ALGEBRA III.

Paul Taylor
August 1984.

## 1. Revision of Algebra II.

This course is a continuation of Algebra II, pursuing the theme of vector spaces and linear maps, although there is a camp (to which the author in part subscribes) which believes that Algebra III should be substantially interchanged with Rings & Modules. From the perspective of a second-year undergraduate, the connections between these courses are easily overlooked, but we shall attempt here to reinforce them.

Recall that a vector-space is the formal-isation of the notion of ("parallelogram") addition and scalar multiplication. The nature of the ground field will therefore become relevant, although the Algebra II course was valid over any field; however you may continue to assume $K = \mathbb{R}$ or $\mathbb{C}$ throughout, and we shall restrict attention to these cases whenever the going gets tough. Also, this course has little to say about infinite-dimensional spaces.

The first major theorem of Algebra II was that every (finite-dimensional) vector space has a basis, the number of elements of which is uniquely determined. This is equivalent to saying that $V \cong K^n$ for some unique $n$, and we shall define an (ordered) basis to be a particular isomorphism. Given an $n$-tuple of values $f_0, \ldots, f_{n-1} \in U$ in another space (of whatever dimension, but of course over the same field) there is a unique linear map $f: V \to U$ taking these values on the basis (namely $(x_0, \ldots, x_{n-1}) \mapsto x_0 f_0 + x_1 f_1 + \ldots + x_{n-1} f_{n-1}$); hence there is a bijection between

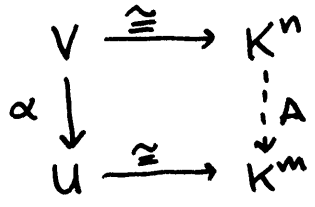
$$\begin{array}{ll} \text{set-functions} & \{0,1,\ldots,n-1\} \longrightarrow U \\ \text{and} \quad \text{linear maps} & V \cong K^n \longrightarrow U \end{array}$$

When we generalise vector-spaces to arbitrary rings (giving a module) we shall say that a module having this property for a particular set is free on that set; hence we may reformulate the result as: every vector space is free on a set unique up to bijection.

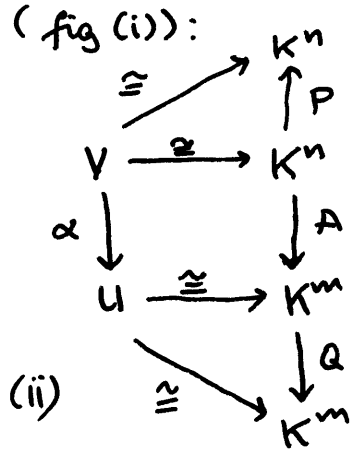
The other major result in Algebra II concerned necessary and sufficient conditions for the existence and uniqueness of solutions of simultaneous linear equations. Since we shall be concerned with endomorphisms (ie where the domain and codomain coincide), this result simplifies to saying that for $\alpha: V \to V$, there is some $x \in V \setminus \{0\}$ with $x\alpha = 0$ iff $\det \alpha \neq 0$. An <u>eigenvector</u> of <u>eigenvalue</u> $\lambda \in K$ is an $x \in V$ with $x\alpha = \lambda x$; hence $\lambda$ is an eigenvalue iff $\det(\alpha - \lambda) = 0$, where $\lambda: V \to V$ means $\lambda$ times the identity.

## 2. Endomorphisms of vector spaces

If $\alpha : V \to U$ is a linear map between vector spaces with bases $V \cong K^n$, $U \cong K^m$, we have a unique linear map $A: K^n \to K^m$ called the <u>matrix</u> of $\alpha$ making the square commute (fig (i)):



(i)



(ii)

Given another pair of bases, there are isomorphisms $P: K^n \to K^n$, $Q: K^m \to K^m$ (ie invertible or nonsingular matrices) given by composition of the new basis with the inverse of the old, and a new matrix, $P^{-1}AQ$

If now $U, V$ coincide, the bases must also coincide, so in the second figure $P = Q$ and the base-change has effect $A \mapsto P^{-1}AP$. this is <u>conjugation</u> of $A$ by $P$ in the group $GL(n, K)$ of invertible $n \times n$ matrices (assuming for the moment that $A$ is) over $K$. The collection of matrices for $\alpha: V \to V$ for various bases is thus the conjugacy class of $\alpha$ and, as in Algebra I, we shall seek canonical representatives of these classes, ie to classify matrices (or endomorphisms) up to conjugacy (or similarity).

