3. Linear algebra

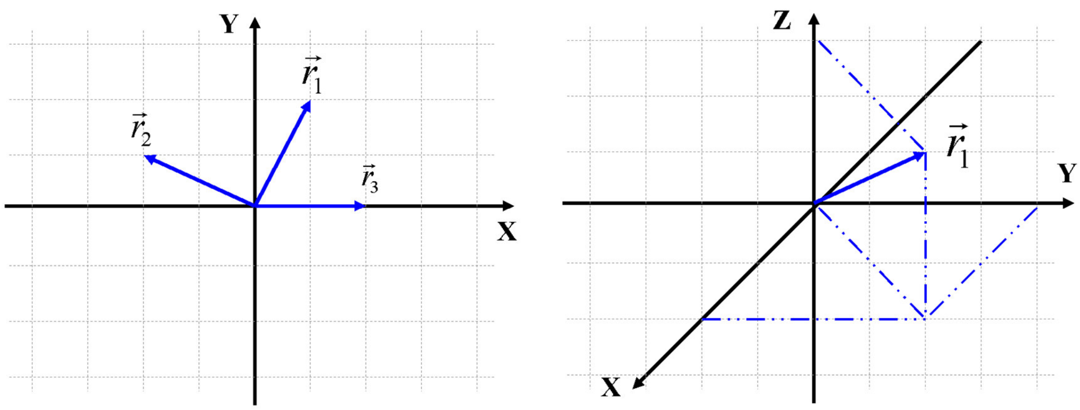
3.1. Vectors

Definition: A **vector in N-dimensional space**, $\vec{a}$, is essentially a collection N elements, organized in a column, $\begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_N \end{pmatrix}$. The elements a_n are called **projections** (or **components**).

Vector projections can be either real numbers (then, it is said that the vector $\vec{a}$ is defined **over the field** of real numbers R) or complex numbers (then, it is said that the vector $\vec{a}$ is defined over the field of complex numbers C). Vector components can be real of complex-valued functions, matrices, other vectors, and so on.

Examples:

One can easily visualize vectors in 2D or in 3D, but it may be not so easy to visualize vectors in higher dimensions. In fact, such visualizing is not possible directly. Instead, one has to recall the definition of the vector. It is a collection of N elements. This means, that a vector in a 4-D space is just a set of 4 numbers (or functions). For many purposes in chemical physics/physical chemistry, the "visualization" is possible. Just imagine all points in lower-dimensional space at a given configuration as a single point in a higher dimensional space. Fig. 1 shows 3 points in 2D space. However, this same picture also illustrates several points in 4D space. For instance, $\vec{a} = \begin{pmatrix} \vec{r}_1 \\ \vec{r}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -2 \\ 1 \end{pmatrix}$ or $\vec{b} = \begin{pmatrix} \vec{r}_1 \\ \vec{r}_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 0 \end{pmatrix}$.



(a) (b)

Figure 1. (a) Vectors in 2D: $\vec{r}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\vec{r}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$, $\vec{r}_3 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$. (b) Vector in 3D: $\vec{r}_1 = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}$.

All coordinates of N particles in 3D spaces form a 3N-dimensional vector:

$\vec{R} = \begin{pmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \dots \\ \vec{r}_N \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & z_1 & \dots & x_N & y_N & z_N \end{pmatrix}^T$. Each such point is called a **configuration**. All possible configurations form (or belong to) a **configurational space** (see for more formal definition of "spaces" below). Figure 2 illustrates 4 72-dimensional vectors. These are 4 configurations from the 72-dimensional configurational space of an azobenzene molecule

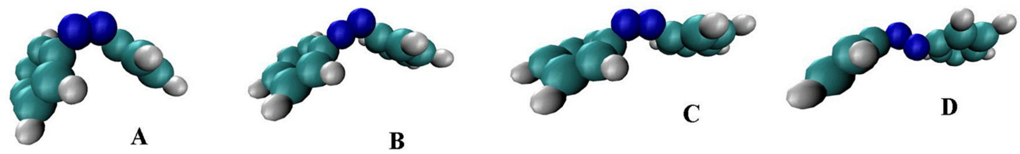


Figure 2. Illustration of 4 points (configurations) in a 72-dimensional configurational space of an azo-benzene molecule.

