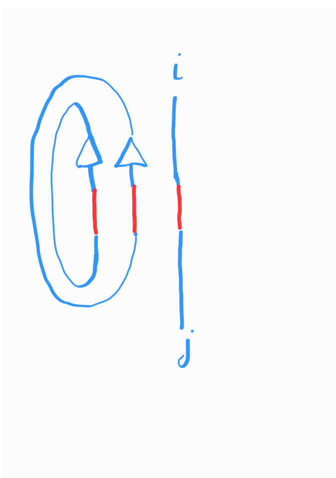
MATH 8450 – LECTURE 5 – FEB 1, 2023

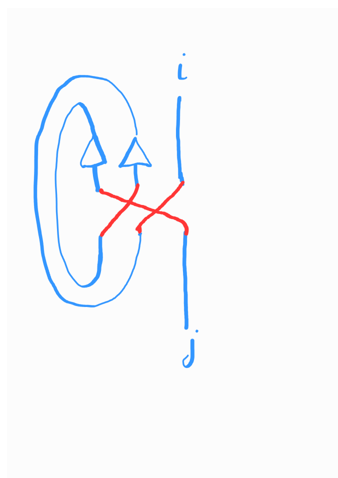
ABDELMALEK ABDESSELAM

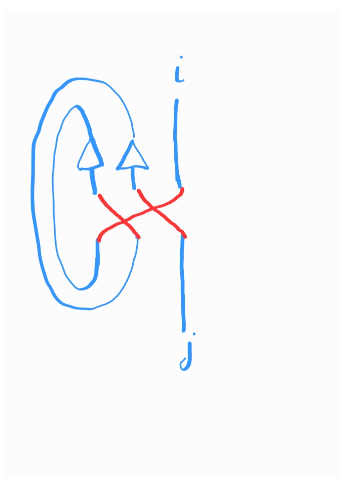
Tensors and diagrams cont'd:

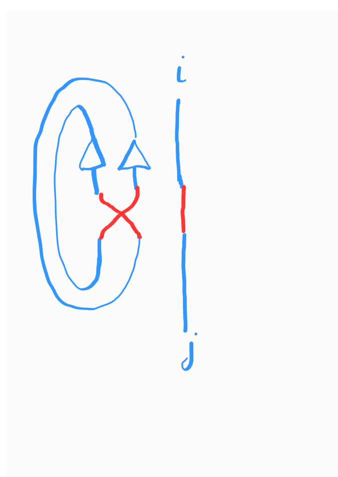
In this lecture we will do a bit of practice with graphical computations. The emphasis is on understanding the evaluation of individual pictures/diagrams. Later we will consider expansions involving infinitely many diagrams which will require more notions from combinatorics, as far as how to properly encode these pictures.

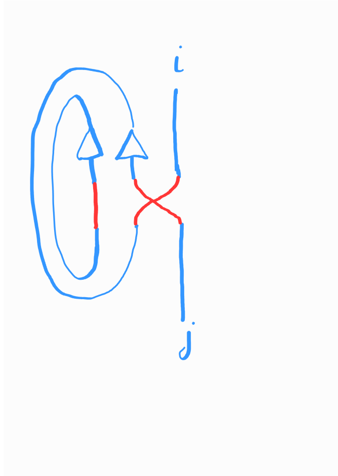
Let us start with the $n = 2$ case of the Cayley-Hamilton Theorem from the last lecture. The LHS of the equation, involving the antisymmetrizer of size 3, multiplied by $3! = 6$ in order to get rid of the denominators, becomes

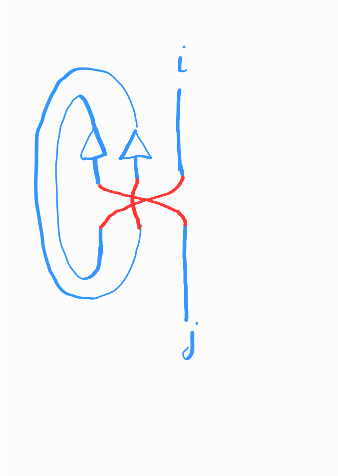
$6 \times \text{LHS} =$ 

$+$ 

$+$ 

$-$ 

$-$ 

$-$ 

after expansion of the antisymmetrizer. Therefore, we have, after evaluating the 6 pictures, in the same order, and stating the equality $\forall i, j$, as an equality of matrices:

$$6 \times \text{LHS} = (\text{tr } A)^2 I + A^2 + A^2 - \text{tr}(A^2) I - (\text{tr } A)A - (\text{tr } A)A \ .$$

