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

8- Wednesday February 1, 2012.

Definition 8.1: If V is an E-vector space, a *quadratic form* on V is a mapping Q from V into E such that there exists a bilinear form B on $V\times V$ (into E) such that $Q(v)=B(v,v)$ for all $v\in V$.[1]

Remark 8.2: If $char(E)\neq 2$ and V is an E-vector space, then every quadratic form Q on V can be written as $B_s(x,x)$ with a symmetric bilinear form B_s on $V\times V$. Indeed, one just has to define $B_s(x,y)=2^{-1}\big(B(x,y)+B(y,x)\big)$. If V has finite dimension n, then it means that $Q(x)=\sum_{i,j=1}^n A_{i,j}x_ix_j$ for a symmetric $n\times n$ matrix A (with entries in E).

If $char(E)=2$, and V has dimension >1, then the result is not true, since V contains a copy of E^2 and on E^2, $Q(x)=x_1x_2$ cannot be written as $B_s(x,x)$, because $B_s(x,y)=\sum_{i,j=1}^2 a_{i,j}x_iy_j$ with $a_{1,2}=a_{2,1}$ implies $B_s(x,x)=a_{1,1}x_1^2+a_{2,2}x_2^2$.

Lemma 8.3: (Gauss's decomposition theorem) If $char(E)\neq 2$ and V is an n-dimensional E-vector space, then every quadratic form Q on V can be written as $Q(x)=\sum_{j=1}^n \kappa_j L_j^2(x)$, where $\kappa_1,\ldots,\kappa_n\in E$, and $L_1,\ldots,L_n$ are linearly independent linear forms (i.e. elements of V^*).[2]
Proof: One uses an induction on n, and the result is clear if $n=1$. One assumes then that $n\geq 2$ and that the result has been proved if the dimension of the space is at most $n-1$, and one uses a basis of V, so that $Q(x)=\sum_{i=1}^n a_ix_i^2+\sum_{i<j}b_{i,j}x_ix_j$, and one may assume that all x_i appear explicitly, since if it not the case the induction hypothesis applies.

If one of the coefficients a_i is $\neq 0$, one defines $L_i(x)=x_i+2^{-1}a_i^{-1}\sum_{j\neq i}b_{i,j}x_j$, so that $Q(x)=a_iL_i(x)^2+Q^*(x')$ where x' denotes the vector with components x_k for $k\neq i$; by the induction hypothesis, Q^* is a combination of $n-1$ squares of linearly independent linear forms, and since they do not use the variable x_i while L_i does, one obtains n linearly independent linear forms by adjoining L_i.

If all coefficients a_i are 0, there exists a coefficient $b_{i,j}\neq 0$ with $i\neq j$, and one defines $\ell_i(x)=x_i+b_{i,j}^{-1}\sum_{k\neq i,j}b_{j,k}x_k$ and $\ell_j(x)=x_j+b_{i,j}^{-1}\sum_{k\neq i,j}b_{i,k}x_k$, so that $Q(x)=b_{i,j}\ell_i(x)\,\ell_j(x)+Q^{**}(x'')$ where x'' denotes the vector with components x_k for $k\neq i,j$; by the induction hypothesis, Q^{**} is a combination of $n-2$ squares of linearly independent linear forms, and since they do not use the variables x_i or x_j while ℓ_i uses x_i but not x_j, and ℓ_j uses x_j but not x_i, one obtains n linearly independent linear forms by adjoining $\ell_i+\ell_j$ and $\ell_i-\ell_j$, noticing that $b_{i,j}\ell_i\ell_j=b_{i,j}4^{-1}\big((\ell_i+\ell_j)^2-(\ell_i-\ell_j)^2\big)$.

Definition 8.4: A quadratic form Q on an $\mathbb{R}$-vector space V is said to be *positive definite* if $Q(x)>0$ for all non-zero $x\in V$ (negative definite if $Q(x)<0$ for all non-zero $x\in V$) and *positive semi-definite* if $Q(x)\geq 0$ for all $x\in V$ (negative semi-definite if $Q(x)\leq 0$ for all $x\in V$).