3.2. Vector spaces

<u>Definition</u>: **Vector (linear) space**, V, is a set of all vectors equipped with the **vector addition** and **scalar multiplication** operations, such that for $\forall x, y, z \in V$ and $\forall a, b \in R(C)$:

- associativity of addition $x + (y + z) = (x + y) + z$,
- commutativity of addition $x + y = y + x$
- existence of zero vector $\exists 0 \in V : 0 + x = x$
- inverse element of addition $\exists (-x) \in V : (-x) + x = 0$

- compatibility of scalar multiplication $a \cdot (b \cdot x) = (a \cdot b) \cdot x$
- existence of identity element of scalar multiplication $\exists 1 : 1 \cdot x = x$
- distributivity w.r.t to vector addition $a \cdot (x + y) = a \cdot x + a \cdot y$
- distributivity w.r.t to scalar addition $(a + b) \cdot x = a \cdot x + b \cdot x$

<u>Note</u>: the vector sign is often omitted in the writings, so instead of $\vec{x}$ we can just use x, but remembering that this is still a vector!

<u>Note</u>: an **element** of a linear space is called **vector**.

Examples:

R^3 - 3D space (Cartesian)

$C[a,b]$- a space of all functions continuous on the $[a,b]$interval

$\{\hat{A}\}$- a space of linear operators

$\{A_n\}$- a space of n-dimensional square matrices

C^n - a space of n complex numbers

$\{P_n\}$- a space of n-order polynomials

Definition: Vectors $\vec{a},\vec{b},\vec{c},...$ are **linearly independent** if $\alpha \cdot \vec{a} + \beta \cdot \vec{b} + \gamma \cdot \vec{c} + ... = 0$is satisfied only for $\alpha = \beta = \gamma = ... = 0$.

Otherwise, they are linearly-dependent, since at least one vector can be expressed as a linear superposition of all other vectors: e.g. if $\alpha \neq 0$, then $\vec{a} = -\frac{\beta}{\alpha}\vec{b} - \frac{\gamma}{\alpha}\vec{c} - ...$

Definition: The **dimensionality** of a linear (vector) space is the maximal number of linearly-independent vectors.

Examples: $C[a,b]$ - dimensionality is infinite, $\{P_n\}$- dimensionality is n.

Definition: **Basis** of an n-dimensional vector space is a set of n linearly-independent vectors: $\{e_1, e_2, ..., e_n\}$

Note: basis is not unique and can be chosen in infinite (most of the times) number of ways

Note: the **significance of the basis** is that any vector f belonging to the n-dimensional space can be represented as a linear combination (**superposition**) of the basis vectors: $f = \sum_{i=1}^{n} c_i e_i$.

My Proposition: Mind – is a superposition of the conceptual basis (basis element = a concept). The dimensionality of mind space is infinite. The more concepts you internalize, the higher the dimensionality of your mind becomes.

3.3. Euclidian, metric, and Hilbert spaces

Definition: The mapping $(\cdot,\cdot): V \times V \to R(C)$ is called a **scalar product** if:

1. $(x, y) = (y, x)$ or $(x, y) = (y, x)^*$ for C

2. $(x + y, z) = (x, z) + (y, z)$

3. $(\alpha \cdot x, y) = \alpha \cdot (x, y)$

4. $(x, x) \geq 0$, such that $(x, x) = 0 \Leftrightarrow x = 0$ (otherwise, it is called a **half-scalar product**)

Definition: Vectors $x, y \in V$ are called **orthogonal** ($x \perp y$) if $(x, y) = 0$

Examples:

$(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3 + ...$ is good for R, but not for C (take a vector $(0 \quad 0 \quad i)^T$ - it will violate the condition 4).

$(x, y) = x_1^* y_1 + x_2^* y_2 + x_3^* y_3 + ...$ is good also for C

$(f, g) = \int_a^b f^*(x) g(x) dx$ is good for $V = C[a, b]$

Definition: A linear space equipped with a scalar product is called **Euclidian space**

Definition: $\rho(\cdot,\cdot): V \times V \to R$ is called a **metric** if:

1. $\rho(x, y) = \rho(y, x)$

2. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$

3. $\rho(x, y) \geq 0$ with $\rho(x, x) = 0 \Leftrightarrow x = 0$

Definition: A linear space equipped with a metric is called **metric space**

Definition: A **norm** of a vector in a Euclidian space is: $\|x\| = \sqrt{(x, x)}$

