1st Year Mathematics
Imperial College London

Autumn 2014

Foundations of Analysis

Lectured by:
Prof. A. Corti

Humbly typed by:
Karim Bacchus

Comments and corrections should be sent to kb514@ic.ac.uk.

Syllabus

An introductory course involving basic material, which will be widely used later.

- Number systems. Integers, rational numbers, real numbers, decimal expansions for rationals and reals.
- Inequalities, complex numbers.
- Induction; examples and applications.
- Sets, functions, countability, logic.
- Permutations and combinations. The Binomial Theorem.
- Equivalence relations and arithmetic modulo n.
- Euclids algorithm.
- Introduction to limits.

Appropriate books

M. Liebeck *A Concise Introduction to Pure Mathematics.*
K. Houston *How to Think Like a Mathematician.*
E. Hurst and M. Gould *Bridging the Gap to University Mathematics.*

Contents

1 Basics

1.1 Sets

Lecture 1

A set S is a collection of objects (called the *elements* of the set)

> **Example 1.1.** A way to specify a set is to list the objects (between curly brackets):
>
> $$S = \{1, 3, 7\}$$
>
> The order of elements is unimportant, as is repetition:
>
> $$\{1, 2\} = \{2, 1\} = \{1, 1, 2\}$$

I say $S_1 \subset S_2$ (S_1 *contained* in S_2) if every element of S_1 is also an element of S_2. I can write this as a *statement*: $x \in S_1 \implies x \in S_2$

I say $S_1 = S_2$ if $S_1 \subset S_2$ & $S_2 \subset S_1$. Elements can be sets: $S = \{1, 2, \{1, 2\}\}$.
But there is one thing you are never allowed to do:

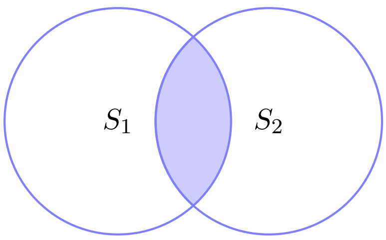
Axiom 1.2 (Foundation Axiom)**.** $S \notin S$

More things about sets:

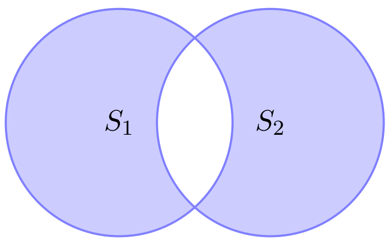
- $a \notin S$ "a is not an element of S"
- $S_1 \cup S_2$ "S_1 union S_2" $= \{x \mid x \in S_1 \text{ or } x \in S_2 \text{ (or both)}\} = \{x : x \in S_1 \text{ or } x \in S_2\}$
- $S_1 \cap S_2$ "S_1 intersection S_2" $= \{x \mid x \in S_1 \text{ and } x \in S_2\}$
- $S_1 \backslash S_2$ "S_1 take away S_2" $= \{x \mid x \in S_1 \& x \notin S_2\}$
- $S_1 \triangle S_2$ "symmetric difference" $= \{x \mid x \in S_1 \text{ or } x \in S_2 \text{ but not both }\}$
  $= (S_1 \cup S_2) \backslash (S_1 \cap S_2)$

When you reason about sets & other mathematical objects, it is useful to draw pictures:

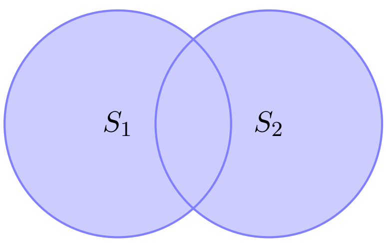
$S_1 \cap S_2$



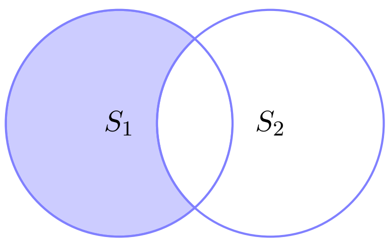
$S_1 \triangle S_2$



$S_1 \cup S_2$



$S_1 \backslash S_2$



Sometimes (often) it is not practical to list all the elements of a set:

Examples 1.3.

