THE STRANGE CASE OF THE CRACOVIAN OPERATORS

A Note on the Early History of Linear Algebra and its Applications in Astronomy.

by Jacques F. Vallee

THE STRANGE CASE OF THE CRACOVIAN OPERATORS

A Question of Common Sense

When first exposed to linear algebra, and learning about matrix operations, most science students must have wondered, as I once did, why on earth mathematicians had invented such complicated ways to multiply together two tables of numbers that were often nothing more than the coefficients in a set of ordinary equations. But linear algebra is intimidating in its elegance and power, and few students have enough curiosity and temerity to raise their hand and ask their professor what would happen if we turned things around and multiplied these tables together in a common sense, natural way, by performing simple column multiplications.

If they did, and if their professors themselves had the curiosity to research the matter, they would find that for most practical applications of matrices the simplest method works just as well and indeed, leads to a lower chance of errors when numerical calculations are performed by hand. Not only is it more natural, but that is exactly the way such operations were used in the early part of this century, and an entire body of theory even existed on the class of mathematical creatures that could happily multiply in this uncomplicated way.

The creatures in question were not called matrices but "cracovians". For all practical purposes they could do all the things matrices did, they were much easier to use, and they had many amusing properties of their own.

There is no way to rediscover cracovian theory today except by accident. I stumbled across it when I was asked to compute stellar motion parameters for a large number of bright stars while serving as a programmer at Dearborn Observatory. Tracking down an obscure astronomical reference I found the formulas oddly unfamiliar, although they gave the correct result. A little more research in foreign astronomical journals of the twenties revealed cracovian theory in all its beauty, a mathematical land that time forgot.

Cracovian Operators in Astronomy

The inventor of cracovians, who must also be credited with the introduction of linear algebra (in cracovian form) into celestial mechanics, is the Polish astronomer, T. Banachiewicz. His fundamental papers appeared in the publications of the Cracov Observatory in 1923 and 1924, and in the Supplemento Internationale (1925).

Determinants were already known in the seventeenth century, but it took a long time for this new branch of mathematics to develop. Although linear algebra had been invented by Cauchy and his colleague Sturm, and although Cayley had introduced the theory, terminology and notation of matrices as early as 1857, it was not until 1925 that physicists Heisenberg, Born and Jordan applied them to problems which were not purely mathematical, and the subject did not become familiar to the engineering profession until the publication of the important paper by Duncan and Collar ("A Method for the Solution of Oscillation Problems by Matrices") in 1934.

Because they were faced with the necessity of performing long and intricate manipulations of spherical coordinates, astronomers had been the first to realize the need to treat problems of transformations in terms of arrays. However, astronomical formulae were by necessity oriented towards numerical computations, and the scheme of matrix multiplication was found cumbersome, error-prone and inconvenient in this respect.

The Polish astronomers felt that the requirements of hand calculations would be met, while mathematical elegance would be preserved, if arrays were written in such a way that their product would involve only column multiplications. This led T. Banachiewicz to develop the theory of <u>cracovian calculus</u>, which represented the first introduction of linear algebra in applied mathematics. His notations were used in many problems, and it was universally admitted that the multiplication of cracovians was more convenient than the multiplication of matrices, predisposing them for all effective <u>calculations</u>. Procedures for checking cracovian operations were given. Soon applications ranged from the derivation of the fundamental formula for Spherical Polygonometry to simpler methods of Least Squares solutions, surveying problems, the theory of the motion of the Moon, and many others.

After World War II, however, as linear algebra lost the character of a discipline exclusively reserved to the pure mathematician, the use of matrices became general. At the same time, the introduction of much more powerful computational techniques seriously limited the usefulness of the cracovian scheme, except in a few specific areas of numerical analysis. When I wrote that cracovians could do all the things matrices did, that was only true for simple calculations. But cracovian product is not associative. Furthermore, from a more theoretical point of view, Russian astronomer Bazhenow pointed out that the isomorphic correspondence of the multiplications between the linear transformations and the matrices did not extend to cracovians.

The cracovian formulae are forgotten today, but they served a useful purpose as the forerunners of the modern treatment of matrix algebra. As numerical instruments, they were of considerable interest, as can be seen from a survey of computing techniques used in the world's observatories in 1935, where 25 out of 27 astronomers who had used it found the technique useful: Lowell Observatory, for example, answered: "We have recently become quite impressed by the elegance and efficiency of the cracovian scheme."

