Polynomial Discriminants

Part 2: Tensor Diagrams

James F. Blinn

Microsoft Research

Several years ago, Jim Kajiya loaned me a copy of a book called *Diagram Techniques in Group Theory*.[1] This book described a graphical representation of the algebra used to solve various problems in mathematical physics. I was only able to understand the first chapter of the book, but even that was enough to excite me tremendously about adapting the technique to the algebra of homogeneous geometry that we're familiar with in computer graphics. I've written up my initial efforts in two columns ("Uppers and Downers," *IEEE CG&A* March 1992 and May 1992, reprinted in the book *Jim Blinn's Corner: Dirty Pixels*.[2]) Recently I've been playing more and more with these diagrammatic ways of doing algebra and have come up with a lot of interesting results. This column presents the first of these—using diagrams to compute discriminants of polynomials and solve a related problem: line-curve tangency. To get into this, I'll briefly review the parts of "Uppers and Downers" that will be useful here.

2D homogeneous geometry

Two-dimensional homogeneous geometry uses three element vectors, 3×3 matrices, $3\times 3\times 3$ tensors, and so on to represent various objects. I'll denote such quantities in uppercase boldface to distinguish them from polynomials discussed later. For example, a homogeneous point **P** is a three-element row vector and a line **L** is a three-element column vector. The point lies on the line if the dot product $\mathbf{P}\cdot\mathbf{L}$ is zero. Figure 1 shows different ways of expressing the dot product. Figures 1a, 1b, and 1c should be familiar to you. I'll explain Figures 1d and 1e shortly. Moving up to curves, the points on a second order (quadratic) curve satisfy the equation in Figure 2a. We can write this in matrix form by arrang-

$$ax + by + cw = 0$$

(a)

$$[x\ y\ w]\begin{bmatrix}a\\b\\c\end{bmatrix} = 0$$

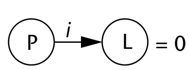
(b)

$$\mathbf{P}\cdot\mathbf{L} = 0$$

(c)

$$P^iL_i = 0$$

(d)



(e)

1 Point on a line.

$$ax^2 + 2Bxy + 2Cxw + Dy^2 + 2Eyw + Fw^2 = 0$$

(a)

$$[x\ y\ w]\begin{bmatrix}A & B & C\\B & D & E\\C & E & F\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = 0$$

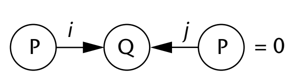
(b)

$$\mathbf{P}\cdot\mathbf{Q}\cdot\mathbf{P}^T = 0$$

(c)

$$P^iQ_{ij}P^j = 0$$

(d)



(e)

2 Point on a quadratic curve.

$$Ax^2 + 3Bx^2y + 3Cxy^2 + Dy^3 + 3Ex^2w + 6Fxyw + 3Gyw^2 + 2Hxw^2 + 3Jyw^2 + Kw^2 = 0$$

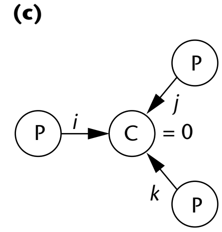
(a)

$$[x\ y\ w]\left\{[x\ y\ w]\begin{bmatrix}\begin{bmatrix}A & B & E\\B & C & F\\E & F & H\end{bmatrix}\\\begin{bmatrix}B & C & F\\C & D & G\\F & G & J\end{bmatrix}\\\begin{bmatrix}E & F & H\\F & G & J\\H & J & K\end{bmatrix}\end{bmatrix}\right\}\begin{bmatrix}x\\y\\z\end{bmatrix} = 0$$

(b)

$$P^iP^jP^kC_{ijk} = 0$$

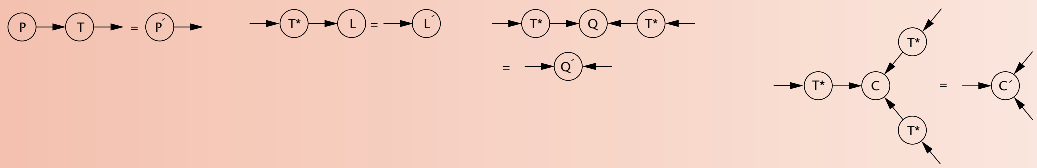
(c)



(d)

3 Point on a cubic curve.

Table 1. Transformations.

