LECTURE 14

Date of Lecture: March 11, 2020

As always, $\mathbf{K} \in \{\mathbf{R}, \mathbf{C}\}$.

The symbol ☡ is for flagging a cautionary comment or a tricky argument. It occurs in the margins and is Knuth's version of Bourbaki's "dangerous bend symbol".

An n-tuple $(x_1, \dots, x_n)$ of symbols (x_i not necessarily real or complex numbers) will also be written as a column vector when convenient. Thus

$$(x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A map $\boldsymbol{f}$ from a set S to a product set $T_1 \times \cdots \times T_n$ will often be written as an n-tuple $\boldsymbol{f} = (f_1, \dots, f_n)$, with f_i a map from S to T_i, and hence, by the above convention, as a column vector

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

(See Remark 2.2.2 of Lecture 5.)

The default norm on Euclidean spaces of the form $\mathbf{R}^n$ is the Euclidean norm $\| \, \|_2$ and we will simply denote it as $\| \, \|$.

☡ Note that $(x_1, \dots, x_n) \neq [x_1 \dots x_n]$. Each side is the transpose of the other.

1. Orthogonal and unitary transformations

We fix a finite dimensional inner product space V over $\mathbf{K}$ as well as a linear transformation

$$T \colon V \to V$$

over $\mathbf{K}$ throughout this section.

1.1. **Another look at the standard inner product on $\mathbf{K}^n$.** Let $\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ be vectors in $\mathbf{R}^n$. The inner product $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$ can be described in terms of matrix multiplication. Indeed, a straightforward application of the formula for matrix multiplication yields

$$\bar{\boldsymbol{v}}^t \boldsymbol{u} = \boldsymbol{u}^t \bar{\boldsymbol{v}} = \sum_{i=1}^{n} u_i \bar{v}_i = \langle \boldsymbol{u}, \boldsymbol{v} \rangle \tag{1.1.1}$$

1.2. **Diagonalisation.** Recall that if T is self-adjoint then V has an orthonormal basis consisting of eigenvectors of T. Now suppose $V = \mathbf{K}^n$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ the standard basis on $\mathbf{K}$, A the matrix of T (which we continue to assume is self-adjoint). We have an orthonormal basis $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ of $\mathbf{K}^n$, consisting of eigenvectors (see Theorem 2.3.2 of of Lecture 13) . Let $\lambda_1 \ldots, \lambda_n \in \mathbf{K}$ be the corresponding eigenvalues, i.e. $T\boldsymbol{f}_i = \lambda_i \boldsymbol{f}_i$, $i = 1, \ldots, n$. We know, by Theorem 2.2.5 Lecture 13, that in fact the λ_i lie in $\mathbf{R}$.

Let $\Lambda \colon \mathbf{K}^n \xrightarrow{\sim} \mathbf{K}^n$ be the isomorphism $\Lambda \mathbf{e}_i = \boldsymbol{f}_i$, $i = 1, \ldots, n$, and Γ the matrix of Λ with respect to $\{\mathbf{e}_i\}$. It is clear that

$$\Gamma^{-1} A \Gamma = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \tag{1.2.1}$$

Since $\Lambda \mathbf{e}_i = \boldsymbol{f}_i$, the matrix Γ has $\boldsymbol{f}_i$ as its i^{th} column. Thus

$$\Gamma = \begin{bmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 & \ldots & \boldsymbol{f}_n \end{bmatrix}.$$

It follows that

$$\overline{\Gamma}^t = \begin{bmatrix} \bar{\boldsymbol{f}}_1^{\,t} \\ \bar{\boldsymbol{f}}_2^{\,t} \\ \vdots \\ \bar{\boldsymbol{f}}_n^{\,t} \end{bmatrix}.$$

