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

6- Friday January 27, 2012.

Definition 6.1: An *Euclidean space* is an $\mathbb{R}$-vector space V equipped with a *symmetric* bilinear form B on $V\times V$ such that $B(v,v)>0$ for all non-zero $v\in V$; one usually writes (v,w) instead of $B(v,w)$, and it is called the *scalar product* of v and w, and one writes $||v||=\sqrt{B(v,v)}$, called the *norm* of v.[1] One says that $u,v\in V$ are *orthogonal* if $(u,v)=0$,[2] and one says that two non-zero vectors $u,v\in V$ make an *angle* $\theta\in[0,\pi]$ if $\cos\theta=\frac{(u,v)}{||u||\,||v||}$, which makes sense by Cauchy–Schwarz inequality, proved below.

Lemma 6.2: One has Cauchy–Schwarz inequality,[3] $|(u,v)|\le ||u||\,||v||$ for all $u,v\in V$. One has $||\lambda\,u||=|\lambda|\,||u||$ for all $\lambda\in\mathbb{R},u\in V$, and one has the *triangle inequality* $||u+v||\le||u||+||v||$ for all $u,v\in V$, so that d defined by $d(x,y)=||x-y||$ for all $x,y\in V$ is a *translation invariant* metric, i.e. it is a metric which satisfies $d(x+a,y+a)=d(x,y)$ for all $x,y,a\in V$.
Proof: One has $||u+v||^2=(u+v,u+v)=(u,u)+(u,v)+(v,u)+(v,v)$ by bilinearity, which is $||u||^2+2(u,v)+||v||^2$ by symmetry, so that the triangle inequality is equivalent to $(u,v)\le||u||\,||v||$ for all $u,v\in V$, which in turn is equivalent to the Cauchy-Schwarz inequality by using it for $\pm v$. Since $0\le||\lambda\,u+v||^2=\lambda^2||u||^2+2\lambda\,(u,v)+||v||^2$ for all $\lambda\in\mathbb{R}$, the discriminant $4(u,v)^2-4||u||^2||v||^2$ is ≤ 0, which is Cauchy-Schwarz inequality. One also has $||\lambda\,u||^2=(\lambda\,u,\lambda\,u)=\lambda^2||u||^2$ by bilinearity, i.e. $||\lambda\,u||=|\lambda|\,||u||$.

Remark 6.3: If V has infinite dimension, a natural question is about the metric space V being complete,[4] i.e. every Cauchy sequence converges, in which case V is called a *real Hilbert space*, while if V is not complete it is only called a (real) pre-Hilbert space.

Definition 6.4: An *orthogonal basis* $e_i,i\in I$, of an Euclidean space V is a basis of V such that $(e_i,e_j)=0$ whenever $i\ne j$; an *orthonormal basis* is an orthogonal basis which is made of vectors of norm 1, i.e. $(e_i,e_j)=\delta_{i,j}$ for all $i,j\in I$.

Remark 6.5: Out of any basis $f_1,\ldots,f_n$ of a finite-dimensional Euclidean space,[5] one can create an orthogonal basis by what is called the *Gram-Schmidt orthogonalization process*,[6,7] although it was used in

[1] A semi-norm on an $\mathbb{R}$-vector space or a $\mathbb{C}$-vector space V is a mapping q from V into $[0,\infty)$ satisfying $q(\lambda\,v)=|\lambda|\,q(v)$ for all $v\in V$ and all scalars λ, and $q(v+w)\le q(v)+q(w)$ for all $v,w\in V$; it is a norm if $q(v)>0$ for all non-zero $v\in V$.

[2] Without an Euclidean structure, one uses the term orthogonal between an element $v\in V$ and an element $v_*\in V^*$, the dual of V, to mean $v_*(v)=0$. In order to distinguish from a possible scalar product, one uses the notation $\langle v,v_*\rangle$ or $\langle v_*,v\rangle$, using the fact that $V\subset V^{**}=(V^*)^*$.

[3] This inequality should also be named after BUNYAKOVSKY.

