, then they are equal.*

*Proof.* If $f$ is integrable, then $\int_Q f$ is the unique $A \in \mathbb{R}$ such that

$$\int_Q s \leq A \leq \int_Q t$$

for all step functions $s \leq f \leq t$.

Let $s, t$ be step functions with $s \leq f \leq t$, and choose a partition $P$ that works for both step functions (i.e. taking the union of the partition divides). Let $R$ be any subrectangle of $P$, and $f$ bounded by $K$, and

$$\widetilde{s} = \begin{cases} s(x) & x \in \operatorname{int} R \\ -K & \text{otherwise} \end{cases}$$
$$\widetilde{t} = \begin{cases} t(x) & x \in \operatorname{int} R \\ K & \text{otherwise} \end{cases}$$

We still have that $\widetilde{s} \leq f \leq \widetilde{t}$, and $\int_Q s = \int_q \widetilde{s}, \int_Q t = \int_Q \widetilde{t}$, but know the iterated integrals of $\widetilde{s}, \widetilde{t}$ always exist.

By the above property of integration, as well as induction, we have that

$$\int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \widetilde{s}(x) dx_n \cdots dx_1 \leq \int \cdots \int f(x) dx_n \cdots dx_1 \leq \int \cdots \int \widetilde{t}(x) dx_n \cdots dx_1$$

and by the properties of step functions, we get

$$\int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \widetilde{s}(x) dx_n = \int_Q \widetilde{s}, \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \widetilde{t}(x) dx_n = \int_q \widetilde{t}$$

We have then the integral of $f$ is exactly the above iterated integral. □

**Corollary.** *(Fubini's theorem) If $f$ is continuous, then iterated integrals don't depend on the order of integration!*

**Definition.** $S \subset \mathbb{R}^n$ is bounded if $\exists$ some closed rectangle $Q \mid S \subseteq Q$.

**Definition.** Let $S \subset Q$. For $\varphi : S \to \mathbb{R}$, define its extension by zero by $\widetilde{\varphi} : Q \to \mathbb{R}$ as

$$\widetilde{\varphi}(x) = \begin{cases} \varphi(x) & x \in S \\ 0 & x \notin S \end{cases}$$

**Definition.** $\varphi$ is integrable on $S$ if $\widetilde{\varphi}$ is integrable on $Q$, and if so let

$$\int_S \varphi := \int_Q \widetilde{\varphi}$$

**Definition.** $S \subset \mathbb{R}^2$ is of graph type I if $\exists f, g : [a, b] \to \mathbb{R}$ with $f \leq g$ and both piecewise $C^1$, and

$$S = \{(x, y) \in \mathbb{R}^2 \mid x \in [a, b], f(x) \leq y \leq g(x)\}$$

or graph type II if

$$S = \{(x, y) \in \mathbb{R}^2 \mid y \in [a, b], f(y) \leq x \leq g(y)\}$$

**Theorem 30.** *$S$ of graph type I, $\varphi : S \to \mathbb{R}$ continuous $\implies$ $\varphi$ is integrable on $S$ and*

$$\int_S \varphi = \int_a^b \int_{f(x)}^{g(x)} \varphi(x, y) dy dx$$

**Lemma.** *Suppose $S$ is of graph type I, $\varphi : S \to \mathbb{R}$, where*

$$S = \{(x, y) : x \in [a, b], f(x) \leq y \leq g(x)\}$$

*Then,*

$$\varphi = \int_{f(x)}^{g(x)} \varphi(x, y) dy$$

*is continuous.*

*Proof.* Let $H(x, s) = (x, h(x, s))$ where $h(x, s) = f(x) + s(g(x) - f(x))$. Then, $H : [a, b] \times [0, 1] \to S$ is a piecewise $C^1$ and continuous bijection. Then,

$$\begin{aligned} \int_{f(x)}^{g(x)} \varphi(x, y) dy &= \int_0^1 \varphi(H(x, s)) \frac{\partial h}{\partial s}(s) ds \\ &= \int_0^1 \varphi(H(x, s))(g(x) - f(x)) ds \\ &= (g(x) - f(x)) \int_0^1 (\varphi \circ H)(s) ds \end{aligned}$$

□

*Proof.* Pick $\epsilon > 0$. $f, g$ are continuous and thus are bounded on $[a, b]$. Then, $\exists c, d$ such that $S \subseteq [a, b] \times [c, d]$. Further, $\varphi$ continuous $\implies \varphi \circ H$ is continuous on $[a, b] \times [0, 1]$, where $H$ is as above. This in turn implies that $\varphi$ is bounded by some $K$ on $S$.

