On Vectors, Points, Rotations, and Rigid Bodies.

Ulrich Mutze, www.ulrichmutze.de

June 9, 2022

1 Introduction

This article was planned to be a pedagogically extended version of Chapter 2 of [17], a work which states that it 'can be considered a translation of the most basic parts of [the program] *PaLa*[1] from commented C++ code to 'normal language' '. This indicates that the perspective on the matter listed in the title is that of computational physics. Computational physics has many flavors. It can deal with modeling factual physical or technical systems aiming at predicting their reactions to modifications which would be more expensive to test with the real system. I worked for some years on a project of that type [2],[17]. It also can be used as an enhanced form of thought experiment (Gedankenexperiment) that lets us test the consistency and implications of hypothetical principles for dynamics, and system evolution [3],[4]. It is more in the spirit of such questions that the present article is written.

The C++ class library [5] which I developed when working on projects of computational physics and constructive mathematics codes all the concepts to be developed in the present article (and much more). In [5] a more systematic approach to geometry is chosen than here. There are different classes for polar vectors and axial vectors, and also the Euclidean group, and reference systems are implemented as classes. The whole geometric framework comes in two versions: two-dimensional and three-dimensional. Presently the main open action items in the geometry part of [5] are: integration of reflections into the Euclidean group, introducing the Lie-algebra of the Euclidean group, and the extension to four dimensions, i.e. implementing relativistic kinematics.

The present article is incomplete since it not yet contains a derivation of the form of rigid body dynamics from the principles of [3]. However, in approaching this aim I discovered a way to merge my favorite description of rigid body kinematics is terms of Euler-Rodrigues parameters with the principles of straightforward (un-constraint) Hamiltonian mechanics. This gives rigid body dynamics a completely regular position within Hamiltonian mechanics. This result presently marks the end of the work, which is planned to be continued.

2 Vectors

When dealing with *space* in classical physics it is natural to employ affine geometry (e.g. [7]), in which *points* and *vectors* are different mathematical entities. This reflects the fact that operations that make sense for points may make no sense for vectors and vice versa: Whereas for two vectors it is meaningful to consider their sum, no such notion exists for points. Further, no single point plays a role different from that of any other single point, whereas for vectors there is the zero vector which is uniquely singled out by its arithmetic properties.

Let $\mathcal{V}$ be an *Euclidean oriented vector space*, i.e. a three-dimensional linear space [1] over the field $\mathbb{R}$ of real numbers for which a scalar product $\cdot : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ and a vector product $\times : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ are defined and have the properties implied by the names. To be completely explicit we list those: For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}, a, b \in \mathbb{R}$ we have the algebraic properties [2]

$$\mathbf{u} \cdot (a\,\mathbf{v} + b\,\mathbf{w}) = a\,(\mathbf{u} \cdot \mathbf{v}) + b\,(\mathbf{u} \cdot \mathbf{w})\,, \quad \mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}\,, \quad \mathbf{v} \cdot \mathbf{v} \geq 0\,, \tag{1}$$

$$\mathbf{u} \times (a\,\mathbf{v} + b\,\mathbf{w}) = a\,(\mathbf{u} \times \mathbf{v}) + b\,(\mathbf{u} \times \mathbf{w})\,, \quad \mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}\,, \tag{2}$$

and the following property which excludes trivial realizations of (1):

$$\mathbf{v} \cdot \mathbf{v} > 0 \;\Leftrightarrow\; \mathbf{v} \neq \mathbf{0}\,. \tag{3}$$

Further, trivial realizations of (2) are excluded by requiring that for each $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ we have for $\mathbf{w} := \mathbf{u} \times \mathbf{v}$ the following equations

$$\begin{aligned} &\mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} = 0\,, \\ &\mathbf{w} \cdot \mathbf{w} = (\mathbf{u} \cdot \mathbf{u})\,(\mathbf{v} \cdot \mathbf{v}) \;-\; (\mathbf{u} \cdot \mathbf{v})^2\,. \end{aligned} \tag{4}$$

If the scalar product is given, there are just two vector products which satisfy (4): Let $\times$ be one of them, then the other is $\times' : (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{v} \times \mathbf{u}$. Each of these vector products provides $\mathcal{V}$ with a notion of *orientation* and the orientations associated with the two different vector products are different.

If we would like to treat 3-dimensional Euclidean vector spaces in general, without singling out an orientation we had to introduce the vector product as an operation $\mathcal{V} \times \mathcal{V} \to \mathcal{A}$ where $\mathcal{A}$ would be a vector space of *axial vectors*. Orientation then would come in as a decision how to interrelate the vector spaces $\mathcal{A}$ and $\mathcal{V}$. It is a unique property of 3-dimensional space that axial vectors (which always are the second order elements of the Grassmann algebra over $\mathcal{V}$) form a vector space of the same dimension as the normal (=polar) vectors (the first order elements of the Grassmann algebra). This is so since for an n-dimensional vector space, the space of axial vectors has dimension $n(n-1)/2$ which equals n only for $n = 3$. There are different approaches to the nature of axial and polar vectors based on the representation theory of the group $O(3)$ and others based on

[1] Here the adjective 'Euclidean' is understood as implying three-dimensionality. This is only terminological convenience.

[2] For $a^{-1}\,\mathbf{u}$ one sometimes writes $\mathbf{u}/a$ or $\frac{\mathbf{u}}{a}$.

specific geometric figures (such as arrows and screws). All are connected and trying to clarify these connections was among the motivations to write this article. The task is far from completed so far. As is clear from what was said so far is that the mathematical structure selected here as a starting point to these questions is the operation of the vector product as an operation $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$.

Since the vector product is an operation $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$, it makes $\mathcal{V}$ an *algebra*, actually one which satisfies the *Jacobi identity*

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \times (\mathbf{w} \times \mathbf{u}) + \mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{0} \tag{5}$$

and thus is a *Lie algebra*. As it will turn out, it is the Lie algebra of the rotation group. For $\mathbf{v} \in \mathcal{V}$ one defines $|\mathbf{v}| := \sqrt{\mathbf{v} \cdot \mathbf{v}}$, which is said to be the *length* of the vector or its *absolute value*. It allows expressing equality in $\mathcal{V}$ as follows

$$\mathbf{u} = \mathbf{v} \quad \Longleftrightarrow \quad |\mathbf{u} - \mathbf{v}| = 0 \; . \tag{6}$$

For all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$, $a \in \mathbb{R}$ we have

$$|a\,\mathbf{u}| = |a|\,|\mathbf{u}| \; , \quad |\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}| \; . \tag{7}$$

From the second equation of (4) we have

$$|\mathbf{u} \cdot \mathbf{v}| = \sqrt{|\mathbf{u}|^2 |\mathbf{v}|^2 - |\mathbf{u} \times \mathbf{v}|^2} \leq |\mathbf{u}||\mathbf{v}| \tag{8}$$

the inequality part of which is known as the *Cauchy-Schwarz inequality*. By definition, $|\mathbf{v}| = 1$ characterizes a *unit vector* and $\mathbf{v} \cdot \mathbf{w} = 0$ characterizes *orthogonal* vectors. For each $\mathbf{v} \in \mathcal{V}$ for which $|\mathbf{v}| \neq 0$ the vector $\hat{\mathbf{v}} := \mathbf{v}/|\mathbf{v}|$ is a unit vector. Setting $d(\mathbf{v}, \mathbf{w}) := |\mathbf{v} - \mathbf{w}|$ defines a *metric*, and thus a *metrical topology* on $\mathcal{V}$.

A triplet $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ of mutually orthogonal unit vectors is said to form an *orthonormal basis* [3] or an *Euclidean system of reference*. It is said to be *positively oriented* if $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$. If this holds, we also have $\mathbf{e}_1 = \mathbf{e}_2 \times \mathbf{e}_3$ and $\mathbf{e}_2 = \mathbf{e}_3 \times \mathbf{e}_1$. For positively oriented Euclidean systems of reference we also use the short name *frame*. The name 'positively oriented' is a bit misleading: actually the list $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ has 'the same orientation as the whole space'. If we would change the orientation of the whole space by adopting the vector product $\times'$ mentioned above, an Euclidean system of reference which was 'positively oriented' before would cease to be positively oriented.

The notion of an orthonormal basis obviously applies also to linear subspaces of $\mathcal{V}$. Such bases then consist of either a single unit vector or two orthogonal unit vectors. Due to its bi-linearity, a scalar product is a convenient tool to generate linear subspaces: For any subset $A \subset \mathcal{V}$ the set

$$A^{\perp} := \{\mathbf{v} \in \mathcal{V} \; : \; \mathbf{v} \cdot \mathbf{a} = 0 \;\; \text{for all } \mathbf{a} \in A\} \tag{9}$$

[3] This (commonly used) contraction of 'orthogonal' and 'normal' has the same funny nature as the abbreviation of 'omnibus' (for all) to 'bus' (not a word, only a syntactical entity). In our case the angle ($\gamma\omega\nu\acute{\iota}\alpha$), to which 'ortho' refers, gets lost and what is left would be ridiculous in German: 'richtig normal(er Wahnsinn, nämlich)'.

is a linear space. And

$$\mathrm{span}(A) := (A^{\perp})^{\perp} \tag{10}$$

is the smallest linear subspace of $\mathcal{V}$ which contains the set A.

How can we construct an orthonormal basis? Assume we know a way to create (specify, select) a vector whenever we want. We make use of this by creating $\mathbf{v}_1$ and $\mathbf{v}_2$. Let us assume that $|\mathbf{v}_1 \cdot \mathbf{v}_2| \neq |\mathbf{v}_1||\mathbf{v}_2| \neq 0$ (otherwise we would try a new selection). Then it is clear that the two vectors

$$\mathbf{u}_1 := \hat{\mathbf{v}}_1 \ , \quad \mathbf{u}_2 := \mathbf{v}_2 - (\mathbf{u}_1 \cdot \mathbf{v}_2)\, \mathbf{u}_1$$

are orthogonal, non-zero, and that the vectors

$$\mathbf{e}_1 := \mathbf{u}_1 \ , \quad \mathbf{e}_2 := \hat{\mathbf{u}}_2 \ , \quad \mathbf{e}_3 := \mathbf{e}_1 \times \mathbf{e}_2$$

form an orthonormal basis.

