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

7- Monday January 30, 2012.

Remark 7.1: The Legendre polynomials, which are the orthogonal polynomials P_n corresponding to $w(x) = 1$ on $[-1,+1]$, are given by a formula found in 1816 by RODRIGUES,[1] in 1824 by IVORY,[2] and in 1827 by JACOBI, and it is now called Rodrigues's formula:[3] $P_n = \frac{n!}{(2n)!}\frac{d^n[(x^2-1)^n]}{dx^n}$, although one may use the normalization $Q_n = \frac{1}{2^n n!}\frac{d^n[(x^2-1)^n]}{dx^n}$ in order to have $Q_n(\pm 1) = (\pm 1)^n$. Verifying Rogrigues's formula means checking $\int_{-1}^{+1}\frac{d^n[(x^2-1)^n]}{dx^n}\,q\,dx = 0$ for all $q \in \mathcal{P}_{n-1}[x]$: by n integration by parts one has $\int_{-1}^{+1}\frac{d^n[(x^2-1)^n]}{dx^n}\,q\,dx = (-1)^n\int_{-1}^{+1}(x^2-1)^n\frac{d^n q}{dx^n}\,dx$, since no term in ± 1 appears, because the derivatives of $(x^2-1)^n$ of order up to $n-1$ have a term x^2-1 as factor, hence they vanish at ± 1; since the nth derivative of a polynomial in $\mathcal{P}_{n-1}[x]$ is 0, one deduces that the integral is 0.

It also tells that for $q = x^n$ the integral is $n!\int_{-1}^{+1}(1-x^2)^n\,dx$, which by taking $x = \cos\theta$ is $2n!\,I_{2n+1}$ with $I_m = \int_0^{\pi/2}\sin^m\theta\,d\theta$, which one computes easily.[4]

Remark 7.2: LEGENDRE introduced the Legendre polynomials P_n in a work on celestial mechanics, probably in relation with writing the Laplacian $\Delta = \sum_{i=1}^3\frac{\partial^2}{\partial x_i^2}$ in spherical coordinates: the Legendre polynomials P_n then appear to be the eigenvectors of the operator A acting on the space of polynomials $\mathbb{R}[x]$ by $A\,P = -\frac{d}{dx}\big((1-x^2)\,\frac{dP}{dx}\big)$.

If one knows that $A\,P_n = \lambda_n P_n$ for some $\lambda_n \in \mathbb{R}$, then by identifying the coefficients of x^n on both sides, one finds that $\lambda_n = n\,(n+1)$ for all $n \in \mathbb{N}$.

One reason why P_n is necessarily an eigenvector of A is that A maps $\mathcal{P}_n[x]$ into itself for all $n \in \mathbb{N}$, and that A is a symmetric operator for the scalar product $(f,g) = \int_{-1}^{+1} f(x)\,g(x)\,dx$ on $\mathbb{R}[x]$, a notion which is defined below, and that an important result is that any symmetric operator on a finite dimensional Euclidean space is diagonalizable on an orthogonal basis of eigenvectors. Since A maps $\mathcal{P}_0[x]$ into itself, P_0 has to be an eigenvector (and it is actually spanning the kernel of A); then, because A maps $\mathcal{P}_1[x]$ into itself, it must have an orthogonal basis of eigenvectors, but the only vector orthogonal to P_0 is P_1, which then has to be an eigenvector of A; one then repeats the argument for $\mathcal{P}_2[x]$, and the only vector orthogonal to P_0 and P_1 is P_2, and so on.

Definition 7.3: If V, W are two E-vector spaces, and A is a linear mapping from V into W, then for $w_* \in W^*$ the mapping $v \mapsto \langle A\,v, w_*\rangle = w_*(A\,v)$ defines an element $A^T w_* \in V^*$, so that $\langle A\,v, w_*\rangle = \langle v, A^T w_*\rangle$; the mapping A^T is linear from W^* into V^* and is called the *transposed* of A.

Remark 7.4: If V_1, V_2, V_3 are three E-vector spaces, $A \in L(V_1, V_2)$ and $B \in L(V_2, V_3$, then $B\,A \in L(V_1, V_3)$ and one has $(B\,A)^T = A^T B^T$, since for all $v \in V_1, w_* \in V_3^*$, one has $\langle (B\,A)\,v, w_*\rangle = \langle A\,v, B^T w_*\rangle = \langle v, A^T(B^T w_*)\rangle$.