$\mathbb{Z}$ = set of integers $= \{0, +1, -1, 2, -2, 3, -3, \dots\}$

$\mathbb{N}$ = set of natural numbers $= \{0, 1, 2, 3, \dots\} = \{n \in \mathbb{Z} \mid n \geq 0\}$

$\mathbb{Q}$ = set of rational numbers $= \{x \mid x = \frac{p}{q},\ p \in \mathbb{Z}, q \in \mathbb{N}\backslash\{0\}\}$

$\mathbb{R}$ = set of real numbers

$\mathbb{C}$ = set of complex numbers

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Lecture 2

2 Induction

> A person who can't tell the difference between strong induction and usual induction, is a sad, demeaning spectacle
>
> \- *Alessio Corti*

3 Relations

Lecture 22

Examples 2.1.

1. $(a,b) \in R \iff a \equiv b \pmod n$
(fix $n \in \mathbb{N}$, a relation on $\mathbb{Z}$)

2. R on $S = \mathbb{Z} \times (\mathbb{Z}\backslash\{0\})$
$((p,q),(N,M)) \in R \iff \exists a,b \in \mathbb{Z}\backslash\{0\}$ such that $pa = Nb$ and $qa = Mb$

Definition. If $a \in S$, the *class of a* is:

$$[a] = \{b \in S \mid (a,b) \in R\}$$

In Example 1: $[a] =$ the "class of a mod n"
In Example 2: $[(p,q)] = \frac{p}{q} \in \mathbb{Q}$

In both examples we have a good way to "picture" the set of classes, which are $\mathbb{Z}\backslash n\mathbb{Z}$ and $\mathbb{Q}$ respectively.

Notation. If $R \subset S \times S$ is an *equivalence relation* on S then we write:

$$a,b \in S\ a\tilde{b} \iff (a,b) \in R$$

We caan tehn re-write (i), (ii), (iii) as:

(i) a

Definition. An *endomorphism* of V is a linear map from V to V. We write $\mathrm{End}(V) = \mathcal{L}(V,V)$ to denote the set of endomorphisms of V:

$$\mathrm{End}(V) = \{\alpha : V \to V, \alpha \text{ linear}\}.$$

The set $\mathrm{End}(V)$ is an algebra: as well as being a vector space over $\mathbb{F}$, we can also multiply elements of it – if $\alpha, \beta \in \mathrm{End}(V)$, then $\alpha\beta \in \mathrm{End}(V)$, i.e. product is *composition* of linear maps.

Recall we have also defined

$$\mathrm{GL}(V) = \big\{\alpha \in \mathrm{End}(V) : \alpha \text{ invertible}\big\}.$$

Fix a basis $b_1, \ldots, b_n$ of V and use it as the basis for both the source and target of $\alpha : V \to V$. Then α defines a matrix $A \in \mathrm{Mat}_n(\mathbb{F})$, by $\alpha(b_j) = \sum_i a_{ij}\, b_i$. If $b_1', \ldots, b_n'$ is another basis, with change of basis matrix P, then the matrix of α with respect to the new basis is PAP^{-1}.

$$\begin{tikzcd}
V \arrow[r, "\alpha"] \arrow[d, "\cong"'] & V \arrow[d, "\cong"] \\
\mathbb{F}^n \arrow[r, "A"'] & \mathbb{F}^n
\end{tikzcd}$$

Hence the properties of $\alpha : V \to V$ which don't depend on choice of basis are the properties of the matrix A which are also the properties of all *conjugate* matrices PAP^{-1}.

These are the properties of the set of $\mathrm{GL}(V)$ orbits on $\mathrm{End}(V) = \mathcal{L}(V, V)$, where $\mathrm{GL}(V)$ acts on $\mathrm{End}(V)$, by $(g, \alpha) \mapsto g\alpha g^{-1}$.

In the next two chapters we will determine the set of orbits. This is called the theory of *Jordan normal forms*, and is quite involved.