Thus

$$3 \times \text{LHS} = A^2 - (\text{tr } A)A + \frac{(\text{tr } A)^2 - \text{tr}(A^2)}{2} \ I \ . \tag{1}$$

For the 2 by 2 matrix A, the characteristic polynomial is

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} \lambda - A_{11} & -A_{12} \\ -A_{21} & \lambda - A_{22} \end{vmatrix} \\ &= \lambda^2 - (\text{tr } A)\lambda + \det(A) \ , \end{aligned}$$

whereas

$$\begin{aligned} \frac{(\text{tr } A)^2 - \text{tr}(A^2)}{2} &= \frac{1}{2}\left[(A_{11} + A_{22})^2 - (A_{11}A_{11} + A_{12}A_{21} + A_{21}A_{12} + A_{22}A_{22})\right] \\ &= A_{11}A_{22} - A_{12}A_{21} \\ &= \det(A) \ . \end{aligned}$$

So (1) is indeed the result of substituting A for λ in the characteristic polynomial. Namely, for $n = 2$, we checked that the graphical equation at the end of Lecture 4 is the statement of the Cayley-Hamilton Theorem.

Homogeneous polynomials: Consider a homogeneous polynomial of degree d in n variables

$$F(x) = F(x_1, \ldots, x_n) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = d}} f_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n} \ .$$

Since we are working over the infinite field $\mathbb{C}$, we will not distinguish the formal polynomial in $\mathbb{C}[x_1, \ldots, x_n]$ from the associated polynomial function $\mathbb{C}^n \to \mathbb{C}$, because we have a bijective correspondence between these two notions. Such homogeneous polynomials F are in bijective correspondence with symmetric tensors:

$$F \longleftrightarrow (F_{i_1, \ldots, i_d})_{i_1, \ldots, i_d \in [n]} \ \in \ \mathbb{C}^{[n]^d} \simeq \mathbb{C}^{n^d} \ .$$

Here symmetric means that the tensor satisfies, $\forall \sigma \in \mathfrak{S}_d$, $\forall (i_1, \ldots, i_d) \in [n]^d$,

$$F_{i_{\sigma(1)},\ldots,i_{\sigma(d)}} = F_{i_1,\ldots,i_d} \ .$$

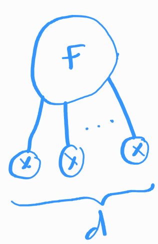
The correspondence is defined by imposing, that for all x,

$$F(x) = \sum_{(i_1,\ldots,i_d)\in[n]^d} F_{i_1,\ldots,i_d} x_{i_1} \cdots x_{i_d} \ .$$

Since we are dealing with a symmetric tensor, we will use a more symmetrical or round graphical representation for the tensor F:

$$:= F_{i_1,\ldots,i_d}$$

As a result



$$= \sum_{(i_1,\ldots,i_d)\in[n]^d} F_{i_1,\ldots,i_d} x_{i_1} \cdots x_{i_d} = F(x).$$

For a given homogeneous polynomial $F(x)$, we need the precise relationship between "the sum of monomials" description

$$F(x) = \sum_{\substack{\alpha\in\mathbb{N}^n \\ |\alpha|=d}} f_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

and the tensorial description

$$F(x) = \sum_{(i_1,\ldots,i_d)\in[n]^d} F_{i_1,\ldots,i_d} x_{i_1} \cdots x_{i_d} \ ,$$

i.e., the relation between the tensor entries $F_{i_1,\ldots,i_d}$ and the monomial coefficients f_α. Let us define a map

$$\begin{array}{rrcl} \mu: & [n]^d & \longrightarrow & \mathbb{N}^n \\ & I=(i_1,\ldots,i_d) & \longmapsto & \mu(I) \end{array}$$

as follows. For all $j\in[n]$, we let the j-th component of the multiindex $\mu(I)$ be

$$\mu(I)_j := |\{\ell\in[d]\mid i_\ell=j\}|\ .$$

We call $\mu(I)$ the multiplicities multiindex of the index sequence I because it simply counts how many times the index value 1 appears, how many times the index value 2 appears, etc. The map is surjective and a section (map given by a choice of preimage) is provided by