[4] A metric space (X,d), is a set X equipped with a metric d, i.e. a mapping from $X\times X$ into $[0,\infty)$ such that $d(y,x)=d(x,y)$ for all $x,y\in X$, $d(x,y)=0$ if and only if $y=x$, and $d(x,z)\le d(x,y)+d(y,z)$ for all $x,y,z\in X$. If (X_1,d_1) and (X_2,d_2) are metric spaces, a mapping f from X_1 into X_2 is called an isometry if $d_2\big(f(a),f(b)\big)=d_1(a,b)$ for all $a,b\in X_1$. A sequence x_n converges to x_∞ (then defined in a unique way) if $d(x_n,x_\infty)\to 0$ as n tends to ∞, and a Cauchy sequence is any sequence x_n such that $d(x_m,x_n)\to 0$ as m and n tend to ∞. A metric space is said to be complete if every Cauchy sequence converges. For every metric space (X,d) there is a complete metric space $(\widehat{X},\widehat{d})$ and an isometry j from X into $\widehat{X}$ such that $j(X)$ is dense in $\widehat{X}$, i.e. every element of $\widehat{X}$ is the limit of a sequence $j(x_n)$ for a sequence $x_n\in X$: applying this result to $\mathbb{Q}$ (with the usual metric $|x-y|$) produces a completion $\widehat{\mathbb{Q}}$ isometric to $\mathbb{R}$, but $\mathbb{R}$ has to be constructed before proving the result, since its proof uses the fact that $\mathbb{R}$ is complete.

[5] The orthogonalization process also applies for a basis $f_n,n\in\mathbb{N}$ of an Euclidean space whose dimension is infinite but countable.

[6] Jorgen Petersen GRAM, Danish mathematician, 1850–1916. He lived in Copenhagen, Denmark. The Gram–Schmidt orthogonalization process is partly named after him, although CAUCHY used it in 1836, and it seems to be a result of LAPLACE.

[7] Erhard SCHMIDT, German mathematician, 1876–1959. He worked in Bonn, Germany, in Zürich, Switzer-

1836 by CAUCHY, and seems to have been devised earlier by LAPLACE. One defines $e_1 = f_1$, and for $i > 1$ one defines $f'_i = f_i - \alpha_i e_1$, where α_i is chosen so that f'_i is orthogonal to e_1, i.e. $\alpha_i ||f_1||^2 = (f_i, f_1)$, and one then repeats the process for $f'_1, \ldots, f'_n$ (for the Euclidean space spanned by these vectors in V). At the end of the process, one has obtained an orthogonal basis where for each i one has $e_i = f_i - \sum_{j<i} \lambda_{i,j} f_j$ for some scalars $\lambda_{i,j}$ with $1 \le j < i \le n$. Of course, one creates an orthonormal basis by dividing e_i by $||e_i||$.

Lemma 6.6: (orthogonal polynomials) If w is a non-negative continuous function on $[\alpha, \beta] \subset \mathbb{R}$ (with $\alpha < \beta$) which is not identically 0, and $P_0 = 1, P_1, \ldots, P_n, \ldots$ are the list of (monic) polynomials obtained from $1, \ldots, x^n, \ldots$ by the Gram–Schmidt orthogonalization process for the scalar product $(P, Q) = \int_\alpha^\beta P(x)\, Q(x)\, w(x)\, dx$, then P_n has n distinct roots in (α, β).
Proof: If the roots of P_n in (α, β) were $a_1, \ldots, a_\ell$ with multiplicity $m_1, \ldots, m_\ell$ and $\ell < n$ (and $m_1 + \ldots + m_\ell \le n$), then taking for Q the product of $(x - a_j)$ for those j for which m_j is odd, the product $P\,Q$ would not change sign at $a_1, \ldots, a_\ell$, hence $P\,Q$ having a constant sign on (α, β) the integral of $P\,Q\,w$ could not be 0, as it should be since the degree of Q is $\le \ell \le n - 1$.