Finally we consider the *automorphisms* of the oriented Euclidean vector space $\mathcal{V}$, i. e. the mappings $\varphi : \mathcal{V} \to \mathcal{V}$ which satisfy for all $\mathbf{v}, \mathbf{w} \in \mathcal{V}$ and all $a, b \in \mathbb{R}$

$$\begin{aligned} \varphi(a\,\mathbf{v} + b\,\mathbf{w}) &= a\,\varphi(\mathbf{v}) + b\,\varphi(\mathbf{w}) \quad \text{linearity} \ , \\ \mathbf{v} \cdot \mathbf{w} &= \varphi(\mathbf{v}) \cdot \varphi(\mathbf{w}) \quad \text{orthogonality} \ , \\ \varphi(\mathbf{v} \times \mathbf{w}) &= \varphi(\mathbf{v}) \times \varphi(\mathbf{w}) \quad \text{conservation of orientation} \ , \end{aligned} \tag{11}$$

and thus *preserve* (leave invariant, are consistent with) the algebraic structures (actually three mappings $\mathbb{R} \times \mathcal{V} \times \mathbb{R} \times \mathcal{V} \to \mathcal{V}, \mathcal{V} \times \mathcal{V} \to \mathbb{R}, \mathcal{V} \times \mathcal{V} \to \mathcal{V}$) appearing in equations (11). The property of bijectivity which is essential for giving the set of automorphisms a group structure (with the composition of mappings as product) has not to be stipulated here since it holds as a consequence of linearity and orthogonality. Mappings $\mathcal{V} \to \mathcal{V}$ which are only known to be linear, need not to be bijective. They are said to be *endomorphisms* of the linear space $\mathcal{V}$. Let the set of these endomorphisms be designated $\mathcal{L}$ (for *linear*). Let us anticipate the name $O(3)$ for the group of those $\varphi \in \mathcal{L}$ which are orthogonal (i.e. satisfy the second equation of (11)), and the name $SO(3)$ for the group of those $\varphi \in O(3)$ which conserve orientation (i.e satisfy the third equation of (11)).

The automorphism of $\mathcal{V}$ will be studied in Sections 4 and 5. They will turn out to be *rotations*. The last equation of (11) plays a particular role: if we assume only the first two equations to hold, then the last one is automatically satisfied at least 'up to the order of the factors' (and, therefore, 'up to a sign') i. e.

$$\varphi(\mathbf{v} \times \mathbf{w}) = \varphi(\mathbf{v}) \times \varphi(\mathbf{w}) \quad \text{or} \quad \varphi(\mathbf{v} \times \mathbf{w}) = \varphi(\mathbf{w}) \times \varphi(\mathbf{v}) \qquad \text{for all } \ \mathbf{v}, \mathbf{w} \in \mathcal{V} \ . \tag{12}$$

For any frame $F = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ and any $\varphi \in SO(3)$ the list $\varphi(F) = (\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), \varphi(\mathbf{e}_3))$ is again a frame. Consider now any two frames F and F'. We will construct $\varphi \in SO(3)$ such that $\varphi(F) = F'$. The basic tool for this is provided by mappings back and forth to $\mathbb{R}^3$ which are associated with a frame: For each frame $F = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ we define the bijective mapping

$$\phi_F \ : \ \mathbb{R}^3 \ \to \ \mathcal{V} \ , \quad (u_1, u_2, u_3) \ \mapsto \ u_1\,\mathbf{e}_1 + u_2\,\mathbf{e}_2 + u_3\,\mathbf{e}_3 \ , \tag{13}$$

which constructs vectors from numerical triplets. In particular it maps the canonical basis $((1,0,0),(0,1,0),(0,0,1))$ of $\mathbb{R}^3$ into F. The inversion of this mapping is given by

$$\phi_F^{-1} \,:\, \mathcal{V} \,\to\, \mathbb{R}^3\,, \quad \mathbf{u} \,\mapsto\, (\mathbf{e}_1\cdot\mathbf{u}, \mathbf{e}_2\cdot\mathbf{u}, \mathbf{e}_3\cdot\mathbf{u})\,. \tag{14}$$

Of course, it maps F into the canonical basis. This allows us to express the desired mapping as $\varphi = \phi_{F'}\circ\phi_F^{-1}$ since this maps first F into the canonical basis of $\mathbb{R}^3$ and then this canonical basis into F'. This suggests to define a *division of frames* by

$$F'/F := \phi_{F'}\circ\phi_F^{-1}\,. \tag{15}$$

When writing the condition for φ as $\varphi\, F = F'$ we are now able to write the solution in a natural fashion as $\varphi = F'/F$. All these relations are special cases of what one finds in *torsors* (see Wikipedia and [26]). Indeed, the set $\mathcal{F}_\mathcal{V}$ of frames in $\mathcal{V}$ with the action

$$SO(3)\times\mathcal{F}_\mathcal{V} \longrightarrow \mathcal{F}_\mathcal{V}\,, \quad (\varphi, F)\mapsto \varphi\, F \tag{16}$$

is a $SO(3)$-torsor (which means what we just showed: that for any pair $(F,F')\in\mathcal{F}_\mathcal{V}\times\mathcal{F}_\mathcal{V}$ there is exactly one $\varphi\in SO(3)$ for which $\varphi\, F = F'$). Another way of expressing this is to say that the action (16) of the group $SO(3)$ on the space $\mathcal{F}_\mathcal{V}$ is *transitive* (each F can be transformed into any F' by some group element) and *effective* (the group element mentioned above is uniquely determined). If one selects in $\mathcal{F}_\mathcal{V}$ an arbitrary element $\tilde{F}$ the mapping

$$\mathcal{F}_\mathcal{V} \longrightarrow SO(3)\,, \quad F\mapsto F/\tilde{F} \tag{17}$$

is bijective; it is, however, not canonical since it depends on the selection of $\tilde{F}$.

So far things were presented without any relation to 'the real world' not even to geometry. Let us sustain this attitude for a short while and even strengthen it in the discussion of how to *construct* an Euclidean oriented vector space. Such a construction can be based on the concept of a list (here of three real numbers) or of a function (here defined on any three element set such as $\{1,2,3\}$ with values in $\mathbb{R}$). In Mathematics we interpret lists as elements of Cartesian product spaces. These are thus functions by definition. In programming, lists (arrays and associative arrays, the latter being also known as 'dictionaries' or 'hashes') and functions are different concepts. They differ mainly by the methods that allow defining particular instances of these kinds of objects. Specifying a list means specifying all of its components (there are only finitely many of those). Specifying a function means specifying an algorithm: how to compute the function value for any given argument. One may hold different opinions whether this is a real difference. From a practical point of view it is a major difference since the definition of algorithms can build on pre-defined algorithms such as the algorithms for defining arithmetic operations, the usual elementary transcendental functions, and random numbers. For the definition of lists it would make sense only in rare cases to build on pre-defined lists. Each programming language provides means to store and otherwise handle algorithms internally much more efficiently than by holding all argument-value pairs in memory. There is no doubt that the latter strategy could provide sufficient functionality for simple applications since it worked for generations of scientists and

engineers in the form of printed function tables, particularly of logarithms. Nevertheless, such a 'memory based function definition' would be a ridiculous project today. For our present aim of modeling Euclidean spaces we thus use arrays, generically written as (x_1, x_2, x_3), (y_1, y_2, y_3) and give the following algorithmic definition of all the operations introduced so far (where a is a generic element of $\mathbb{R}$):

$$\begin{aligned} a\,(x_1, x_2, x_3) &:= (a\,x_1,\ a\,x_2,\ a\,x_3) \\ (x_1, x_2, x_3) + (y_1, y_2, y_3) &:= (x_1 + y_1,\ x_2 + y_2,\ x_3 + y_3) \\ (x_1, x_2, x_3) \cdot (y_1, y_2, y_3) &:= x_1\,y_1 + x_2\,y_2 + x_3\,y_3 \\ (x_1, x_2, x_3) \times (y_1, y_2, y_3) &:= (x_2\,y_3 - x_3\,y_2,\ x_3\,y_1 - x_1\,y_3,\ x_1\,y_2 - x_2\,y_1) \end{aligned} \tag{18}$$

making use of the algorithms for addition, subtraction, and multiplication in $\mathbb{R}$. As is obvious, the first three operations are not influenced by the indexing of the components in the sense that for each *permutation* σ of the set $\{1, 2, 3\}$ the map $P_\sigma : (x_1, x_2, x_3) \mapsto (x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)})$ is an automorphism for the algebraic structure defined by the first three operations. It is an automorphism for the whole structure (i.e. for the Euclidean oriented vector space) if σ is an *even* permutation, i.e. the identity, or $1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1$, or $1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 2$. Of course, this purely numerical model allows us to code the concept of an oriented Euclidean spaces as a *type* in any object oriented programming language. In most such languages the convenient syntactic category is named *class* and the close relation between classes in this sense and classes in the sense of set theory is an interesting topic.

> Taking into account the role that this structure plays in modeling physical space — the background of all physical, chemical, and biological processes — it is tempting to think of this structure as an invention of Nature (to be compared with the inventions of Nature which led to molecular evolution and finally to life [8]). One may imagine a way of physical existence of things in which the modality of happening at spatially separated places [4] was not yet available (since the logical connections enabling 'spaciality' were not yet made, comparable in the biochemical analogue to a state where atoms did not yet find the opportunity to combine to any amino-acid molecule). With 'spatiality' becoming possible, a considerable part of energy would flow into this new modality of existence and space would form at an increasing rate — leading to the Big Bang. Of course this is not more than speculation as long as no predictive models of such a space formation process have been constructed. With respect to our understanding of 'the inner workings of Nature' we probably are in a similar situation as Schrödinger was with respect to the 'inner working of life' when he was writing (in 1943) *What is Life ?*. He reflected about the possible physical nature of genetic information, clearly seeing that the striking constancy of physiological organization over generations poses the hardest problem for physical understanding. He saw this nature as that of an 'aperiodic crystal' which is surprisingly to the point, although not as specific as we know it today. With respect to Nature in general (and in much smaller dimensions than those of the bio-molecules) one may see the hardest problem in the kind of

[4] As a model for the logical implications of spatial separation one may compare parallel computing and normal serial computing.

physical organization which enables a behavior that is so closely related to the behavior of real numbers and arrays of those. Knowing that numbers can be coded as finite sequences of binary quantities we may wonder where and under what circumstances binary quantities come together and form chains that then act on their environment in a way that that a numerical value can be associated with that interaction.

The role of space $\mathcal{V}$ (and also of the point space $\mathcal{P}$ to be considered in the next section) in physics is strongly influenced by its relation to *time*. Since Einstein, we know that *Nature* determines the concept of *equality* only for spatiotemporal points [5] , also known as *events* or *world points*. As will be discussed soon, equality is well defined in all structures which can be represented as *sets* [6] . This indicates that it is a sound concept to define *Minkowski space* as a set — the 'set of all world points'. Starting from this set, one then may define $\mathcal{V}$ as a subset (a space-like hyperplane with suitable structure) or as a set of suitably defined equivalence classes. Space-like hyperplanes (see [11] for their relativistic transformation properties) can be used to define systems of reference as $\mathbb{R}^4$-valued functions on the Minkowski space. Given a system of reference, we may meaningfully say that two events happened at the same place in space if their spatial coordinates were found to agree, even if their time coordinates were registered as different.