$$\alpha\longmapsto 1^{\alpha_1}2^{\alpha_2}\cdots n^{\alpha_n} := (1,\ldots,1,2,\ldots,2,\ldots,n\ldots,n)$$

where 1 appears α_1 times, followed by 2 appearing α_2 times, and so on. We now have, with perhaps too much detail,

$$\begin{aligned} F(x) &= \sum_{(i_1,\ldots,i_d)\in[n]^d} F_{i_1,\ldots,i_d}x_{i_1}\cdots x_{i_d} \\ &= \sum_{I\in[n]^d} 1\times F_{i_1,\ldots,i_d}\ x_{i_1}\cdots x_{i_d} \\ &= \sum_{I\in[n]^d}\left(\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=d}} \mathbb{1}\{\mu(I)=\alpha\}\right)\times F_{i_1,\ldots,i_d}\ x_{i_1}\cdots x_{i_d} \\ &= \sum_{I\in[n]^d}\sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=d}} \mathbb{1}\{\mu(I)=\alpha\}\times F_{i_1,\ldots,i_d}\ x_{i_1}\cdots x_{i_d} \\ &= \sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=d}}\sum_{I\in[n]^d} \mathbb{1}\{\mu(I)=\alpha\}\times F_{i_1,\ldots,i_d}\ x_{i_1}\cdots x_{i_d} \\ &= \sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=d}}\sum_{I\in[n]^d} \mathbb{1}\{\mu(I)=\alpha\}\times F_{1^{\alpha_1}2^{\alpha_2}\cdots n^{\alpha_n}}\ x^\alpha \\ &= \sum_{\substack{\alpha\in\mathbb{N}^n\\|\alpha|=d}} F_{1^{\alpha_1}2^{\alpha_2}\cdots n^{\alpha_n}}\ x^\alpha\left(\sum_{I\in[n]^d}\mathbb{1}\{\mu(I)=\alpha\}\right)\ . \end{aligned}$$

Now we have

$$\sum_{I\in[n]^d}\mathbb{1}\{\mu(I)=\alpha\} = \binom{d}{\alpha_1}\binom{d-\alpha_1}{\alpha_2}\cdots\binom{d-\alpha_1\cdots-\alpha_{n-1}}{\alpha_n} = \frac{d!}{\alpha_1!\cdots\alpha_n!}$$

which is called the multinomial coefficient usually denoted by

$$\binom{d}{\alpha_1,\ldots,\alpha_n}\ .$$

We will also use the notation $\binom{d}{\alpha} = \frac{d!}{\alpha!}$ where $\alpha! := \alpha_1! \cdots \alpha_n!$. The above count corresponds to choosing where to put the 1's among d spots, where to put the 2's , etc. when forming an index sequence I with imposed multiplicity multiindex α. Since the coefficients in the monomial expansion are uniquely determined, we get the relation

$$\binom{d}{\alpha} F_{1^{\alpha_1} 2^{\alpha_2} \cdots n^{\alpha_n}} = f_\alpha$$

Equivalently, because of the symmetry of the tensor, for all $I = (i_1, \ldots, i_d) \in [n]^d$,

$$F_I = \frac{1}{\binom{d}{\alpha}} f_\alpha \tag{2}$$

where $\alpha = \mu(I)$. With this elementary but important relation in hand we will now do some practice with graphical computations while revisiting some early childhood mathematics.

Solving quadratic equations: When $a \neq 0$, the solutions to the quadratic equation

$$ax^2 + bx + c = 0$$

are of course given by the formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

which features the fundamental quantity $\Delta_2 := b^2 - 4ac$ or discriminant. The latter detects root collision.