Pick partition

$$a = x_0 < x_1 < \cdots < x_m = b$$

such that $\operatorname{span} f|_{[x_{i-1},x_i]} \leq \epsilon_1$, and $\operatorname{span} g|_{[x_{i-1},x_i]} \leq \epsilon_1$ where

$$\epsilon_1 = \frac{\epsilon}{16K(b-a)} > 0$$

Let $I_i = [x_{i-1}, x_i]$ and $c_i = \min f|_{I_i}, c_i' = \min g|_{I_i}$, $d_i = \max f|_{I_i}, d_i' = \max g|_{I_i}$. Then, $\varphi|_{I_i \times [d_i, c_i']}$ is continuous and the small span theorem yields a partition of $I_i \times [d_i, c_i']$ into subrectangles $R$ such that

$$\operatorname{span} \varphi|_R \leq \epsilon_2 = \frac{\epsilon}{4(b-a)(d-c)} > 0$$

Define

$$s(x, y) = \begin{cases} 0 & (x, y) \in I_i \times [c, c_i), (x, y) \in I_i \times (d_i', d] \\ \min \varphi|_R & (x, y) \in \operatorname{int} R \\ -K & \text{otherwise} \end{cases}$$

$$t(x, y) = \begin{cases} 0 & (x, y) \in I_i \times [c, c_i), (x, y) \in I_i \times (d_i', d] \\ \max \varphi|_R & (x, y) \in \operatorname{int} R \\ K & \text{otherwise} \end{cases}$$

Then,

$$\begin{aligned} \int_{[a,b]\times[c,d]} (t - s) &= \sum_R \operatorname{span} \varphi|_R \operatorname{vol}(R) + \sum_{i=1}^{m} 2K(x_i - x_{i-1})(d_i' - c_i' + d_i - c_i) \\ &\leq \epsilon_2(b-a)(d-c) + 4K\epsilon_1(b-a) \\ &= \frac{\epsilon}{2} < \epsilon \end{aligned}$$

That shows integrability; consider now

$$\int_c^d \widetilde{\varphi}(x, y)dy = \int_{f(x)}^{g(x)} \varphi(x, y)dy$$

The comparison theorem yields

$$\int_c^d s(x, y)dy \leq \int_c^d \widetilde{\varphi}(x, y)dy \leq \int_c^d t(x, y)dy$$

$$\int_{[a,b]\times[c,d]} s \leq \int_a^b \int_{f(x)}^{g(x)} \varphi(x,y)dydx \leq \int_{[a,b]\times[c,d]} t$$

Thus, for any $\epsilon > 0$,

$$\left| \int_S \varphi - \int_a^b \int_{f(x)}^{g(x)} \varphi(x,y)dydx \right| < \epsilon$$

and we are done. □

**Definition.** $S \subset \mathbb{R}^n$ is of graph type if $\exists T \subset \mathbb{R}^{n-1}$ of graph type and continuous piecewise $C^1$ functions $f, g : T \to \mathbb{R}$ such that $f \leq g$ and for some $i$,

$$S = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \mathbf{x} = (x_1, \ldots, \hat{x}_i, \ldots, x_n) \in T, f(\mathbf{x}) \leq g(\mathbf{x})\}$$

**Theorem 31.** *Let $S \subset \mathbb{R}^n$ be graph type, $\varphi : S \to \mathbb{R}$ continuous. Then $\varphi$ is continuous and*

$$\int_S \varphi = \int_T \left( \int_{f(\mathbf{x})}^{g(\mathbf{x})} \varphi(x_1, \ldots, x_n)dx_i \right)$$

*Proof.* Inductive version of previous. □

**Definition.** The volume of a graph type region $S$ is $\int_S 1$.

**Definition.** The boundary of $S$ is the curve

$$\partial S := C_1 \# C_2 \# - C_3 \# - C_4$$

where they are parameterized by

$$\begin{aligned}
&\gamma_1 : [a,b] \to \mathbb{R} && \gamma_1(x) = (x, f(x)) \\
&\gamma_2 : [f(b), g(b)] \to \mathbb{R} && \gamma_2(x) = (b, x) \\
&\gamma_3 : [a,b] \to \mathbb{R} && \gamma_3(x) = (x, g(x)) \\
&\gamma_4 : [f(a), g(a)] \to \mathbb{R} && \gamma_4(x) = (a, x)
\end{aligned}$$