As noted before, $\lambda \in K$ is an eigenvalue of $\alpha: V \to V$ iff $\det(\alpha - \lambda) = 0$, ie iff $\lambda$ is a root of

the <u>characteristic polynomial</u> $\chi_\alpha(t) \equiv \det(\alpha - t)$. In practice the latter may be found for a matrix $A$ (with respect to a particular basis), but it is of course basis-independent and of degree $n$ since $t^n$ occurs in the diagonal. $\chi_\alpha$ is therefore an <u>invariant</u> of (the conjugacy of) the endomorphism $\alpha$ (although it does not determine it up to conjugacy). It also follows that any endomorphism (or matrix) over $\mathbb{C}$ has an eigenvalue (and an eigenvector) since every polynomial has a root.

Since by definition the domain and codomain of an endomorphism coincide, it may be composed with itself, and in general if $f(t) \in K[t]$ is any polynomial over the ground field we may form $f(\alpha): V \to V$ for any endomorphism $\alpha: V \to V$. Clearly the $0, \ldots, n^2$ powers of $\alpha$ cannot be linearly independent (where $n = \dim V$) so there is some polynomial of degree at most $n^2$ satisfied by $\alpha$. By the Euclidean algorithm (which should be familiar to you but will be discussed in Rings & Modules) there is a <u>minimal polynomial</u> $m_\alpha(t)$ (wlog with highest coefficient $1 \in K$) such that the polynomials satisfied by $\alpha$ are exactly the multiples of $m_\alpha$.

It is an overemphasised fact (since we have trivially obtained the bound $n^2$ for the degree of $m_\alpha$) that every endomorphism satisfies its own characteristic equation (the <u>Cayley-Hamilton theorem</u>), so $m_\alpha \mid \chi_\alpha$, and despite the author's contempt for this result we shall nevertheless give three proofs of it.

## 3. Diagonal matrices

A matrix is <u>diagonal</u> if it is zero off the diagonal, and an endomorphism is <u>diagonalisable</u> if it has a diagonal matrix wrt some basis. Given the freedom to choose arbitrary bases independently "at each end", any linear map has a diagonal matrix (with 1's & 0's on the diagonal), but we must have the same basis at either end. There is, as may easily be determined, no diagonal matrix conjugate to

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Of course working with a diagonal matrix is very easy – a polynomial may be applied to each (diagonal) coefficient independently. The Cayley-Hamilton theorem is therefore trivial for diagonalisable endomorphisms, and may be proved for all real endomorphisms by a straightforward perturbation/continuity argument (ie the diagonalisable matrices are dense; in fact by a combination of analytic, algebraic and logical sledgehammers even a competent Part II student could wield, the theorem can be proved ... easily in this way for finitely-generated modules over an arbitrary ring).

Diagonalising an endomorphism means finding $\alpha$-invariant subspaces with $V \cong V_0 \oplus V_1 \oplus \dots \oplus V_{n-1}$, so that we can work on each separately. Diagonalisable endomorphisms are therefore direct sums of 1-D spaces (notice we are now confusing $V, \alpha$ and the pair $(V, \alpha)$: we are thinking of a space as equipped with a particular endomorphism). More generally, what do indecomposable spaces look like, and when can we find a decomposition, and with what uniqueness? ...

We shall only attempt to answer this question over $\mathbb{C}$ (ie an algebraically closed field), and all you need to know is the answer – the Jordan (canonical form – because it makes other results easier to remember.

