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

4- Monday January 23, 2012.

Remark 4.1: As Marcel BERGER wrote it,[1] "an affine space is nothing more than a vector space whose origin we try to forget about, by adding translations to the linear maps". In other words, an *affine space* A comes with an underlying E-vector space V, and there is a mapping from $A\times V$ onto A, denoted $(a,v)\mapsto a+v$, such that

i) $a+0=a$ for all $a\in A$,

ii) $(a+v)+w=a+(v+w)$ for all $a\in A$ and all $v,w\in V$,

iii) for all $v\in V$, the mapping $a\mapsto a+v$, called the *translation by* v is a bijection of A.

Said otherwise, V acts (as an Abelian group) on A by translation, so that the only v acting with a fixed point is 0, and there is only one orbit.

An affine space is said to have dimension n if the underlying vector space V has dimension n; an affine subspace is called a *line* if it has dimension 1, a *plane* if it has dimension 2, an *hyperplane* if it has dimension $n-1$ (or *co-dimension* 1, the co-dimension being n minus the dimension).

Remark 4.2: In an affine space A, the finite sum $\sum_{i\in I}\lambda_i a_i$ (with $\lambda_i\in E, a_i\in A$ for all $i\in I$) has a meaning in only two cases, if either $\sum_i\lambda_i=1$ and it gives a point in A, or $\sum_i\lambda_i=0$ and it gives a vector in V. In the case where $\sum_i\lambda_i=1$, then for any $b\in A$ one may write $\sum_{i\in I}\lambda_i a_i=b+\sum_{i\in I}\lambda_i(a_i-b)$, which is a point in A, independent of the choice of b, which is called the *barycenter* of the points $a_i, i\in I$ for the weights $\lambda_i, i\in I$. In the case where $\sum_i\mu_i=0$, then for any $b\in A$ one may write $\sum_{i\in I}\mu_i a_i=\sum_{i\in I}\mu_i(a_i-b)$, which is a vector in V, independent of the choice of b.

In an affine space A of dimension n, one may decide to take any point M as origin, and any basis $e_1,\ldots,e_n$ of V, and then $(x_1,\ldots,x_n)\mapsto M+\sum_{i=1}^n x_i e_i$ defines a bijection from E^n onto A. One may then talk about a polynomial in A by considering a polynomial in $E[x_1,\ldots,x_n]$, and changing origin in A is related to using Taylor's expansion formula.

Lemma 4.3: If a polynomial $P\in E[x_1,x_2]$ has restriction 0 on a line of equation $Q=0$ (with $Q=\alpha\,x_1+\beta\,x_2+\gamma$ with $\alpha,\beta,\gamma\in E$ and α and β not both 0), and E has at least $degree(P)+1$ elements, then $P=Q\,P_1$ for some $P_1\in E[x_1,x_2]$.

Proof: Since α or β is non-zero, there exists a point M on the line $Q=0$, and a vector $e_1\neq 0$ with $M+e_1$ on the line $Q=0$, and one chooses $e_2\neq 0$ such that e_1,e_2 is a basis of V. In this new set of coordinates one has $P=\sum_{i,j\geq 0}a_{i,j}x_1^i x_2^j$, and the equation of Q is $x_2=0$. By hypothesis the polynomial $P_0=\sum_{i\geq 0}a_{i,0}x_1^i\in E[x_1]$ vanishes at every element of E, and since E has more elements than $degree(P_0)$, one has $P_0=0$, so that $a_{i,0}=0$ for all i; said otherwise, $a_{i,j}\neq 0$ implies $j\geq 1$, so that $P_0=\sum_{i\geq 0,j\geq 1}a_{i,j}x_1^i x_2^j=x_2P_1$ with $P_1=\sum_{i\geq 0,j\geq 1}a_{i,j}x_1^i x_2^{j-1}$.

