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

3- Friday January 20, 2012.

Remark 3.1: If V and W are E-vector spaces with V having finite dimension, then for $A \in L(V,W)$, one has $dim\big(ker(A)\big) + dim\big(im(A)\big) = dim(V)$: if $X = ker(A)$ and Y is a complement of X in V, then $dim(X) + dim(Y) = dim(V)$ (by adjoining a basis of X to a basis of Y, which makes a basis of V), and then the restriction $A|_Y$ of A to Y is injective, so that $A|_Y$ is an isomorphism of Y onto $im(A|_Y)$, hence $dim(Y) = dim\big(im(A|_Y)\big)$; then, $im(A|_Y) = im(A)$ because each $v \in V$ has a unique decomposition $v = x+y$ (with $x \in X, y \in Y$), which implies $A\,v = A\,y = A|_Y y$.

A simple application is that if $dim(W) = dim(V) < \infty$, then $A \in L(V,W)$ is injective if and only if it is surjective, and this means that a way to prove that $A\,v = w$ has a unique solution for all $w \in W$ is to prove that V and W have the same dimension, and that $A\,x = 0$ implies $x = 0$, which is often more easy than finding an explicit solution for $A\,v = w$. Sometimes, one may be able to check a formula which for each $w \in W$ gives one solution $v \in V$ of $A\,v = w$, and then (since one has proved that A is surjective, hence injective) such a solution is unique.

Lemma 3.2: (Lagrange's interpolation polynomial) Let $a_1, \ldots, a_m$ be distinct elements of a field E (with $m \geq 1$),[1] then for $\alpha_1, \ldots, \alpha_m \in E$ there exists a unique *interpolation polynomial* $P \in E[x]$ of degree $\leq m-1$ such that $P(a_i) = \alpha_i$ for $i = 1, \ldots, m$.
Proof: One takes for V the E-vector space $\mathcal{P}_{m-1}[x]$ of polynomials $P \in E[x]$ of degree $\leq m-1$; one has $dim(V) = m$, since a basis of V is $\{1, x, \ldots, x^{m-1}\}$. One takes $W = E^m$, so that $dim(W) = m$, and A is the linear mapping defined by $A(P) = \big(P(a_1), \ldots, P(a_m)\big)$, and one then checks that A is injective: indeed, $A(P) = 0$ means $P(a_i) = 0$ for $i = 1, \ldots, m$, so that P has m distinct zeros, and since $deg(P) < m$ one deduces that $P = 0$.

Remark 3.3: It is easy to write explicitly the interpolation polynomial, and it is what LAGRANGE must have done (in the case $E = \mathbb{R}$, I suppose). One has $P = \sum_{i=1}^m \alpha_i \Pi_i$, where, for $i = 1, \ldots, m$, Π_i is the particular interpolation polynomial satisfying $\Pi_i(a_j) = \delta_{i,j}$ for $j = 1, \ldots, m$. Then, since Π_i vanishes at all a_j for $j \neq i$, it must be a multiple of $Q_i = \prod_{j \neq i}(x - a_j)$, and since $deg(Q_i) = m-1$, one has $\Pi_i = c\,Q_i$ for a (non-zero) scalar c, and evaluating both sides at a_i gives $c = \big(Q_i(a_i)\big)^{-1} = \prod_{j \neq i}(a_i - a_j)^{-1}$.

Lemma 3.4: (Hermite's interpolation polynomial) Let $a_1, \ldots, a_m$ be distinct elements of a field E (with $m \geq 1$), then for $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m \in E$ there exists a unique interpolation polynomial $Q \in E[x]$ of degree $\leq 2m-1$ such that $Q(a_i) = \alpha_i$ and $Q'(a_i) = \beta_i$ for $i = 1, \ldots, m$.
Proof: One takes for V the E-vector space $\mathcal{P}_{2m-1}[x]$ of polynomials $P \in E[x]$ of degree $\leq 2m-1$; one has $dim(V) = 2m$, since a basis of V is $\{1, x, \ldots, x^{2m-1}\}$. One takes $W = E^{2m}$, so that $dim(W) = 2m$, and A is the linear mapping defined by $A(P) = \big(P(a_1), \ldots, P(a_m), P'(a_1), \ldots, P'(a_m)\big)$, and one then checks that A is injective: indeed, $A(P) = 0$ means $P(a_i) = 0$ and $P'(a_i)$ for $i = 1, \ldots, m$, so that P has m distinct double zeros, and since $deg(P) < 2m$ one deduces that $P = 0$.