	Point	Line	Quadratic	Cubic
a	$PT=P'$	$(T^*)L=L'$	$(T^*)Q(T^*)^T=Q'$	messy
b	$P^iT_i^j=(P')^j$	$(T^*)_j^i\,L_i=(L')_j$	$(T^*)_k^i\,Q_{ij}(T^*)_l^j=(Q')_{kl}$	$(T^*)_l^i\,(T^*)_m^j\,(T^*)_n^k\,C_{ijk}=(C')_{lmn}$
c				

ing the coefficients into the 3 × 3 symmetric matrix of Figure 2b. Next up, the points on a third order (cubic) curve satisfy Figure 3a. We can also write this by arranging the coefficients into a 3 × 3 × 3 symmetric generalization of a matrix. Doing this with conventional matrix notation is a bit weird. About the best we can do is to show it as a vector of matrices as in Figure 3b.

Now let's talk about transformations. We geometrically transform points by postmultiplying by a 3 × 3 matrix: **PT** = **P′**, and we transform lines by premultiplying by the adjoint of the matrix: **T*L**=**L′**. Table 1 shows various ways to write these expressions, as well as those for transforming curves.

Finally, the cross product of two point-vectors **P** and **R** gives the line passing through them: **P**×**R**=**L** (see Figure 4). In a dual fashion, the cross product of two line-vectors **L** and **M** gives their point of intersection: **L**×**M**=**P**.

$$\begin{bmatrix}P^1 & P^2 & P^3\end{bmatrix}\times\begin{bmatrix}R^1 & R^2 & R^3\end{bmatrix}=\begin{bmatrix}L_1\\ L_2\\ L_3\end{bmatrix}$$

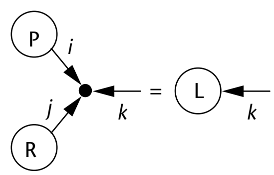
(a)

$$\mathbf{P}\times\mathbf{R}=\mathbf{L}$$

(b)

$$P^iR^j\varepsilon_{ijk}=L_k$$

(c)



(d)

4 The cross product.

The problem

In looking over these expressions, we see that our notation has two problems. The first is the need to take the transpose of **P** when multiplying by **Q**. This is very fishy. Column matrices are supposed to represent lines, not points. In fact, there's something fundamentally different about matrices that represent transformations and matrices that represent quadratic curves. We can't, however, distinguish between them with standard vector notation. The second problem is the inability to conveniently represent entities with more than two indices. Our attempt to arrange the coefficients of a cubic polynomial into a triply indexed "cubical matrix" is an example of the problem.

Fortunately, we can adapt two notational schemes from the world of theoretical physics to alleviate these shortcomings—Einstein Index Notation (EIN) and the diagram notation I referred to in my opening monologue. I originally called these Feynman diagrams but they differ enough to give them the more appropriate name tensor diagrams. They're more like the diagrams from Kuperberg.[3]

2DH tensor diagrams

EIN differentiates between two types of indices for vector or matrix elements: the point-like ones (which we'll call contravariant and write as superscripts) and line-like ones (which we will call covariant and write as subscripts). Thus an element of a point-vector is P^i and an element of a line-vector is L_i. (Note that superscript indices aren't the same as exponents. Mathematicians ran out of places to put indices and started overloading their notation. Live with it.). Dot products happen only between matching pairs of covariant and contravariant indices. Thus the dot of a point and a line is

$$\mathbf{P}\cdot\mathbf{L}=\begin{bmatrix}P^1 & P^2 & P^3\end{bmatrix}\cdot\begin{bmatrix}L_1 & L_2 & L_3\end{bmatrix}$$
$$=\sum_i P^iL_i$$

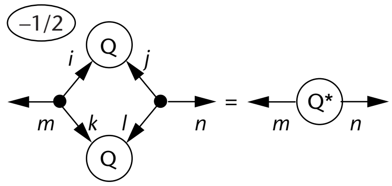
We simplify further by omitting the sigma and stating that any superscript or subscript pair that has the same letter implicitly implies a summation over that letter. The EIN form of a dot product is then simply P^iL_i. A more complicated expression may have many tensors and superscripts and subscripts, and will implicitly be summed over all pairs of identical upper and lower indices. (These summations are also called tensor contractions.) We can see this in the EIN for higher order curves in Figures 2d and 3c. Note that the expression for EIN is basically a model for the terms that are summed. Each individual factor in the notation is just a number, so the factors can be rearranged in any order, as Figure 3c shows.