Remark 4.4: The condition that E has at least $degree(P)+1$ elements is necessary for the factorization to hold: if E is a finite field with $q=p^k$ elements then x^q-x vanishes at all elements of E, so that for $P=x_1^q-x_1$, E has $degree(P_0)$ elements, and P is 0 at any point of a line Q, but unless the line has an equation $x_1=c$ for some $c\in E$, P is not a multiple of Q.

Remark 4.5: If $E=\mathbb{R}$, and instead of a line one considers $Q=x_1^2+x_2^2$, so that the zero set of Q is the origin, it is not true that one can factor $x_1^2+x_2^2$ for all polynomials P vanishing at the origin.

However, in the case $E=\mathbb{C}$, one can factor $x_1^2+x_2^2$ for all polynomials P vanishing on the zero set of $Q=x_1^2+x_2^2$, since $x_1^2+x_2^2=(x_1+i\,x_2)\,(x_1-i\,x_2)$ and the zero set of Q is two intersecting lines, and one can apply Lemma 4.3 twice.

[1] Marcel BERGER, French mathematician, born in 1927. He worked at Université Denis Diderot (Paris VII) in Paris, France, and he was for a few years director of IHES (Institut des Hautes Études Scientifiques) in Bures sur Yvette, France.

If E is algebraically closed, and $P, Q \in E[x_1, \ldots, x_n]$, it is a consequence of Hilbert's nullstellensatz that if $Q = 0$ implies $P = 0$, then a power of P is a multiple of Q.[2]

Lemma 4.6: Given a non-degenerate triangle with vertices a_1, a_2, a_3 in an affine plane (i.e. such that a_1, a_2, a_3 are not aligned), then for $\alpha_1, \alpha_2, \alpha_3 \in E$ there is a unique $P \in \mathcal{P}_1[x_1, x_2]$ such that $P(a_i) = \alpha_i$ for $i = 1, 2, 3$.
Proof: The linear mapping $P \mapsto \big(P(a_1), P(a_2), P(a_3)\big)$ from $\mathcal{P}_1[x_1, x_2]$ into E^3 is a bijection if and only if it is injective, i.e. $P(a_1) = P(a_2) = P(a_3) = 0$ implies $P = 0$. Since $P(a_1) = P(a_2) = 0$ implies that the restriction of P to the line L_{12} through a_1 and a_2 is 0, because it a polynomial of degree ≤ 1 with two distinct zeros, one has $P = Q_{12}\, P_1$ where $Q_{12} = 0$ is the equation of L_{12}. Then P_1 must be a constant, and writing $P(a_3) = 0$ shows that this constant is 0.

Remark 4.7: The corresponding finite element is called Courant's triangle,[3] but it may have been used before COURANT by SYNGE.[4]

If for $i = 1, 2, 3$, one denotes λ_i the particular interpolation polynomial in $\mathcal{P}_1[x_1, x_2]$ such that $\lambda_i(a_j) = \delta_{i,j}$ for $i, j = 1, 2, 3$, then one has $P = \sum_i P(a_i)\, \lambda_i$ for all $P \in \mathcal{P}_1[x_1, x_2]$, since both sides of the equality are polynomials in $\mathcal{P}_1[x_1, x_2]$ which take the same values at the three vertices of the triangle. In particular, one has $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $M = \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3$ for all points M (with $\lambda_1, \lambda_2, \lambda_3$ evaluated at M, of course), and $\lambda_1(M), \lambda_2(M), \lambda_3(M)$ are called the *barycentric coordinates* of M (with respect to a_1, a_2, a_3, of course). For example, if (in some coordinates) $a_1 = (0, 0), a_2 = (0, 1), a_3 = (1, 0)$, then $\lambda_1 = 1 - x_1 - x_2$, $\lambda_2 = x_2$, and $\lambda_3 = x_1$.