**Definition.** For $U \subset \mathbb{R}^2$ open, $F : U \to \mathbb{R}^2$ a $C^1$ vector field with $F = (P, Q)$. Let the "mini-curl" of $F$ be $\operatorname{curl} F : U \to \mathbb{R}$ and

$$\operatorname{curl} F = D_1 Q - D_2 P$$

*Remark.* Note that $F$ closed $\iff$ $\operatorname{curl} F = 0$.

**Theorem 32.** *(Green) Let $S \in \mathbb{R}^2$ be of graph type I, given by $f, g : [a, b] \to \mathbb{R}$, and $S \subset U$ for some $U$ open. Then, let $F : U \to \mathbb{R}^2$ by a $C^1$ vector field. Then,*

$$\int_S \operatorname{curl} F = \oint_{\partial S} F \cdot d\gamma$$

*Proof.* First, reduce piecewise $C^1$ case to the $C^1$ case. $\exists$ partition $a = x_0 < x_1 < \cdots < x_n = b$ such that $f|_{(x_{i-1}, x_i)}$ and $g|_{(x_{i-1}, x_i)}$ are $C^1$.
Let $S_i = \{(x, y) \in S \mid x \in [x_{i-1}, x_i]\}$.
So we can assume $f, g$ are $C^1$. Let $F = (P, Q) = (P, 0) + (0, Q)$.
By linearity, it suffices to show

$$-\int_S D_2 P = \oint_{\partial S} (P, 0) \cdot d\gamma \tag{1}$$

$$\int_S D_1 Q = \oint_{\partial S} (0, Q) \cdot d\gamma \tag{2}$$

For $C_2, C_4$, we have that $\gamma'(t) = (0, 1)$ and $(P(\gamma(t)), 0) \cdot \gamma'(t) = 0$.
Then, we get

$$\begin{aligned}
\oint_{\partial S} (P, 0) \cdot d\gamma &= \int_{C_1} (P, 0) d\gamma_1 + \int_{C_2} (P, 0) d\gamma_2 - \int_{C_3} (P, 0) d\gamma_3 - \int_{C_4} (P, 0) d\gamma_4 \\
&= \int_a^b (P(\gamma_1(x)), 0) \cdot (1, f'(x)) dx - \int_a^b (P(\gamma_3(x)), 0) \cdot (1, g'(x)) dx \\
&= \int_a^b (P(x, f(x)) - P(x, g(x))) dx \\
&= -\int_a^b P(x, y)|_{y=f(x)}^{y=g(x)} \\
&= -\int_a^b \int_{f(x)}^{g(x)} D_2 P dy dx
\end{aligned}$$

Now we need to show (2): let $h(x, s) = f(x) + s(g(x) - f(x))$. Then,

$$\begin{aligned}
\int_S D_1 Q &= \int_a^b \int_{f(x)}^{g(x)} D_1 Q dy dx \\
&= \int_a^b \int_0^1 D_1 Q(x, h(x, s)) D_2 h(x, s) ds dx
\end{aligned}$$

Returning to $\partial S$,

$$\oint_{\partial S} (0, Q) \cdot d\gamma = \int_{C_1} (0, Q) d\gamma_1 + \int_{C_2} (0, Q) d\gamma_2 - \int_{C_3} (0, Q) d\gamma_3 - \int_{C_4} (0, Q) d\gamma_4$$