The previous discussion suggests the need for a homogeneous polynomial in order to bring graphical calculus into play. So we change the name of the variable x to x_1 and introduce a new variable x_2 to be used as padding material, in order to make all the terms of $f(x) = ax^2 + bx + c$ become of total degree 2. Namely, we define the polynomial

$$F(x_1, x_2) := ax_1^2 + bx_1x_2 + cx_2^2 \ .$$

This is called the homogeneization of f. Its tensorial representation is

$$F(x_1, x_2) := F_{11}x_1^2 + 2F_{12}x_1x_2 + F_{22}x_2^2$$

and therefore the entries of the corresponding symmetric tensor are

$$\begin{aligned} F_{11} &= a \\ F_{12} &= \frac{b}{2} \\ F_{21} &= \frac{b}{2} \\ F_{22} &= c \ . \end{aligned}$$

We will need a new elementary building block for our diagrammatic computations corresponding to the tensor (or matrix rather)

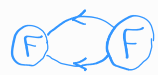
$$\varepsilon = (\varepsilon_{ij})_{i,j \in [2]} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ .$$

We let

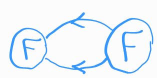
$:= \varepsilon_{ij}$

We now have the following result.

Proposition 1. *The discriminant is given by*

$$\Delta_2 = -2 \times$$


Proof: By definition,



$$= \sum_{i,j,k,\ell=1}^{2} F_{ij}\varepsilon_{ik}\varepsilon_{j\ell}F_{k\ell}$$

since there are four elementary pieces assembled (two F's and two ε's) and four sutures/junction points (and therefore four indices to be summed over). The last sum written in full (ignoring terms giving zero when $i = k$ or $j = \ell$) is

$$F_{11}F_{22} + F_{12}F_{21} \times (-1) + F_{21}F_{12} \times (-1) + F_{22}F_{11}(-1)^2$$
$$= 2F_{11}F_{22} - 2F_{12}^2 = 2ac - \frac{b^2}{2}$$

after substituting the values of the F tensor entries, and the proposition follows. □

Solving cubic equations: Let us consider the case of cubic equations

$$ax^3 + bx^2 + cx + d = 0$$

with $a \neq 0$. By dividing by a, one can reduce the general situation to the case $a = 1$ which we now assume. Then one can "complete the cube" by writing the equation as

$$\left(x + \frac{b}{3}\right)^3 + \cdots = 0$$

and also change variables to $x + \frac{b}{3}$ instead of x (a particular case of a Tschirnhaus transformation). This brings us to the case of depressed or reduced cubic equations $f(x) = 0$ with

$$f(x) = x^3 + px + q \ .$$

The solutions are then given by the so-called Cardano formula

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \ .$$

A priori, there are three choices for the cube roots, so a total of 9 possibilities, however, one has to pick the cube roots such that their product is equal to $-\frac{p}{2}$. Note the conspicuous quantity

$$\Delta_3 = \frac{q^2}{4} + \frac{p^3}{27}$$

which up to a constant multiple is the discriminant which here too detects root collision. Like before, we introduce the homogeneization of f given by

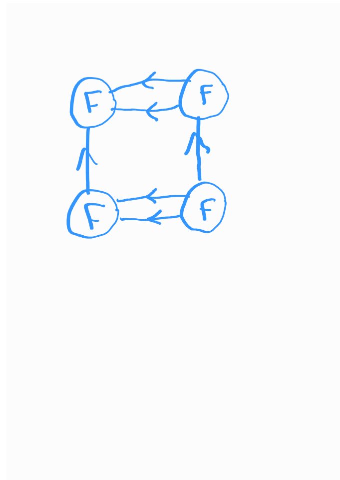
$$F(x_1, x_2) := x_1^3 + px_1x_2^2 + qx_2^3 \ ,$$

and the corresponding symmetric tensor $(F_{ijk})_{1 \le i,j,k \le 3}$, as well as its graphical representation by a round "blob" with three legs. From (2) we read off the tensor entries which are given explicitly by

$$\begin{aligned} F_{111} &= 1 \\ F_{112} &= 0 \\ F_{122} &= \frac{p}{3} \\ F_{222} &= q \ , \end{aligned}$$

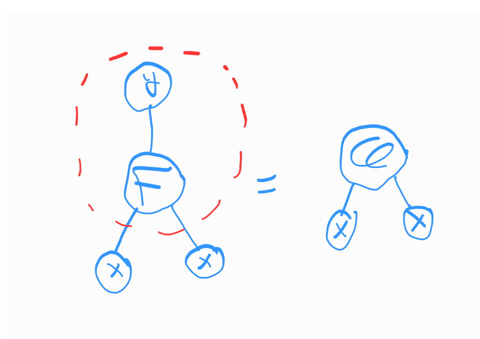
while the missing entries can be deduced by changing the position of indices, since the tensor is symmetric.