Tensor diagram notation is a translation of EIN into a graph. We represent a point as a node with an outward arrow indicating a covariant index. A line, with its covariant index, is a node with an inward arrow. The adj **Q** = **Q***

(a)

$\frac{1}{2}\, Q_{ij} Q_{kl}\, \varepsilon^{ikm} \varepsilon^{jln} = (Q^*)^{mn}$

(b)



(c)

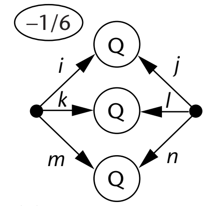
5 **The adjoint of a matrix.**

det **Q**

(a)

$\frac{1}{6}\, Q_{ij} Q_{kl} Q_{mn}\, \varepsilon^{ikm} \varepsilon^{jln}$

(b)



(c)

6 **Determinant of a 3 × 3 matrix.**

dot product—that is, the summation over the covariant-contravariant pair—is an arc connecting two nodes. See Figures 1e, 2e, and 3d for the diagram notation of the expressions we've seen so far. For many of the diagrams I have labeled the arcs with the index they correspond to in EIN. Some later, more complex, diagrams will not need this.

Transformations

A transformation matrix has one contravariant and one covariant index. Multiplying a point by such a matrix will "annihilate" its contravariant index leaving a result that has a free covariant index, making the result be a point. Table 1, row b, shows the EIN form of the transformation of various quantities. Row c of the table shows how this translates into diagram notation. Now we can see the difference between the two types of matrices. A transformation matrix has one of each type of index (denoted with one arrow out and one arrow in); a quadratic matrix has two covariant indices (denoted with both arrows in). In Figure 2d the two covariant-contravariant index pairs annihilate each other to produce a scalar.

Cross products and adjoints

We abbreviate the algebra for cross products and matrix adjoints by defining a three-index 3 × 3 × 3 element anti-symmetric tensor called the Levi-Civita epsilon. The elements of epsilon are defined to be

$$\begin{aligned} \varepsilon_{123} &= \varepsilon_{231} = \varepsilon_{312} = +1 \\ \varepsilon_{321} &= \varepsilon_{132} = \varepsilon_{213} = -1 \\ \varepsilon_{ijk} &= 0 \quad \text{otherwise} \end{aligned} \tag{1}$$

Multiplying two vectors by epsilon forms their cross product. Since epsilon has three subscript indices, multiplying in two points with superscript indices will result in a vector with one remaining subscript index (a line). The diagram form of epsilon is a node with three inward pointing arcs. I show this node as a small dot as in Figure 4d. You can imagine a similar figure for the dual form, the cross product of two lines: $\mathbf{L} \times \mathbf{M} = \mathbf{P}$. Just use a contravariant form of epsilon ε^{ijk}, so that $L_i M_j \varepsilon^{ijk} = P^k$ and flip the direction of all arrows in the diagram.

We must be careful about how the anti-symmetry of epsilon translates into a diagram. The convention is to label the arcs counterclockwise around the dot. A mirror reflection of an epsilon diagram will reverse the order of its indices and therefore flip its algebraic sign.

Epsilon is also useful to form matrix adjoints. Figure 5 shows various ways to denote the adjoint. The raw EIN expression $Q_{ij} Q_{kl} \varepsilon^{ikm} \varepsilon^{jln}$ gives twice the adjoint, so I had to insert a factor of 1/2 to get the correct answer. I also mirrored the diagram for the first epsilon in the EIN (and introduced a minus sign) to make the whole diagram a bit prettier. These factors and signs clutter things up a bit but are necessary to get right.

Now that we have the adjoint, the determinant is not far behind—multiply the adjoint by **Q** and take the trace (connect the two dangling arcs). The result is three times the determinant. Various notations appear in Figure 6.

Homogeneous polynomials

Now let's go down a dimension and take a look at one-dimensional homogeneous geometry. This is effectively the study of homogeneous polynomials. Basically we have the same thing as before, but everything is now composed of two-element vectors, 2 × 2 matrices and 2 × 2 × 2 tensors, which I'll write as lower case boldface. A homogeneous linear equation is written in various notations as Figure 7 illustrates. Figure 8 shows a homogeneous quadratic equation, and Figure 9 shows a homogeneous cubic equation. (Unfortunately I find that I have to use the letter C in two contexts, once as a coefficient and once as a tensor name. Live with it.)