We propose to provide some details about cracovian calculus, and to show how it was used in practice in the computation of stellar velocities.

Elements of Cracovian Calculus

To satisfy the requirements stated above, we are led to write linear transformations in the form of arrays which involve only column multiplications, for example:

$$\left\{\begin{matrix} x \\ y \\ z \end{matrix}\right\} \left\{\begin{matrix} a11 & a21 & a31 \\ a12 & a22 & a32 \\ a13 & a23 & a33 \end{matrix}\right\} = \left\{\begin{matrix} x' \\ y' \\ z' \end{matrix}\right\} \qquad (1)$$

with:

$$x' = a11.x + a12.y + a13.z$$
$$y' = a21.x + a22.y + a23.z \qquad (2)$$
$$z' = a31.x + a32.y + a33.z$$

While we are used to representing the same linear transformation in the form:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a11 & a12 & a13 \\ a21 & a22 & a23 \\ a31 & a32 & a33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad (3)$$

By comparing the arrays in expressions (1) and (3) we recognize that the cracovian of the linear transformation is the transpose of the usual matricial notation. Now, let cracovian multiplication be defined by the simple rule:

$$c_{ij} = \sum_k a_{kj} \; b_{ki} \qquad (4)$$

as opposed to the rule for matrix multiplication:

$$c_{ij} = \sum_k a_{ik} \; b_{kj} \qquad (5)$$

Thus, given two arrays:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$$

Their product as matrices is:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix} \qquad (6)$$

while their product as cracovians is:

$$\begin{Bmatrix} a & b \\ c & d \end{Bmatrix} \begin{Bmatrix} p & q \\ r & s \end{Bmatrix} = \begin{Bmatrix} ap+cr & bp+dr \\ aq+cs & bq+ds \end{Bmatrix} \qquad (7)$$

In the resulting array of expression (7) we recognize the matrix product where the second array has been transposed:

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ap+cr & bp+dr \\ aq+cs & bq+ds \end{pmatrix} = {}^{t}B \times A \qquad (8)$$

Thus, if we denote by @ the cracovian product, we have the formula:

$$\boxed{A \mathbin{@} B = {}^{t}B \times A} \qquad (9)$$

Now, let T be the cracovian unity:

$$T \mathbin{@} A = \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} @ \begin{Bmatrix} a & b \\ c & d \end{Bmatrix} = \begin{Bmatrix} a & c \\ b & d \end{Bmatrix} \qquad (10)$$

Thus, cracovians exhibit a property which is without equivalent in matrix algebra: premultiplication by the cracovian unity directly provides the transpose.

Coming back to the cracovian product, we deduce from (8) the converse property:

$$\boxed{A \times B = B \mathbin{@} {}^{t}A} \qquad (11)$$

Having computed A@B, let us now compute B@A:

$$B \mathbin{@} A = \begin{Bmatrix} p & q \\ r & s \end{Bmatrix} @ \begin{Bmatrix} a & b \\ c & d \end{Bmatrix} = \begin{Bmatrix} ap+cr & aq+cs \\ bp+dr & bq+ds \end{Bmatrix} \qquad (12)$$

Comparing this result with expression (6) we observe that,

if $C = A@B$, then $B@A = T@C$.

In other words,

$$\boxed{A \; @ \; B = {}^{t}(B \; @ \; A)} \qquad (13)$$

So far, we have used only square arrays, which are simpler for illustration purposes. But in the more general case we would observe that, while the product of two matrices AxB is not necessarily defined whenever BxA is defined, both left and right multiplication are always defined together for the cracovians, which is another interesting advantage over matrices.

<u>The cracovians as rotation operators.</u>

Consider figure 1 and call (a,b,c) the coordinates of point M before rotation of the axes, while (a′, b′, c′) are its coordinates after rotation. If Ox is the axis of the rotation of angle α, we can write:

$$\begin{aligned} a' &= a \\ b' &= b \cos\alpha + c \sin\alpha \\ c' &= c \cos\alpha - b \sin\alpha \end{aligned} \qquad (14)$$

The matricial notation of this transformation is well known. In terms of cracovians, however, we will use formula (1) and write:

$$\begin{Bmatrix} a' \\ b' \\ c' \end{Bmatrix} = \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} @ \begin{Bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{Bmatrix} \qquad (15)$$

This cracovian operator will be denoted by $p(\alpha)$.

Similarly, we shall denote by $q(\alpha)$ and $r(\alpha)$ the cracovian operators of the rotations of angle α with Oy (resp. Oz) as the axis.