## 14. Jordan canonical form

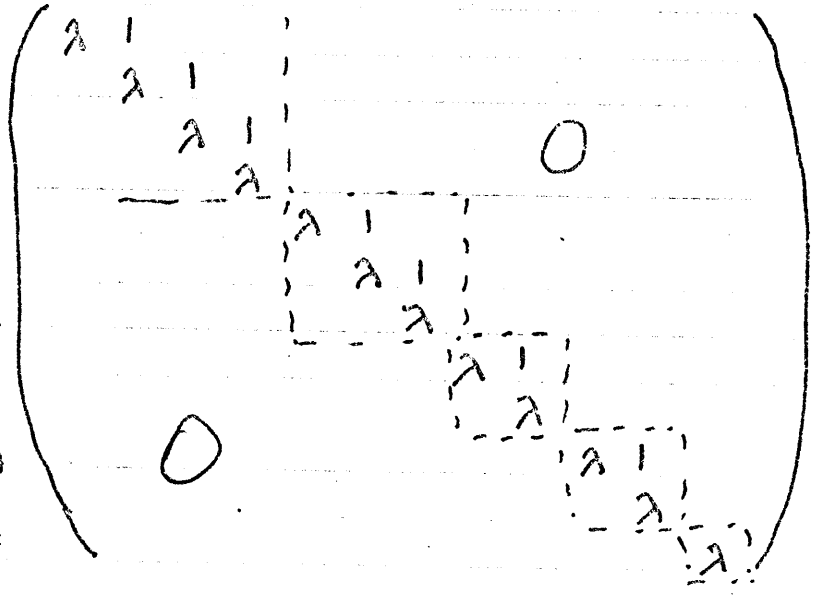
Clearly there is a decomposition for finite-dimensional spaces [why?]; moreover the characteristic polynomial (like the determinant) of a direct sum is the product of that of the components, whilst for the minimal polynomial it is the least common multiple. Conversely, if the minimal polynomial of $(V,\alpha)$ is a product of coprime factors then the space can be decomposed as a direct sum of subspaces on which the minimal polynomial is each of the factors. For if $m = m_1 m_2$ with $m_1, m_2$ coprime, there are (by the Euclidean algorithm) polynomials $r_1, r_2$ with $r_1 m_1 + r_2 m_2 = 1$, so for $v \in V$ put $v_1 = v r_2(\alpha) m_2(\alpha)$, $v_2 = v r_1(\alpha) m_1(\alpha)$ so that $v = v_1 + v_2$, $v_1 m_1 = 0 = v_2 m_2$ and $v_i \in V_i = \ker m_i(\alpha)$ which is $\alpha$-invariant (since $\alpha$ commutes with any polynomial in it)

5

Now (assuming algebraic closure) let $m(t) = (t-\lambda_1)^{a_1}(t-\lambda_2)^{a_2}\dots(t-\lambda_r)^{a_r}$ with $\lambda_1,\dots,\lambda_r$ distinct. Put $m_i(t) = (t-\lambda_i)^{a_i}$ and $V_i = \ker m_i(\alpha)$ so that $V = V_1 \oplus \dots \oplus V_r$ uniquely [why?]. The $V_i$ are, however, not yet indecomposable (consider a diagonal matrix with two equal coefficients). We are left to study a space $V$ with an endomorphism $\alpha$ whose minimal polynomial is $(t-\lambda)^a$ for some $a \in \mathbb{N}, \lambda \in K$; replacing $\alpha$ by $\alpha - \lambda$, wlog $\lambda = 0$: such an $\alpha$ is said to be nilpotent of class $a$.

Let $(V,\alpha)$ be an indecomposable space of eigenvalue $\lambda$. Let $e_{1,1}, e_{2,1}, \dots, e_{i,1}$ be a basis of $\ker(\alpha-\lambda)$. Let $x \in V$ be a vector with $x(\alpha-\lambda)^{i-1} = e_{i,1}$ so $x(\alpha-\lambda)^i = 0$; what is the inverse image of $e_{i,1}$? $V_{1,i} = (\ker e_{i,1})(\alpha-\lambda)^{i-1} = \ker(\alpha-\lambda) \cap U_{1,i}$ for some subspace $U_{1,i}$ (and $\alpha$-invariant, of course) of dimension $\le 1$, for if $x(\alpha-\lambda) = e_{1,i} = y(\alpha-\lambda)$ then $(x-y) \in \ker(\alpha-\lambda)$, so let $e_{2,i} \in U_{1,i}$ (with $e_{2,i}(\alpha-\lambda) = e_{1,i}$ (unless $U_{1,i} = 0$). Similarly for $\{x \in V : x(\alpha-\lambda)^2 = 0\}$ and so on by induction. We then have a basis $\{e_{j,i}\}$ with $e_{j+1,i}(\alpha-\lambda) = e_{j,i}$ $(i,j \ge 1)$ and $e_{1,i}(\alpha-\lambda) = 0$. Thus $\alpha$ has matrix like

$j,i =$ 4,1 / 3,1 / 2,1 / 1,1 / 3,2 / 2,2 / 1,2 / 2,3 / 1,3 / 2,4 / 1,4 / 1,5




Of course the squares in this matrix are indecomposable components, so in fact if $(V,\alpha)$ is itself indecomposable, there was only one eigenvalue in the first place. However we have found the components of the Jordan canonical form for this root of the minimal polynomial.