It is possible to translate such a mathematical construction into operational directions since world points are straightforward idealizations of real world entities. Space points are more indirect idealizations since they assume for operational definition the existence of a stable environment (such as a metrology laboratory with optical tables and coordinate measuring machines in action).

Set theory and programming

For any two mathematical objects x and y which are defined in the framework of set theory the meaning of the statement $x = y$ is clear: it says $\forall z(z \in x \Leftrightarrow z \in y)$. In this context it looks quite natural to resort to the elements of mathematical objects x and y. There is another context in which it seems not to be common to consider these elements although it would be enlightening: If we introduce a function in the style 'let $f : X \to Y$ be continuous' then, the mental association for the traditionally educated mathematician when seeing $f(x)$ is that there is a $y \in Y$ which is 'well-defined' by saying $y = f(x)$. Respecting the set theoretical framework suggests a more detailed association: Inspecting the elements of x (and recursively the elements of those elements) [7] we are confronted with a wealth of concrete information which, so our idea, should be considered as input for an algorithm which determines the elements of y, and thus y itself. If we hold a sufficiently wide idea of an algorithm (one which allows all operations which the axioms of set theory allow, such as forming the union of the elements of any set) this defines not a new concept

[5] As Einstein says, 'Die Naturgesetze sind Aussagen über zeiträumliche Koinzidenzen' (The laws of nature are propositions concerning spatiotemporal coincidences) Prinzipielles zur allgemeinen Relativitätstheorie, Annalen der Physik Vierte Folge Band 55, 1918

[6] In object oriented programming it is very common to define classes without defining equality for them.

[7] If nothing special is known about the definition of a set X, there is no such procedure as 'inspecting the elements of X', the only fact we are given is that for any arbitrary set it is decided (in the Platonic world, not necessarily decidable by any mathematician) whether it is an element of X or not. In all sets that are reasonably considered in mathematical physics there is more information available, and access to the elements, at least to the elements of a representative (probably finite) subset is possible.

of a function, but only an explication of the set theoretical function concept. It also brings the mathematical view and a programmer's view very closely together. In programming, the information that can be obtained from being given some 'instance x of class X' can always be traced down, probably over a cascade of user-defined classes, to numbers and character strings, and — finally — to arrays of bits. The definition of class X specifies the structure of this information explicitly and completely. However, not all details of this structure normally are visible to a programmer who uses this class in his programs. Although absolutely necessary for letting the class function according to specification, these details may contain arbitrary or preliminary elements so that the creator of the class may wish to change some internals without affecting the programs which meanwhile were written by others and which build on the functionality of X. One speaks of *data hiding* in this case; one then could view a typical set theoretic presentation of a mathematical task as hiding data perfectly: we treat numbers without mentioning their elements, even ordered pairs (x, y) without mentioning the elements $\{x\}$ and $\{x, y\}$ which make up the set (x, y) according to the view of Wiener and Kuratowski. The reason for this form of data hiding is very similar to that in programming. There are various set theoretical representations of numbers and ordered pairs and as 'users' of these mathematical concepts we only need their 'behavior' which manifests itself — just as in programming — in the availability ('definedness') of pertinent functions (the canonical projections for ordered pairs and the arithmetic operations for numbers). The set-based formalism in mathematics can be viewed as a powerful programming language with a very simple and uniform syntax. A concept or construction that is hard to formulate in this language, will be even harder to formulate in C++ . This analogy of mathematical formulation and C++ formulation allows translating sets and functions defined in this article into classes and methods (member functions) of C++ directly.

Let us return to the meaning of $f(x)$. If functions act as input to an algorithm, e.g. an algorithm which defines for given functions f and g the sum function $f + g$, we don't need a description of the algorithm for computing $f(x)$ and $g(x)$; the result is sufficient: $(f + g)(x) := f(x) + g(x)$. If we have to define a function of a real argument, e.g. the real square root, the algorithm, e.g. $f(x) := \{x_1 = x, y = x_1 + 1, \text{while}(y \neq x_1)\,\{x_1 = y, y = (x_1 + x/x_1)/2\}, y\}$ needs the functions addition, subtraction, and division of real numbers. However, these functions enter as generators of function values irrespective of the underlying algorithm. On the other hand, it is clear that the internals of number representation and implementation of the arithmetic functions is indispensable for the algorithm to work.

Interestingly the concept of a reference system, which — as we have just seen — is required when we are to speak about Euclidean space in normal mathematical language, has a representation also in our numerical model, namely as a list of three orthonormal vectors from which the coordinates of individual vectors can be obtained by forming scalar products. One could interpret the numerical lists which characterize a vector in our numerical model from the very beginning as values of coordinates with respect to some coordinate system. One then would read (18) as a definition of the vector space operations 'in terms of coordinates'. From a physical point of view it is rather evident, however, that coordinates can't play a role in the process of forming the sum of two vectors (e.g. of electrical fields from two distinct sources) since maintaining a coordinate system is itself a physical process (involving a net of synchronized clocks) which may not yet be completed at the space-time point where the vector sum first is needed. Rather, this numerical triplet is a 'state variable' which allows the arithmetic operations be defined autonomously without interacting in any sense with a system of reference. Actually, instead of a triplet of real numbers, a triplet of three integer numbers together

with a single common real factor fixing the order of magnitude would be a more natural data structure for the internal state.

In summary, we see that the concept of an oriented Euclidean space links several concepts for which it is not completely clear which ones one should consider as primary and which as derived. That everything fits together and forms a logical entity independent of the question of appropriate order of introduction is a noteworthy fact.

As mentioned above, one may turn the present conceptual discussion into a concrete definition of classes in an object oriented programming language. I am fascinated by the guidance which the strict rules of C++ provide in situations where classes have to be defined which depend on each other (e.g. one of the classes defines functions which take arguments from the other class). Even in situations where there is a choice, one finds often that clear rules of syntactic simplicity and straightforwardness single out one of the possible ways. In any way, the C++ compiler guarantees that any class edifice which compiles avoids all 'vicious circles' (impredicative definitions).

2.1 Orientation of lists of vectors

So far, we know orientation only as a property of orthonormal bases (and, in a sense, for the whole of $\mathcal{V}$). To extend the concept to less special entities one defines the *determinant function*

$$\det \,:\, \mathcal{V}\times\mathcal{V}\times\mathcal{V} \,\to\, \mathbb{R}\,, \quad (\mathbf{u},\mathbf{v},\mathbf{w}) \,\mapsto\, (\mathbf{u}\times\mathbf{v})\cdot\mathbf{w} \tag{19}$$

which is a linear function in each of its arguments, takes the value 0 if the arguments are linearly dependent (i.e. a non-trivial linear combination of them exists which has the value 0), but does not vanish identically. An interesting result in linear algebra says that two functions which enjoy these properties are proportional of each other and are *alternating* functions in the sense of satisfying:

$$\begin{aligned} \det(\mathbf{u},\mathbf{v},\mathbf{w}) = \det(\mathbf{v},\mathbf{w},\mathbf{u}) = \det(\mathbf{w},\mathbf{u},\mathbf{v}) = \\ -\det(\mathbf{u},\mathbf{w},\mathbf{v}) = -\det(\mathbf{w},\mathbf{v},\mathbf{u}) = -\det(\mathbf{v},\mathbf{u},\mathbf{w})\,. \end{aligned} \tag{20}$$

In addition, such a general determinant function on any linear space of finite dimension vanishes if and only if the arguments are linearly dependent (see [9], pp. 43-46). Apart from determining orientation, this function also determines the volume of the parallelepiped defined by the argument vectors.

If we start from a three-dimensional real linear space which has no direct relation to geometry (e.g. the space of solutions $f \,:\, \mathbb{R} \,\to\, \mathbb{R}$ of the ordinary linear differential equation $f''' = 0$, or the set of functions $\{1,2,3\} \,\to\, \mathbb{R}$) one may ask which degrees of freedom exist to turn this space into an oriented Euclidean space according to the previous definition. The answer is quite simple: Select any triplet of linearly independent elements of this linear space. Then there exist a uniquely determined scalar product and a uniquely determined vector product which turns this triplet into a positively oriented Euclidean system of reference. This makes clear that 'being positively oriented' is not an intrinsic property but a conventional one. Orientation (also referred to as *screw sense*

or *chirality*) is not only associated with mutually orthogonal triplets of vectors but with arbitrary triplets according to the function

$$\mathrm{ori} \,:\, \mathcal{V}\times\mathcal{V}\times\mathcal{V} \,\to\, \{-1,0,1\}\,, \quad (\mathbf{u},\mathbf{v},\mathbf{w}) \,\mapsto\, \mathrm{sign}(\det(\mathbf{u},\mathbf{v},\mathbf{w})) \tag{21}$$

Notice $\mathrm{ori}(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3) = -\mathrm{ori}(\mathbf{b}_3,\mathbf{b}_2,\mathbf{b}_1)$ which makes clear that it is not a *set* of vectors which determines the orientation but a *list* of those. A triplet $(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)$ of vectors is linear independent (and thus a basis of $\mathcal{V}$) *iff* (i.e. if and only if) it is oriented (i.e. $\mathrm{ori}(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3) \neq 0$). Let this be the case. Then, for any $\mathbf{u}\in\mathcal{V}$ there exist $\alpha_i\in\mathbb{R}$ for which $\mathbf{u} = \alpha_1\,\mathbf{b}_1 + \alpha_2\,\mathbf{b}_2 + \alpha_3\,\mathbf{b}_3$ holds. Scalar product and vector product provide convenient means to compute these development coefficients: We define a new basis $(\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3)$, said to be the basis *dual* (or *reciprocal*) to $(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)$, by

$$\mathbf{c}_1 := \frac{\mathbf{b}_2\times\mathbf{b}_3}{\det(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)}\,, \quad \mathbf{c}_2 := \frac{\mathbf{b}_3\times\mathbf{b}_1}{\det(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)}\,, \quad \mathbf{c}_3 := \frac{\mathbf{b}_1\times\mathbf{b}_2}{\det(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)} \tag{22}$$

and find

$$\alpha_1 = \mathbf{c}_1\cdot\mathbf{u}\,, \quad \alpha_2 = \mathbf{c}_2\cdot\mathbf{u}\,, \quad \alpha_3 = \mathbf{c}_3\cdot\mathbf{u} \tag{23}$$

and

$$\mathbf{b}_i\cdot\mathbf{c}_j = \delta_{ij}\,, \quad \det(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)\det(\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3) = 1\,. \tag{24}$$