We are thus led to write:

$$q(\alpha) = \begin{Bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{Bmatrix} \qquad (16)$$

and

$$r(\alpha) = \begin{Bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{Bmatrix} \qquad (17)$$

These operators are basic in the computation of stellar velocity components.

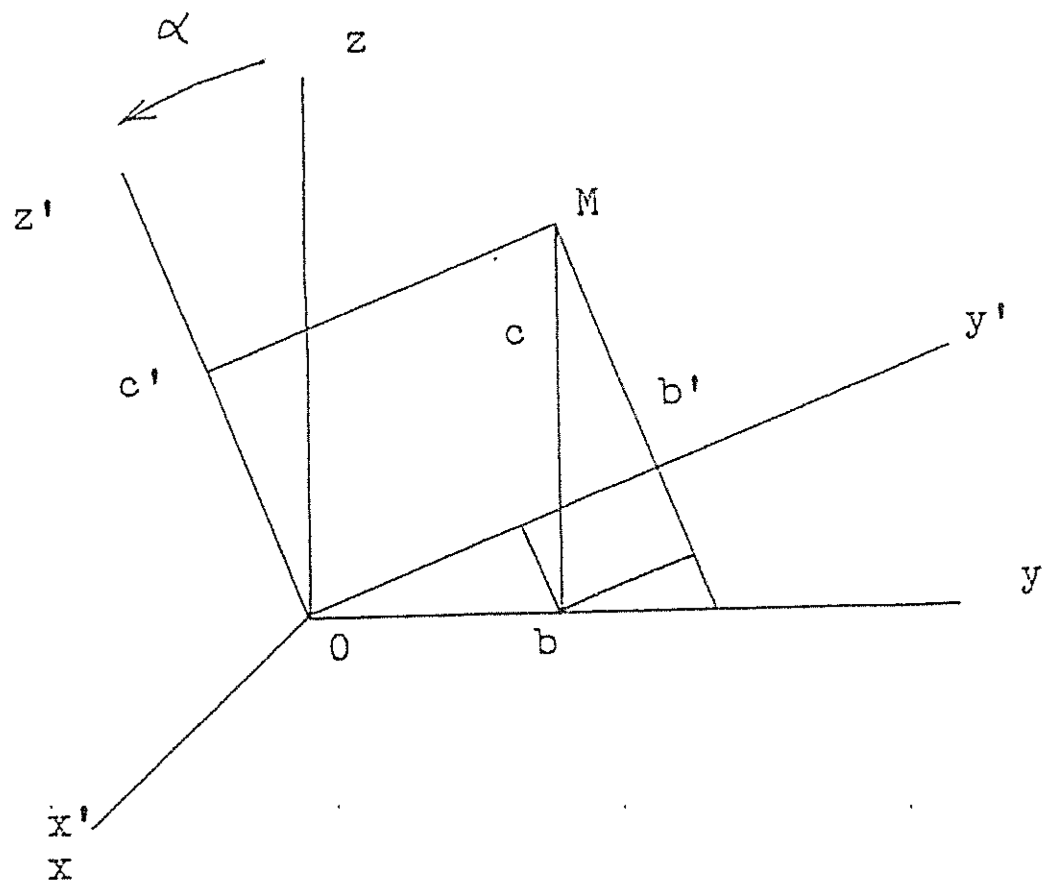


Figure 1. Rotation of angle α with Ox as axis. (Cracovian p)

It should be noted that, because cosines occur only on the diagonals of all three operators p,q and r, we have immediately:

$$q(-\alpha) = T \; @ \; q(\alpha) = \text{transpose of } q(\alpha). \qquad (18)$$

<u>The use of cracovians in the computation of space velocity.</u>

The knowledge of four parameters is essential in determining the direction and amplitude of the velocity vector of a star: its distance from the sun (which, if expressed in parsecs, is the reciprocal of its parallax π in seconds of arc); its radial velocity R, and the components μ_α and μ_δ of its tangential velocity. These quantities being known, the problem arises of transforming these four parameters of motion (after correction for the earth's own motion) into velocity components defined by reference to the equatorial system, and then to the galactic standard of rest.