Lemma 4.8: One assumes that $char(E) \ne 2$. Given a_1, a_2, a_3 not aligned in an affine plane, and defining $a_{ij} = 2^{-1}(a_i + a_j)$ for $i < j$, then for $\alpha_1, \alpha_2, \alpha_3, \alpha_{12}, \alpha_{13}, \alpha_{23} \in E$ there is a unique $P \in \mathcal{P}_2[x_1, x_2]$ such that $P(a_i) = \alpha_i$ for $i = 1, 2, 3$, and $P(a_{ij}) = \alpha_{ij}$ for $i < j$.
Proof: The linear mapping $P \mapsto \big(P(a_1), P(a_2), P(a_3), P(a_{12}), P(a_{13}), P(a_{23})\big)$ from $\mathcal{P}_2[x_1, x_2]$ into E^6 is a bijection if and only if it is injective, i.e. $P(a_1) = P(a_2) = P(a_3) = P(a_{12}) = P(a_{13}) = P(a_{23}) = 0$ implies $P = 0$. Since $P(a_1) = P(a_{12}) = P(a_2) = 0$ implies that the restriction of P to the line L_{12} through a_1 and a_2 is 0, because it is a polynomial of degree ≤ 2 with three distinct zeros, one has $P = \lambda_3 Q$ (since $\lambda_3 = 0$ is the equation of L_{12}). Then $Q \in \mathcal{P}_1[x_1, x_2]$ is 0 at a_3, a_{13}, a_{23} which are not aligned, so that $Q = 0$ by Lemma 4.6, hence $P = 0$.

Remark 4.9: Using the barycentric coordinates $\lambda_1, \lambda_2, \lambda_3$, one can write explicitly what the interpolation polynomial P is: one has $P = \sum_i P(a_i)\, \lambda_i(2\lambda_i - 1) + \sum_{i<j} P(a_{ij}) 4\lambda_i\lambda_j$

Lemma 4.10: One assumes that $char(E) \ne 2, 3$. Given a_1, a_2, a_3 not aligned in an affine plane, and defining $a_{iij} = 3^{-1}(2a_i + a_j)$ and $a_{ijj} = 3^{-1}(a_i + 2a_j)$ for $i < j$, and $a_{123} = 3^{-1}(a_1 + a_2 + a_3)$, then for any list $\alpha_i \in E$ for all i, $\alpha_{iij}, \alpha_{ijj} \in E$ for all $i < j$, and $\alpha_{123} \in E$ there is a unique $P \in \mathcal{P}_3[x_1, x_2]$ such that $P(a_i) = \alpha_i$ for $i = 1, 2, 3$, $P(a_{iij}) = \alpha_{iij}$ and $P(a_{ijj}) = \alpha_{ijj}$ for $i < j$, and $P(a_{123}) = \alpha_{123}$.
Proof: The linear mapping from $\mathcal{P}_3[x_1, x_2]$ (which has dimension 10) into E^{10} associating to P the list of its values at the ten points, is a bijection if and only if it is injective, i.e. P vanishing at the ten points implies $P = 0$. Since $P(a_1) = P(a_{112}) = P(a_{122}) = P(a_2) = 0$ implies that the restriction of P to the line L_{12} through a_1 and a_2 is 0, because it is a polynomial of degree ≤ 3 with four distinct zeros, one has $P = \lambda_3 Q$. Then $Q \in \mathcal{P}_2[x_1, x_2]$ is 0 at the six remaining points, and $Q = 0$ by Lemma 4.8, hence $P = 0$.

Remark 4.11: Using the barycentric coordinates $\lambda_1, \lambda_2, \lambda_3$, one can write explicitly what the interpolation polynomial P is. The polynomial which is 1 at a_i and 0 at the nine other points is proportional to $\lambda_i(\lambda_i -$

---

[2] Since $E[x_1, \ldots, x_n]$ is an UFD, one can then deduce that P is a multiple of Q if $Q = r_1 \cdots r_m$ with $r_1, \ldots, r_m$ irreducible elements, and r_i and r_j are not associates for $i \ne j$.

