

Matrix Algebra

Matrices, Tensors, and Coordinate Transformations

Basics

For sake of clarity we do not write variables in *italics* in this module

This is not a course in matrix algebra (including vector and tensor calculus), but a quick reminder, assuming you know the basic facts of life here.

- We also cut a lot of corners, not distinguishing much between matrices (a mathematical object) and tensors (a physical object), "true" vectors and "polar" vectors, Cartesian and non-Cartesian coordinate systems and the like.
- We will deal with some topics of matrix algebra roughly in the sequence they come up in the backbone chapters

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

- A matrix then is an assembly of nine numbers arranged as shown on the left.

In a simplified way of speaking, a **matrix** (or better **tensor**) allows to correlate **vectors** in a simple linear way.

- Every component of the vector $\underline{r} = (r_1, r_2, r_3)$ can be expressed as a linear function of all components of a second vector $\underline{t} = (t_1, t_2, t_3)$ by the equations

$$\begin{aligned} r_1 &= a_{11} \cdot t_1 + a_{12} \cdot t_2 + a_{13} \cdot t_3 \\ r_2 &= a_{21} \cdot t_1 + a_{22} \cdot t_2 + a_{23} \cdot t_3 \\ r_3 &= a_{31} \cdot t_1 + a_{32} \cdot t_2 + a_{33} \cdot t_3 \end{aligned}$$

- In matrix notation we simple write

$$\underline{r} = \mathbf{A} \cdot \underline{t}$$

- With \mathbf{A} being the symbol for the matrix defined above.
- We then have already defined how a matrix is multiplied with a vector and that a new vector is the result of the multiplication.

The matrix \mathbf{A} , if interpreted as an entity that relates two vectors with each other, must have *certain properties* that are not required for a general matrix (that might express, e.g., the coefficients of a linear system of equations with several unknowns).

- If we change the coordinate system in which we express the vectors, the components of the vectors will be different numbers, but the vectors themselves (the arrows) stay unchanged. This imposes some conditions on the set of nine numbers - the matrix - connecting the components of the vectors and *any matrix meeting these conditions we call a tensor*.
- A *tensor* thus is a set of nine numbers, and the numerical value of these numbers depends on the coordinate system in which the tensor is expressed. If we do a coordinate transformation, the numerical values of the nine components must then transform in a specific way.

Transforming a coordinate system into another one is done by matrices as follows:

- If the first vector \underline{r} is chosen to be one of the *unit vectors* defining some Cartesian coordinate system, the second vector \underline{r}' obtained by multiplying \underline{r} with the transformation matrix \mathbf{T} , can be interpreted as the *unit vector* of some new coordinate system
- The set of unit vectors \underline{r}_i with $i = x, y, z$ will be changed to a new set \underline{r}'_i by

$$\underline{r}'_i = \mathbf{T} \cdot \underline{r}_i$$

- and **T** is called the **transformation matrix**. It is clear that **T** must have certain properties if the \underline{r}'_i are also supposed to be unit vectors
- While this is clear, it is not so clear what we have to do if we want to reverse the transformation. The simple thing is to write

$$\underline{r}_i = (\mathbf{T}^{-1}) \cdot \underline{r}'_i$$

- and defining (\mathbf{T}^{-1}) to be the **inverse matrix** to **T** so that the operation can be reversed.

But how do we calculate the numerical values of the components of \mathbf{T}^{-1} if we know the numerical values of the components of **T** ??

- In order to be able to give a simple formula, we first have to introduce something else, the **determinant of a matrix**

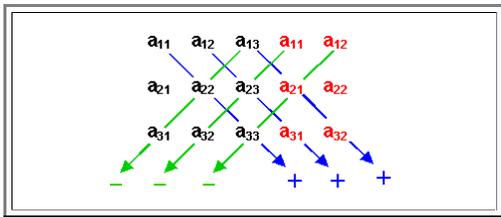
The Determinant of a Matrix

The determinant **|A|** of a matrix **A** is a *single number* calculated by summing up the diagonal products in a special fashion.