Definition: **Hilbert space** - an infinitely-dimensional Euclidian space

Example: $H = \{\psi \mid \psi \in C[a, b], \|\psi\| < \infty\}$

3.4. Inequalities

Schwarz: $\forall x, y \in V$, $\|(x, y)\| \le \|x\| \cdot \|y\|$

Proof: Consider a scalar product of form $(x+ty, x+ty)$, with an arbitrary real number $\forall t \in R$. Using the properties of the scalar product, we can simplify it:
$(x+ty, x+ty) = (x,x) + 2t(x, y) + t^2(y, y) \ge 0$. Here we also used the non-negativity property of the scalar product. Since this inequality holds for any parameter t, there must be additional restriction connecting scalar products (x,x), (x, y), and (y, y). Consider two cases:

a) if $(y, y) = 0$, then, according to properties of the scalar product, we have: $y = 0 \Rightarrow (x, y) = 0$, so the inequality $(x,x) \ge 0$ holds true.
b) if $(y, y) \ne 0$, then, using the temporary variables $a = (y, y)$, $b = (x, y)$, and $c = (x,x)$, the inequality can be transformed: $at^2 + 2bt + c = a\left(t + \frac{b}{a}\right)^2 + c - \frac{b^2}{a} \ge 0$. According to properties of the scalar product, $a = (y, y) \ge 0$, so $a\left(t + \frac{b}{a}\right)^2 \ge 0$ is always satisfied. To satisfy the inequality for any value of the parameter t, we require that $c - \frac{b^2}{a} \ge 0 \Leftrightarrow a \cdot c - b^2 \ge 0 \Leftrightarrow a \cdot c \ge b^2$. Recalling the definition of the auxiliary variables and the definition of a vector norm, we obtain: $\|(x, y)\| \le \|x\| \cdot \|y\|$

Cauchy-Bunyakovsky:

These inequalities are the direct consequences of the Schwarz inequality, when applied to specific types of vector spaces.

$V = R^n$ then $\left|\sum_{i=1}^{n} x_i \cdot y_i\right| \le \sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2}$

$V = C[a,b]$ then $\left|\int_a^b f^*(x)g(x)dx\right| \le \sqrt{\int_a^b f^*(x)f(x)dx} \cdot \sqrt{\int_a^b g^*(x)g(x)dx}$

Pythagoras theorem: If $x \perp y \Leftrightarrow (x, y) = 0$ then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$

Proof: is straightforward, comes from definitions:
$\|x + y\|^2 = (x + y, x + y) = (x, y) + 2(x, y) + (y, y) =| (x, y) = 0 |= \|x\|^2 + \|y\|^2$

Triangle inequality: $\forall x, y \in V \Rightarrow \|x+y\| \le \|x\| + \|y\|$

Proof: $\|x+y\|^2 = (x+y, x+y) = (x,x) + 2(x,y) + (y,y)$. Recalling that $\|(x,y)\| \le \|x\| \cdot \|y\|$, we obtain:

$\|x+y\|^2 = (x+y, x+y) = (x,x) + 2(x,y) + (y,y) \le \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$. Taking square roots of the both sides, we prove the triangle inequality.

3.5. Matrices

Matrix multiplication: $X = AB \Leftrightarrow X_{ij} = \sum_k A_{ik} B_{kj}$. Algorithm: take all elements of a row i of the matrix A and all elements of a column j of the matrix B. Multiply them, one by one. Add all products up to a single sum. Put the sum into i,j-th element of the product matrix.

Multiplication is possible only if the dimensions of the matrices match (Fig. 3a). Also, depending on the order of matrices, you may get either a number of a whole matrix (Fig. 3b).