[3] Richard COURANT, German-born mathematician, 1888–1972. He worked at Georg-August-Universität, Göttingen, Germany, and at NYU (New York University), New York, NY. The department of mathematics of NYU is named after him, the Courant Institute of Mathematical Sciences.

[4] John Lighton SYNGE, Irish mathematician, 1897–1995. He started and finished his career in Dublin, Ireland, but he also worked in Toronto, Canada, at OSU (Ohio State University), Columbus, OH, and he was from 1946 to 1948 the head of the mathematics department at Carnegie Tech (Carnegie Institute of Technology), now CMU (Carnegie Mellon University), Pittsburgh, PA.

$3^{-1})\,(\lambda_i - 2\cdot 3^{-1})$, and it is 1 for $\lambda_i = 1$, so that it is $2^{-1}\lambda_i(3\lambda_i - 1)\,(3\lambda_i - 2)$. The polynomial which is 1 at a_{iij} and 0 at the nine other points is proportional to $\lambda_i\lambda_j(\lambda_i - 3^{-1})$, and it is 1 for $\lambda_i = 2\cdot 3^{-1}$ and $\lambda_j = 3^{-1}$, so that it is $9\cdot 2^{-1}\lambda_i\lambda_j(3\lambda_i - 1)$; similarly, the polynomial which is 1 at a_{ijj} and 0 at the nine other points is $9\cdot 2^{-1}\lambda_i\lambda_j(3\lambda_j - 1)$. The polynomial which is 1 at a_{123} and 0 at the nine other points is $27\lambda_1\lambda_2\lambda_3$.

Remark 4.12: For discussing multi-dimensional Hermite interpolation polynomials, one uses the notation $DP(M)$ for the total derivative of P at M, defined by $DP(M)\cdot v = \sum_i \frac{\partial P}{\partial x_i} v_i$ for all v in the underlying vector space V, i.e. it is an element of V^* (and the partial derivatives are evaluated at M, of course).

In the case $V = \mathbb{R}^2$ with an Euclidean structure for defining a normal vector n to the edge through a_i and a_j, one uses $\frac{\partial P}{\partial n}$ for $DP(M)\cdot n$, and it is usually evaluated at the middle $M = a_{ij}$.

One uses $D^2P(M)$ for the second derivative, which belongs to $B_{sym}(V,V)$, the E-vector space of symmetric bilinear forms on $V\times V$ (into E), defined by $D^2P(M)\cdot(v,w) = \sum_{i,j}\frac{\partial^2 P}{\partial x_i\partial x_j} v_i w_j$ for all $v,w\in V^*$.

Lemma 4.13: One assumes that $char(E) \neq 3$. Given a_1,a_2,a_3 not aligned in an affine plane (with underlying vector space V), and defining $a_{123} = 3^{-1}(a_1+a_2+a_3)$, then for any list $\alpha_i\in E, \beta_i\in V^*$ for all i, $\alpha_{123}\in E$, there is a unique $P\in\mathcal{P}_3[x_1,x_2]$ such that $P(a_i) = \alpha_i$ and $DP(a_i) = \beta_i$ for $i = 1,2,3$, and $P(a_{123}) = \alpha_{123}$.
Proof: The linear mapping from $\mathcal{P}_3[x_1,x_2]$ (which has dimension 10) into E^{10} associating to P the list of the values $P(a_i)$ and $DP(a_i)$ at the three vertices, and $P(a_{123})$ is a bijection if and only if it is injective, i.e. all the ten quantities for P vanishing imply $P = 0$. Since $P(a_1) = DP(a_1)\cdot(a_2-a_1) = P(a_2) = DP(a_2)\cdot(a_1-a_2) = 0$ implies that the restriction of P to the line L_{12} through a_1 and a_2 is 0, because it is a polynomial of degree ≤ 3 with two distinct double zeros, one has $P = \lambda_3 Q$. Similarly, P is a multiple of λ_1 and of λ_2, so that $P = c\,\lambda_1\lambda_2\lambda_3$ for a constant c, and $P(a_{123}) = 0$ implie $c = 0$, hence $P = 0$.