From this it is clear that a positively oriented Euclidean system of reference is a self-dual basis and vice versa. For preparing the definition of angles between vectors we observe that for $\mathbf{v},\mathbf{w}\in\mathcal{V}$ the quantities $x := \mathbf{v}\cdot\mathbf{w}$, $y := |\mathbf{v}\times\mathbf{w}|$, $r := |\mathbf{v}|\,|\mathbf{w}|$, as a consequence of (4) satisfy $x^2+y^2 = r^2$ and $y\geq 0$ to the effect that there is a uniquely defined angle $\varphi\in[0,\pi]$ for which $x = r\,\cos(\varphi), y = r\,\sin(\varphi)$ holds, and which is said to be the *polar angle* determined by (x,y) and is sometimes written as $\arg(x,y)$. Many programming languages (e.g. C, Java, Ruby) represent the polar angle φ as $\mathrm{atan2}(y,x)$ in reminiscence of the incomplete representation $\varphi = \arctan(y/x)$ which in too many formularies is recommended even for $y<0$ where it holds only up to adding or subtracting π. Notice that the mathematically correct representation $\varphi = \arccos(y)$ is practically useless for small values of φ. A useful unrestricted (as far as $(x,y)\neq(0,0)$) representation is $\varphi = 2\arctan(\frac{y}{x+\sqrt{x^2+y^2}})$. In pure mathematics polar angles enter the stage with the complex plane, where Euler's famous formula

$$\mathrm{e}^{i\,\varphi} = \cos(\varphi) + i\,\sin(\varphi)$$

implies the representation $\varphi = \arg(x,y) = \mathrm{Im}\log(x+iy)$. The following function summarizes the definition of the angle between vectors:

$$\angle \,:\, \mathcal{V}\times\mathcal{V} \,\to\, [0,\pi]\,, \quad (\mathbf{v},\mathbf{w}) \,\mapsto\, \arg(\mathbf{v}\cdot\mathbf{w}, |\mathbf{v}\times\mathbf{w}|)\,. \tag{25}$$

Obviously $\angle(\mathbf{v},\mathbf{w}) = \angle(\mathbf{w},\mathbf{v})$ and $|\mathbf{v}+\mathbf{w}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 + 2\,|\mathbf{v}|\,|\mathbf{w}|\cos(\angle(\mathbf{v},\mathbf{w}))$. It is interesting to note that angles between vectors in oriented two-dimensional space behave quite differently. Here it makes sense to ask for the rotation needed to move one vector into the other when rotating in the sense of the given orientation. Then angles

in the range $[0, 2\pi)$ are possible. To make this explicit, I use the natural transcription of the concept of a vector product to Euclidean two-dimensional space (denoted $\mathcal{V}_2$) by replacing the binary operation $\times : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ by a unary one: $\times : \mathcal{V}_2 \to \mathcal{V}_2$ (as done in the well-known program system *Mathematica* by Stephen Wolfram). This operation is a linear orthogonal mapping (with linearity and orthogonality defined in analogy to (11)) which assigns to any vector $\mathbf{v}$ a vector $\mathbf{v}'$ which is orthogonal to $\mathbf{v}$ and has the same length. Since the mapping itself is assumed to be orthogonal the choice of selecting such a vector or its negative is determined for each vector if it has been defined for a single non-zero vector. Then $\angle(\mathbf{v}, \mathbf{w}) := \varphi$ where φ is uniquely determined by

$$\varphi \in [0, 2\pi) \quad \text{and} \quad \mathbf{w} = \cos(\varphi)\, \mathbf{v} + \sin(\varphi)\, (\times \mathbf{v}) \; . \tag{26}$$

3 Points

An oriented Euclidean point space is a triplet $(\mathcal{P}, \mathcal{V}, +)$ where $\mathcal{V}$ is an oriented Euclidean vector space, and the mapping $+ : \mathcal{P} \times \mathcal{V} \to \mathcal{P}$ satisfies:

1. For all $\mathbf{v}, \mathbf{w} \in \mathcal{V}, p \in \mathcal{P}$ we have $(p + \mathbf{v}) + \mathbf{w} = p + (\mathbf{v} + \mathbf{w})$.

2. For any two points $p, q \in \mathcal{P}$ there is a uniquely defined element of $\mathcal{V}$, conveniently written as $q - p$, for which $p + (q - p) = q$. For $p + \frac{1}{2}(q - p)$ one also writes $\frac{1}{2}(p + q)$.

One often finds $\overrightarrow{p\,q}$ written for $q - p$. This notation lacks the algorithmic cleverness of the difference notation.

The notions of differences and mean-values of points can easily be extended: Let $(p_i)_{i=1}^n$ be a family of points and $(w_i)_{i=1}^n$ a family of real numbers. For arbitrarily selected origin $o \in \mathcal{P}$ we consider the $\mathcal{P}$-valued expression

$$o + w_1\, (p_1 - o) + \ldots + w_n\, (p_n - o)$$

which is (as is easily seen) independent of o iff the sum of the w_i is 1; it then represents a well-defined point. Further, the expression

$$w_1\, (p_1 - o) + \ldots + w_n\, (p_n - o)$$

is independent of o iff the sum of the w_i is 0; it the represents a well-defined vector. In both cases one writes the well-defined object (be it a point or a vector) as

$$w_1\, p_1 + \ldots + w_n\, p_n \; . \tag{27}$$

This is a symbolic notation in so far as the subterms $w_i\, p_i$ are not defined objects, only the expression as a whole is defined. (Also in the notation $q - p$ the subterm $-p$ is not defined.)

Let an oriented Euclidean vector space $\mathcal{V}$ is given. Then all completions to an oriented Euclidean point space $(\mathcal{P}, \mathcal{V}, +)$ can easily be described: Let X be any set for which there is a bijection $\phi : X \to \mathcal{V}$. Then define $\mathcal{P} := X$ and $p + \mathbf{v} := \phi^{-1}(\phi(p) + \mathbf{v})$ for all

$p \in X$ and $\mathbf{v} \in \mathcal{V}$. Notice that on the left-hand side of this definition there appears $+ : \mathcal{P} \times \mathcal{V} \to \mathcal{P}$, whereas on the right-hand side we only need $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ which is defined as a part of the definition of $\mathcal{V}$.

An (affine) point space in relation to its vector space is a special case of structures of wide-spread use in mathematical physics. First, it is a special case of a transformation group: The abelian group $\mathcal{V}_+$ underlying $\mathcal{V}$ acts transitively and effectively (see the text following (16)) on $\mathcal{P}$ thus making $\mathcal{P}$ a $\mathcal{V}_+$*-torsor*. Second, it is a *differential manifold* with the special structure that the tangent space at any point is the same vector space $\mathcal{V}$. If we insist in having disjoint tangent spaces for different points we can achieve this but get a natural bijection between any pair of local tangent spaces, which then is the simplest and most comfortable case of an *affine connection* on our manifold. This derives from a *Riemannian metric* which makes $\mathcal{P}$ a *metric space*. This metric has here a particularly simple form: The distance $d(p, q)$ between points p and q is given by the length of the connecting vector $p - q$

$$d(p, q) = |p - q| \ . \tag{28}$$

The property of $\mathcal{P}$ as a metric space turns out to be of great organizing power. Any mapping $\varphi : \mathcal{P} \to \mathcal{P}$ which preserves the metric (i.e. is an *isometry*):

$$|\varphi(p) - \varphi(q)| = |q - p| \ \text{ for all } p, q \in \mathcal{P} \tag{29}$$

is bijective and *affine* (see [7], Chapter XII, Corollary on p. 445 and Theorem 15). Being affine means for a mapping to be linear in the same hidden way as $\mathcal{P}$ is linear. This will be explained in detail in subsection 3.2.

An isometry $\mathcal{P} \to \mathcal{P}$ is also called *rigid motion* [8]. If it preserves orientation, it is said to be a *proper rigid motion*. Both types of mappings form a group. As will be explained in subsection 3.2 they are 'inhomogeneous versions' of the groups $O(3)$ and $SO(3)$ and thus will be designated $IO(3)$ and $ISO(3)$. Due to their close connection to Euclidean geometry and the structure of Euclidean spaces they often are referred to as the *(full) Euclidean group* and the *(proper) Euclidean group*.

3.1 Orientation of lists of points

A list of three points (p_1, p_2, p_3) defines a closed directed polygon $p_1 \to p_2 \to p_3 \to p_1$ and it is tempting to associate the orientation $\mathrm{ori}(p_2 - p_1, p_3 - p_2, p_1 - p_3)$ with this situation. However, the vectors under consideration are linearly dependent (their sum is $\mathbf{0}$) and the orientation (as defined by (21)) is 0. This may surprise since following this directed path, fixes the same kind of directedness which characterizes the motion of a rotating body. However, if we are to judge whether the 'sense' of this motion matches that of the clock hand, we come to different conclusions depending on the viewing direction. Applying this observation to the clock itself, we see that it is an important feature of a

[8] This is 'motion in a geometrical sense' where time plays no role. Later we will have to deal with motion in the sense of physics too.

traditional clock that it allows us to see the hands only when looking to the front side of it, otherwise there would be nothing like a clockwise sense of rotation.

To define non-zero orientations we need one point more: Thus for any list (p_0, p_1, p_2, p_3) of four points we define

$$\mathrm{ori}(p_0, p_1, p_2, p_3) := \mathrm{ori}(p_1 - p_0, p_2 - p_0, p_3 - p_0) \tag{30}$$

where on the right-hand side we refer to the orientation of list of vectors as defined in (21). The intuition behind this is not simply the motion of a rotating body represented by the closed polygon $p_1 \to p_2 \to p_3 \to p_1$ but also the vector $(p_1-p_0)+(p_2-p_0)+(p_3-p_0)$ as a kind of directed rotation axis. An example would be a wheel rotatable around an axis the ends of which carrying engravings 'begin' and 'end' respectively. If we spin the wheel relative to the axis it defines a certain value of orientation. If we fix the axis, stop the wheel, and make it spin again but in the opposite direction, then we also have changed the sign of the orientation (which is a number $\in \{-1, 0, 1\}$ according to the present definition). So the wheel could be used as a materialized definition of a particular orientation only if it would be prevented to rotate in the opposite direction (e.g. by a ratchet and pawl mechanism). A more elegant 'orientation normal' is a screw and a nut. If we rotate the screw within the nut (or the other way round) it performs a linear motion in addition to the rotation. In this case the sense of rotation and the direction of translation are in fixed relation: if we change the sense of rotation, the sense of translation changes too. So the system defines a constant value of orientation in all its operating modes. The normal screws in engineering are the so called right screws (in the few occasions that left screws are needed too, these have to be particularly marked in engineering drawings). Their orientation as defined above is just what is positive orientation $+1$ if the vector product is defined according to this convention. Right screws and left screws are in a natural relation to right hands and left hands in primates: The (non-thumb) fingers in a fist define a sense of rotation (following fingers in the direction towards the finger tips) and the protruding thumb (in the direction towards its tip) defines a direction.