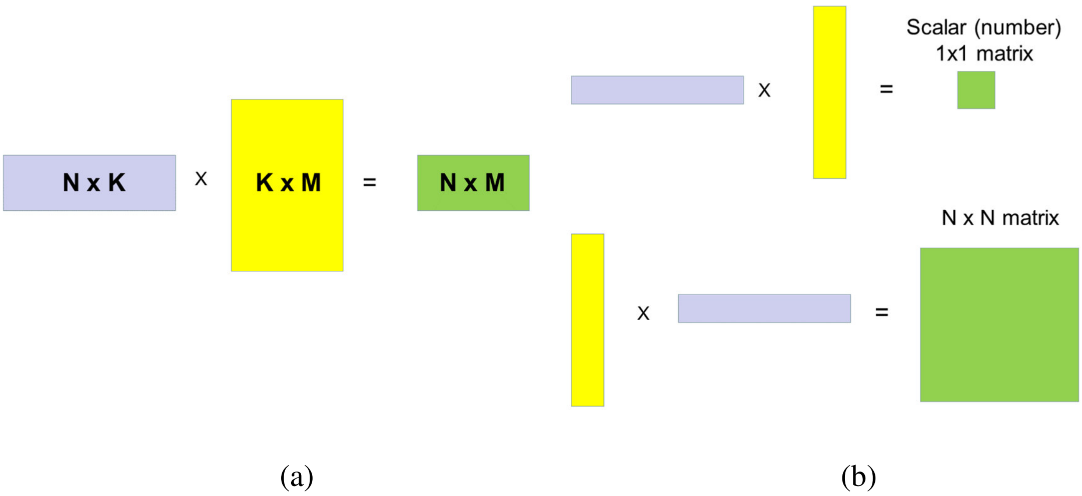


Figure 3. Illustration of the matrix multiplication.

<u>Note</u>: A vector can be considered a N x 1 matrix (a column-vector).

Operations on matrices:

Transpose: A^T Elements are defined as: $(A^T)_{ij} = A_{ji}$

Complex conjugation: A^* Elements are defined as: $(A^*)_{ij} = (A_{ij})^*$

Hermitian conjugation: $A^{+} = \left(A^{T}\right)^{*} = \left(A^{*}\right)^{T}$ Elements are defined as: $\left(A^{+}\right)_{ij} = \left(A_{ji}\right)^{*}$

Useful relationships: $(AB)^{-1} = B^{-1}A^{-1}$, $(AB)^{T} = B^{T}A^{T}$, $(AB)^{+} = B^{+}A^{+}$, but $(AB)^{*} = A^{*}B^{*}$

3.6. Elements of the field theory

Definition: If $u = u(x,y,z)$ is a scalar field (a "normal" function), then **gradient** is defined as:

$$grad(u) = \vec{\nabla}u = \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial u}{\partial z} \end{pmatrix}.\ \text{Here, } \vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \text{ is a } \textbf{gradient operator}.$$

Note: gradient is a vector

Definition: If $\vec{a} = \left(a_x(x,y,z), a_y(x,y,z), a_z(x,y,z)\right)^{T}$ is vector field (a multi-component function), then the **divergence** is defined as:

$$div(\vec{a}) = \left(\vec{\nabla}, \vec{a}\right) = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}$$

Note: divergence is a scalar

Definition: If $\vec{a} = \left(a_x(x,y,z), a_y(x,y,z), a_z(x,y,z)\right)^{T}$ is vector field (a multi-component function), then the **rotor** is defined as:

$$rot(\vec{a}) = \vec{\nabla} \times \vec{a} = \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \\ i & j & k \end{vmatrix}$$

Note: rotor is a vector

3.7. Linear operators

Definition: A mapping $\hat{F} : X \to X'$, that is $\forall \varphi \in X, \psi = \hat{F}\varphi \in X'$ is called an **operator**.

Definition: An operator $\hat{F} : X \to X'$ that obeys the relationship $\hat{F}(c_1\psi_1 + c_2\psi_2) = c_1\hat{F}(\psi_1) + c_2\hat{F}(\psi_2)$, where $\psi_1, \psi_2 \in X; c_1, c_2 \in R(C)$ is called a **linear operator**

Linear operators have the following properties:

$\left(\hat{F}+\hat{G}\right)\psi=\hat{F}\psi+\hat{G}\psi$

$\left(c\hat{F}\right)\psi=c\left(\hat{F}\psi\right)$

$\left(\hat{F}\hat{G}\right)\psi=\hat{F}\left(\hat{G}\psi\right)$

The last property simply states that the action of a composite operator (a "product" of two) is computing from the right: the rightmost operator ($\hat{G}$) first acts on the vector to which the overall operator is applied (ψ) to produce a new vector ($\varphi=\hat{G}\psi$). After that, the next operator from the left ($\hat{F}$) is acting on the resulting vector (φ) to produce a final result.