$$
\begin{aligned}
&= \int_{C_1} (0,Q) d\gamma_1 + \int_a^b (0,Q)\cdot(0,1)ds - \int_{C_3} (0,Q)d\gamma_3 - \int_b^a (0,Q)\cdot(0,1)ds \\
&= \int_{C_1} (0,Q) d\gamma_1 + \int_{f(b)}^{g(b)} Q(b,s)ds - \int_{C_3} (0,Q)d\gamma_3 - \int_{f(a)}^{g(a)} Q(a,s)ds \\
&= \int_{C_1} (0,Q) d\gamma_1 - \int_{C_3} (0,Q)d\gamma_3 + \int_0^1 Q(x,h(x,s))D_2h(x,s)|_{x=a}^{x=b} ds \\
&= \int_{C_1} (0,Q) d\gamma_1 - \int_{C_3} (0,Q)d\gamma_3 + \int_0^1 \int_a^b \frac{\partial}{\partial x} Q(x,h(x,s))D_2h(x,s)dxds \\
&= \int_a^b (0,Q)\cdot(1,f'(x))dx - \int_a^b (0,Q)\cdot(1,g'(x)) + \int_0^1 \int_a^b \frac{\partial}{\partial x} Q(x,h(x,s))D_2h(x,s)dxds \\
&= \int_a^b (Q(x,f(x))f'(x) - Q(x,g(x)g'(x)))dx + \int_0^1 \int_a^b \frac{\partial}{\partial x} Q(x,h(x,s))D_2h(x,s)dxds \\
&= -\int_a^b Q(x,h(x,s))D_1h(x,s)|_{s=0}^{s=1} + \int_0^1 \int_a^b \frac{\partial}{\partial x} Q(x,h(x,s))D_2h(x,s)dxds \\
&= -\int_a^b \int_0^1 \frac{\partial}{\partial s}(Q(x,h,(x,s))D_1h(x,s))dsdx + \int_0^1 \int_a^b \frac{\partial}{\partial x} Q(x,h(x,s))D_2h(x,s)dxds \\
&= \int_0^1 \int_a^b (\frac{\partial}{\partial x} Q(x,h(x,s))D_2h(x,s) - \frac{\partial}{\partial s}(Q(x,h,(x,s))D_1h(x,s)))dxds
\end{aligned}
$$

We will now show that the integrands are equal.

$$
\begin{aligned}
&\frac{d}{dx}(Q(x,h(s,x))\frac{d}{ds}(x,s)) - \frac{d}{ds}(Q(x,h(x,s))\frac{d}{dx}h(x,s)) \\
=&\frac{\partial Q}{\partial x}D_2h + QD_1D_2g - \frac{\partial Q}{\partial s}D_1h - QD_1D_2h \\
=&D_1QD_2h + D_2QD_1hD_2h - D_2QD_2hD_1h = D_1QD_2h
\end{aligned}
$$

which was what we wanted. □

*Remark.* Note that this is a generalization of the idea that closed equals conservative on star-shaped domains; in particular, that case is when both sides vanish everywhere!

**Theorem 33.** *(Change of variables) Green's theorem lets us use change of variables in two-dimensions. Let* $S, T \subseteq \mathbb{R}^2$ *be of graph type,* $\varphi : S \to T$ *a* $C^1$ *bijection. Suppose that* $\forall (x,y) \in S$, $\det(D\varphi(x,y)) > 0$ *(in some sense, this condition preserves the direction of* *the map). Then, if $f$ continuous $T \to \mathbb{R}$,*

$$\int_T f = \int_S (f \circ \varphi) \det D\varphi$$

*Proof.* Use Green's to write both sides as a line integral. See Apostol for details. □

*Remark.* This still holds in higher dimensions! But this won't be proved, and the generalization of Green's needed will also never be proved in this class.

**Definition.** Take a $C^1$ parametric surface (in $\mathbb{R}^3$) is a region $T \subset \mathbb{R}^2$ of graph type and a map $r = (x, y, z) : T \to \mathbb{R}^3$ that is continuous and $C^1$ on $T - \partial T$. It's boundary is $r(\partial T)$, parameterized by $r(\gamma_i)$ where $\gamma_i$ are as before.

*Example.* Any graph of $f : T \to \mathbb{R}$ gives a parameterized surface,

$$r(u, v) = (u, v, f(u, v))$$

**Definition.** The cross product gives a vector perpedicular to two given vectors in $\mathbb{R}^3$ (and only $\mathbb{R}^3$).

$$A = (a_1, a_2, a_3)$$
$$B = (b_1, b_2, b_3)$$
$$A \times B = (a_2b_3 - b_2a_3, b_1a_3 - a_1b_3, a_1b_2 - b_1a_2) = \det\left(\begin{bmatrix} \vec{e_1} & \vec{e_2} & \vec{e_3} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\right)$$