Remark 3.5: It is not difficult to write explicitly the interpolation polynomial, and with $Q_i = \prod_{j \neq i}(x - a_j)$ for $i = 1, \ldots, m$, one has $P = \sum_{i=1}^m (\lambda_i x + \mu_i) Q_i^2$, and one then has to solve m linear systems of two equations: if $c_i = Q_i(a_i) \neq 0$ and $d_i = Q_i'(a_i)$, then $P(a_i) = \alpha_i$ means $(\lambda_i a_i + \mu_i)\, c_i^2 = \alpha_i$ and $P'(a_i) = \beta_i$ means $\lambda_i c_i^2 + 2(\lambda_i a_i + \mu_i)\, c_i d_i = \beta_i$, which one solves easily for finding λ_i and μ_i.

Remark 3.6: One generalizes easily to the case where at the point a_i one imposes the value of P and the values of the derivatives of P up to order $\kappa_i - 1$, with $\kappa_i \geq 1$ for $i = 1, \ldots, m$, and the degree of P is $\leq \kappa_1 + \ldots + \kappa_m - 1$.

However, if for three distinct values $a_1, a_2, a_3 \in E$, one looks for a polynomial P of degree ≤ 2 such that $P(a_1) = \alpha_1, P(a_2) = \alpha_2$ and $P'(a_3) = \beta_3$ (i.e. without imposing any condition on $P(a_3)$), then the situation is different, since the uniqueness question consists in wondering if $P(a_1) = P(a_2) = P'(a_3) = 0$

[1] If E is a finite field, it has size $q = p^k$ for a prime p, and m then satisfies $m \leq q$, of course.

implies $P = 0$, and since one must have $P = c\,(x - a_1)\,(x - a_2)$, it gives $P'(a_3) = c\,(2a_3 - a_1 - a_2) = 0$, and one may then have $P \neq 0$ if $2a_3 = a_1 + a_2$ (which is not possible if E has characteristic 2).

Remark 3.7: Having understood the question of interpolation in dimension 1, one then wonders about higher dimensions, and one first computes the dimension of the E-vector space $\mathcal{P}_\ell[x_1, \ldots, x_k]$ of polynomials $P \in E[x_1, \ldots, x_k]$ of degree $\leq \ell$, but one may also consider the larger E-vector space $\mathcal{Q}_\ell[x_1, \ldots, x_k]$ of polynomials $Q \in E[x_1, \ldots, x_k]$ which have degree $\leq \ell$ in each variable.[2] From $dim\big(\mathcal{P}_\ell[x_1]\big) = \ell + 1$, one deduces that $dim\big(\mathcal{P}_\ell[x_1, x_2]\big) = 1 + \ldots + (\ell + 1) = \frac{(\ell+1)\,(\ell+2)}{2} = \binom{\ell+2}{2}$, and (using formulas on binomial coefficients) one then finds by induction on k that $dim\big(\mathcal{P}_\ell[x_1, \ldots, x_k]\big) = \binom{\ell+k}{k}$.

Considering problems in the plane (i.e. $k = 2$), the case $\ell = 1$ corresponds to $\mathcal{P}_1[x_1, x_2]$, which has the basis $\{1, x_1, x_2\}$, and one may then expect that for any three distinct points a_1, a_2, a_3 in $E \times E$ there exists an interpolation polynomial $P \in E[x_1, x_2]$ of degree ≤ 1 satisfying $P(a_i) = \alpha_i$ for $i = 1, 2, 3$ whatever $\alpha_1, \alpha_2, \alpha_3 \in E$ are, but except for $E = \mathbb{Z}_2$,[3] there is a supplementary condition, and one must avoid the case where the three points a_1, a_2, a_3 are on the same line (in other words, interpolation works for the vertices of a non-degenerate triangle). Indeed, if P has degree ≤ 1 and vanishes at a_1, a_2, a_3, the restriction of P to the line through a_1 and a_2, parametrized by $v = a_1 + t\,(a_2 - a_1)$ for $t \in E$, is a polynomial of degree ≤ 1 in t, has two zeros (at $t = 0$ and $t = 1$), so that the restriction of P to the line is 0, and if a_3 happens to also be on this line there is a non-zero P satisfying these constraints. In order to go further with this method, we shall have to learn what can be deduced for a polynomial whose restrictions to a few lines are 0.