If $V_2 = V_1$ and $A = I$, then $A^T = I$, but one should not be misled by the notation, since one writes I for the identity on any vector field V, but for more precision, one writes id_X for the identity mapping on (any set) X. If $A \in L(V_1, V_2)$ is invertible, so that $A^{-1} \in L(V_2, V_1)$, then $(A^{-1})^T = (A^T)^{-1}$ (which one sometimes writes A^{-T}), since $A^{-1}A = id_{V_1}$ gives $A^T(A^{-1})^T = id_{V_1^*}$ and $A\,A^{-1} = id_{V_2}$ gives $(A^{-1})^T A^T = id_{V_2^*}$.

Remark 7.5: In order to be able to compare A and A^T, they should have the same domain of definition, so that one needs to consider $A \in L(W^*, W)$ for a vector space W, hence $A^T \in L(W^*, W^{**})$, and one uses the

[1] Benjamin Olinde RODRIGUES, French banker and mathematician, 1795–1851. Rodrigues's formula is named after him. In 1840 he published a result on transformation groups, which amounted to a discovery of the quaternions, three years before HAMILTON.

[2] Sir James IVORY, Scottish mathematician, 1765–1842.

[3] After being called the Ivory–Jacobi formula, it was HERMITE who pointed out in 1865 that RODRIGUES had discovered it first.

[4] After $I_0 = \frac{\pi}{2}$, and $I_1 = 1$, one computes I_m for $m \geq 2$ by induction: by integration by parts, $I_m = -\int_0^{\pi/2}\sin^{m-1}\theta\,d\cos\theta = \int_0^{\pi/2}\cos\theta\,d\sin^{m-1}\theta = (m-1)\int_0^{\pi/2}\cos^2\theta\,\sin^{m-2}\theta\,d\theta = (m-1)\,(I_{m-2} - I_m)$.

fact that $W \subset W^{**}$, with equality if W is finite dimensional. In the latter case, one says that A is symmetric if $A^T = A$, and that A is skew-symmetric if $A^T = -A$, and this is consistent with the next remark if one uses a basis $e_1, \ldots, e_n$ of W and the dual basis $e^1, \ldots, e^n$ of W^*.

In order to define $A^T A$, one needs to to have $W^* = W$, and this will mean considering an Euclidean space and using an orthogonal basis $e_1, \ldots, e_n$, because in that case $e^i = e_i$.[5]

Remark 7.6: If $e_k, k \in K$, is a basis of V and $f_\ell, \ell \in L$, is a basis of W, and $A \in L(V, W)$, then $A_{i,j}$ is the entry in row i and column j, and since column j contains the image of $A\, e_j$, it means that one has $A\, e_j = \sum_i A_{i,j} f_i$. Since $A^T \in L(W^*, V^*)$, one assumes that V and W are finite dimensional and one uses the dual basis $e^k, k \in K$, for V^*, and the dual basis $f^\ell, \ell \in L$, for W^*, and one has $A^T f^j = \sum_i A^T_{i,j} e^i$. For identifying $A^T_{i,j}$ one applies this equality to e_k, and the right side is $\sum_i A^T_{i,j} \langle e^i, e_k \rangle = A^T_{k,j}$, while the left side is $\langle A^T f^j, e_k \rangle = \langle f^j, A\, e_k \rangle = \langle f^j, \sum_i A_{i,k} f_i \rangle = A_{j,k}$, giving $A^T_{k,j} = A_{j,k}$ for all j, k in the corresponding sets of indices.

Remark 7.7: If V is a finite dimensional Euclidean space, there is a natural isomorphism of V^* onto V which to $v^* \in V^*$ associates the element $v_* \in V$ such that $v^*(v) = (v_*, v)$ for all $v \in V$; then, if one uses an orthonormal basis $e_1, \ldots, e_n$ of V, and the corresponding dual basis $e^1, \ldots, e^n$ of V^*, the choice $v^* = e^i$ gives $v_* = e_i$.

If $A \in L(V, V)$, then $A^T \in L(V, V)$ is defined by $(A\, u, v) = (u, A^T v)$ for all $u, v \in V$, and the same computation than in Remark 7.6 then shows that, if one uses an orthonormal basis, one has $A^T_{i,j} = A_{j,i}$ for $i, j = 1, \ldots, n$.