Contrast this with the properties of a linear map $\alpha : V \to W$ which don't depend on the choice of basis of both V and W; that is, the determination of the $\mathrm{GL}(V) \times \mathrm{GL}(W)$ orbits on $\mathcal{L}(V, W)$. In chapter 1, we've seen that the only property of a linear map which doesn't depend on the choices of a basis is its rank – equivalently that the set of orbits is isomorphic to $\{i \mid 0 \leq i \leq \min(\dim V, \dim W)\}$.

We begin by defining the determinant, which is a property of an endomorphism which doesn't depend on the choice of a basis.

3.1 Determinants

Definition. We define the map $\det : \mathrm{Mat}_n(\mathbb{F}) \to \mathbb{F}$ by

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{1,\sigma(1)} \dots a_{n,\sigma(n)}.$$

Recall that S_n is the group of permutations of $\{1, \dots, n\}$. Any $\sigma \in S_n$ can be written as a product of transpositions (ij).

Then $\epsilon : S_n \to \{\pm 1\}$ is a group homomorphism taking

$$\epsilon(\sigma) = \begin{cases} +1 & \text{if number of transpositions is even,} \\ -1 & \text{if number of transpositions is odd.} \end{cases}$$

In class, we had a nice interlude here on drawing pictures for symmetric group elements as braids, composition as concatenating pictures of braids, and how $\epsilon(w)$ is the parity of the number of crossings in any picture of w. This was just too unpleasant to type up; sorry!

Example 2.2. We can calculate det by hand for small values of n:

$$\det \begin{pmatrix} a_{11} \end{pmatrix} = a_{11}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ \quad -a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \end{array}$$

The complexity of these expressions grows nastily; when calculating determinants it's usually better to use a different technique rather than directly using the definition.

Lemma 2.3. *If A is upper triangular, that is, if $a_{ij} = 0$ for all $i > j$, then* $\det A = a_{11} \, \dots \, a_{nn}$.

Proof. From the definition of determinant:

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{1,\sigma(1)} \dots a_{n,\sigma(n)}.$$

If a product contributes, then we must have $\sigma(i) \le i$ for all $i = 1, \dots, n$. Hence $\sigma(1) = 1$, $\sigma(2) = 2$, and so on until $\sigma(n) = n$. Thus the only term that contributes is the identity, $\sigma = \mathrm{id}$, and $\det A = a_{11} \, \dots \, a_{nn}$. □

Lemma 2.4. $\det A^T = \det A$, *where* $(A^T)_{ij} = A_{ji}$ *is the* transpose.

Proof. From the definition of determinant, we have

$$\begin{aligned}\det A^T &= \sum_{\sigma\in S_n} \epsilon(\sigma)\, a_{\sigma(1),1}\cdots a_{\sigma(n),n}\\ &= \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}\end{aligned}$$

Now $\prod_{i=1}^n a_{\sigma(i),i} = \prod_{i=1}^n a_{i,\sigma^{-1}(i)}$, since they contain the same factors but in a different order. We relabel the indices accordingly:

$$= \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{k=1}^{n} a_{k,\sigma^{-1}(k)}$$

Now since ϵ is a group homomorphism, we have $\epsilon(\sigma\cdot\sigma^{-1}) = \epsilon(\iota) = 1$, and thus $\epsilon(\sigma) = \epsilon(\sigma^{-1})$. We also note that just as σ runs through $\{1,\dots,n\}$, so does σ^{-1}. We thus have

$$= \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{k=1}^{n} a_{k,\sigma(k)} = \det A.$$ □

Writing v_i for the ith column of A, we can consider A as an n-tuple of column vectors, $A = (v_1,\dots,v_n)$. Then $\mathrm{Mat}_n(\mathbb{F}) \cong \mathbb{F}^n\times\cdots\times\mathbb{F}^n$, and det is a function $\mathbb{F}^n\times\cdots\times\mathbb{F}^n \to \mathbb{F}$.