Assume that $\mathrm{ori}(p_0, p_1, p_2, p_3) \neq 0$ then the mapping

$$\begin{aligned} \beta \,:\, \{r \in \mathbb{R}^4 \,:\, r_0 + r_1 + r_2 + r_3 = 1\} \;&\to\; \mathcal{P} \\ (r_0, r_1, r_2, r_3) \;&\mapsto\; r_0\, p_0 + r_1\, p_1 + r_2\, p_2 + r_3\, p_3 \end{aligned} \tag{31}$$

is bijective and defines what is called *barycentric coordinates*. See (27) for the symbolic linear combination of points.

A mapping $\varphi : \mathcal{P} \to \mathcal{P}$ preserves orientation iff for any list (p_0, p_1, p_2, p_3) of points we have

$$\mathrm{ori}(\varphi(p_0), \varphi(p_1), \varphi(p_2), \varphi(p_3)) = \mathrm{ori}(p_0, p_1, p_2, p_3)\ . \tag{32}$$

3.2 Affine mappings in point space

A mapping $\varphi : \mathcal{P} \to \mathcal{P}$ is called affine iff there is a linear mapping $A : \mathcal{V} \to \mathcal{V}$ such that

$$\varphi(p) = \varphi(q) + A(p - q) \;\text{ for all } p, q \in \mathcal{P}\ . \tag{33}$$

In the situation considered above, that φ was an isometry (rigid motion), A obviously is orthogonal. If it also preserves orientation (proper rigid motion), A is an automorphism of the oriented Euclidean space $\mathcal{V}$, and thus a rotation (see below (11)). The linear mapping $A \in \mathcal{L}$ is uniquely determined by the affine mapping φ and is designated $\varphi^\sharp$ and said to be the *trace* of φ ([7], p. 427; this notion is not directly related to the notion of the trace of a linear mapping as defined in (48)). It has a deeper sense in the framework of *categories*. There one also states [7] $\mathcal{P}^\sharp = \mathcal{V}$ and that $\#$ is a *functor* which connects the category of affine spaces with the category of linear spaces. Upon selection of an arbitrary origin $o \in \mathcal{P}$ the affine mapping φ determines also a vector $\mathbf{a} \in \mathcal{V}$ such that

$$\varphi(p) = o + \mathbf{a} + A(p - o) =: (o, \mathbf{a}, A)\, p \qquad \text{for all } p \in \mathcal{P}\ . \tag{34}$$

Obviously, $\mathbf{a} = \varphi(o) - o$. Since the origin o is arbitrary the terms

$$(o, \mathbf{a}, A) \quad o \in \mathcal{P},\ \mathbf{a} \in \mathcal{V},\ A \in \mathcal{L}$$

defined in (34) do not provide a unique parametrization of affine maps. Actually we have

$$(o, \mathbf{a}, A) = (o', \mathbf{a}', A') \quad \text{iff} \quad \big(A = A' \ \text{ and } \ A(o' - o) = o' - o + \mathbf{a}' - \mathbf{a}\big)\ , \tag{35}$$

and thus the behavior with respect to a shift of the origin is as follows

$$(o + \mathbf{s}, \mathbf{a}, A) = (o, \mathbf{a} + A(\mathbf{s}) - \mathbf{s}, A) \qquad \text{for all } \mathbf{a}, \mathbf{s} \in \mathcal{V}, A \in \mathcal{L}\ . \tag{36}$$

If A is bijective, so is $(o, \mathbf{a}, A)$, and we have

$$(o, \mathbf{a}, A)^{-1} = (o, A^{-1}(-\mathbf{a}), A^{-1})\ . \tag{37}$$

The composition of two affine maps is again an affine map which can be represented reasonably simply in the notation of (34) if everything refers to the same origin:

$$(o, \mathbf{a}, A) \circ (o, \mathbf{a}', A') = (o, \mathbf{a} + A(\mathbf{a}'), A \circ A')\ . \tag{38}$$

The identity mapping $\mathcal{P} \to \mathcal{P}$ has the representation $(o, \mathbf{0}, \mathbf{1})$, where $\mathbf{1}$ is the identity mapping $\mathcal{V} \to \mathcal{V}$.

For any subset $\mathcal{S}$ of $\mathcal{L}$ one defines for arbitrary $o \in \mathcal{P}$

$$\mathcal{S}^\flat := \{(o, \mathbf{a}, A)\ :\ \mathbf{a} \in \mathcal{V},\ A \in \mathcal{S}\} \tag{39}$$

getting a set of affine maps $\mathcal{P} \to \mathcal{P}$ which is independent of the choice of the origin o. Here the symbol $\flat$ in its relation to symbol $\sharp$ is borrowed from music where these symbols stand respectively for dividing and multiplying the frequency of a tone by $2^{\frac{1}{12}}$. If $\mathcal{S}$ is a group (with respect to composition of mappings), so is $\mathcal{S}^\flat$. The mappings $(o, \mathbf{a}, \mathbf{1}),\ \mathbf{a} \in \mathcal{V}$ are called *translations* and are independent of $o \in \mathcal{P}$. Translations are said to be *inhomogeneous* since they don't map 'zero to zero' (actually origin to origin). Therefore $S^\flat$ as resulting from S by allowing for translations is sometimes said to be the

inhomogeneous form of S and is designated IS. This is the origin of the names $ISO(3)$ and $IO(3)$ for the Euclidean groups.

The *homogeneous* mappings $(o, \mathbf{0}, A),\ A \in \mathcal{S}$ do depend on the choice of the origin $o \in \mathcal{P}$: o is a fixed point of the mapping. Although we always have

$$(o, \mathbf{a}, A) = (o, \mathbf{a}, \mathbf{1}) \circ (o, \mathbf{0}, A) \tag{40}$$

and the translation is independent of o, the second factor depends on o just as does the left-hand side. Put differently: Each affine mapping other than the identity can be written as a product of a translation and a homogeneous mapping in many ways. Given the affine mapping φ we obtain $A = \varphi^\sharp$. But only upon selection of o we get a representation $\varphi = (o, \mathbf{a}, A)$. As given explicitly in (36) the translation vector $\mathbf{a}$ depends on o. So the mapping $(o, \mathbf{a}, \mathbf{1})$ depends also on o despite the above stated fact that it does not depend on o for $\mathbf{a}$ given.

By definition, an *Euclidean system of reference* in $\mathcal{P}$ is a list (p_0, p_1, p_2, p_3) of points such that the three vectors $\mathbf{e}_i := p_i - p_0$ form an Euclidean system of reference in $\mathcal{V}$. The term $\mathrm{ori}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ (which equals $\mathrm{ori}(p_0, p_1, p_2, p_3)$) defines the orientation of the system. For the most frequently used case that the orientation is positive we use the short name *frame* for such an Euclidean system of reference. These frames form the set $\mathcal{F}_\mathcal{P}$. Instead of writing $F = (p_0, p_1, p_2, p_3)$ for some $F \in \mathcal{F}_\mathcal{P}$ one may put the orthonormal vectors into evidence writing

$$F = (p_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3),\ (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \in \mathcal{F}_\mathcal{V}\ . \tag{41}$$

For analyzing the frames in $\mathcal{P}$ we proceed in close analogy to the treatment of frames in $\mathcal{V}$. For any frame $F = (p_0, p_1, p_2, p_3)$ and any $\varphi \in ISO(3)$ the list

$$\varphi(F) = (\varphi(p_0), \varphi(p_1), \varphi(p_2), \varphi(p_3))$$

is again a frame. Consider now any two frames F and F'. We will construct $\varphi \in ISO(3)$ such that $\varphi(F) = F'$. The basic tool for this is provided by mappings back and forth to $\mathbb{R}^3$ which are associated with a frame: For each frame $F = (p_0, p_1, p_2, p_3)$ we define the bijective mapping

$$\phi_F : \mathbb{R}^3 \to \mathcal{P}\ , \quad (u_1, u_2, u_3) \mapsto p_0 + u_1\,(p_1 - p_0) + u_2\,(p_2 - p_0) + u_3\,(p_3 - p_0)\ , \tag{42}$$

which constructs points from of numerical triplets. In particular it maps the 'canonical frame'

$$((0,0,0), (1,0,0), (0,1,0), (0,0,1))$$

of $\mathbb{R}^3$ into F. The inversion of this mapping is given by

$$\phi_F^{-1} : \mathcal{P} \to \mathbb{R}^3\ , \quad q \mapsto (\,(p_1 - p_0)\cdot(q - p_0),\ (p_2 - p_0)\cdot(q - p_0),\ (p_3 - p_0)\cdot(q - p_0)\,)\ . \tag{43}$$

Of course, it maps F into the 'canonical frame' of $\mathbb{R}^3$. This allows us to express the desired mapping as $\varphi = \phi_{F'} \circ \phi_F^{-1}$ since this maps first F into the canonical basis of $\mathbb{R}^3$ and this canonical basis into F'. Again, we define a *division of frames* by

$$F'/F := \phi_{F'} \circ \phi_F^{-1} \tag{44}$$

and we write the condition for φ as $\varphi F = F'$ and are now able to write the solution as $\varphi = F'/F$ which is natural. Now the set of frames is an $ISO(3)$-torsor. If one selects in $\mathcal{F}_{\mathcal{P}}$ an arbitrary element $\tilde{F}$ the mapping

$$\mathcal{F}_{\mathcal{V}} \longrightarrow ISO(3) \ , \quad F \mapsto F/\tilde{F} \tag{45}$$

is bijective; it is, however, not canonical since it depends on the selection of $\tilde{F}$.

4 Linear maps

Linear maps $\mathcal{V} \to \mathcal{V}$ (endomorphisms of the linear space $\mathcal{V}$) arise in physics mainly in two contexts which are not directly related:

1. Under the name of *tensors* of rank two they describe properties of individual physical systems. An example of this, to which we will come later in detail, is the *tensor of inertia* which connects angular velocity and angular momentum of a rigid body.

2. Under the name of *active transformations* they map states of physical systems into other potential states of the same systems thereby providing a framework for defining and analyzing *symmetries* of physical laws. The mapping $\varphi^\sharp$ associated with any isometry φ of $\mathcal{P}$ is an example of this; it provides the most natural approach to the notion of angular velocity in rigid body motion.