- For a **3 x 3** matrix we have

$$|A| = a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} - a_{13} \cdot a_{22} \cdot a_{31} - a_{11} \cdot a_{23} \cdot a_{32} - a_{12} \cdot a_{21} \cdot a_{33}$$

- Look at the written matrix **A** above and you see that you start by doing the products by going down diagonally from left to right, adding the products of the three possible diagonals - always completing a diagonal by repeating the matrix if necessary. Then you subtract the product you obtain by going down the diagonal from right to left.
- This sounds more complicated as it is; graphically it looks like this:



The determinant of a matrix obtained in this way is a number that comes up a lot in all kinds of matrix operation; the same is true for a related quantity, the **subdeterminant A_{ik}** of the matrix **A**

- There are as many subdeterminants as there are elements in the matrix. A_{ik} is obtained by
 1. Erasing the line and the row that contains the element a_{ik} and calculating the determinant of the matrix that remains, and
 2. multiplying the number obtained by $(-1)^{i+k}$

With the concept of a subdeterminant, we can also define the **rank of a matrix**:

- The rank of a matrix is the number of row (or columns, resp.) of the determinant or largest subdeterminant with non-zero value. In other words, the rank of a **3 x 3** matrix **A** is **rank(A) = 3** if **|A| ≠ 0**; if **|A| = 0**, you look for the largest subdeterminant

With determinant and subdeterminant, the *inverse matrix is easy to formulate* :

- The inverse matrix \mathbf{A}^{-1} to **A** has the elements $(a_{ik})^{-1}$ given by

$$(a_{ik})^{-1} = \frac{A_{ki}}{|A|}$$

- i.e. the value of the respective subdeterminant divided by the value of the determinant. Note that the indexes are interchanged ("**ik**" ⇒ "**ki**"); and that the "**-1**" must be read as "*inverse*", it is not an exponent!!!
- We will not prove it here; but it is not too difficult - just solve the system of equations given above for the \underline{t}_i .

Two more important points follow directly:

- An inverse matrix (A^{-1}) to A only exists if the determinant of A is not zero!
- The product of A^{-1} and A results in the **identity matrix I**

$$A^{-1} \cdot A = I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The last claim is unproved, we first need the multiplication rule for matrices to prove it.

- Multiplication of the matrix A with the matrix B gives a new matrix C ; and the element c_{ik} of C is obtained by taking the scalar product of the "line" or "row" vector in row i of matrix A times the column vector of column k of matrix B . This is best seen in a kind of graph:

$$\begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & c_{32} & \times \end{pmatrix} = \begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ a_{13} & a_{12} & a_{32} \end{pmatrix} \cdot \begin{pmatrix} \times & b_{12} & \times \\ \times & b_{22} & \times \\ \times & b_{32} & \times \end{pmatrix}$$

- Now it is still fairly messy, but straightforward to prove the claim from above - you may want to try it.

A useful relation is that the multiplication of any matrix with the identity matrix I doesn't change anything.

$$I \cdot A = A$$

- And this is also true for multiplying a vector with I :

$$I \cdot \underline{r} = \underline{r}$$

From the various definitions you may get the feeling, that signs are important and possibly tricky. Well, that's true.

- Matrix multiplication, in general is not commutative, i.e. $A \cdot B \neq B \cdot A$ - you must watch out if you multiply from the left or from the right.
- Still, we now can solve "mixed" vector - matrix equations. Take, for example

$$\underline{r}_0 = A^{-1} \cdot \underline{r} + \underline{I}(l)$$

- Multiplying from the right with I yields.

$$I \cdot \underline{r}_0 = I \cdot A^{-1} \cdot \underline{r}_0 + I \cdot \underline{I}(l)$$

$$I \cdot \underline{r}_0 = A^{-1} \cdot \underline{r}_0 + I \cdot \underline{I}(l)$$

- That looks a bit stupid, but with this cheap trick we now we have only tensors in connection with \underline{r}_0 , which means we now can combine the "factors" of \underline{r}_0 , giving

$$(I - A^{-1}) \cdot \underline{r}_0 = \underline{I}(l)$$

our **O**-lattice theory master equation.

One last important property of transformation matrices is that their determinant gives directly the volume ratio of the unit cells:

$$|A| = \frac{V(\text{after the transformation})}{V(\text{before the transformation})}$$

This is not particularly easy to see, but simply consider two points:

- 1. The base vector $\underline{a}(I)$ is transformed to the base vector $\underline{a}(II)$ via

$$\underline{a}_i(II) = A \cdot \underline{a}_i(I)$$

- 2. The volumes V of elementary cells is given by

$$V = (\underline{a}_1 \times \underline{a}_2) \cdot \underline{a}_3$$

- Since we produce the O -lattice from a crystal lattice with the matrix $I - A^{-1}$, the volume V_O of an O -lattice cell (in units of the volume of a crystal unit cell) is

$$\frac{1}{V_O} = |I - A^{-1}|$$

- Again, as remarked above; watch out for signs.