The subspace decomposition is clearly not unique, but by considering $\dim \ker (\alpha-\lambda)^k$ we may see that the sizes of the subblocks are uniquely determined. The degree of the factor $(x-\lambda)^2$ in the minimal polynomial is the size of the largest subblock of eigenvalue $\lambda$ whilst that in the characteristic polynomial is the total size of the $\lambda$-block. It's easy to prove the Cayley-Hamilton theorem for an indecomposable space (or subblock), whence it follows in general. Finally, an endomorphism is diagonalisable iff the roots of the minimal polynomial are distinct; in particular this holds if some power of it is the identity.

## 5. Bilinear forms

A bilinear form is a function $f: V \times V \to K$ linear in each variable separately. So for each $x \in V$, $y \longmapsto f(x,y)$ is linear (call it $f(x,-)$), and moreover $x \longmapsto f(x,-)$ is a linear map $\bar{f}: V \to V^*$ (recall $V^*$, the dual of $V$, is the space of linear maps $V \to K$ with pointwise addition and scalar multiplication and that in the finite-dimensional case, $\dim V = \dim V^*$ but they are not *naturally* isomorphic). $\bar{f}$ is thus an element of the space $[V \to V^*]$ or a linear map $[V \to V^*]^* \to K$ (where $W \cong W^{**}$ naturally): in Rings and Modules we shall construct the tensor product $V \otimes V$ such that bilinear maps from $V \times V$ correspond to linear maps from $V \otimes V$, so $V \otimes V \cong [V \to V^*]^*$ (naturally).

$f: V \times V \to K$ is symmetric or skew-symmetric (or antisymmetric) if $f(y,x) = f(x,y)$ or $-f(x,y)$ respectively. For $K = \mathbb{C}$ we also have Hermitean if $f(y,x) = \overline{f(x,y)}$: there's nothing to be gained by studying skew-Hermitean forms because they can be reduced to the Hermitean case by multiplying by $i = \sqrt{-1}$. Of course if $K$ has characteristic 2 (so $-1 = 1$) symmetric and antisymmetric coincide and most of the other definitions and results in the remainder of the course fall apart, so we shall henceforward exclude this case.

Choosing a basis for $V$ we may represent a bilinear form by a matrix $f_{ij} = f(e_i, e_j)$. However there is considerable danger of confusion here, because the meaning of this kind of matrix is quite different from that corresponding to an endomorphism or other linear map *between* vector spaces. This problem is faced squarely in General Relativity, where there are conventions of lower (covariant) and upper (contravariant) indices; familiarity with this convention is perhaps the best way of understanding this point. A particular consequence is that whereas a change of base replaces the matrix $A$ of an endomorphism by $P^{-1}AP$, that of a *bilinear form* becomes $P^TAP$ (where $^T$ denotes transpose).

A <u>quadratic form</u> $\varphi : V \to K$ is a function (not linear) such that $\varphi(x) = f(x,x)$ for some (wlog symmetric) bilinear form $f : V \times V \to K$. This is determined uniquely by $\varphi$ since

$$f(x,y) = \tfrac{1}{2}\left[\, \varphi(x+y) - \varphi(x) - \varphi(y) \,\right]$$

where of course we must have $\operatorname{char} K \neq 2$. Quadratic and symmetric bilinear forms are thus equivalent. In the complex case we're interested in <u>Hermitean forms</u>.

## 6. <u>Diagonalisation of bilinear forms.</u>

As with endomorphisms, representing a form by a diagonal matrix is desirable, and it is possible iff the form is symmetric (Hermitean in the complex case, where $f_{ij} = f(e_i, \bar{e}_j)$). We prove this by induction on the dimension, observing that if $\varphi(x) = f(x,x) = 0$ for all $x \in V$ then $f = 0$ and the result is trivial. Choose therefore some $e_1$ with $f(e_1, e_1) \neq 0$ and consider $V_1 = e_1^\perp = \{ x \in V : f(e_1, x) = 0 \}$: this is a complement to $\langle e_1 \rangle$ by the isomorphism theorem for vector spaces (also known as the rank-nullity formula) since the image of $f(e_1, -) : V \to K$ is the whole of $K$. Restricting $f$ (or $\varphi$) to $V_1$ gives a symmetric bilinear form of lower dimension, which is therefore diagonalisable, say with basis $e_2, \ldots, e_n$. It's easy to check that $e_1, e_2, \ldots, e_n$ gives a diagonalising basis for ($f$ and) $V$.