Now the $(i,j)^{\text{th}}$-entry of $\overline{\Gamma}^t \Gamma$ is $\bar{\boldsymbol{f}}_i^{\,t} \boldsymbol{f}_j$. Since $\bar{\boldsymbol{f}}_i^{\,t} \boldsymbol{f}_j = \langle \boldsymbol{f}_j, \boldsymbol{f}_i \rangle = \delta_{ij}$ we conclude that

$$\overline{\Gamma}^t \Gamma = I \tag{1.2.2}$$

where $I = I_n$ is the identity $n \times n$ matrix. This is equivalent to each of the following equalities:

$$\overline{\Gamma}^t = \Gamma^{-1} \quad \text{and} \quad \Lambda^* = \Lambda^{-1}. \tag{1.2.3}$$

Definition 1.2.4. A matrix Γ over $\mathbf{K}$ satisfying $\Gamma^{-1} = \overline{\Gamma}^t$ is called a *unitary* matrix. A matrix is said to be *orthogonal* if it is unitary and has real entries.

Note that an orthogonal matrix is the same as a real matrix Γ satisfying $\Gamma^{-1} = \Gamma^t$.

1.3. **Orthogonal and unitary operators.** Orthogonal and unitary matrices can be characterised in a number of ways. Here is the main result.

Theorem 1.3.1. *Let $\Lambda \colon V \to V$ be a $\mathbf{K}$-linear transformation. The following are equivalent.*

(a) *Λ is invertible and $\Lambda^{-1} = \Lambda^*$.*
(b) *Λ transforms every orthonormal basis $f_1, \ldots, f_n$ of V into an orthonormal basis $\Lambda f_1, \ldots, \Lambda f_n$.*
(c) *Λ transforms some orthonormal basis $f_1, \ldots, f_n$ of V into an orthonormal basis $\Lambda f_1, \ldots, \Lambda f_n$.*
(d) *Λ preserves inner products, i.e. $\langle \Lambda v, \Lambda w \rangle = \langle v, w \rangle$ for $v, w \in V$.*

Proof. (a) $\implies$ (b): Let $f_1, \ldots, f_n$ be an orthonormal basis. Then $\langle \Lambda f_i, \Lambda f_j \rangle = \langle f_i, \Lambda^* \Lambda f_j \rangle = \langle f_i, \Lambda^{-1} \Lambda f_j \rangle = \langle f_i, f_j \rangle = \delta_{ij}$ and hence we are done.

(b) $\implies$ (c): Obvious.

(c) $\implies$ (d): Suppose $f_1, \ldots, f_n$ is an orthonormal basis which gets transformed by Λ to an orthonormal basis $g_1, \ldots, g_n$. Let $\Phi_f \colon V \xrightarrow{\sim} \mathbf{K}^n$ and $\Phi_g \colon V \to \mathbf{K}^n$ be the isomorphisms of $\mathbf{K}$ vector spaces given by $\Phi_f(f_i) = \mathbf{e}_i$ and $\Phi_g(g_i) = \mathbf{e}_i$, $i = 1, \ldots, n$, where, as always $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is the standard basis on $\mathbf{K}^n$. We have a commutative diagram of isomorphisms

$$\begin{tikzcd}
V \arrow[rr, "\Lambda", "\sim"'] \arrow[dr, "\Phi_f"', "\wr"] & & V \arrow[dl, "\Phi_g", "\wr"'] \\
& \mathbf{K}^n &
\end{tikzcd}$$

According to Proposition 1.1.1 of Lecture 13, Φ_f and Φ_g preserve inner products. It follows that so does Λ.

(d) $\implies$ (a): Suppose Λ preserves inner products. Then $\|\Lambda v\| = \|v\|$ for every $v \in V$. It follows that $\Lambda v = 0$ if and only if $v = 0$, whence Λ is invertible. Next, if v and w are in V then $\langle \Lambda v, w \rangle = \langle \Lambda v, \Lambda \Lambda^{-1} w \rangle = \langle v, \Lambda^{-1} w \rangle$, the last equality following from the hypothesis that Λ preserves inner products. Thus $\Lambda^* = \Lambda^{-1}$. ☐

Definition 1.3.2. Let $\Lambda \colon V \to V$ be a linear transformation satisfying any of the equivalent conditions of Theorem 1.3.1. Λ is said to be *orthogonal* if $\mathbf{K} = \mathbf{R}$. It is said to be *unitary* if $\mathbf{K} = \mathbf{C}$.