The 2D Epsilon

The only slightly subtle item is the form of the two-element epsilon. Instead of having three indices, each with three values, the two-element epsilon has two indices (making it a simple matrix) each with two values. By analogy to Equation 1, the contravariant form of epsilon is

$$\begin{aligned} \varepsilon^{12} &= +1 \\ \varepsilon^{21} &= -1 \\ \varepsilon^{ij} &= 0 \quad \text{otherwise} \end{aligned}$$

In other words,

$$\varepsilon = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$Ax + bw = 0$

(a)

$$[x\ \ w]\begin{bmatrix} A \\ B \end{bmatrix} = 0$$

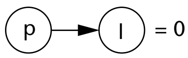
(b)

$\mathbf{p} \cdot \mathbf{l} = 0$

(c)

$p^i l_i = 0$

(d)



(e)

7 **Homogeneous linear equation.**

$Ax^2 + 2Bbw + Cw^2 = 0$

(a)

$$[x\ \ w]\begin{bmatrix} A & B \\ B & C \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix} = 0$$

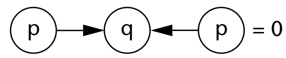
(b)

$\mathbf{pqp}^T = 0$

(c)

$p^i p^j p_{ij} = 0$

(d)



(e)

8 **Homogeneous quadratic equation.**

$Ax^3 + 3Bx^2w + 3Cxw^2 + Dx^3 = 0$

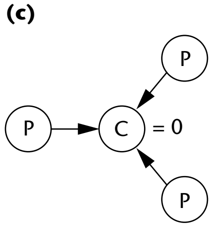
(a)

$$[x\ \ w]\left\{[x\ \ w]\begin{bmatrix} \begin{bmatrix} A & B \\ B & C \end{bmatrix} \\ \begin{bmatrix} B & C \\ C & D \end{bmatrix} \end{bmatrix}\right\}\begin{bmatrix} x \\ w \end{bmatrix} = 0$$

(b)

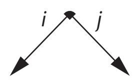
$p^i p^j p^k p_{ijk} = 0$

(c)

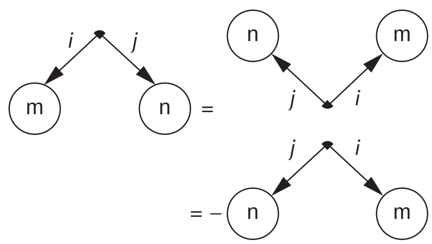


(d)

9 **Homogeneous cubic equation.**



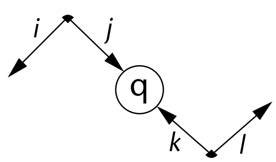
10 **The 2D epsilon.**



11 **Rotations and reflections of the 2D epsilon.**

$(q^*)^{il} = \varepsilon^{ij} q_{jk}\, \varepsilon^{ik}$

(a)

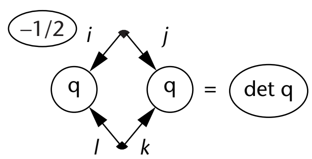


(b)

12 **Adjoint of a 2 × 2 matrix.**

$\frac{1}{2}\, q_{jk} q_{li}\ \varepsilon^{ij} \varepsilon^{lk} = \det \mathbf{q}$

(a)



(b)

13 **Determinant of a 2×2 matrix.**

The Einstein notation is simply ε_{ij} or ε^{ij} and the diagram notation looks like Figure 10.

I purposely constructed Figure 10 to be asymmetrical. The convention is that when the diagram points down (as above) the first index is on the left. A mirror reflection of this diagram will perform a sign flip on the diagram's value. If the diagram were not asymmetrical, a mirror flip would not be detectable.

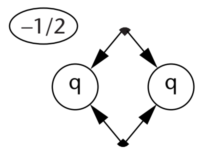
To drive this home, compare the EIN expressions

$$m^i \varepsilon_{ij} n^j = n^j \varepsilon_{ij} m^i = -n^j \varepsilon_{ji} m^i$$

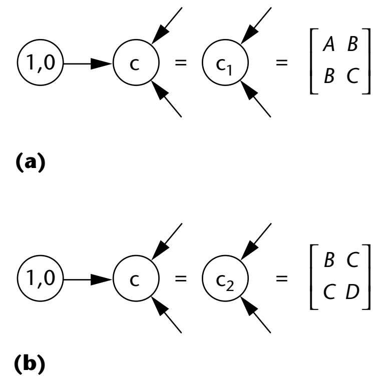
with their diagram counterparts. The first equality represents a rotation, the second has a reflection. See Figure 11.