The freedom with which we chose $e_i$ shows that this basis is far from unique, and we are left to consider what we can say about the coefficients on the diagonal. For a general field or ring (and $\mathbb{Q}, \mathbb{Z}$ are the most difficult cases) this is a major question of number theory, but we shall answer it in the cases of $\mathbb{R}$ and $\mathbb{C}$. Here straightforward scaling reduces the choice of coefficients to just $\{-1,0,+1\}$ for $\mathbb{R}$ and $\{0,1\}$ for $\mathbb{C}$ and the distribution of such values, $(m,z,p)$, is called the <u>signature</u> of the form. Clearly $m+z+p = \dim V$ is an invariant, as is $m+p$ since this is the rank of the matrix, solving the problem for $\mathbb{C}$. This leaves us with one parameter in the real case (sometimes the signature is defined as the number $p-m$)

There are many traps in chasing the correct observation to prove this result, due frequently to a subconscious assumption that $\phi$ (the quadratic form) is somehow linear. However, there are subspaces of dimension $p$, $m$ on which the form is respectively positive definite and negative definite (see below), and these are the largest such dimensions, for any space of dimension $r > p$ on which the form were positive definite would intersect the given space of $m$ dimensions on which it is negative-definite in a space of dimension $\geq r+m-(p+m+z) = r-p-z \geq 1$ on which the form is both positive and negative definite, which is impossible. Similarly $m$.

A form is <u>positive semi-definite</u> if $f(x,x) \geq 0$ for all $x$ and <u>positive definite</u> if $f(x,x) > 0$ for $x \neq 0$; equivalently $m=z=0$ and $m=0$ respectively. Similarly negative. It is <u>nonsingular</u> if $f(x,-) \neq 0$ for $x \neq 0$, ie $\bar{f}: V \to V^*$ is an isomorphism; equivalently $z=0$. A nonsingular bilinear form is thus a choice of a particular isomorphism $V \cong V^*$; we shall pursue this point, particularly in the positive definite case, in the remainder of the course.

Indefinite nonsingular forms cannot be dismissed, since the Lorentz metric is an example. They must be treated with considerable care as regards signs and they can behave very peculiarly.

## 7. Inner Product Spaces.

Vector spaces will now come equipped with a positive-definite (and hence nonsingular) quadratic form (called an inner product), which is probably what you thought a vector space was before doing Algebra II. The diagonalisation algorithm of the previous section is called the Gramm-Schmidt orthogonalisation process and a base in which the form is diagonal is called orthogonal; it may be scaled to give 1's on the diagonal (the "identity matrix" or $\delta_{ij}$), in which case the base is orthonormal. In this case the space may as well be $\mathbb{R}^n$ or $\mathbb{C}^n$ with the standard basis and inner product.

The basis is of course still not the unique one making the form diagonal with 1's: it may still be "rotated". The group of linear automorphisms preserving a real quadratic (resp. complex Hermitean) form is called the orthogonal (resp. unitary) group, denoted $O_n$ (resp. $U_n$); the group of all linear automorphisms whatever of a vector space is called the general linear group, denoted $GL_n$ and the group of linear automorphisms preserving a skew-symmetric form is called the symplectic group ($Sp_n$). The preservation of a quadratic form requires the modulus of the determinant to be 1 but an orthogonal (unitary) map may have determinant $-1$ (any value on the unit circle in $\mathbb{C}$); the groups of linear maps with determinant 1 are called $SO_n$, $SU_n$ & $SL_n$, where "S" stands for "special". These groups will be studied in greater detail in Topological Groups.

We are in a position to prove Euler's theorem: every rotation (element of $SO_3$) in $\mathbb{R}^3$ has an axis (eigenvector of eigenvalue 1). $O_n$ (resp. $U_n$) consists of those matrices in $GL_n$ satisfying $AA^T = 1$ (resp. $A\overline{A}^T = 1$), so a straightforward calculation with the Jordan canonical form shows that an element of $SO_3$ is diagonalisable over $\mathbb{C}$ with eigenvalues $\lambda, \mu, \nu$ such that $\lambda\mu\nu = 1 = |\lambda| = |\mu| = |\nu|$. Since these are roots of a cubic equation with real coefficients they must either all be real (hence either $\{1,1,1\}$ or $\{1,-1,-1\}$) or one real (and positive, hence 1) and two conjugate complex.