Lemma 4.14: (Argyris)[5] One assumes that $E = \mathbb{R}$. Given a_1,a_2,a_3 not aligned in a real affine plane, and defining $a_{ij} = \frac{a_i+a_j}{2}$ for $i<j$, then for any list $\alpha_i\in\mathbb{R}, \beta_i\in V^*, \gamma_i\in B_{sym}(V,V)$ for all i, $\alpha_{ij}\in E$ for $i<j$, there is a unique $P\in\mathcal{P}_5[x_1,x_2]$ such that $P(a_i) = \alpha_i$, $DP(a_i) = \beta_i$, $D^2P(a_i) = \gamma_i$ for $i = 1,2,3$, and $\frac{\partial P}{\partial n}(a_{ij}) = \alpha_{ij}$ for $i<j$.
Proof: The linear mapping from $\mathcal{P}_5[x_1,x_2]$ (which has dimension 21) into $\mathbb{R}^{21}$ associating to P the list of the values $P(a_i)$, $DP(a_i)$, $D^2P(a_i)$ at the three vertices, and $\frac{\partial P}{\partial n}(a_{ij})$ at the three middle points is a bijection if and only if it is injective, i.e. all the twenty-one quantities for P vanishing imply $P = 0$. Since $P(a_1) = DP(a_1)\cdot(a_2-a_1) = D^2P(a_1)\cdot(a_2-a_1,a_2-a_1) = 0$ as well as $P(a_2) = DP(a_2)\cdot(a_1-a_2) = D^2P(a_2)\cdot(a_1-a_2,a_1-a_2) = 0$ implies that the restriction of P to the line L_{12} through a_1 and a_2 is 0, because it is a polynomial of degree ≤ 5 with two distinct triple zeros, one has $P = \lambda_3 Q$. This implies that the restriction of $\frac{\partial P}{\partial n}$ to the line L_{12} is $c\,Q$ (with $c = \frac{\partial\lambda_3}{\partial n}\neq 0$), and since one then deduces that $Q(a_1) = DQ(a_1)\cdot(a_2-a_1) = 0$, $Q(a_2) = DQ(a_2)\cdot(a_1-a_2) = 0$, and $Q(a_{12}) = 0$, then the restriction of Q to the line L_{12} is 0, because it is a polynomial of degree ≤ 4 with two distinct douple zeros and a distinct single zero, one has $Q = \lambda_3 R$, hence $P = \lambda_3^2 R$. This implies that P is a multiple of $\lambda_1^2\lambda_2^2\lambda_3^2$, which has degree 6, hence $P = 0$.

Remark 4.15: In Lemmas 4.6, 4.8, 4.10, and 4.13, the restriction of P to an edge of the triangle is defined by the degrees of freedom associated to this edge, either at the two vertices or at interior points of the edge. Consequently, in a triangulation, the piecewise polynomial interpolated function will be of class C^0 (continuous even at the interfaces).

In Lemma 4.14, the restriction of P and of $\frac{\partial P}{\partial n}$ to an edge of the triangle is defined by the degrees of freedom associated to this edge, either at the two vertices or at the middle of the edge. Consequently, in a triangulation, the piecewise polynomial interpolated function will be of class C^1 (continuously differentiable even at the interfaces).

Additional footnotes: Carathéodory.[6]

[5] John Hadji Argyris, Greek-born engineer, 1913–2004. He worked in London, England, and in Stuttgart, Germany. He was a nephew of Carathéodory.

[6] Constantin Carathéodory, German mathematician (of Greek origin), 1873–1950. He worked at Georg-August-Universität, Göttingen, in Bonn, in Hanover, Germany, in Breslau (then in Germany, now Wrocław, Poland), in Berlin, Germany. After World War I, he worked in Athens, Greece and in Smyrna (then in Greece, now Izmir, Turkey), and in München (Munich), Germany.