**Prop.** *The cross product has a few nice properties, including being bilinear, anti-commutative, and that $A \times B$ is perpedicular to $A, B$.*
*Further, $A \times \lambda A = 0$, and $(A \times B) \cdot C = \det(A|B|C)$ where $A, B, C$ are expressed as column vectors in the right-hand side.*

**Definition.** Let $r : T \to \mathbb{R}^3$ parameterize a surface (assumed $C^1$). Let

$$N = \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}$$

be the outward normal vector at $(u, v)$.
Let $S = r(T), S \subset U$ and open set and $F : U \to \mathbb{R}^3$, and

$$\int_S F \cdot dr^2 := \int_T F(r(u, v)) \cdot N(u, v) dudv$$

**Prop.** *This is independent of the parameterization. For $r : T \to \mathbb{R}^3$ and $q : \widetilde{T} \to T$ a continuous bijection that is $C^1$ on $\widetilde{T} \setminus \partial\widetilde{T}$ with $\det(Dq) > 0$ everywhere, let $\widetilde{r} : \widetilde{T} \to \mathbb{R}^3$ be $\widetilde{r} = r \circ q$. Then,*

$$\int_T F(r(u,v)) \cdot \left(\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}\right) dudv = \int_{\widetilde{T}} F(\widetilde{r}(\widetilde{u},\widetilde{v})) \cdot \left(\frac{\partial \widetilde{r}}{\partial \widetilde{u}} \times \frac{\partial \widetilde{r}}{\partial \widetilde{v}}\right)$$

*Proof.* Show that

$$\widetilde{N} = \det(Dq) \cdot N$$

and use the change of variables formula. □

*Remark.* If $\det(Dq) < 0$, the integral switches signs. Then, an orientation can be considered a "consistent" choice of outward normal vectors.

**Definition.** A piecewise $C^1$ parameterized surface is a finite set $S_1, \ldots, S_n$ of $C^1$ parameterized surfaces, which are nonintersecting except on the boundary. We will also need to pick the outward normal consistently.

If $S$ is a piecewise $C^1$ parameterized surface with pieces $S_1, \ldots, S_n$, let

$$\int_S F \cdot dr^2 := \sum_{i=1}^{n} \int_{S_i} F \cdot dr^2$$

*Remark.* It is not always possible to pick an orientation for a piecewise $C^1$ parameterized surface (e.g. a Mobius strip).

**Definition.** For $U \subset \mathbb{R}^3$ open and $F : U \to \mathbb{R}^3$ be $C^1$ given by $F = (P, Q, R)$, the curl of $F$, $\operatorname{curl} F : U \to \mathbb{R}^3$ is the vector field given by

$$\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) = \det\left(\begin{bmatrix} \vec{e_1} & \vec{e_2} & \vec{e_3} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial x} & \frac{\partial}{\partial x} \\ P & Q & R \end{bmatrix}\right) = \nabla \times F$$

Note that $\operatorname{curl} F = 0 \iff F$ is closed.

**Theorem 34.** *(Stokes's) Let $S$ be a $C^2$ parameterized surface in $\mathbb{R}^3$, $F : U \to \mathbb{R}^3$, and $S \subset U$. Then,*

$$\int_S \operatorname{curl} F \cdot dr^2 = \oint_{\partial S} F \cdot d\gamma$$

*where $\delta S = r(\delta T)$ where $r : T \to \mathbb{R}^3$ is a parameterization of $S$.*

**Definition.** For $U \subset \mathbb{R}^3$ open, $F : U \to \mathbb{R}^3$ a $C^1$ vector field, and $F = (P, Q, R)$. The divergence of $F$ is

$$\operatorname{div} F = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = \nabla \cdot F$$

*Remark.* Note that $\operatorname{div}(\operatorname{curl} F) = 0$ if $F$ is $C^2$.

**Theorem 35.** *If $U$ is star-shaped,* $\operatorname{div}(F) = 0$ *yields that $F$ is a* curl.

**Theorem 36.** *(Divergence) Given a region $W \subset \mathbb{R}^3$ of graph type, there is a notion of $\partial W$ as a piecewise parameterized surfaces with outward normal vectors $N$. Then, for such a $W \subset U$ with $F : U \to \mathbb{R}^3$ a $C^1$ vector field,*

$$\int_W \operatorname{div} F = \int_{\delta W} F \cdot dr^2$$