<u>Examples</u>:

- $\hat{x}: f \to xf$ (multiplication by number)
- $\frac{d}{dx}: f \to f'$ (derivative)
- But the operator $\hat{F}: f \to f^2$ is not linear, because: $\hat{F}:\left(af+bg\right)\to\left(af+bg\right)^2 \neq af^2+bg^2$
- Lets assume we have an operator $\hat{G}=x\frac{d}{dx}$. To compute the action of the square operator $\hat{G}^2$ on a function $f(x)$, we utilize the algorithm to compute the action of a composite operator:

  $\hat{G}^2 f=\left(\hat{G}\hat{G}\right)f=\hat{G}\left(\hat{G}f\right)=\hat{G}\left(x\frac{d}{dx}f\right)=\hat{G}\left(x\frac{df}{dx}\right)=x\frac{d}{dx}\left(x\frac{df}{dx}\right)=x\left(x'f'+xf''\right)=$

  $=\left(xx'f'+x^2f''\right)=\left(x\frac{d}{dx}+x^2\frac{d^2}{dx^2}\right)f$

This example also demonstrates how one can do algebraic operations on operators. One just have to consider an action of the operator on an arbitrary function f. So, in this case we have:

$$\hat{G}^2=\left(x\frac{d}{dx}+x^2\frac{d^2}{dx^2}\right)=\hat{G}+x^2\frac{d^2}{dx^2}$$

Note also that $\left(x\frac{d}{dx}\right)^2$ is NOT just $\left(x^2\frac{d^2}{dx^2}\right)$

<u>Note</u>: In general, $\hat{F}\hat{G}\neq\hat{G}\hat{F}$

<u>Definition</u>: Let $\hat{F},\hat{G}: X \to X'$ are two linear operators. Then $\left[\hat{F},\hat{G}\right]=\hat{F}\hat{G}-\hat{G}\hat{F}$ is called a **commutator**, and $\left[\hat{F},\hat{G}\right]_+=\hat{F}\hat{G}+\hat{G}\hat{F}$ is called an **anti-commutator**.

<u>Note</u>: If a commutator of two operators is zero, the operators are said to commute.

Example: To compute a commutator of two operators, one needs to consider an explicit action of each product in the sum on an arbitrary function – similar to what we have done above to compute a composite operator. For instance, lets compute a commutator of two operators $\hat{A} = x^2$ and $\hat{B} = \frac{d^2}{dx^2}$:

$$[\hat{A}, \hat{B}]f = (\hat{A}\hat{B} - \hat{B}\hat{A})f = \left(x^2 \frac{d^2}{dx^2} - \frac{d^2}{dx^2} x^2 \right) f = x^2 f'' - (x^2 f)'' = x^2 f'' - (2f + 4xf' + x^2 f'') =$$

$$= -(2f + 4xf') = -\left(2 + 4x\frac{d}{dx} \right) f$$

Therefore, $[\hat{A}, \hat{B}] = -\left(2 + 4x\frac{d}{dx} \right)$

Definition: Let $\hat{F} : V \to V'$ is a linear operator. The scalars $\lambda \in R(C)$ and the vectors $\psi \in V$ that obey $\hat{F}\psi = \lambda\psi$ are called **eigenvalues** and **eigenvectors** of the operator $\hat{F}$, respectively.

Example: The function $\psi = \cos(4x)$ is an eigenvector of the operator $\hat{F} = -\frac{d^2}{dx^2}$, because:

$\hat{F}\psi = -\frac{d^2}{dx^2}\cos(4x) = 16\cos(4x) = 16\psi$, meaning that the eigenvalue corresponding to this eigenvector is 16.

Definition: An operator $\hat{F}^+ : Y \to Y'$ is **conjugate** to the operator $\hat{F} : X \to X'$ if $\forall \varphi \in X, \forall \psi \in Y$ $(\psi, \hat{F}\varphi) = (\hat{F}^+\psi, \varphi)$.

Note: for the vectors out of a Hilbert space $C[a,b]$ the scalar product is defined as integrals (see section 3.3), so the condition above means: $\int_a^b \psi^* (\hat{F}\varphi) dx = \int_a^b (\hat{F}^+\psi)^* \varphi dx$.

Definition: An operator $\hat{F}$ such that $\hat{F}^+ = \hat{F}$ is called **Hermitian**

Theorem: Eigenvalues of Hermitian operators are real

Proof: Consider the eigenvalue problem $\hat{F}\psi = \lambda\psi$. Compute a scalar product of the left and right-hand sides with the eigenvalues:

$$(\psi, \hat{F}\psi) = (\psi, \lambda\psi) = \lambda(\psi, \psi)$$

On the other hand, since the operator is Hermitian, $(\psi, \hat{F}\psi) = (\hat{F}\psi, \psi) = (\lambda\psi, \psi) = \lambda^*(\psi, \psi)$

So we obtain that $\lambda = \lambda^*$, which is the case when $\lambda \in R$