Remark 8.5: If V is an n-dimensional Euclidean space, and Q is a quadratic form on V, it can be written as $Q(x)=\sum_{i,j=1}^n A_{i,j}x_ix_j$ for a real symmetric $n\times n$ matrix A, and since there exists an orthonormal basis of eigenvectors $e_i, i=1,\ldots,n$ of A, with eigenvalues $\lambda_1,\ldots,\lambda_n$, one has $Q(x)=\sum_{i=1}^n\lambda_i(e_i,x)^2$: one deduces that Q is positive definite if and only if $\lambda_i>0$ for all i, and that it is positive semi-definite if and only if $\lambda_i\geq 0$ for all i.

Lemma 8.6: (Sylvester's law of inertia) If V is an n-dimensional Euclidean space, and Q is a quadratic form on V, then all decompositions $Q(x)=\sum_{i=1}^n\kappa_iL_i(x)^2$ with $L_1,\ldots,L_n$ linearly independent have the same number of positive κ_i (corresponding to $i\in I$), the same number of zero κ_j (corresponding to $j\in J$), and the same number of negative κ_k (corresponding to $k\in K$, so that I,J,K is a partition of $\{1,\ldots,n\}$).

[1] If $e_i, i\in I$, is a basis of V, one may put a total order on I, and then Q can be written as $\sum_{i\leq j}q_{i,j}v_iv_j$ (where $v=\sum_i v_ie_i$). If V has dimension n, it is then any polynomial function in $v_1,\ldots,v_n$ of degree ≤ 2 which has no terms of degree 0 or 1.

[2] If E is a field in which every element is a square, then one could replace L_j by ℓ_jL_j with $\ell_j^2=\kappa_j$, and write Q as a sum of squares of linear forms, but since some κ_j may be 0, one must change the statement of independence, and say that the non-zero L_j are linearly independent.

Proof: One write $Q(x)=(A\,x,x)$ for a symmetric A, and one denotes V_+ the direct sum of the eigen-spaces of A with positive eigenvalues and d_+ its dimension, V_0 the kernel of A and d_0 its dimension, and V_- the direct sum of the eigen-spaces of A with negative eigenvalues and d_- its dimension. Let $W_+=\{x\in V\mid L_j(x)=0, j\in J, L_k(x)=0, k\in K\}$ (having dimension $|I|$), $W_0=\{x\in V\mid L_i(x)=0, i\in I, L_k(x)=0, k\in K\}$ (having dimension $|J|$), and $W_-=\{x\in V\mid L_i(x)=0, i\in I, L_j(x)=0, j\in J\}$ (having dimension $|K|$). On W_+, the restriction of Q is positive definite, so that W_+ cannot intersect $V_0\oplus V_-$ on which Q is negative semi-definite, hence $|I|+d_0+d_-\le n$, i.e. $|I|\le d_+$. On $W_+\oplus W_0$, the restriction of Q is positive semi-definite, so that $W_+\oplus W_0$ cannot intersect V_- on which Q is negative definite, hence $|I|+|J|\le d_++d_0$. On W_-, the restriction of Q is negative definite, so that W_- cannot intersect $V_+\oplus V_0$ on which Q is positive semi-definite, hence $|K|\le d_-$. Since $|I|+|J|+|K|=n=d_++d_0+d_-$, one deduces that $|I|=d_+$, $|J|=d_0$, and $|K|=d_-$.

Definition 8.7: If V_1,V_2,W are $\mathbb{C}$-vector spaces, a mapping f from V_1 into W is said to be *anti-linear* if $f(x+y)=f(x)+f(y)$ for all $x,y\in V_1$, and $f(\lambda\,x)=\overline{\lambda}\,f(x)$ for all $x\in V,\lambda\in\mathbb{C}$. A mapping g from $V_1\times V_2$ into W is said to be *sesqui-linear* if $x\mapsto g(x,y)$ is linear from V_1 into W for all $y\in V_2$, and $y\mapsto g(x,y)$ is anti-linear from V_2 into W for all $x\in V_1$. A sesqui-linear mapping h from $V_1\times V_1$ into W is said to be *Hermitian symmetric* if $h(y,x)=\overline{h(x,y)}$ for all $x,y\in V_1$.[3]

An *Hermitian space* V is a $\mathbb{C}$-vector space equipped with a Hermitian symmetric *scalar product* $B(x,y)$, usually simply denoted (x,y), such that $(x,x)>0$ for all non-zero $x\in V$, and the norm of $v\in V$ is $||v||=\sqrt{(v,v)}$. One says that x is orthogonal to y if $(x,y)=0$; an orthogonal basis is a basis $e_i,i\in I$, such that $(e_i,e_j)=0$ whenever $i\ne j$; an orthonormal basis is a basis $e_i,i\in I$, such that $(e_i,e_j)=\delta_{i,j}$ for all $i,j\in I$.