Examples 1.3.3. Let

$$S^1 := \{ z \in \mathbf{C} \mid |z| = 1 \}$$

be the unit circle in $\mathbf{C}$, centred at 0.

1. A unitary $n \times n$ matrix Γ must have its determinant in S^1. Indeed $|\det \Gamma|^2 = \det \overline{\Gamma} \det \Gamma = \det \overline{\Gamma}^t \det \Gamma = \det I = 1$. In particular if Γ is orthogonal, then $\det \Gamma = \pm 1$.

2. If Λ is unitary, and $v \in V$ an eigenvector for Λ with eigenvalue λ, then $|\lambda| \|v\| = \|\lambda v\| = \|\Lambda v\| = \|v\|$. Hence $|\lambda| = 1$. Thus all eigenvalues of Λ are in S^1. In particular, all the real eigenvalues of Λ are in $\{1, -1\}$.

3. Rotations and reflections (in $\mathbf{R}^2$ and $\mathbf{R}^3$) preserve inner products and hence are orthogonal.

Let $n = 2$ and Γ an orthogonal matrix with $\Gamma \neq I$. Let $\boldsymbol{u}$ be its first column, and $\boldsymbol{v}$ its second column. Since $\boldsymbol{u}$ is a unit vector, there is a unique $\theta \in [0, , 2\pi)$ such that $\boldsymbol{u} = (\cos\theta, \sin\theta)$. A little thought shows that since $\boldsymbol{v}$ is orthogonal to $\boldsymbol{u}$, and of unit length, we either have $\boldsymbol{v} = (-\sin\theta, \cos\theta)$ or $\boldsymbol{v} = (\sin\theta, -\cos\theta)$. In the former case we have

$$\Gamma = \Gamma_\theta := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

which is rotation by an angle θ. Unless $\theta = 0$ such a rotation cannot fix any vector, and if $\theta \neq 0$ then the only rotation which has eigenvectors is rotation by π radians, i.e., $\Gamma = -I$.

Exchanging the two columns of Γ_θ gives the matrix $\Gamma'_\theta = \left[\begin{smallmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{smallmatrix}\right]$. This represents a reflection. Indeed, its eigenvalues are $\lambda = 1, -1$, and the eigenspace corresponding to $\lambda = 1$, namely the line $(\sin\theta + 1)x - (\cos\theta)y = 0$, gives the axis of the reflection (the axis of the "mirror") Γ'_θ. Check this. Check also that the rotation Γ_θ has no real eigenvalues, unless it is the identity matrix.

Finally check that if

$$\Gamma := \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

then again Γ has an axis of reflection, i.e., it has two eigenvalues, 1 and -1, and the eigenspace corresponding to 1 is a (two way) mirror for reflection.

4. Let $n = 3$. Suppose Γ is orthogonal. Let us first treat the case where $\det\Gamma = 1$. Transposing two adjacent columns will give us an orthogonal matrix with determinant -1. Let λ_1, λ_2, λ_3 be the roots of the characteristic polynomial of Γ (counted with repetition in case of multiple roots). We know that non-real roots occur conjugate pairs, and so either all the roots are real or only one root is real. Suppose only one root is real, say λ_1. Then $\lambda_3 = \bar{\lambda}_2$ and hence $1 = \det\Gamma = \lambda_1\lambda_2\lambda_3 = \lambda_1|\lambda_2|^2 = \lambda_1$. Thus $\lambda_1 = 1$ in this case. Suppose all roots are real. Then they are all ± 1. Not all of them can be -1 since $\det\Gamma = 1$. So one of them is 1. Thus in either case, without loss of generality, we may assume $\lambda_1 = 1$.