Lemma 6.7: (Gauss's quadrature formula) If w is a non-negative continuous function on $[\alpha, \beta] \subset \mathbb{R}$ (with $\alpha < \beta$) which is not identically 0, there exists a unique quadrature formula with n (distinct) points $a_1, \ldots, a_n$ in (α, β) and weights $w_1, \ldots, w_n$, for which the quadrature formula of approximating $\int_\alpha^\beta P(x)\, w(x)\, dx$ by $\sum_{i=1}^n w_i P(a_i)$ is exact on $\mathcal{P}_{2n-1}[x]$. This quadrature formula has positive weights, and $a_1, \ldots, a_n$ are the (n distinct zeros) of the (monic) polynomial P_n of degree n obtained from $1, \ldots, x^n, \ldots$ by the Gram–Schmidt orthogonalization process for the scalar product $(P, Q) = \int_\alpha^\beta P(x)\, Q(x)\, w(x)\, dx$.
Proof: One defines $\Pi_n = \prod_{i=1}^n (x - a_i)$. If a quadrature formula with points $a_1, \ldots, a_n$ is exact on $\mathcal{P}_{2n-1}[x]$, then one must have $\int_\alpha^\beta \Pi_n(x)\, x^k w(x)\, dx = 0$ for $k = 0, \ldots, n-1$ since $P = \Pi_n\, x^k$ has degree $\le 2n - 1$, hence the integral is given by the quadrature formula, which gives 0: since Π_n is monic and orthogonal to $\mathcal{P}_{n-1}[x]$, it is equal to P_n, which has n distinct zeros in (α, β) by Lemma 6.6, and $a_1, \ldots, a_n$ must then be the zeros of P_n. Let $w_1, \ldots, w_n$ be the uniquely defined weights giving a formula exact on $\mathcal{P}_{n-1}[x]$. For $P \in \mathcal{P}_{2n-1}[x]$, the Euclidean division of P but P_n gives $P = P_n\, q + r$ with $q, r \in \mathcal{P}_{n-1}[x]$, and since the quadrature formula is exact for $P_n q$ by the choice of points a_n and exact on r by the choice of weights, the formula is then exact for P. Notice that no formula can be exact on $\mathcal{P}_{2n}[x]$ since the integral of P_n^2 is > 0, while the quadrature formula gives 0. For showing that $w_i > 0$, one applies the formula to $Q_i = \prod_{j \ne i} (x - a_j)^2 \in \mathcal{P}_{2n-2}[x]$, and $\int_\alpha^\beta Q(x)\, w(x)\, dx > 0$, and is equal to the result of the quadrature formula, which is $w_i \prod_{j \ne i} (a_i - a_j)^2$.

Remark 6.8: The results are also true in the case of an infinite interval, if w decays fast enough at infinity, and for example, the case $w(x) = e^{-x}$ on $[0, \infty)$ leads to the Laguerre polynomials,[8] whose zeros are called the Gauss–Laguerre points, and the case $w(x) = e^{-x^2}$ on $(-\infty, +\infty)$ leads to the Hermite polynomials, whose zeros are called the Gauss–Hermite points.

The case $w = 1$ on $[-1, +1]$ leads to the Legendre polynomials, whose zeros are called the Gauss–Legendre points.

The case $w = 1$ on $[-1, +1]$ when one imposes $a_1 = -1$ and $a_n = +1$ leads to quadrature points called the Gauss–Lobatto points.[9]

---

land, in Erlangen, Germany, in Breslau (then in Germany, now Wrocław, Poland), and in Berlin, Germany. The Gram–Schmidt orthogonalization process is partly named after him, although it was used by CAUCHY in 1836, and seems to be a result of LAPLACE. Hilbert–Schmidt operators are partly named after him.

[8] Edmond Nicolas LAGUERRE, French mathematician, 1834–1886.

[9] Rehuel LOBATTO, Dutch mathematician, 1797–1866. He worked in Delft, The Netherlands. The Gauss–Lobatto quadrature method is named after him.