Proposition 2. *We have*

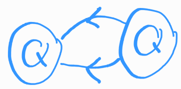
$$\Delta_3 = -\frac{1}{8} \times$$ 

Proof: By definition, the evaluation of the picture is given by a sum over twelve indices which results in 4096 terms which are product of four F tensor entries and six ε matrix entries. Note that the antisymmetry of the matrix ε forces the value of the second index if we know that of the first. So this reduces the sum to $2^6 = 64$ terms which is still a lot. We will therefore try to use a more efficient way of evaluating the diagram. This will show us along the way some tricks one can do like substituting complex diagrammatic structures inside blobs.

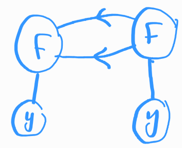
We start by introducing a new vector $y = (y_1, y_2)$ together with its graphical representation and write



as a definition of the y-dependent quadratic form Q in the x variables. The blob of Q is equal to the part of the picture in the round dotted box. Using the previous section on quadratics, we have



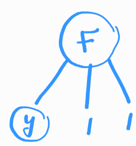
$$=$$



$$= 2(Q_{11}Q_{22} - Q_{12}^2)\ .$$

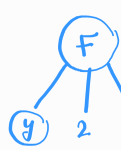
We now have the easy calculations with single index contractions:

$$Q_{11} =$$



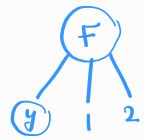
$$= F_{111}\ y_1 + F_{112}\ y_2 = y_1$$

$$Q_{22} =$$



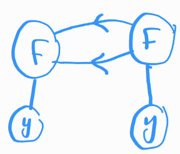
$$= F_{122}\ y_1 + F_{222}\ y_2 = \frac{p}{3} y_1 + q y_2$$

$$Q_{12} =$$



$$= F_{112}\ y_1 + F_{122}\ y_2 = \frac{p}{3} y_2 \ .$$

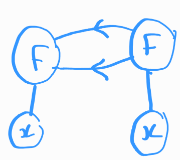
As a result, we have



$$= 2\left[y_1 \left(\frac{p}{3} y_1 + q y_2 \right) - \left(\frac{p}{3} y_2 \right)^2 \right]$$

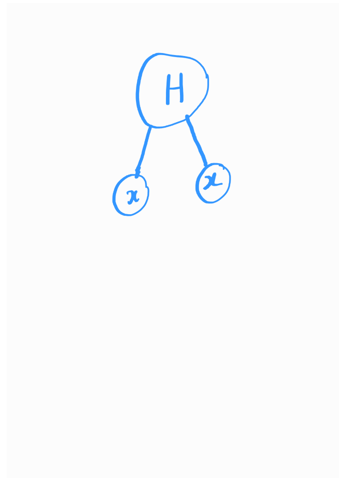
$$= \frac{2p}{3} y_1^2 + 2q y_1 y_2 - \frac{2p^2}{9} y_2^2 \ .$$

We now trade the y's for the original x variables, i.e., do the substitution $y := x$ in the previous equation. This gives

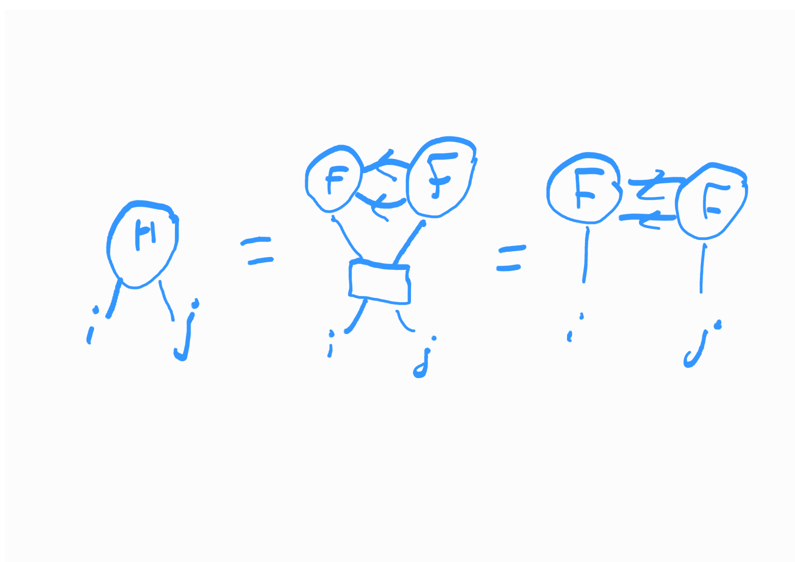


$$= \frac{2p}{3} x_1^2 + 2q x_1 x_2 - \frac{2p^2}{9} x_2^2 =: H(x)$$

which defines the so-called Hessian $H(x)$ of the cubic F. The Hessian also has a graphical representation