As introduced below (11) we use the symbol $\mathcal{L}$ for the set of linear maps $\mathcal{V} \to \mathcal{V}$ and refer sometimes to its elements as tensors, when 'element of $\mathcal{L}$' is considered annoying. These tensors form a real associative involution algebra with unity (the identity map $\mathbf{1}$; the zero-element O of the algebra is given by $O\,\mathbf{v} = \mathbf{0}$). Here it is understood that the linear operations are carried over from $\mathcal{V}$ and that the composition $\circ$ of maps is taken as the product, and that the involution is given by $L \mapsto L^*$, where L^* is the map adjoint to L. For the application of $L \in \mathcal{L}$ to a $\mathbf{v} \in \mathcal{V}$ we mostly write $L\,\mathbf{v}$ instead to the orthodox notation $L(\mathbf{v})$. To be sure, these definitions imply that we have

$$\begin{aligned} L \circ L'\,\mathbf{v} = L\,L'\,\mathbf{v} \ , \quad (\alpha\,L + \alpha'\,L')\,\mathbf{v} = \alpha\,L\,\mathbf{v} + \alpha'\,L'\,\mathbf{v} \ , \\ L\mathbf{v} \cdot \mathbf{w} = \mathbf{v} \cdot L^*\mathbf{w} \ , \quad L = L^{**} \ , \quad (L \circ L')^* = L'^* \circ L^* \end{aligned} \tag{46}$$

for all $L, L' \in \mathcal{L}, \mathbf{v}, \mathbf{w} \in \mathcal{V}, \alpha, \alpha' \in \mathbb{R}$. For $\alpha^{-1}\,L$ one also writes L/α or $\frac{L}{\alpha}$.

The product operation deserves a remark: For general mappings $\mathcal{V} \to \mathcal{V}$ the composition is computationally a difficult operation. Such a mapping has to be defined as an algorithm, and thus also products of such mappings have to be defined as an algorithms for which the the algorithms which represent the factors are function arguments. In the language C++ algorithms can be coded as classes, the instances of which are called *function objects* and coding the composition operation is straightforward. However, iterating the composition such as required in studies of dynamical systems lets the algorithm grow in a way that provokes stack overflow. Of course, the execution time will grow linearly with the number of iterations, which is hardly acceptable in dynamical simulations.

With linear mappings we are in a better position since these have a representation in terms of numerical parameters (*matrix elements*) which allow us to express the operations (46) in terms of these elements. Actually, for any selected orthonormal triplet $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ the mapping

$$\Phi_{\mathbf{e}} \,:\, \mathcal{L} \,\to\, M(3,\mathbb{R})\;, \quad L \,\mapsto\, (\mathbf{e}_i \cdot L\,\mathbf{e}_j)_{i,j\in\{1,2,3\}} =: (L_{ij})_{i,j\in\{1,2,3\}} \tag{47}$$

is an isomorphism to the ring $M(3,\mathbb{R})$ of real 3×3 matrices. (As is well-known, also non-orthonormal bases in $\mathcal{V}$ define similar isomorphism but for them one needs the dual basis too in order to express the matrix elements.) Therefore each product or power of linear mappings will be again be representable as such a matrix. Only rounding errors are reminiscent of the chain of operations which led to this matrix. If a linear map is known to have particular properties, such as being orthogonal, one may characterize it by fewer parameters than 9 real matrix elements and one may be able to express the parameters of the product in terms of the parameters of the factors. Even families of non-linear transformations may allow such a representation of their product in terms of real parameters. This observation led to the theory of *Lie groups*. For the various types of special linear maps to be introduced in the rest of this chapter we will always point out how their product can be expressed directly in terms of these special types.

4.1 Determinant, trace, and the operation of inversion

Let $(\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathcal{V}\times\mathcal{V}\times\mathcal{V}$ such that $\det(\mathbf{u},\mathbf{v},\mathbf{w}) \neq 0$, then we define for any $L\in\mathcal{L}$ the numbers $\det(L)$ and $\mathrm{tr}(L)$ (*trace* of L) by

$$\begin{aligned} \det(L) &:= \frac{\det(L\mathbf{u},L\mathbf{v},L\mathbf{w})}{\det(\mathbf{u},\mathbf{v},\mathbf{w})}\;, \\ \mathrm{tr}(L) &:= \frac{\det(L\mathbf{u},\mathbf{v},\mathbf{w}) + \det(\mathbf{u},L\mathbf{v},\mathbf{w}) + \det(\mathbf{u},\mathbf{v},L\mathbf{w})}{\det(\mathbf{u},\mathbf{v},\mathbf{w})}\;. \end{aligned} \tag{48}$$

It turns out that these definitions are independent of the determinant function and the vector triplet. They satisfy for all $L, L' \in \mathcal{L}$

$$\begin{aligned} &\det(L\circ L') = \det(L)\det(L')\;, \quad \det(L) = \det(L^*)\;, \quad \det(\mathbf{1}) = 1\;, \\ &\mathrm{tr}(L+L') = \mathrm{tr}(L) + \mathrm{tr}(L')\;, \quad \mathrm{tr}(L\circ L') = \mathrm{tr}(L'\circ L)\;, \quad \mathrm{tr}(L) = \mathrm{tr}(L^*)\;, \quad \mathrm{tr}(\mathbf{1}) = 3 \end{aligned} \tag{49}$$

The trace operation allows to define a scalar product in the linear space $\mathcal{L}$:

$$L\cdot L' := \mathrm{tr}(L^*\circ L')\;. \tag{50}$$

The absolute value $\sqrt{L\cdot L}$, associated with this scalar product is the *Hilbert-Schmidt norm*, which is always larger or equal to the usual operator norm

$$\|\,L\,\| := \sup\{|L\mathbf{v}| \;:\; \mathbf{v}\in\mathcal{V},\, |\mathbf{v}| = 1\}\;. \tag{51}$$

In terms of matrix elements (47) one has the standard formulas for the determinant and the trace

$$\begin{aligned} \mathrm{tr}(L) &= L_{11} + L_{22} + L_{33} \ , \\ \det(L) &= \sum_{\sigma \in Perm(\{1,2,3\})} \mathrm{sign}(\sigma)\ L_{1\sigma(1)}\ L_{2\sigma(2)}\ L_{3\sigma(3)} \ . \end{aligned} \tag{52}$$

The famous *characteristic polynomial* is given by

$$\begin{aligned} \det(\lambda\,\mathbf{1} - L) &= \lambda^3 - \mathrm{tr}(L)\,\lambda^2 \\ &\quad + (L_{11}\,L_{22} - L_{12}\,L_{21} + L_{11}\,L_{33} - L_{13}\,L_{31} + L_{22}\,L_{33} - L_{23}\,L_{32})\,\lambda - \det(L) \ . \end{aligned} \tag{53}$$

Each $L \in \mathcal{L}$ determines two linear subspaces of $\mathcal{V}$: the *null space* or *kernel* $\mathcal{N}(L) := L^{-1}(\{\mathbf{0}\})$ and the *range* of values $L(\mathcal{V})$. L is *injective* iff the first of those equals $\{\mathbf{0}\}$ and *surjective* iff the second of those equals $\mathcal{V}$. Both properties have a natural algebraic characterization:

$$L \ \text{injective} \quad \Longleftrightarrow \quad \exists\, L' \ \ L' \circ L = \mathbf{1} \tag{54}$$

$$L \ \text{surjective} \quad \Longleftrightarrow \quad \exists\, L' \ \ L \circ L' = \mathbf{1} \tag{55}$$

These statements hold in any linear space. In finite-dimensional spaces, as for our $\mathcal{V}$, we easily see

$$\dim(\mathcal{N}(L)) + \dim(L(\mathcal{V})) = \dim(\mathcal{V}) \ , \tag{56}$$

$$\det(L) = 0 \quad \Longleftrightarrow \quad \dim(L(\mathcal{V})) < \dim(\mathcal{V}) \quad \Longleftrightarrow \quad \dim(\mathcal{N}(L)) > 0 \tag{57}$$

which implies

$$L \ \text{surjective} \quad \Longleftrightarrow \quad L \ \text{injective} \quad \Longleftrightarrow \quad L \ \text{bijective} \quad \Longleftrightarrow \quad \det(L) \neq 0 \ . \tag{58}$$

For each bijective $L \in \mathcal{L}$ one writes L^{-1} for its *inverse* and has

$$L^{-1} \in \mathcal{L} \ , \quad L^{-1} \circ L = L \circ L^{-1} = \mathbf{1} \ , \quad (L^{-1})^{-1} = L \ , \quad \det(L^{-1}) = \det(L)^{-1} \ . \tag{59}$$

Of course, for bijective $L, L' \in \mathcal{L}$ the product is bijective, as is L^*, and we have

$$(L \circ L')^{-1} = (L')^{-1} \circ L^{-1} \ , \quad (L^*)^{-1} = (L^{-1})^* \ . \tag{60}$$

Therefore, the bijective elements of $\mathcal{L}$ form a *group* with unity, multiplication, and inversion as defined above. This group is normally named the *general linear group in three dimensions* $GL(3,\mathbb{R})$. The subset of elements L satisfying $\det(L) = 1$ is the *special linear group in three dimensions* $SL(3,\mathbb{R})$. An element L of $GL(3,\mathbb{R})$ which satisfies $L^{-1} = L^*$ is said to be *orthogonal*; it also satisfies $L \circ L^* = L^* \circ L = \mathbf{1}$ and

$$L\,\mathbf{v} \cdot L\,\mathbf{w} = \mathbf{v} \cdot \mathbf{w} \quad \forall\, \mathbf{v}, \mathbf{w} \in \mathcal{V} \ . \tag{61}$$

The orthogonal elements form a subgroup $O(3)$ of $GL(3,\mathbb{R})$. There is no need to express reference to the real numbers in the notation since the corresponding group in complex

spaces, the unitary group, is named by a different letter: $U(3)$. Within $O(3)$ we have a further subgroup singled out by satisfying any of the two properties

$$\det(L) = 1\ , \quad L\,\mathbf{v} \times L\,\mathbf{w} = L(\mathbf{v} \times \mathbf{w}) \quad \forall\, \mathbf{v}, \mathbf{w} \in \mathcal{V}\ . \tag{62}$$

which are equivalent on $O(3)$. This is the *special orthogonal group* $SO(3)$ and its elements are called *rotations*. Since the rotations thus are compatible with all operations which define an Euclidean oriented vector space and (since they are invertible) they are the *isomorphisms* of such spaces.

4.2 Eigenvalues and eigenvectors

By definition, an *eigenvalue* of $L \in \mathcal{L}$ is a number $\lambda \in \mathbb{R}$ for which there is a vector $\mathbf{x} \in \mathcal{V}, \mathbf{x} \neq \mathbf{0}$, such that $L\,\mathbf{x} = \lambda\,\mathbf{x}$. Each vector having all the properties mentioned here for $\mathbf{x}$ is called an *eigenvector* of L, and λ is said to be the eigenvalue belonging to $\mathbf{x}$. An interesting fact concerning $\mathcal{L}$ is: *Each $L \in \mathcal{L}$ has at least one eigenvector.* As is well known, and is an easy consequence of (57), λ is an eigenvalue of L iff $\det(\lambda\,\mathbf{1} - L) = 0$. Equation (53) shows that this means for λ that it be the root of a polynomial of third order. Such a root exists for sure since every polynomial of an odd order can easily be shown to attain positive as well as negative values and is continuous. This has an interesting implication for elements of $SO(3)$: For each element of $O(3)$ we see from (61) that all eigenvalues belong to $\{+1, -1\}$. To see that $+1$ is an eigenvalue we thus need only to show that there is an eigenvalue > 0: The characteristic polynomial of L has the value $-\det(L)$ for $\lambda = 0$ and tends to $+\infty$ for λ tending to $+\infty$. For $L \in SO(3)$ the term $-\det(L)$ is negative, and we have a sign change in $(0, +\infty)$ which implies the existence of a positive root. The eigenvector to which the eigenvalue $+1$ belongs is an invariant vector which represents an *axis of rotation*. The plane formed by the vectors orthogonal to the axis thus is also invariant. In this plane a mapping is induced which necessarily belongs to $SO(2)$. (Definition and properties of $SO(2)$ are so simple that they are taken for granted.) Unless their rotation angle is π there are no eigenvalues of this map. For π-rotation obviously every vector is an eigenvector to eigenvalue -1.