Now let's use the epsilon. Figure 12 shows the adjoint of a 2 × 2 matrix, by analogy to Figure 5.

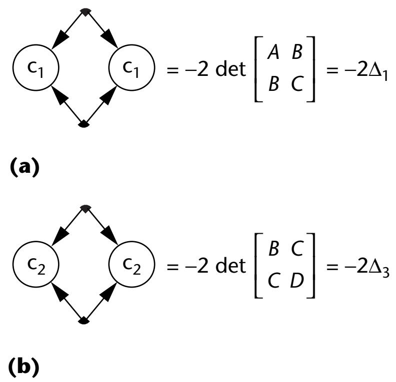
We get the determinant by analogy to Figure 6. Multiply Figure 12 by $\mathbf{q}$ and take the trace. This gives twice the determinant. Flip one of the epsilons to make the diagram neater. Figure 13 shows what we get.



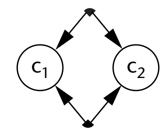
14 Discriminant of a quadratic polynomial.



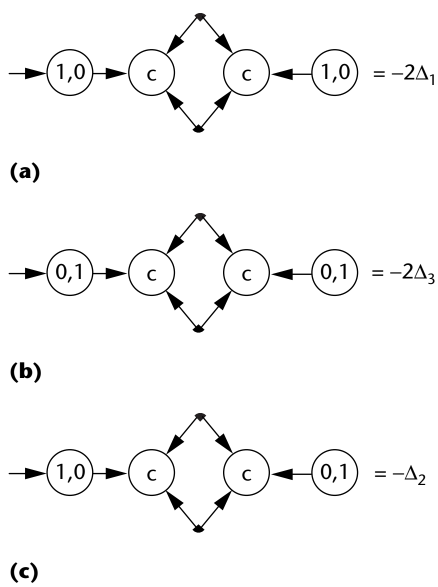
15 Slices of the c tensor.



16 Determinants of the matrices in Figure 15.



17 Cross determinant.



18 Values of each matrix element.

A 1DH application: Discriminants

The discriminant of a polynomial is a condition on the coefficients that guarantees that the polynomial has a double root. In our last installment (November/December 2000 *IEEE CG&A*) we learned how to write this quantity in matrix terms. Now let's see how this looks in diagram form. The discriminant of the quadratic polynomial from Figure 4a is

$$B^2 - AC = -\det\begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

Figure 14 shows this in diagram form.

The discriminant of the cubic polynomial from Figure 5a is

$$\Delta = \det\begin{bmatrix} 2\Delta_1 & \Delta_2 \\ \Delta_2 & 2\Delta_3 \end{bmatrix} \qquad (2)$$

where the matrix elements are defined as

$$\Delta_1 = AC - B^2 = \det\begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

$$\Delta_2 = AD - BC = \det\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$\Delta_3 = BD - C^2 = \det\begin{bmatrix} B & C \\ C & D \end{bmatrix}$$

What does this look like in diagram form? Let's look at the individual "slices" of the **c** tensor. We form these by multiplying one index by a "basis vector" like (1, 0) or (0, 1). See Figure 15. Figure 16 shows the determinants of these two matrices. Now what happens if we mash together $\mathbf{c}_1$ and $\mathbf{c}_2$ as a sort of "cross determinant" with the diagram form shown in Figure 17? The value of this diagram is, in conventional matrix form

$$trace\left\{\begin{bmatrix} A & B \\ B & C \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} B & C \\ C & D \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\} =$$

$$trace\left\{\begin{bmatrix} BC - AD & AC - B^2 \\ C^2 - BD & 0 \end{bmatrix}\right\} =$$

$$BC - AD = -\Delta_2$$

Now, remembering the definitions of $\mathbf{c}_1$ and $\mathbf{c}_2$, we have just shown the relations in Figure 18. What we have just done is to find expressions for each of the elements of the matrix in Equation 2. Figure 19 puts these back together into a matrix. One interesting thing about this demonstration is that it shows why there are factors of two for the Δ_1 and Δ_3 terms, but not for the Δ_2. Anyway, the final step is easy. The discriminant of the cubic **c** equals the determinant of this matrix (with the appropriate minus sign and scale factor) as Figure 20 shows. You can see this as a nice generalization of the discriminant diagram for the quadratic polynomial in Figure 14.