## 8. Forms and endomorphisms coincide.

In section 5 we remarked that a form $f: V \times V \to K$ gives rise to a linear map $\bar{f}: V \to V^*$, which is an isomorphism in the case of a nonsingular finite-dimensional form. By the theory of dual spaces there is a dual map $\bar{f}^*: V \cong V^{**} \to V^*$ which coincides with $\bar{f}$ iff the form is positive definite (which is a source of pitfalls in Special Relativity). In our case, then, we may identify $V$ and $V^*$ as in Vector Calculus, and (other) bilinear forms correspond naturally to endomorphisms whilst the matrix of the dual of an endomorphism is the transpose of the original matrix (or Hermitean transpose in the complex case).

Returning to a basis-free setting, let us denote the form by $\langle -,- \rangle$. Let $\alpha: V \to V$ be any endomorphism: we aim to define its transpose $\alpha^*$ such that $\langle x\alpha, y \rangle = \langle x, y\alpha^* \rangle$ for all $x, y \in V$. By the isomorphism $V \cong V^*$ induced by $\langle -,- \rangle$, $y\alpha^*$ will correspond to the linear map $\langle -, y\alpha^* \rangle: V \to K$, which is of course $x \mapsto \langle x, y\alpha^* \rangle = \langle x\alpha, y \rangle$; this defines $\alpha^*$ uniquely. If $\alpha = \alpha^*$, it is said to be <u>self-conjugate</u>: endomorphisms of this kind are of considerable importance in Quantum Mechanics and Mathematical Methods. Finally, if $\alpha$ is such that $\langle x\alpha, y\alpha \rangle = \langle x, y \rangle$ for all $x, y \in V$ then it (is nonsingular and) preserves the form; equivalently $\alpha\alpha^* = \alpha^*\alpha = 1$ as in the previous section. Of course here the confusion between $P^{-1}AP$ and $P^TAP$ disappears.

The last topic in the course concerns the simultaneous diagonalisation of symmetric bilinear forms, which one showmanship-minded lecturer appears to regard as a conjuring trick, although it's really just a case of the applied mathematicians' technique of making the best use of symmetry and choosing an appropriate base. One of the forms must be positive-definite, so we use it to make an inner-product space in which to diagon-alise the other as (if we wish) an endomorphism. That this is possible is precisely the statement that there's an orthogonal (or orthonormal) base of eigenvectors (over $\mathbb{C}$, although in fact the eigenvalues are real).

Consider the second form as a self-conjugate

endomorphism $\alpha = \alpha^*$, and let $e$ be an eigenvector of eigenvalue $\lambda$. Then $\bar{e}$ is an eigenvector of eigenvalue $\bar{\lambda}$, whence $\bar{\lambda}\langle e, \bar{e}\rangle = \langle e, \bar{e}\alpha^*\rangle = \langle e\alpha, \bar{e}\rangle = \lambda\langle e, \bar{e}\rangle$ so since $e \neq 0$, $\lambda = \bar{\lambda}$ is real. Put $e^\perp = \{x \in V : \langle e, \bar{x}\rangle = 0\}$; this is a complement to the space spanned by $e$ as in §6, and $V = \langle e\rangle \oplus e^\perp$ with [check this!] both $\langle -, -\rangle$ and $\alpha$ the direct sum. If $e, f$ are eigenvectors of eigenvalues $\lambda \neq \mu$ then $\lambda\langle e, \bar{f}\rangle = \mu\langle e, \bar{f}\rangle$ whence $e, f$ are orthogonal.

Examples of this in classical mechanics include principal axes for an ellipsoid, and for the inertia tensor of a rigid body, and the stress tensor in a fluid or in an electromagnetic field. In Quantum Mechanics this method is used to find eigenstates ("observable states") indexed by quantum numbers.

## 9. Simultaneous diagonalisation of endomorphisms

On the subject of simultaneous diagonalisation but returning to the first part of the course, when can two diagonalisable endomorphisms be simultaneously diagonalised? (In Quantum Mechanics this is simultaneous observability). If an endomorphism $\alpha$ is diagonalisable, the space is the direct sum of eigenspaces: $V = \oplus A_\lambda$ where $A_\lambda = \{x \in V : x\alpha = \lambda x\}$. Hence $\alpha, \beta$ are diagonalisable on each $A_\lambda \cap B_\mu$; the condition, therefore, is that $\oplus_{\lambda\mu} A_\lambda \cap B_\mu = V$, which is clearly equivalent to $\alpha\beta = \beta\alpha$, as is well-known in Quantum Mechanics!