21-238, Math Studies Algebra 2, Department of Mathematical Sciences, Carnegie Mellon University
Spring 2012: Monday, Wednesday, Friday, 10:30 am, Doherty Hall 1211.
Luc Tartar, University Professor of Mathematics, Wean Hall 6212, tartar@cmu.edu

10- Monday February 6, 2012.

Remark 10.1: If $B \in L(V,V)$ is nilpotent, and $Y_i = ker(B^i)$, and $d_i = dim(Y_i)$ for $i = 1, \ldots, k$, so that $0 < d_1 < \ldots < d_k = m = dim(V)$, the construction of Jordan blocks is actually related to the fact that $d_1 - 0 \geq d_2 - d_1 \geq d_3 - d_2 \geq \ldots \geq d_{k-1} - d_k$, in the casse $k \geq 2$, of course, since $k = 1$ means $B = 0$, which is diagonal in any basis of V. If $j \geq 2$ and Z_j is a complement of Y_{j-1} in Y_j, then the restriction of B to Z_j is injective, since the kernel of B is $Y_1 \subset Y_{j-1}$; by choosing $d_j - d_{j-1}$ linearly independent vectors spanning Z_j, their images are linearly independent and span a subspace which is in Y_{j-1} but only intersect Y_{j-2} at $\{0\}$, so that $d_j - d_{j-1}$ is $\leq d_{j-1} - d_{j-2}$ (denoting $d_0 = 0$).

Starting with $j = k$, let $f_1, \ldots, f_a$ (with $a = m - d_{k-1}$) be a basis of a complement Z_k of Y_{k-1} in $Y_k = V$, one then adds the images $f_{a+1} = B\,f_a, \ldots, f_{2a} = B\,f_a$ and one eventually completes with $f_{2a+1}, \ldots, f_b$ for having a basis of a complement Z_{k-1} of Y_{k-2} in V (with $b = m - d_{k-2}$); one continues with the images of the newly added vectors $f_{a+1}, \ldots, f_b$, completing eventually for having a basis of a complement Z_{k-2} of Y_{k-3} in V, and one repeats until one has a basis of V. Each initial vector (from $\{f_1, \ldots, f_a\}$) together with its successive images correspond to a Jordan block of size k, and the vectors added later correspond to smaller size Jordan blocks.[1]

Definition 10.2: If V is an E-vector space, two linear mappings $A, B \in L(V,V)$ are called *similar* if there exists $P \in GL(V)$ (i.e. an invertible linear mapping from V into V) such that $B = P^{-1}A\,P$. Similarly, two $n \times n$ matrices A, B with entries in E are called similar if there exists an invertible $n \times n$ matrix P such that $B = P^{-1}A\,P$.

Remark 10.3: One should consider that P serves in changing basis, and that the columns of P are the elements of a new basis: if a vector v corresponds to a column vector X in the initial basis and to a column vector X' in the new basis, one has $X = P\,X'$; the same relation $Y = P\,Y'$ must then hold for the images, i.e. $Y = A\,X$ using the initial basis, and $Y' = B\,X'$ using the new basis, so that since $Y' = P^{-1}Y = P^{-1}A\,X = P^{-1}A\,P\,X'$, which should be $B\,X'$ for any X', and one deduces the relation $B = P^{-1}A\,P$.

Of course, similarity is an equivalence relation, reflexivity follows from choosing $P = I$, symmetry corresponds to writing $A = Q^{-1}B\,Q$ with $Q = P^{-1}$, and transitivity follows from $R^{-1}(P^{-1}A\,P)\,R = (P\,R)^{-1}A\,(P\,R)$. Two similar matrices have the same characteristic polynomial, since $P^{-1}A\,P - \lambda\,I = P^{-1}(A - \lambda\,I)\,P$, and $det\big(P^{-1}(A - \lambda\,I)\,P\big) = det(P^{-1})\,det(A - \lambda\,I)\,det(P)$, which is $det(A - \lambda\,I)$ because $det(P^{-1})\,det(P) = det(P^{-1}P) = det(I) = 1$. However, two matrices with the same characteristic polynomials are not necessarily similar, except if this polynomial splits over E with simple roots

The analysis made for putting a matrix in Jordan form shows that two matrices which have the same characteristic polynomial which splits over E are similar if and only if for each eigenvalue λ and for all $j \geq 1$ the number of Jordan blocks of size j is the same for the two matrices.

Remark 10.4: If a quadratic form q on a finite dimensional E-vector space is $\sum_{i,j} Q^1_{i,j}X_iX_j$ (with $Q^1_{j,i} = Q^1_{i,j}$ for all i,j) in the initial basis, and $\sum_{i,j} Q^2_{i,j}X'_iX'_j$ (with $Q^2_{j,i} = Q^2_{i,j}$ for all i,j) in the new basis, then, using $X = P\,X'$, one deduces that $Q^2 = P^TQ^1P$ (where $P^T_{i,j} = P_{j,i}$ for all i,j). Actually, since $P \in L(V,V)$, one has $P^T \in L(V^*,V^*)$, which is consistent with having $Q^1, Q^2 \in L(V,V^*)$, and when choosing a basis for V one uses the dual basis for V^*, and it makes sense to talk about symmetry of a linear mapping from V into V^*; indeed, one has $q(v) = \langle Q\,v, v\rangle$ for a symmetric $Q \in L(V,V^*)$ (even if V is infinite-dimensional), and then if one chooses a basis of V, and the dual basis on V^*, Q is represented by a symmetric matrix.

If V is an Euclidean space, then using only orthonormal basis corresponds to having $P^T = P^{-1}$, consequence of $P^TP = I$ for expressing that P is an isometry.