A 2DH Application: Tangency

Now let's use this 1DH result to solve a 2DH geometry problem: tangency.

Quadratic with line

Figure 2 gave us the condition of a point being on a quadratic curve. How can we generate an expression that determines if a line **L** is tangent to curve **Q**? Let's start by assuming that we have two points, **R** and **S**, on **L**. (We don't need to know how we found these two points. In fact, they will disappear shortly.) Then a general point on the line is

$$\mathbf{P}(\alpha,\beta)=\alpha\mathbf{R}+\beta\mathbf{S}$$

In matrix notation

$$\mathbf{P}=\begin{bmatrix}\alpha & \beta\end{bmatrix}\begin{bmatrix}R^1 & R^2 & R^3\\ S^1 & S^2 & S^3\end{bmatrix}$$

$$\mathbf{P}=\mathbf{aV}$$

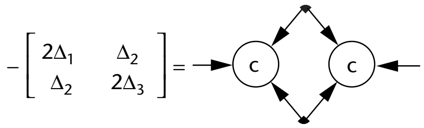
The 2×3 matrix is a sort of conversion from the world of 2D (1DH) vectors (homogeneous polynomials) to the world of 3D (2DH) vectors (homogeneous curves). Figure 21 shows this in diagram form. (For these mixed mode diagrams I made thicker arrows for the three-element summations and thinner arrows for the two-element summations.) If we plug this into the quadratic curve equation we get a homogeneous polynomial in (α, β) that evaluates the quadratic function at each point on the line. So, plug Figure 21 into Figure 2e and you get the results shown in Figure 22. This turns the 3×3 symmetric quadratic curve matrix **Q** into a 2×2 symmetric quadratic polynomial matrix **q**. Figure 23 shows just **q** by itself. The condition of the line being tangent to the curve is the same as the condition that the polynomial has a double root. That polynomial has a double root iff its determinant is zero. Plugging this into the diagram form of the determinant gives us the condition that the polynomial has a double root, and thus that the line hits the curve at exactly one point (Figure 24).

Now look at the diagram fragment shown in Figure 25. Write this as a matrix product:

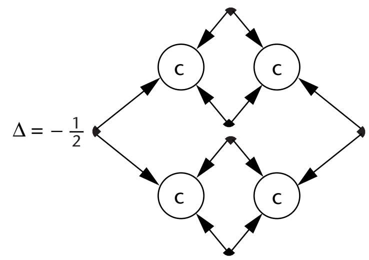
$$\begin{bmatrix}R^1 & S^1\\ R^2 & S^2\\ R^3 & S^3\end{bmatrix}\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}R^1 & R^2 & R^3\\ S^1 & S^2 & S^3\end{bmatrix}=$$

$$\begin{bmatrix}0 & S^2R^1-S^1R^2 & S^3R^1-S^1R^3\\ R^2S^1-S^2R^1 & 0 & S^3R^2-S^2R^3\\ R^3S^1-S^3R^1 & S^2R^3-S^3R^2 & 0\end{bmatrix}$$

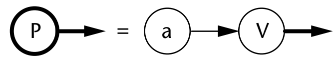
You can recognize the elements of this matrix as the components of the cross product of the two points **R** and **S**. But these are just the elements of the line-vector **L** arranged into an anti-symmetric matrix. Figure 26 shows this in diagram form.



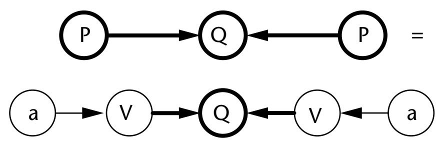
19 Putting together Figure 18 into a matrix.



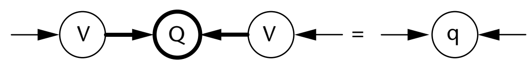
20 The diagram form of the cubic discriminant.



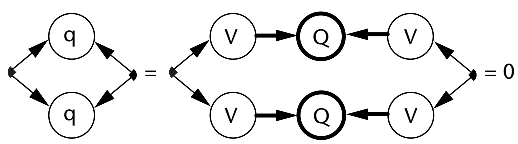
21 Diagram for a point on line L.



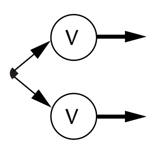
22 Diagram for the point also being on curve Q.