The ratios μ_α/π and μ_δ/π represent the tangential velocity components in astronomical units per year. If k = 4.738 is the speed in km/s that corresponds to 1 A.U./year, then the tangential velocity components are obtained from the proper motion components by the formulae:

$$T_\alpha = k\mu_\alpha/\pi \qquad \text{and} \qquad T_\delta = k\mu_\delta/\pi \qquad (19)$$

These quantities, like radial velocity R, are now expressed in km/s and are referred to the star's own system. It is brought into coincidence with the equatorial system by two rotations (α , δ) giving the equatorial linear components (x,y,z) of the stellar velocity. But two more rotations of the coordinate system are required in order to bring this system to coincide with the galactic axes, yielding the three galactic components u,v and w.

We need only say a word of the classical trigonometric method, which involves the computation of numerous intermediate parameters, in particular the galactic coordinates of the star, and which lends itself with difficulty to high-speed computation: If μ_b and μ_ℓ are computed along with the galactic coordinates b and ℓ by suitable changes of spherical coordinates, we can calculate $T_\ell = k \mu_\ell / \pi$ and $T_b = k \mu_b / \pi$, then the projection:

$$V_\rho = R \cos b - V \sin b \tag{20}$$

and finally:

$$\left[\begin{array}{l} u = V_\rho \cos \ell - V_\ell \sin \ell \\ v = V_\rho \sin \ell + V_\ell \cos \ell \\ w = R \sin b + V_\rho \cos b \end{array}\right. \tag{21}$$

In practice, a great number of intermediate steps (where precision often deteriorates) are necessary to evaluate μ_ℓ , μ_b , ℓ and b by this process, and the successive projections and changes of coordinates, although mathematically convenient, are very cumbersome.

On the contrary, methods using linear algebra have great elegance and lend themselves easily to high-speed calculation using digital computers. In the approach A. Przybylski has proposed in Acta Astronomica vol. 12 (1962) no.4, we perform four rotations with appropriate axes to bring the x-axis in the direction of the galactic center C; the y-axis in the direction of galactic rotation; and the z-axis in the plane Galactic Center - Galactic North Pole - earth. The corresponding sequence of cracovian operations is, with the notations of formulae (15,16 and 17):

$$\begin{Bmatrix} u \\ v \\ w \end{Bmatrix} = \begin{Bmatrix} R \\ T\alpha \\ T\delta \end{Bmatrix} @ \; q(\delta) \; @ \; r(\alpha_c - \alpha) \; @ \; [p(\eta) \; @ \; q(\delta_c)] \qquad (22)$$