Thus we know that Γ fixes every point on a line through the origin. In greater detail, let $\boldsymbol{u}$ be an eigenvector of unit length for λ_1. Let ℓ be the line $\ell = \{t\boldsymbol{u} \mid t \in \mathbf{R}\}$. Then Γ acts as the identity on ℓ. Since Γ preserves inner products, if $P = \ell^\perp$, then $\Gamma(P) = P$. Clearly $\Gamma|_P \colon P \to P$ is an orthogonal linear transformation. Moreover, it is clear that $1 = \det\Gamma = (\det\Gamma|_\ell)(\det\Gamma|_P) = \det\Gamma|_P$. Pick any orthonormal basis $\{\boldsymbol{v}, \boldsymbol{w}\}$ of P in such a way that $\det[\boldsymbol{u} \quad \boldsymbol{v} \quad \boldsymbol{w}] = 1$ (by interchanging $\boldsymbol{v}$ and $\boldsymbol{w}$ if necessary). Then, with respect to the ordered basis $\boldsymbol{v}, \boldsymbol{w}$, the operator $\Gamma|_P$ has matrix $\left[\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right]$ for some θ. Now Γ is the identity if and only if $\Gamma|_P$ is the identity on P, and the latter happens if and only if 1 is an eigenvalue of $\Gamma|_P$. Indeed $\left[\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right]$ has eigenvalues $e^{i\theta}$ and $e^{-i\theta}$ and either both are 1 or neither is 1. Or, put in a more transparent way, if a rotation of the plane fixes one non-zero point, then it fixes all points.

The upshot is this. *If* $\Gamma \neq I$ *then* Γ *has exactly one eigenvalue* λ_1 *equal to* 1. The corresponding eigenspace ℓ is the "axis of rotation" for Γ.

Using this we can produce all 3×3 orthogonal matrices Γ. We leave aside the trivial case of $\Gamma = 1$. Pick any line through the origin as the axis of rotation. This is equivalent to picking a point on the unit sphere S^2 in $\mathbf{R}^3$. Say this point is $\boldsymbol{u}$. Let P be the plane orthogonal to $\boldsymbol{u}$. Then $P \cap S^2$ is a circle C. Pick a second point $\boldsymbol{v}$ on C. Let Q be the plane defined by $\boldsymbol{u}$ and $\boldsymbol{v}$, i.e. the linear span of $\boldsymbol{u}$ and $\boldsymbol{v}$. The line ℓ in $\mathbf{R}^3$ perpendicular to Q and passing through $\mathbf{0}$ must lie on P, since it is orthogonal to $\boldsymbol{u}$ and P is the plane orthogonal to $\boldsymbol{u}$. Since C is a circle in P, centred at $\mathbf{0}$, the line ℓ meets C in two points. Pick $\boldsymbol{w}$ as any of those two points. The matrix $\Gamma = \begin{bmatrix} \boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w} \end{bmatrix}$ is orthogonal, and a little thought show that all orthogonal matrices can be obtained this way. For those who know cross products, $\boldsymbol{w}$ can be taken to be $\boldsymbol{u}\times\boldsymbol{v}$ or its negative (these are the two points of intersection of ℓ with C). See below for the formula for $\boldsymbol{u}\times\boldsymbol{v}$.