Remark 3.8: The two-dimensional interpolation theory was developed for approximating functions on domains Ω of the plane $\mathbb{R}^2$, for example by defining a *triangulation* of Ω,[4] and considering functions whose restriction to each triangle is a polynomial of some kind, and the contact between adjacent triangles may be imposed to give a continuous function, or a continuously differentiable function.

It is usual for mathematicians to wonder about generalizations, and after having studied a situation in a real plane to wonder which properties have been used: a first observation is to replace $\mathbb{R}$ by an arbitrary field E so that one develops linear algebra in a general context, and a plane will mean working in $E^2 = E \times E$. In some situations, one considers the normal derivative at the middle of an edge, and since normal means an angle of $\frac{\pi}{2}$ one must notice that there is no notion of angles in E^2, and we shall study later the notion of an *Euclidean structure* (where $E = \mathbb{R}$) and of an *Hermitian structure* (where $E = \mathbb{C}$), where angles make sense, but there is also a difficulty about the middle of an edge, which has no meaning if E has characteristic 2, since 2^{-1} does not exist in such a field. Suppose then that one uses a field E of characteristic $\neq 2$, and one considers a non-degenerate triangle with vertices a_1, a_2, a_3, so that the middle of the edges are defined by $a_{i,j} = 2^{-1}(a_i + a_j)$, and one can then consider the three medians, but the fact that the three medians intersect at the centre of gravity G of the triangle only holds if a field E of characteristic $\neq 3$, since $G = 3^{-1}(a_1 + a_2 + a_3)$, and we shall have to define the notion of *barycenter* and *barycentric coordinates* in a non-degenerate triangle.

If E has characteristic 3, let us consider the triangle with $a_1 = (0, 0), a_2 = (1, 0), a_3 = (0, 1)$, and call x, y the coordinates in E^2. With $a_{2,3} = (2^{-1}, 2^{-1})$ the equation of the median through a_1 and $a_{2,3}$ is $x - y = 0$, and with $a_{1,3} = (0, 2^{-1})$, the equation of the median through a_2 and $a_{1,3}$ has the form $\alpha\, x + \beta\, y = \gamma$, with $\alpha = \gamma$ and $\beta\, 2^{-1} = \gamma$, so that if one takes $\gamma = 1$ the equation is $x + 2y = 1$, but since $2 = -1$ in characteristic 3, it is the same as $x - y = 1$; similarly, the equation of the median through a_3 and $a_{2,3}$ is $2x + y = 1$, i.e. $x - y = -1$, and the three medians are parallel! After having developed projective geometry, there will be a line at infinity and one point at infinity in each direction, so that in characteristic 3 the three medians intersect at a common point at infinity (and in characteristic 2 the middle of the edges are at infinity, in three different directions).

---

[2] The dimension of $\mathcal{Q}_\ell[x_1, \ldots, x_k]$ is $(\ell + 1)^k$.

[3] In $\mathbb{Z}_2 \times \mathbb{Z}_2$, which is a plane with 4 points, there are only 6 lines, and each line only has 2 points.

[4] One asks that two adjacent triangles share the same edge, i.e. a vertex of a triangle cannot be an interior point of the edge of another triangle. When one imposes the value of a polynomial and some of its partial derivatives at a point, the information can only be used for this triangle if the point is interior to the triangle, but if the point is interior to an edge, it can be used for the two adjacent triangles sharing the edge (unless the edge is on the boundary of Ω), and if the point is a vertex, it can be used for all triangles having this vertex.