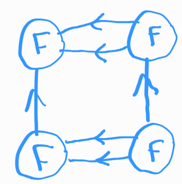
$H(x) =$ 

but identifying the blob of H as some encapsulation of a structure made of F's requires some care. We have



where we inserted a symmetrizer (because one side made of x's is symmetric) at first, so H is symmetric, only to realize as a second step that here, by accident, the inner F structure is already symmetric. This is because reversing an ε arrow produces a (-1) factor, and here we would reverse two arrows which results in no net change.

Finally, again using the section on quadratics we a have



$=$

$= 2(H_{11}H_{22} - H_{12}^2)$.

From the equation above where $H(x)$ was defined we immediately read off the tensor/matrix entries

$$\begin{aligned} H_{11} &= \frac{2p}{3} \\ H_{12} &= q \\ H_{22} &= -\frac{2p^2}{9} \end{aligned}$$

which upon substitution in the expression $2(H_{11}H_{22} - H_{12}^2)$ and multiplication by $-\frac{1}{8}$ gives the expression $\frac{q^2}{4} + \frac{p^3}{27}$ and proves the proposition. □

Remark 1. *The above computations come from 19th century invariant theory, in particular that of binary ($n = 2$) forms (homogeneous polynomials). The first fundamental theorem of classical invariant theory for SL_2 says that every polynomial in the coefficients of a form F which is invariant under linear change of coordinates by an element of SL_2 must be a linear combination of pictures made of F blobs and ε arrows. This generalizes to SL_n. Note also that the entire theory of angular momentum in quantum mechanics can be done with this kind of graphical calculus. In particular one can write formulas with pictures for SU_2 Clebsch-Gordan coefficients $\langle j_1, m_1, j_2, m_2 | J, M\rangle$, but that is another story.*

Remark 2. *(As per Diana's question) Discriminants generalize to higher dimensions. If $F_1(x), \ldots, F_n(x)$ are n homogeneous polynomials in n variables $x_1, \ldots, x_n$, of respective degrees $d_1, \ldots, d_n$, then there is a unique polynomial $\mathrm{Res}(F_1, \ldots, F_n)$ in the coefficients of $F_1, \ldots, F_n$ which satisfies the following properties.*

(1) $\mathrm{Res}(F_1, \ldots, F_n) = 0$ *iff* $\exists x \in \mathbb{C}^n \backslash \{0\}, \forall i \in [n], F_i(x) = 0$.

(2) $\forall i \in [n]$, $\mathrm{Res}(F_1, \ldots, F_n)$ *is homogeneous of degree* $\prod_{j \neq i} d_j$ *in the coefficients of* F_i.

(3) $\mathrm{Res}(F_1, \ldots, F_n) = 1$ *when* $F_1(x) = x_1^{d_1}, \ldots, F_n(x) = x_n^{d_n}$ *(the "diagonal case").*

This is called the (multidimensional) resultant. If $(d_1, \ldots, d_n) = (1, \ldots, 1)$ then the resultant is just the determinant of the matrix formed by the coefficients of the linear forms $F_1, \ldots, F_n$. Now if $H(x)$ is a homogeneous polynomial of degree d in n variables, one can define $F_i = \frac{\partial H}{\partial x_i}$ and then take the resultant of these F_i's. By definition, this is the discriminant of H, and *it detects if the hypersurface* $\{H(x) = 0\}$ *is singular (discriminant is zero) or smooth (discriminant is nonzero). The case* $n = 2$ *corresponds to hypersurfaces in a space of dimension 1, which are just collections of points with possible multiplicities. Smooth means the points are distinct, while singular means there are point/root collisions. This puts the elementary* $\Delta_2 = b^2 - 4ac$ *under the same roof as much deeper and more complicated objects, i.e., discriminants of hypersurfaces in arbitrary dimension.*