Lemma 7.8: Let V be an E-vector space, and $v_0^*, v_1^*, \ldots, v_m^* \in V^*$ (with $m \geq 1$) be such that $v_1^*(v) = \ldots = v_m^*(v) = 0$ imply $v_0^*(v) = 0$, then there exist $\lambda_1, \ldots, \lambda_m \in E$ such that $v_0^* = \lambda_1 v_1^* + \ldots + \lambda_m v_m^*$.
Proof: If $m = 1$, and $v_1^* \neq 0$,[6] one picks a non-zero w with $a = v_1^*(w) \neq 0$, and since $v_1^*\big(v - a^{-1} v_1^*(v)\, w\big) = 0$, one has $v_0^*\big(v - a^{-1} v_1^*(v)\, w\big) = 0$, i.e. one takes $\lambda_1 = a^{-1} v_0^*(w)$.

For $m \geq 2$, one uses induction on m: if $W = \{v \in V \mid v_1^*(v) = 0\}$, then the induction applied to W implies the existence of $\mu_2, \ldots, \mu_n \in E$ such that $z^* = v_0^* - \mu_2 v_2^* - \ldots - \mu_n v_n^*$ is 0 on W; the case $m = 1$ then implies that $z = \mu_1 v_1^*$ for some $\mu_1 \in E$.

Lemma 7.9: If V is a finite dimensional Euclidean space, and $A \in L(V, V)$ is *symmetric*, then there exists an orthonormal basis of eigenvectors of A.
Proof: If $e \in V$ is an eigenvector of A, i.e. $A\, e = \lambda\, e$ with $\lambda \in \mathbb{R}$, then A maps the orthogonal $e^\perp = \{v \in V \mid (v, e) = 0\}$ into itself, since $(v, e) = 0$ implies $(A\, v, e) = (v, A^T e) = (v, A\, e) = \lambda\, (v, e) = 0$. The lemma is proved by induction on the dimension, if one shows that A necessarily has a real eigenvalue, since one normalizes an eigenvector e, and then the problem is the same for the restriction of A to $e^\perp$, which is symmetric, since the property $(A\, u, v) = (A\, v, u)$ for all $u, v \in V$ stays true for every subspace W which A maps into itself.

One minimizes $F(u) = (A\, u, u)$ for u on the unit sphere $||u|| = 1$, and the minimum is attained at an element e by an argument of compactness, and then one wants to write that the derivative of F at e in any tangent direction is 0. For any v orthogonal to e, one defines $u(\varepsilon) = \frac{e + \varepsilon\, v}{||e + \varepsilon\, v||}$, and one writes that $F\big(u(\varepsilon)\big) \geq F(e)$ for ε small: since $F\big(u(\varepsilon)\big) = \frac{(A\, [e + \varepsilon\, v], [e + \varepsilon\, v])}{||e + \varepsilon\, v||^2}$, and one notices that $\big(A\, (e + \varepsilon\, v), (e + \varepsilon\, v)\big) = (A\, e, e) + 2\varepsilon\, (A\, e, v) + o(|\varepsilon|)$ by using the hypothesis that $A^T = A$, and that $||e + \varepsilon\, v||^2 = ||e||^2 + 2\varepsilon\, (e, v) + o(|\varepsilon|) = 1 + o(|\varepsilon|)$ by the assumption that $(e, v) = 0$, one deduces that $F\big(u(\varepsilon)\big) = F(e) + 2\varepsilon\, (A\, e, v) + o(|\varepsilon|)$, hence $(A\, e, v) = 0$ by taking ε small of either sign; since this is true for all v satisfying $(e, v) = 0$, one deduces by Lemma 7.8 that $A\, e = \lambda\, e$ for some $\lambda \in \mathbb{R}$.

---

[5] If V has dimension n, then V^* has dimension n, so that V and V^* are isomorphic, and for each basis $e_1, \ldots, e_n$ of V and each basis $f_1, \ldots, f_n$ of V^* there is an isomorphism of V onto V^* sending e_i to f_i, but it has no intrinsic character; one may impose to consider the dual basis $e^1, \ldots, e^n$ on V^*, and consider the isomorphism mapping e_i to e^i, but again it has no intrinsic character. In an Euclidean space, there is an intrinsic choice, since a unit vector e_1 determines the subspace spanned by the other elements of an orthonormal basis, which is $e_1^\perp$, the subspace orthogonal to e_1, and e^1 is then determined uniquely in terms of e_1, and this isomorphism of V onto V^* is called canonical, since it is independent of which basis is used.

[6] If $v_1^* = 0$, then all v satisfy $v_1^*(v) = 0$, so that $v_0^*(v) = 0$, hence $v_0^* = 0$, which is $\lambda_1 v_1^*$ for any λ_1.