Proposition 2.5. *The function* $\det : \mathrm{Mat}_n(\mathbb{F}) \to \mathbb{F}$ *is* multilinear*; that is, it is linear in each column of the matrix separately, so:*

$$\begin{aligned}\det(v_1,\dots,\lambda_i\, v_i,\dots,v_n) &= \lambda_i \det(v_1,\dots,v_i,\dots,v_n)\\ \det(v_1,\dots,v_i'+v_i'',\dots,v_n) &= \det(v_1,\dots,v_i',\dots,v_n) + \det(v_1,\dots,v_i'',\dots,v_n).\end{aligned}$$

We can combine this into the single condition

$$\begin{aligned}\det(v_1,\dots,\lambda_i'\, v_i' + \lambda_i''\, v_i'',\dots,v_n) = {}&\lambda_i' \det(v_1,\dots,v_i',\dots,v_n)\\ &+ \lambda_i'' \det(v_1,\dots,v_i'',\dots,v_n).\end{aligned}$$

Proof. Immediate from the definition: $\det A$ is a sum of terms $a_{1,\sigma(1)},\dots,a_{n,\sigma(n)}$, each of which contains only one factor from the ith column: $a_{\sigma^{-1}(i),i}$. If this term is $\lambda_i'\, a_{\sigma^{-1}(i),i} + \lambda_i''\, a''_{\sigma^{-1}(i),i}$, then the determinant expands as claims. □

Example 2.6. If we split a matrix along a single column, such as below, then $\det(A) = \det A' + \det A''$.

$$\det\begin{pmatrix}1&7&1\\3&4&1\\2&3&0\end{pmatrix} = \det\begin{pmatrix}1&3&1\\3&2&1\\2&1&0\end{pmatrix} + \det\begin{pmatrix}1&4&1\\3&2&1\\2&2&0\end{pmatrix}$$

Observe how the first and third columns remain the same, and only the second column changes. (Don't get confused: note that $\det(A+B) \neq \det A + \det B$ for general A and B.)

Corollary 2.7. $\det(\lambda A) = \lambda^n A$.

Proof. This follows immediately from the definition, or from applying the result of proposition **??** multiple times. □

Proposition 2.8. *If two columns of A are the same, then* $\det A = 0$.

Proof. Suppose v_i and v_j are the same. Let $\tau = (i\,j)$ be the transposition in S_n which swaps i and j. Then $S_n = A_n \coprod A_n\tau$, where $A_n = \ker \epsilon : S_n \to \{\pm 1\}$. We will prove the result by splitting the sum

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

into a sum over these two cosets for A_n, observing that for all $\sigma \in A_n$, $\epsilon(\sigma) = 1$ and $\epsilon(\sigma\tau) = -1$.

Now, for all $\sigma \in A_n$ we have

$$a_{1,\sigma(1)} \dots a_{n,\sigma(n)} = a_{1,\tau\sigma(1)} \dots a_{n,\tau\sigma(n)},$$

as if $\sigma(k) \notin \{i, j\}$, then $\tau\sigma(k) = \sigma(k)$, and if $\sigma(k) = i$, then

$$a_{k,\tau\sigma(k)} = a_{k,\tau(i)} = a_{k,j} = a_{k,i} = a_{k,\sigma(k)},$$

and similarly if $\sigma(k) = j$. Hence

$$\det A = \sum_{\sigma \in A_n} \left(\prod_{i=1}^{n} a_{\sigma(i),i} - \prod_{i=1}^{n} a_{\sigma\tau(i),i} \right) = 0.$$ □

Proposition 2.9. *If I is the identity matrix, then* $\det I = 1$

Proof. Immediate.

Theorem 2.10

These three properties characterise the function det.

Before proving this, we need some language.

Definition. A function $f : \mathbb{F}^n \times \dots \times \mathbb{F}^n \to \mathbb{F}$ is a *volume form* on $\mathbb{F}^n$ if

(i) It is *multilinear*, that is, if

$$f(v_1, \dots, \lambda_i v_i, \dots, v_n) = \lambda_i f(v_1, \dots, v_i, \dots, v_n)$$
$$f(v_1, \dots, v_i + v_i', \dots, v_n) = f(v_1, \dots, v_i, \dots, v_n) + f(v_1, \dots, v_i', \dots, v_n).$$

We saw earlier that we can write this in a single condition:

$$f(v_1, \dots, \lambda_i v_i + \lambda_i' v_i', \dots, v_n) = \lambda_i f(v_1, \dots, v_i, \dots, v_n) + \lambda_i' f(v_1, \dots, v_i', \dots, v_n).$$

(ii) It is *alternating*; that is, whenever $i \neq j$ and $v_i = v_j$, then $f(v_1, \ldots, v_n) = 0$.