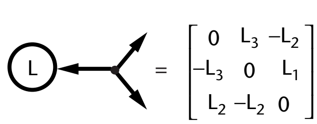
23 The symmetric quadratic polynominal matrix q.



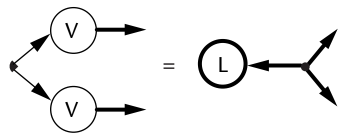
24 Diagram for a double root to q, and a single point of intersection for L.



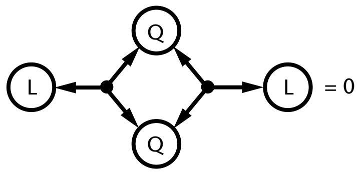
25 Extracting from Figure 24, the diagram fragment that represents line L.



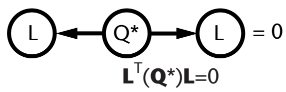
26 Converting L to an anti-symetric matrix.



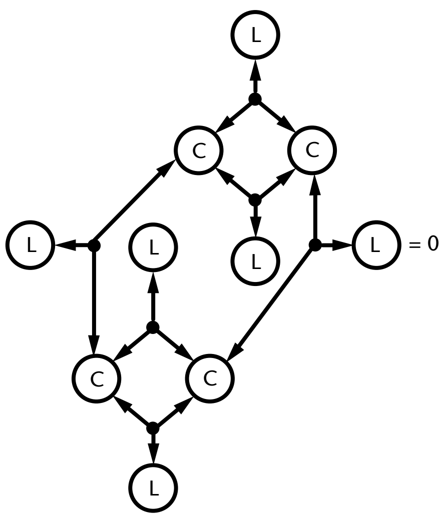
27 Two equivalent diagrams for a line.



28 The condition that line L is tangent to curve Q.



29 The test for line tangency uses of the adjoint of Q.



30 The condition that line L is tangent to curve C.

Figure 27 sums it up. Note that the right hand side of this doesn't require explicit points on **L**, so if all you have are the **L** components you don't need to explicitly find points on **L**. Putting Figures 24 and 27 together we get that the line **L** is tangent to curve **Q** if Figure 28 holds true.

Figure 28, without the **L** nodes, is just the expression of the adjoint of the matrix **Q** from Figure 8 (times minus two). In other words, while we use **Q** to test for point incidence, we use **Q*** to test for line incidence (tangency) as shown in Figure 29

Cubic with line

So, going up an order, what's the condition of line **L** being tangent a cubic curve **C**? That is, we want an expression involving the vector **L** and the cubic coefficient tensor **C** that's zero if **L** is tangent to **C**. With the groundwork we've laid, this is easy. First, compare Figures 14 and 28 to see how I converted the quadratic discriminant into a quadratic curve tangency equation. I just replaced each 2D epsilon with a 3D epsilon attached to a copy of **L**, and replaced **q** with **Q**. Now do the same thing with the discriminant of a cubic polynomial (Figure 20). Figure 30 shows what we get. Figure 30 represents a polynomial expression that is fourth order in **C** and sixth order in **L**. Since it has 18 arcs, the EIN version of this would require 18 index letters. All in all, it's something that would be rather difficult to arrive at in any other, nondiagram, way.

Notation, notation, and notation

A lot of the notational language of mathematics consists of the art of creative abbreviation. For example, a vector-matrix product **PT** is an abbreviation for a lot of similar looking algebraic expressions. However, clunky expressions like Figure 3b showed that this notation isn't powerful enough to allow us to easily manipulate the sort of expressions that we're encountering here. EIN has this power, but often gets buried under an avalanche of index letters. The tensor diagram method of drawing EIN is a better way to handle the index bookkeeping. What I've shown here is only the tip of the iceberg. There are a lot of other things that diagram notation can do well that I'll cover in future columns.

Our languages help form how we think. I believe that this notation can help us think about these and similar problems and allow us to come up with solutions that we wouldn't find any other way. ■

References

1. G.E. Stedman, *Diagram Techniques in Group Theory*, Cambridge University Press, Cambridge, England, 1990.
2. J. Blinn, *Jim Blinn's Corner: Dirty Pixels*, Morgan Kaufmann, San Francisco, 1998.
3. G. Kuperberg, "Involutory Hopf Algebras and 3-Manifold Invariants," *Int'l J. Mathematics*, vol. 2, no. 1, 1991, pp. 41-66.

Readers may contact Blinn by email at blinn@microsoft.com.