[1] Practically, one only talks about Jordan blocks for sizes ≥ 2. Here the eigenvalue is 0, but a Jordan block for an eigenvalue λ is then a $d \times d$ block with λ in the diagonal, 1s in the diagonal above, and 0s elsewhere.

Definition 10.5: One says that two E-vector spaces V, W are *in duality*, if there is a bilinear form B from $V \times W$ into E such that for each non-zero $v \in V$ there exists a (non-zero) $w \in W$ such that $B(v,w) \neq 0$, and for each non-zero $w \in W$ there exists a (non-zero) $v \in V$ such that $B(v,w) \neq 0$: V may then be identified to a subspace of W^*, W may be identified to a subspace of V^*. Instead of $B(v,w)$, one writes either $\langle v, w\rangle_{V,W}$ or $\langle w, v\rangle_{W,V}$, and if there is no possible confusion, one does not write the names of the spaces used.

Definition 10.6: If V, W are two E-vector spaces, then for $v \in V, w \in W$, the *tensor product* $v \otimes w$ is the mapping which to $(v_*, w_*) \in V^* \times W^*$ associates $\langle v, v_*\rangle \langle w, w_*\rangle$, and the tensor product $V \otimes W$ is the E-vector space of finite combinations of such elements. If $e_i, i \in I$, is a basis of V, and $f_j, j \in J$, is a basis of W, then $e_i \otimes f_j$ is a basis of $V \otimes W$.

More generally, if $V_1, \ldots, V_m$ are E-vector spaces, the tensor product $V_1 \otimes \cdots \otimes V_m$ is the space of finite linear combinations of tensors like $v_1 \otimes \cdots \otimes v_m$, which sends $(w_1, \ldots, w_m) \in V_1^* \times \cdots \times V_m^*$ on $\langle v_1, w_1\rangle \cdots \langle v_m, w_m\rangle$.

Remark 10.7: Often, one restricts attention to the case when one uses a finite-dimensional E-vector space V, and that each V_i is either V or V^*, for $i = 1, \ldots, m$; in this case, one uses the same basis $e_1, \ldots, e_n$ for all copies of V, and the same dual basis $e^1, \ldots, e^n$ for all copies of V^*, of course. There are two important conventions in this framework, one which is about the position of the indices for reflecting if one refers to a copy of V or to a copy of V^*, and the second which is *Einstein's convention*,[2] of summing on repeated indices, and one should be in the bottom and the other should be in the top, without having to write a sum sign.

A vector $x \in V$ has a decomposition $x = \sum_i x^i e_i$ on the basis, which one writes $x = x^i e_i$ using Einstein's convention, and the position of the index i in x^i is consistent with the fact that $x^i = e^i(x)$, and such a formula implicitly means "for all i" since the index i only appears once on either side, and the occurrence on both sides should be at the same level.

A vector $\xi \in V^*$ has a decomposition $\xi = \sum_i \xi_i e^i$ on the basis, which one writes $\xi = \xi_i e^i$ using Einstein's convention, and the position of the index i in ξ_i is consistent with the fact that $\xi_i = \xi(e_i)$.

A linear mapping P from V into V has entries $P^i{}_j$, so that $y = P\,x$ is written as $y^i = P^i{}_j x^j$ using Einstein's convention. The Kronecker symbol $\delta^i{}_j$ then corresponds to the identity matrix I mapping V onto V, while $\delta_i{}^j$ corresponds to the identity matrix I mapping V^* onto V^*.

A quadratic form q on V is written as $q_{ij} x^i x^j$ using Einstein's convention, and it means that one avoids writing both $\sum_i$ and $\sum_j$; of course, q_{ij} corresponds to entries of a linear mapping from V into V^*.

Remark 10.8: In $\mathbb{R}^3$, the cross product $c = a \times b$ means $c^i = \sum_{j,k=1}^{3} \varepsilon^i{}_{jk} a^j b^k$ for $i = 1, 2, 3$, where the *completely antisymmetric tensor* is defined by $\varepsilon^i{}_{jk} = 0$ if two of the indices i, j, k, are equal, and $\varepsilon^i{}_{jk}$ is the signature of the permutation $(123) \mapsto (ijk)$ if the three indices i, j, k, are distinct. However, not only does the formula requires to only use an orthonormal basis, but it must also have the same orientation than the canonical basis, i.e. the matrix P for changing basis does not just belong to the *orthogonal group* $\mathbb{O}(3)$, but to the *special orthogonal group* $S\mathbb{O}(3)$ of such orthogonal matrices having determinant $+1$ (called *rotations*).

Additional footnotes: Planck.[3]

[2] Albert Einstein, German-born physicist, 1879–1955. He received the Nobel Prize in Physics in 1921, for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect. He worked in Bern, in Zürich, Switzerland, in Prague, now capital of the Czech Republic, at ETH (Eidgenössische Technische Hochschule), Zürich, Switzerland, in Berlin, Germany, and at IAS (Institute for Advanced Study), Princeton, NJ. The Max Planck Institute for Gravitational Physics in Potsdam, Germany, is named after him, the Albert Einstein Institute.

[3] Max Karl Ernst Ludwig Planck, German physicist, 1858–1947. He received the Nobel Prize in Physics in 1918, in recognition of the services he rendered to the advancement of Physics by his discovery of energy quanta. He worked in Kiel and in Berlin, Germany. There is a Max Planck Society for the Advancement of the Sciences, which promotes research in many institutes, mostly in Germany (I spent my sabbatical year 1997–1998 at the Max Planck Institute for Mathematics in the Sciences in Leipzig, Germany).