Example 2.11. We have seen that $\det : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F}$ is a volume form. It is a volume form f with $f(e_1, \ldots, e_n) = 1$ (that is, $\det I = 1$).

Remark. Let's explain the name "volume form".

Let $\mathbb{F} = \mathbb{R}$, and consider the volume of a rectangular box with a corner at 0 and sides defined by $v_1, \ldots, v_n$ in $\mathbb{R}^n$. The volume of this box is a function of $v_1, \ldots, v_n$ that almost satisfies the properties above. It doesnt quite satisfy linearity, as the volume of a box with sides defined by $-v_1, v_2, \ldots, v_n$ is the same as that of the box with sides defined by $v_1, \ldots, v_n$, but this is the only problem. (Exercise: check that the other properties of a volume form are immediate for voluems of rectangular boxes.)

You should think of this as saying that a volume form gives a *signed* version of the volume of a rectangular box (and the actual volume is the absoulute value). In any case, this explains the name. You've also seen this in multi-variable calculus, in the way that the determinant enters into the formula for what happens to integrals when you change coordinates.

Theorem 2.12: Precise form

The set of volume forms forms a vector space of dimension 1. This line is called the *determinant line*.

Lecture 9

Proof. It is immediate from the definition that volume forms are a vector space. Let $e_1, \ldots, e_n$ be a basis of V with $n = \dim V$. Every element of V^n is of the form

$$\left(\sum a_{i1}\, e_i, \sum a_{i2}\, e_i, \ldots, \sum a_{in}\, e_i\right),$$

with $a_{ij} \in \mathbb{F}$ (that is, we have an isomorphism of sets $V^n \xrightarrow{\sim} \mathrm{Mat}_n(\mathbb{F})$). So if f is a volume form, then

$$\begin{aligned} f\left(\sum_{i_1=1}^{n} a_{i_1 1}\, e_{i_1}, \ldots, \sum_{i_n=1}^{n} a_{i_n n}\, e_{i_n}\right) &= \sum_{i_1=1}^{n} a_{i_1 1}\, f\left(e_{i_1}, \sum_{i_2=1}^{n} a_{i_2 1}\, e_{i_2}, \ldots, \sum_{i_n=1}^{n} a_{i_n n}\, e_{i_n}\right) \\ &= \cdots = \sum_{1 \leq i_1, \ldots, i_n \leq n} a_{i_1 1} \ldots a_{i_n n}\, f(e_{i_1}, \ldots, e_{i_n}), \end{aligned}$$

by linearity in each variable. But as f is alternating, $f(e_{i_1}, \ldots, e_{i_n}) = 0$ unless $i_1, \ldots, i_n$ is $1, \ldots, n$ in some order; that is,

$$(i_1, \ldots, i_n) = \big(\sigma(1), \ldots, \sigma(n)\big)$$

for some $\sigma \in S_n$.

Claim. $f(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = \epsilon(\sigma)\, f(e_1, \ldots, e_n)$.

Given the claim, we get that the sum above simplifies to

$$\sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n}\, \epsilon(w)\, f(e_1, \ldots, e_n),$$

and so the volume form is determined by $f(e_1, \dots, e_n)$; that is, $\dim(\{\text{vol forms}\}) \leq 1$. But $\det : \mathrm{Mat}_n(\mathbb{F}) \to \mathbb{F}$ is a well-defined non-zero volume form, so we must have $\dim(\{\text{vol forms}\}) = 1$.

Note that we have just shown that for any volume form

$$f(v_1, \dots, v_n) = \det(v_1, \dots, v_n)\, f(e_1, \dots, e_n).$$

So to finish our proof, we just have to prove our claim.