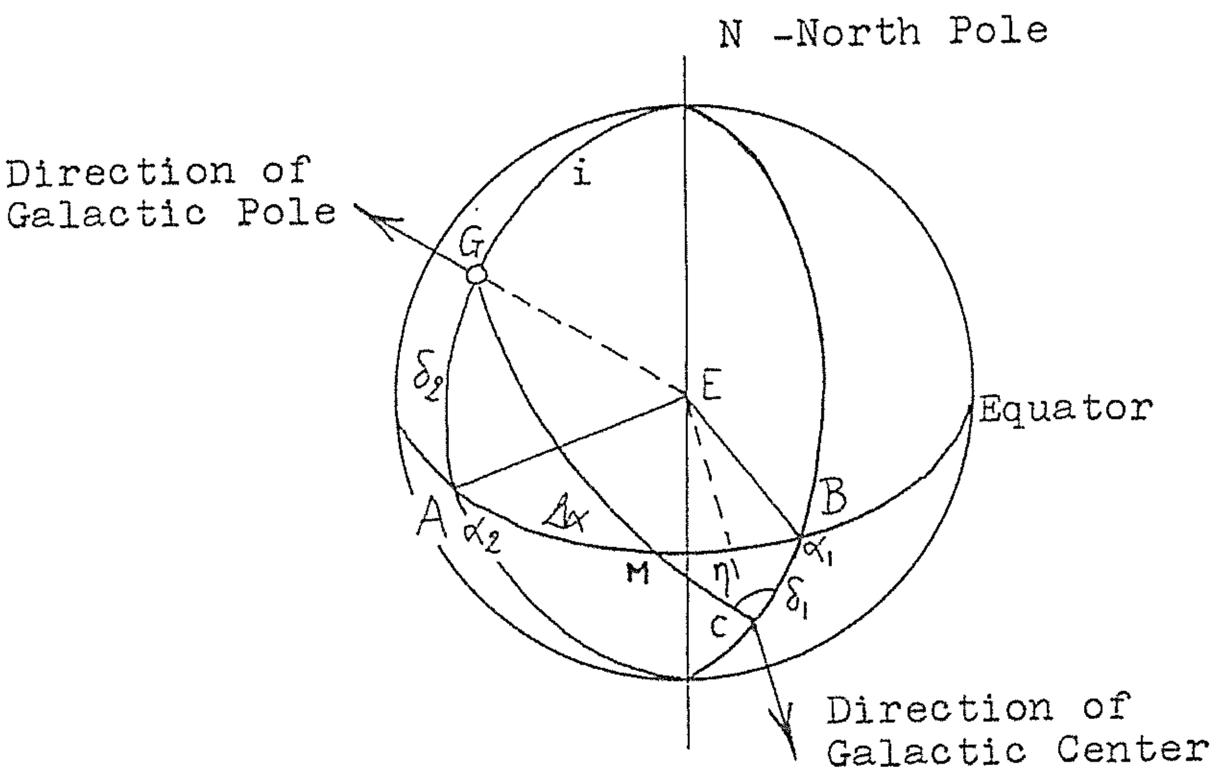
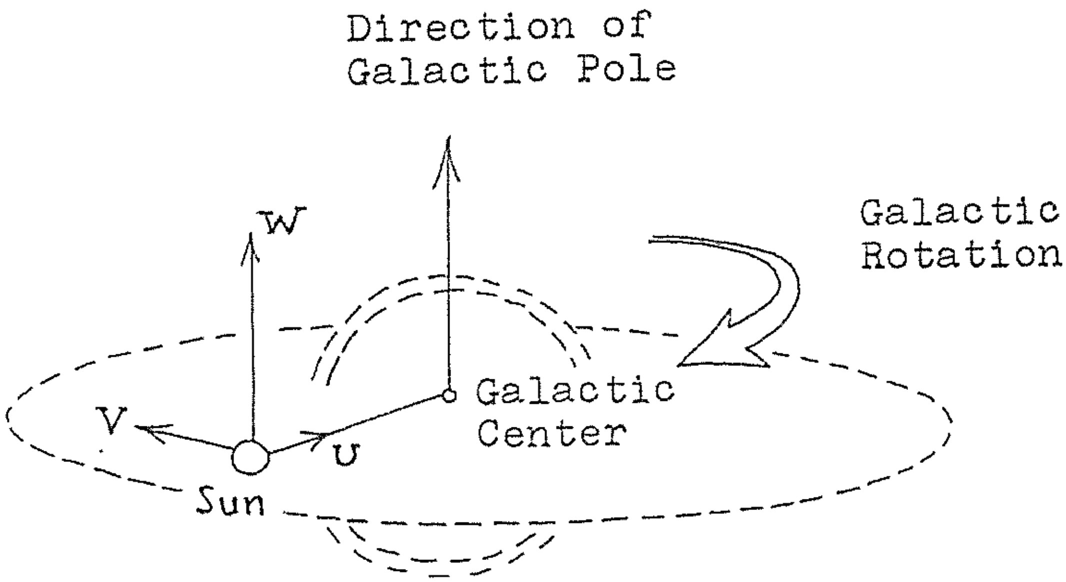


Figure 2: The Galactic System - Definition.



where α_c, δ_c and η are constants, respectively: the right ascension and declination of the Galactic Center and the parallactic angle (which is the angle Galactic North Pole - Galactic Center - North Pole as shown on Figure 2 and as defined in Monthly Notices 121, 123, 1960).

A. Przybylski remarked in his paper that, while the classical formulae for the computation of these quantities (for instance, those given by Smart in Stellar Dynamics, 1938, p.14) require the computation of the galactic coordinates of the star and its galactic parallactic angle, the cracovian scheme made the computation of intermediate data unnecessary and required only sixteen multiplications and nine additions. The last two cracovians in formula (22), it will be noted, contain only constant values and can be computed once and for all. No wonder cracovian operators were popular with astronomers before the Second World War. The technique brought significant savings in time, fewer opportunities for mistakes and greater accuracy.

When the electronic computer became available, matrix computation was readily reduced to software and the need for cracovians vanished. It is unfortunate that their history was forgotten at the same time, if only because knowledge of this amusing alternative to matrix algebra might help make the teaching of this branch of mathematics more interesting to students.

In my own work I went ahead and wrote the stellar velocity reduction program using cracovians instead of matrices. The calculations posed no problems, but anyone stumbling on the program source code without warning and trying to understand it may have experienced serious puzzlement and perhaps an unfortunate headache.