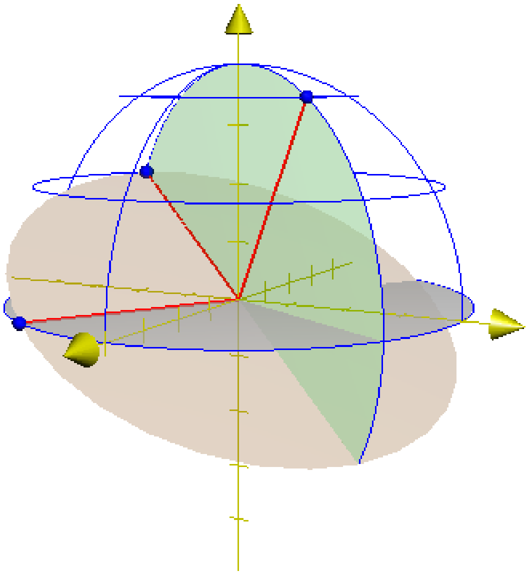
Here is one way this can be done. If $(1,\theta,\phi)$ are the spherical coordinates of $\boldsymbol{u}$ (say $\theta \notin \mathbf{Z}\pi$) then in cartesian terms $\boldsymbol{u} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Since θ is the angle the vector $\boldsymbol{u}$ makes with the positive z-axis, the point $\boldsymbol{v}$ with spherical coordinates $(1, \frac{\pi}{2}-\theta, \phi+\pi)$ lies in the plane containing $\boldsymbol{u}$ and the z-axis, and is at right angles to $\boldsymbol{u}$. In cartesian terms, $\boldsymbol{v} = (-\cos\theta\cos\phi, -\cos\theta\sin\phi, \sin\theta)$. An easy computation of the cross product $\boldsymbol{u}\times\boldsymbol{v}$ yields $\boldsymbol{w} = (\sin\phi, -\cos\phi, 0)$. Thus

$$\Gamma = \Gamma_{\theta,\phi} = \begin{bmatrix} \sin\theta\cos\phi & -\cos\theta\cos\phi & \sin\phi \\ \sin\theta\sin\phi & -\cos\theta\sin\phi & -\cos\phi \\ \cos\theta & \sin\theta & 0 \end{bmatrix}$$

is an orthogonal matrix, cooked up via the recipe given above this paragraph. One can then "rotate" the last two column vectors, namely $\boldsymbol{v}$ and $\boldsymbol{w}$, by an angle ψ in the plane P (see above) to get another orthogonal matrix,

$$\Gamma_{\theta,\phi,\psi} = \begin{bmatrix} \boldsymbol{u} & \boldsymbol{v}_\psi & \boldsymbol{w}_\psi \end{bmatrix}$$

where $\boldsymbol{v}_\psi = (\cos\psi)\boldsymbol{v} + (\sin\psi)\boldsymbol{w}$ and $\boldsymbol{w}_\psi = -(\sin\psi)\boldsymbol{v} + (\cos\psi)\boldsymbol{w}$. One checks that $\det\Gamma_{\theta,\phi,\psi} = 1$. Permuting the columns of $\Gamma_{\theta,\phi,\psi}$ we get every orthogonal matrix (for various choices of θ, ϕ and ψ), including those with determinant -1. Note that $\Gamma_{\theta,\phi} = \Gamma_{\theta,\phi,\psi}$ with $\psi = 0$. In the following picture, $\theta = \pi/6$, $\phi = \pi/3$ and $\psi = 0$ and the three vectors are $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$.



The blue dot (and the attached red line segment, if you wish to think of a vector as a directed line segment) which is closer to the viewer on the green plane is $\boldsymbol{u}$. The vector on the green plane away from the viewer is $\boldsymbol{v}$. The remaining vector is $\boldsymbol{w}$. The inclined plane is orthogonal to $\boldsymbol{u}$ and is the plane P. We have only shown a portion of it, namely the unit disc in P centred at $\mathbf{0}$. The boundary of the disc is C, and note that $\boldsymbol{v}$ and $\boldsymbol{w}$ lie on it. The horizontal (equatorial) plane that you see is of course $z = 0$. Note that $\boldsymbol{w}$ lies on the intersection of the horizontal plane and P.