Proof of claim. First, for any $v_1, \dots, v_n \in V$, we show that

$$f(\dots, v_i, \dots, v_j, \dots) = -f(\dots, v_j, \dots, v_i, \dots),$$

that is, swapping the ith and jth entries changes the sign. Applying multilinearity is enough to see this:

$$\begin{aligned}\underbrace{f(\dots, v_i + v_j, \dots, v_i + v_j, \dots)}_{=0 \text{ as alternating}} &= \underbrace{f(\dots, v_i, \dots, v_i, \dots)}_{=0 \text{ as alternating}} + \underbrace{f(\dots, v_j, \dots, v_j, \dots)}_{=0 \text{ as alternating}} \\ &\quad + f(\dots, v_i, \dots, v_j, \dots) + f(\dots, v_j, \dots, v_i, \dots).\end{aligned}$$

Now the claim follows, as an arbitrary permutation can be written as a product of transpositions, and $\epsilon(w) = (-1)^{\#\text{ of transpositions}}$. ☐

Remark. Notice that if $\mathbb{Z}/2 \not\subset \mathbb{F}$ is not a subfield (that is, if $1 + 1 \neq 0$), then for a multilinear form $f(x, y)$ to be alternating, it suffices that $f(x, y) = -f(y, x)$. This is because we have $f(x, x) = -f(x, x)$, so $2f(x, x) = 0$, but $2 \neq 0$ and so 2^{-1} exists, giving $f(x, x) = 0$. If $2 = 0$, then $f(x, y) = -f(y, x)$ for any f and the correct definition of alternating is $f(x, x) = 0$.

If that didn't make too much sense, don't worry: this is included for mathematical interest, and isn't essential to understand anything else in the course.

Remark. If $\sigma \in S_n$, then we can attach to it a matrix $P(\sigma) \in \mathrm{GL}_n$ by

$$P(\sigma)_{ij} = \begin{cases} 1 & \text{if } \sigma^{-1} i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Exercises 2.13. Show that:

(i) $P(w)$ has exactly one non-zero entry in each row and column, and that entry is a 1. Such a matrix is called a *permutation matrix*.
(ii) $P(w)\, e_i = e_j$, hence
(iii) $P : S_n \to \mathrm{GL}_n$ is a group homomorphism;
(iv) $\epsilon(w) = \det P(w)$.

Theorem 2.14

Let $A, B \in \mathrm{Mat}_n(\mathbb{F})$. Then $\det AB = \det A \det B$.

Slick proof. Fix $A \in \mathrm{Mat}_n(\mathbb{F})$, and consider $f : \mathrm{Mat}_n(\mathbb{F}) \to \mathbb{F}$ taking $f(B) = \det AB$. We observe that f is a volume form. (Exercise: check this!!) But then

$$f(B) = \det B \cdot f(e_1, \ldots, e_n).$$

But by the definition,

$$f(e_1, \ldots, e_n) = f(I) = \det A. \qquad \square$$

Corollary 2.15. *If* $A \in \mathrm{Mat}_n(\mathbb{F})$ *is invertible, then* $\det A^{-1} = 1/\det A$.

Proof. Since $AA^{-1} = I$, we have

$$\det A \det A^{-1} = \det AA^{-1} = \det I = 1,$$

by the theorem, and rearranging gives the result. $\square$

Corollary 2.16. *If* $P \in \mathrm{GL}_n(\mathbb{F})$, *then*

$$\det(PAP^{-1}) = \det P \det A \det P^{-1} = \det A.$$

> **Definition.** Let $\alpha : V \to V$ be a linear map. Define $\det \alpha \in \mathbb{F}$ as follows: choose *any* basis $b_1, \ldots, b_n$ of V, and let A be the matrix of α with respect to the basis. Set $\det \alpha = \det A$, which is well-defined by the corollary.

Remark. Here is a coordinate free definition of $\det \alpha$.