4.3 The exponential function

These operations allow us to associate with each $L \in \mathcal{L}$ and each power series with real coefficients and with infinite radius of convergence (e.g. a polynomial) a tensor by replacing the variable of the the power series by L. Studying the set of polynomials p for which $p(L) = O$ is the usual algebraic approach to the *spectral theory* of $\mathcal{L}$. However, also power series of transcendental functions play a useful role, particularly that of the *exponential function*

$$\exp(L) := \mathbf{1} + L + \frac{1}{2}L \circ L + \frac{1}{2 \cdot 3}L \circ L \circ L + \cdots + \frac{1}{n!}L^n \cdots\ . \tag{63}$$

For each L this term represents a *linear* map. On the other hand, this term depends on L in a *non-linear* manner. In quantum theory a corresponding mixed non-linear/linear

structure reconciles the linear state structure (superposition principle) with the non-linearity associated with interaction (roughly: 'the whole is more than the sum of its parts'). The following properties hold:

$$\exp(O) = \mathbf{1}\,, \quad (\exp(L))^* = \exp(L^*)\,, \quad \det(\exp(L)) = \exp(\mathrm{tr}(L))\,, \tag{64}$$

$$L \circ \exp(L') \circ L^{-1} \;=\; \exp(L \circ L' \circ L^{-1})\,, \tag{65}$$

$$L \circ L' = L' \circ L \quad \Longleftrightarrow \quad \exp(L) \circ \exp(L') = \exp(L + L')\,. \tag{66}$$

$L \in \mathcal{L}$ is said to be *symmetric* iff $L = L^*$ and *skew-symmetric* iff $L = -L^*$. Each $L \in \mathcal{L}$ is the sum of a symmetric element $L_s := \frac{1}{2}(L + L^*)$ and a skew-symmetric (anti-symmetric) one $L_a = \frac{1}{2}(L - L^*)$. From (64) and (66) one easily finds that $\exp(L)$ is orthogonal iff L is skew-symmetric, and that $\exp(L)$ is symmetric iff L is symmetric. If L and L^* *commute* (i.e. $L \circ L^* = L^* \circ L$) statement (66) yields a factorization of $\exp(L)$

$$\exp(L) = \exp(L_s + L_a) = \exp(L_s) \circ \exp(L_a) \tag{67}$$

into a symmetric and an orthogonal factor which is similar to the polar decomposition in 4.7. It is, however, not identical to it since in the present case the eigenvalues of the symmetrical factor need not to be positive.

4.4 Homotheties, dyads, and projectors

There are two types of tensors which play a prominent role in clarifying the structure of $\mathcal{L}$: *multiples of the identity*, also named *homotheties*, and the *dyads*. We start with the multiples of the identity and set for all $\alpha \in \mathbb{R}$

$$\underline{\alpha} \,:\, \mathcal{V} \,\to\, \mathcal{V}\,, \quad \mathbf{v} \,\mapsto\, \alpha\,\mathbf{v}\,. \tag{68}$$

Obviously $\underline{\alpha}$ is symmetric and satisfies $\det(\underline{\alpha}) = \alpha^3$ and $\mathrm{tr}(\underline{\alpha}) = 3\alpha$. Further, it commutes with every $L \in \mathcal{L}$: $\underline{\alpha} \circ L = L \circ \underline{\alpha}$. Interestingly, the converse is also true: If L_0 satisfies $L_0 \circ L = L \circ L_0$ for all $L \in \mathcal{L}$, there is an $\alpha \in \mathbb{R}$ such that $L_0 = \underline{\alpha}$. The mapping

$$\_ \,:\, \mathbb{R} \,\to\, \mathcal{L}\,, \quad \alpha \,\mapsto\, \underline{\alpha} \tag{69}$$

is a homomorphism of rings. In particular, the zero-element O in the linear space $\mathcal{L}$ equals $\underline{0}$, the unit element $\mathbf{1}$ of the associative algebra $\mathcal{L}$ equals $\underline{1}$.

Universal building blocks in $\mathcal{L}$ are the *dyadic products* of vectors $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ which are defined as follows:

$$|\,\mathbf{a}\,\rangle\langle\,\mathbf{b}\,| \,:\, \mathcal{V} \,\to\, \mathcal{V}\,, \quad \mathbf{v} \,\mapsto\, (\mathbf{b}\cdot\mathbf{v})\,\mathbf{a}\,. \tag{70}$$

We have $\det(|\,\mathbf{a}\,\rangle\langle\,\mathbf{b}\,|) = 0$ and $\mathrm{tr}(|\,\mathbf{a}\,\rangle\langle\,\mathbf{b}\,|) = \mathbf{a}\cdot\mathbf{b}$. Let $(\mathbf{e}_1, \mathbf{e}_1, \mathbf{e}_1)$ be an *ortho-normal* triplet (i.e. a list of three mutually orthogonal unit vectors) and $L \in \mathcal{L}$ arbitrary, then one has the representation

$$L = |\,L\,\mathbf{e}_1\,\rangle\langle\,\mathbf{e}_1\,| + |\,L\,\mathbf{e}_2\,\rangle\langle\,\mathbf{e}_2\,| + |\,L\,\mathbf{e}_3\,\rangle\langle\,\mathbf{e}_3\,| \tag{71}$$

as a sum of three *dyads*. The algebraic properties of dyads are simple:

$$|\,\mathbf{u}\,\rangle\langle\,\mathbf{v}\,|\circ|\,\mathbf{w}\,\rangle\langle\,\mathbf{x}\,| = \mathbf{v}\cdot\mathbf{w}\,|\,\mathbf{u}\,\rangle\langle\,\mathbf{x}\,|\,, \quad (|\,\mathbf{u}\,\rangle\langle\,\mathbf{v}\,|)^* = |\,\mathbf{v}\,\rangle\langle\,\mathbf{u}\,|\,, \tag{72}$$

$$|\,\mathbf{u}\,\rangle\langle\,\mathbf{v}\,|\cdot|\,\mathbf{w}\,\rangle\langle\,\mathbf{x}\,| = (\mathbf{u}\cdot\mathbf{w})\,(\mathbf{v}\cdot\mathbf{x})\,. \tag{73}$$

For each unit vector $\mathbf{u}$ the dyad

$$P_{\mathbf{u}} := |\,\mathbf{u}\,\rangle\langle\,\mathbf{u}\,| \tag{74}$$

is a *projector*. This means that it is *idempotent* (i.e. $P_{\mathbf{u}}\circ P_{\mathbf{u}} = P_{\mathbf{u}}$) and symmetric [9] . These properties are simple consequences of (72). An arbitrary projector $P\in\mathcal{L}$ is characterized by its range $P(\mathcal{V})$: Let $\mathbf{e_1},\ldots,\mathbf{e_r}$ be an orthonormal basis in this space, then the following representation $P = P_{\mathbf{e_1}}+\ldots+P_{\mathbf{e_r}}$ — actually a *spectral representation* (see 4.6) — holds. For any linear subspace $S\subset\mathcal{V}$ we write P_S for the projector which is uniquely determined by the property that $P_S(\mathcal{V}) = S$. With P_S also $\mathbf{1}-P_S$ is a projector. Its range is $S^{\perp}$.

Equation (73) says that the map

$$\otimes\,:\,\mathcal{V}\times\mathcal{V}\,\to\,\mathcal{L}\,,\quad (\mathbf{v},\mathbf{w})\,\mapsto\,|\,\mathbf{v}\,\rangle\langle\,\mathbf{w}\,| \tag{75}$$

is a *tensor product* of Euclidean spaces. This means that the function is bilinear and that scalar products of tensors and vectors are related by

$$(\mathbf{v}\otimes\mathbf{w})\cdot(\mathbf{v}'\otimes\mathbf{w}') = (\mathbf{v}\cdot\mathbf{v}')\,(\mathbf{w}\cdot\mathbf{w}')\,. \tag{76}$$

4.5 Linear maps induced by the vector product

Just as the scalar product allowed us to define the dyads, also the vector product allows defining an interesting subset of $\mathcal{L}$:

$$A_{\mathbf{a}}\,:\,\mathcal{V}\,\to\,\mathcal{V}\,,\quad \mathbf{v}\,\mapsto\,\mathbf{a}\times\mathbf{v}\,. \tag{77}$$

For each $\mathbf{a}\in\mathcal{V}$ we have $A_{\mathbf{a}}\in\mathcal{L}$ and the association $\mathbf{a}\mapsto A_{\mathbf{a}}$ sets up a bijection between vectors and skew-symmetric tensors. Some authors (e.g. [23], p. 211) write $\hat{\mathbf{a}}$ for that $A_{\mathbf{a}}$, a notation which we have used up to denote unit vector formation. One easily shows $\det(A_{\mathbf{a}}) = \mathrm{tr}(A_{\mathbf{a}}) = 0$. The well-known identity

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = (\mathbf{u}\cdot\mathbf{w})\,\mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\,\mathbf{w} \tag{78}$$

shows that the product of maps of type (77) is a linear combination of a dyad and a homothety:

$$A_{\mathbf{u}}\circ A_{\mathbf{v}} = |\,\mathbf{v}\,\rangle\langle\,\mathbf{u}\,| - (\mathbf{u}\cdot\mathbf{v})\,\mathbf{1} \tag{79}$$

and the Jacobi identity (5) implies

$$A_{\mathbf{u}}\circ A_{\mathbf{v}} - A_{\mathbf{v}}\circ A_{\mathbf{u}} = A_{\mathbf{u}\times\mathbf{v}}\,. \tag{80}$$

[9] There is another terminology (which will not be used here) according to which idempotent linear maps are projectors and symmetric projectors are *orthogonal projectors* or *ortho-projectors*.