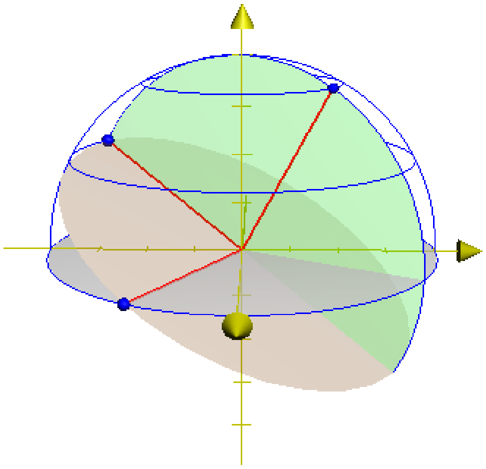
Finally, if $\boldsymbol{u} = (u_1, u_2, u_3)$ and $\boldsymbol{v} = (v_1, v_2, v_3)$, then

$$\boldsymbol{u} \times \boldsymbol{v} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).$$

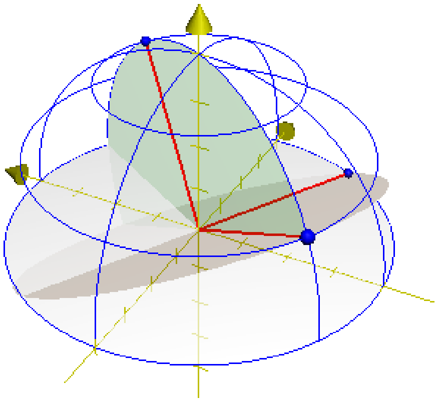
Cramer's rule shows that $\boldsymbol{u} \times \boldsymbol{v}$ is orthogonal to $\boldsymbol{u}$ and $\boldsymbol{v}$. It is not hard to show that its length is $|\boldsymbol{u}||\boldsymbol{v}|\sin\alpha$ where α is the angle "between" $\boldsymbol{u}$ and $\boldsymbol{v}$ and the direction is such that $\det\begin{bmatrix}\boldsymbol{u} & \boldsymbol{v} & \boldsymbol{u} \times \boldsymbol{v}\end{bmatrix} \geq 0$.

One doesn't really need cross products for the specific example we did, once we pick $\boldsymbol{v}$ the way we have (in the plane containing the z-axis and $\boldsymbol{u}$, and orthogonal to $\boldsymbol{u}$). If one does that, then clearly $\boldsymbol{w}$ has to be chosen orthogonal to this plane, which is why we took the cross product $\boldsymbol{u} \times \boldsymbol{v}$, which is a vector orthogonal to this plane. However, here is another way (and there are many other ways) of thinking. The vector $\boldsymbol{w}$ must be orthogonal to the z-axis, since the z-axis is in the plane of $\boldsymbol{u}$ and $\boldsymbol{v}$. So it must be in the xy-plane. Next, $\boldsymbol{w}$ has to be orthogonal to the projection of $\boldsymbol{u}$ on to the xy-plane, since it is anyway orthogonal to the projection of $\boldsymbol{u}$ to the z-axis. The projection of $\boldsymbol{u}$ to the xy-plane makes an angle ϕ with the positive x-axis, by definition of spherical coordinates. So a vector orthogonal to that either makes an angle $\phi - \pi/2$ or $\phi + \pi/2$ with the positive x-axis. The unit vector in the xy-plane making an angle $\phi - \pi/2$ with the positive x-axis is $(\cos(\phi - \pi/2), \sin(\phi - \pi/2), 0)$, i.e. we can pick $\boldsymbol{w} = (\sin\phi, -\cos\phi, 0)$ as we did. The other choice (namely the unit vector making an angle $\phi + \pi/2$ with the positive x-axis) gives the negative of this vector, and that too is acceptable.

Here is a view from another angle of the previous picture. The x-axis points (more or less) towards the viewer. Mark out θ and ϕ in the picture, and see how $\boldsymbol{v}$ and $\boldsymbol{w}$ were constructed.



And for variety, here is yet another view. (Where are the x and y axes?)



About these notes. These course notes are a reasonably faithful record of the lectures given at the Chennai Mathematical Institute (CMI) in the January-April semester of 2019-20. The course is Analysis II, a core course for first year undergraduates at CMI. For more material related to this course, visit `https://www.cmi.ac.in/~pramath/teaching.html#ANA2`. If you have comments on these notes, or on related course material, please send an email to pramath@cmi.ac.in.