Pick f any volume form for V, $f \neq 0$. Then

$$(x_1, \ldots, x_n) \mapsto f(\alpha x_1, \ldots, \alpha x_n) = (f\alpha)(x_1, \ldots, x_n)$$

is also a volume form. But the space of volume forms is one-dimensional, so there is some $\lambda \in \mathbb{F}$ with $f\alpha = \lambda f$, and we define

$$\lambda = \det \alpha$$

(Though this definition is independent of a basis, we haven't gained much, as we needed to choose a basis to say anything about it.)

Proof 2 of $\det AB = \det A \det B$. We first observe that it's true if B is an elementary column operation; that is, $B = I + \alpha E_{ij}$. Then $\det B = 1$. But

$$\det AB = \det A + \det A',$$

where A' is A except that the ith and jth column of A' are the same as the jth column of A. But then $\det A' = 0$ as two columns are the same.

Next, if B is the permutation matrix $P((i\ j)) = s_{ij}$, that is, the matrix obtained from the identity matrix by swapping the ith and jth columns, then $\det B = -1$, but $A\, s_{ij}$ is A with its ith and jth columns swapped, so $\det AB = \det A \det B$.

Finally, if B is a matrix of zeroes with r ones along the leading diagonal, then if $r = n$, then $B = I$ and $\det B = 1$. If $r < n$, then $\det B = 0$. But then if $r < n$, AB has some columns which are zero, so $\det AB = 0$, and so the theorem is true for these B also.

Now any $B \in \mathrm{Mat}_n(\mathbb{F})$ can be written as a product of these three types of matrices. So if $B = X_1 \cdots X_r$ is a product of these three types of matrices, then

$$\begin{aligned}
\det AB &= \det\big((AX_1 \cdots X_{m-1})\, X_m\big) \\
&= \det(AX_1 \cdots X_{m-1}) \det X_m \\
&= \cdots = \det A \det X_1 \cdots \det X_m \\
&= \cdots = \det A \det (X_1 \cdots X_m) \\
&= \det A \det B.
\end{aligned}$$

□

Remark. That determinants behave well with respect to row and column operations is also a useful way for humans (as opposed to machines!) to compute determinants.

Proposition 2.17. *Let $A \in \mathrm{Mat}_n(\mathbb{F})$. Then the following are equivalent:*

(i) *A is invertible;*
(ii) $\det A \neq 0$*;*
(iii) $r(A) = n$*.*

Proof. (i) $\implies$ (ii). Follows since $\det A^{-1} = 1/\det A$.

(iii) $\implies$ (i). From the rank-nullity theorem, we have

$$r(A) = n \iff \ker \alpha = \{0\} \iff A \text{ invertible}.$$

Finally we must show (ii) $\implies$ (iii). If $r(A) < n$, then $\ker \alpha \neq \{0\}$, so there is some $\Lambda = (\lambda_1, \ldots, \lambda_n)^T \in \mathbb{F}^n$ such that $A\Lambda = 0$, and $\lambda_k \neq 0$ for some k. Now put

$$B = \begin{pmatrix}
1 & & & \lambda_1 & & \\
& 1 & & \lambda_2 & & \\
& & \ddots & \vdots & & \\
& & & \lambda_k & & \\
& & & \vdots & \ddots & \\
& & & \lambda_n & & 1
\end{pmatrix}$$

Then $\det B = \lambda_k \neq 0$, but AB is a matrix whose kth column is 0, so $\det AB = 0$; that is, $\det A = 0$, since $\lambda_k \neq 0$. □

This is a horrible and unenlightening proof that $\det A \neq 0$ implies the existence of A^{-1}. A good proof would write the matrix coefficients of A^{-1} in terms of $(\det A)^{-1}$ and the matrix coefficients of A. We will now do this, after some showing some further properties of the determinant.

We can compute $\det A$ by expanding along any column or row.

> **Definition.** Let A^{ij} be the matrix obtained from A by deleting the ith row and the jth column.

Theorem 2.18

(i) Expand along the jth column:

$$\begin{aligned}\det A &= (-1)^{j+1}\, a_{1j} \det A^{1j} + (-1)^{j+2}\, a_{2j} \det A^{2j} + \cdots + (-1)^{j+n}\, a_{nj} \det A^{nj} \\ &= (-1)^{j+1} \left[a_{1j} \det A^{1j} - a_{2j} \det A^{2j} + a_{3j} \det A^{3j} - \cdots + (-1)^{n+1}\, a_{nj} \det A^{nj} \right]\end{aligned}$$

(the thing to observe here is that the signs alternate!)