Remark 8.8: As for an Euclidean space, one has $|(x,y)|\le||x||\,||y||$ for all $x,y\in V$,[4] since if $(x,y)=r\,e^{i\,\theta}$, so that $(y,x)=r\,e^{-i\,\theta}$, then for $t\in\mathbb{R}$ one has $0\le(x+t\,e^{i\,\theta}y,x+t\,e^{i\,\theta}y)=||x||^2+2t\,r+t^2||y||^2$, and because it is true for all $t\in\mathbb{R}$, one deduces that $r^2\le||x||^2||y||^2$. As a consequence, $d(x,y)=||x-y||$ defines a (translation invariant) metric, since the triangle inequality means $||a+b||\le||a||+||b||$, and $||a+b||^2=||a||^2+||b||^2+2\Re(a,b)$ while $(||a||+||b||)^2=||a||^2+||b||^2+2||a||\,||b||$.

Remark 8.9: An Hermitian space is also called a (complex) pre-Hilbert space, and it is called a Hilbert space if the space is complete, i.e. if every Cauchy sequence converges.[5]

One should pay attention to a difference in notation with physicists, who use DIRAC's notation: mathematicians write $(x,y)\in\mathbb{C}$, which is linear in x and anti-linear in y, while physicists write $\langle b\,|\,a\rangle$, which is linear in a and anti-linear in b; it means that the *ket* $|\,a\rangle$ is an element of a Hilbert space H, while the *bra* $\langle b\,|$ is an element of the dual H'. Then the notation $|\,a\rangle\,\langle b\,|$ denotes a linear operator from H into itself, which mathematicians write $a\otimes b$ (and which is the mapping $x\mapsto(b,x)\,a$).

[3] If h is Hermitian symmetric, one has $h(x,x)\in\mathbb{R}$ for all $x\in V_1$. Conversely, a sesqui-linear mapping h from $V_1\times V_1$ into W which satisfies $h(x,x)\in\mathbb{R}$ for all $x\in V_1$ is Hermitian symmetric: for all $x,y\in V_1$, one has $h(x,y)+h(y,x)=h(x+y,x+y)-h(x,x)-h(y,y)\in\mathbb{R}$, and replacing x by $\lambda\,x$ with $\lambda\in\mathbb{C}$, one deduces that $\lambda\,h(x,y)+\overline{\lambda}\,h(y,x)\in\mathbb{R}$ for all $\lambda\in\mathbb{C}$, which implies $h(y,x)=\overline{h(x,y)}$.

[4] However, $|\cdot|$ denotes the modulus of a complex number, since $(x,y)\in\mathbb{C}$.

[5] The space ℓ^0 of complex sequences with only a finite number of non-zero terms is a Hermitian space with the scalar product $(x,y)=\sum_n x_n\overline{y_n}$, but it is not complete, and its completion is isometric to ℓ^2, the space of square integrable complex sequences, i.e. $||x||^2=\sum_{n=1}^\infty|x_n|^2<+\infty$, with the scalar product $(x,y)=\sum_{n=1}^\infty x_n\overline{y_n}$. The space $C([0,1])$ of continuous complex functions on $[0,1]$ with the scalar product $(u,v)=\int_0^1 u(x)\,\overline{v(x)}\,dx$ (where the integral is the Riemann integral), is a (complex) pre-Hilbert space but it is not complete; however, describing its completion requires inventing the Lebesgue integral, since the completion is isometric to $L^2\big((0,1)\big)$, the space of (equivalence classes) of square integrable complex functions, i.e. $||u||^2=\int_0^1|u(x)|^2\,dx<+\infty$, but where the integral is the Lebesgue integral.