(ii) Expanding along the ith row:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij} \det A^{ij}.$$

The proof is boring book keeping.

Proof. Put in the definition of A^{ij} as a sum over $w \in S_{n-1}$, and expand. We can tidy this up slightly, by writing it as follows: write $A = (v_1 \cdots v_n)$, so $v_j = \sum_i a_{ij}\, e_i$. Then

$$\begin{aligned}\det A = \det(v_1, \ldots, v_n) &= \sum_{i=1}^{n} a_{ij} \det(v_1, \ldots, v_{j-1}, e_i, v_{j+1}, \ldots, v_n) \\ &= \sum_{i=1}^{n} (-1)^{j-1}\, a_{ij} \det\left(e_i, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\right).\end{aligned}$$

as $\epsilon(1\,2\,\ldots\,j) = (-1)^{j-1}$ (in class we drew a picture of this symmetric group element, and observed it had $j-1$ crossings.) Now $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$, so we pick up $(-1)^{i-1}$ as the sign of the permutation $(1\,2\,\ldots\,i)$ that rotates the 1st through ith rows, and so we get

$$\sum_i (-1)^{i+j-2}\, a_{ij} \det \begin{pmatrix} 1 & * \\ 0 & A^{ij} \end{pmatrix} = \sum_i (-1)^{i+j}\, a_{ij} \det A^{ij}. \qquad \square$$

Definition. For $A \in \mathrm{Mat}_n(\mathbb{F})$, the *adjugate matrix*, denoted by $\operatorname{adj} A$, is the matrix with

$$(\operatorname{adj} A)_{ij} = (-1)^{i+j} \det A^{ji}.$$

Example 2.19.

$$\operatorname{adj} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \qquad \operatorname{adj} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & -2 & -3 \\ 1 & 0 & -1 \\ -2 & 1 & 2 \end{pmatrix}.$$

Theorem 2.20: Cramer's rule

$(\operatorname{adj} A) \cdot A = A \cdot (\operatorname{adj} A) = (\det A) \cdot I.$

Proof. We have

$$\big((\operatorname{adj} A)\, A\big)_{jk} = \sum_{i=1}^{n} (\operatorname{adj} A)_{ji}\, a_{ik} = \sum_{i=1}^{n} (-1)^{i+j} \det A^{ij} a_{ik}$$

Now, if we have a diagonal entry $j = k$ then this is exactly the formula for $\det A$ in (i) above. If $j \neq k$, then by the same formula, this is $\det A'$, where A' is obtained from A by replacing its jth column with the kth column of A; that is A' has the j and kth columns the *same*, so $\det A' = 0$, and so this term is zero. $\square$

Corollary 2.21. $A^{-1} = \dfrac{1}{\det A} \operatorname{adj} A$ *if* $\det A \neq 0$.

The proof of Cramer's rule only involved multiplying and adding, and the fact that they satisfy the usual distributive rules and that multiplication and addition are commutative. A set in which you can do this is called a *commutative ring*. Examples include the integers $\mathbb{Z}$, or polynomials $\mathbb{F}[x]$.

So we've shown that if $A \in \operatorname{Mat}_n(R)$, where R is any commutative ring, then there exists an inverse $A^{-1} \in \operatorname{Mat}_n(R)$ if and only if $\det A$ has an inverse in R: $(\det A)^{-1} \in R$. For example, an integer matrix $A \in \operatorname{Mat}_n(\mathbb{Z})$ has an inverse with integer coefficients if and only if $\det A = \pm 1$.

Moreover, the matrix coefficients of $\operatorname{adj} A$ are polynomials in the matrix coefficients of A, so the matrix coefficients of A^{-1} are polynomials in the matrix coefficients of A and the inverse of the function $\det A$ (which is itself a polynomial function of the matrix coefficients of A).

That's very nice to know.