Although the trace of any $A_{\mathbf{u}}$ vanishes, it allows recovering the scalar product:

$$\operatorname{tr}(A_{\mathbf{u}} \circ A_{\mathbf{v}}) = \operatorname{tr}(|\,\mathbf{v}\,\rangle\langle\,\mathbf{u}\,|) - (\mathbf{u}\cdot\mathbf{v})\operatorname{tr}(\mathbf{1}) = -2\,\mathbf{u}\cdot\mathbf{v}\ . \tag{81}$$

For the skew-symmetric part of a dyad the associated vector is therefore given by

$$|\,\mathbf{u}\,\rangle\langle\,\mathbf{v}\,| - |\,\mathbf{v}\,\rangle\langle\,\mathbf{u}\,| = A_{\mathbf{v}\times\mathbf{u}}\ . \tag{82}$$

In discussing the motion of rigid bodies we will encounter the following vector equation, which is written here followed by its solution:

$$\mathbf{x} + \mathbf{a}\times\mathbf{x} = \mathbf{y} \quad\Longleftrightarrow\quad \mathbf{x} = \frac{\mathbf{y} - \mathbf{a}\times\mathbf{y} + (\mathbf{a}\cdot\mathbf{y})\mathbf{a}}{1 + \mathbf{a}\cdot\mathbf{a}}\ . \tag{83}$$

In terms of inversion of linear maps this may be written as

$$(\underline{1} + A_{\mathbf{a}})^{-1} = \frac{\underline{1} - A_{\mathbf{a}} + |\,\mathbf{a}\,\rangle\langle\,\mathbf{a}\,|}{1 + \mathbf{a}\cdot\mathbf{a}}\ . \tag{84}$$

A more general version of this, the *resolvent* $(\underline{\lambda} - A_{\mathbf{a}})^{-1}$, $\lambda \in \mathbb{R}$ can, due to the obvious property

$$A_{\alpha\,\mathbf{a}} = \alpha\, A_{\mathbf{a}} \tag{85}$$

be expressed in terms of (84)

$$(\underline{\lambda} - A_{\mathbf{a}})^{-1} = \frac{\left(\underline{1} + A_{-\frac{\mathbf{a}}{\lambda}}\right)^{-1}}{\lambda}\ . \tag{86}$$

Many writers, especially in physics, prefer matrices over linear maps and so represent the matrix belonging to $A_{\mathbf{a}}$ (with respect to a positively oriented orthonormal basis) in terms of *spin matrices* (for spin 1)

$$S_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\ ,\quad S_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}\ ,\quad S_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{87}$$

as $(A_{\mathbf{a}})_{ij} = \sum_{k=1}^{3} a_k (S_k)_{ij}$. It is instructive to note that this holds in any positively oriented basis with the same matrices S_i. With the Levi-Civita epsilon function

$$\epsilon : \{1,2,3\}\times\{1,2,3\}\times\{1,2,3\} \to \{-1,0,1\}\ ,\quad (i,j,k) \mapsto -\frac{1}{2}(i-k)(j-i)(k-j) \tag{88}$$

the matrices S_i can be written as

$$(S_i)_{jk} = -\epsilon_{ijk}\ . \tag{89}$$

They enjoy the property $[S_i, S_j] := S_i S_j - S_j S_i = \sum_{k=1}^{3} \epsilon_{ijk} S_k$ which mirrors (80). A beneficial property of the spin matrices is that they provide an algorithm for finding $\mathbf{a}$ in the representation $S = A_{\mathbf{a}}$ of an arbitrary skew-symmetric $S \in \mathcal{L}$.

$$a_i = -\frac{1}{2}\operatorname{tr}(S\,S_i)\ ,\quad \forall i \in \{1,2,3\}\ , \tag{90}$$

where S under the trace is the matrix of S with respect to the basis to which the components a_i refer.

4.6 Spectral representation of symmetric tensors

With symmetric tensors there is more structure associated than with the skew-symmetric ones: For any symmetric $L \in \mathcal{L}$ there is a triplet $(\lambda_1, \lambda_2, \lambda_3)$ of real numbers (the *eigenvalues* of L) and an Euclidean system of reference $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ such that

$$L = \lambda_1 \,|\, \mathbf{e}_1 \,\rangle\langle\, \mathbf{e}_1 \,| + \lambda_2 \,|\, \mathbf{e}_2 \,\rangle\langle\, \mathbf{e}_2 \,| + \lambda_3 \,|\, \mathbf{e}_3 \,\rangle\langle\, \mathbf{e}_3 \,| \,. \tag{91}$$

Let us call this a *spectral representation* of L. It is uniquely determined by L if the eigenvalues are mutually different and are indexed according to falling value. If, for instance $\lambda_2 = \lambda_3$, then $\mathbf{e}_2, \mathbf{e}_3$ may be replaced by $\cos\varphi\,\mathbf{e}_2 + \sin\varphi\,\mathbf{e}_3, -\sin\varphi\,\mathbf{e}_2 + \cos\varphi\,\mathbf{e}_3$ for any $\varphi \in \mathbb{R}$ and the expression (91) remains a spectral representation of L. If all eigenvalues are equal, $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ may be replaced by any Euclidean system of reference and the expression (91) remains a spectral representation of L (then L is a homothety: $L = \underline{\lambda_1} = \underline{\lambda_2} = \underline{\lambda_3}$). For the smallest eigenvalue $\lambda := \inf\{\lambda_1, \lambda_2, \lambda_3\}$ and the largest one $\Lambda := \sup\{\lambda_1, \lambda_2, \lambda_3\}$ we have the representation

$$\lambda = \inf\{\mathbf{v} \cdot L\,\mathbf{v} \;:\; \mathbf{v} \in \mathcal{V},\, |\mathbf{v}| = 1\}\,, \quad \Lambda = \sup\{\mathbf{v} \cdot L\,\mathbf{v} \;:\; \mathbf{v} \in \mathcal{V},\, |\mathbf{v}| = 1\}\,. \tag{92}$$

and the vectors for which the extrema are attained are the eigenvectors for the respective values. For each function $f : \mathbb{R} \to \mathbb{R}$ one sees that

$$f(L) := f(\lambda_1) \,|\, \mathbf{e}_1 \,\rangle\langle\, \mathbf{e}_1 \,| + f(\lambda_2) \,|\, \mathbf{e}_2 \,\rangle\langle\, \mathbf{e}_2 \,| + f(\lambda_3) \,|\, \mathbf{e}_3 \,\rangle\langle\, \mathbf{e}_3 \,| \,. \tag{93}$$

is a definition, i.e. the linear map $f(L)$ is independent of the spectral representation used in (93). Remarkably, functions of L defined in this way and functions defined via power series (see (63)) coincide if both construction methods work. A symmetric operator L whose smallest eigenvalue is positive (in the sense of ≥ 0) is said itself to be *positive*. For any such operator there is a well defined *square root* $L^{1/2}$ and actually any power L^p is defined with positive p. If the smallest eigenvalue of L is strictly positive, then L^p is defined for any real p.

Typically the knowledge of spectral representations of two symmetric tensors L_1 and L_2 does not entail a straightforward spectral representation of the sum (which, of course, also is a symmetric tensor). In the three-dimensional case it is easy to obtain the spectral representation by explicit formulas so that reference to the spectral properties of the addends will hardly be necessary. For the corresponding problem for linear operators in higher dimensional spaces a good tool is the *second resolvent equation*

$$(\underline{\lambda} - (L_1 + L_2))^{-1} = (\underline{\lambda} - L_1)^{-1} + (\underline{\lambda} - (L_1 + L_2))^{-1}\, L_2\, (\underline{\lambda} - L_1)^{-1} \tag{94}$$

which can be iterated to yield

$$\begin{aligned}(\underline{\lambda} - (L_1 + L_2))^{-1} = (\underline{\lambda} - L_1)^{-1} + \sum_{i=1}^{n} (\underline{\lambda} - L_1)^{-1}\, {L_2}^i\, (\underline{\lambda} - L_1)^{-1} \\ + (\underline{\lambda} - (L_1 + L_2))^{-1}\, {L_2}^{n+1}\, (\underline{\lambda} - L_1)^{-1}\end{aligned} \tag{95}$$

and thus arrives at an expression which no longer contains $L_1 + L_2$ if the final term can be shown to vanish in the limit of growing n. A similar elimination of the sum in a function expression is provided by the famous *Kato-Trotter formula*

$$\exp(L_1 + L_2) = \lim_{n\to\infty} \left(\exp(L_1/n)\, \exp(L_2/n)\right)^n \ . \tag{96}$$

The corresponding question for the product instead of the sum is less intricate. It can have an answer only if $L_1\, L_2$ is symmetric and this is the case iff L_1 and L_2 commute. Then the situation is canonical: There exists a orthonormal triplet $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ of common eigenvectors so that we have the spectral representations

$$\begin{aligned} L_1 &= \lambda_1\, |\, \mathbf{e}_1\, \rangle\langle\, \mathbf{e}_1\, | + \lambda_2\, |\, \mathbf{e}_2\, \rangle\langle\, \mathbf{e}_2\, | + \lambda_3\, |\, \mathbf{e}_3\, \rangle\langle\, \mathbf{e}_3\, | \\ L_2 &= \mu_1\, |\, \mathbf{e}_1\, \rangle\langle\, \mathbf{e}_1\, | + \mu_2\, |\, \mathbf{e}_2\, \rangle\langle\, \mathbf{e}_2\, | + \mu_3\, |\, \mathbf{e}_3\, \rangle\langle\, \mathbf{e}_3\, | \end{aligned} \tag{97}$$

and sum and product have the following obvious spectral representations

$$\begin{aligned} L_1 + L_2 &= (\lambda_1 + \mu_1)\, |\, \mathbf{e}_1\, \rangle\langle\, \mathbf{e}_1\, | + (\lambda_2 + \mu_2)\, |\, \mathbf{e}_2\, \rangle\langle\, \mathbf{e}_2\, | + (\lambda_3 + \mu_3)\, |\, \mathbf{e}_3\, \rangle\langle\, \mathbf{e}_3\, | \\ L_1 L_2 &= (\lambda_1\, \mu_1)\, |\, \mathbf{e}_1\, \rangle\langle\, \mathbf{e}_1\, | + (\lambda_2\, \mu_2)\, |\, \mathbf{e}_2\, \rangle\langle\, \mathbf{e}_2\, | + (\lambda_3\, \mu_3)\, |\, \mathbf{e}_3\, \rangle\langle\, \mathbf{e}_3\, | \ . \end{aligned} \tag{98}$$

4.7 Singular value decomposition and polar decomposition

For tensors which are not symmetric one achieves something very similar to a spectral representation by considering the symmetric tensor $L^* \circ L$, which equals $\underline{0}$ iff $L = \underline{0}$. Since we need nothing to learn about the zero tensor, we assume $L \neq \underline{0}$. Obviously, all eigenvalues of $L^* \circ L$ are non-negative and so one writes them conveniently as squares of numbers in writing down a spectral representation

$$L^* \circ L = \lambda_1^2\, |\, \mathbf{e}_1\, \rangle\langle\, \mathbf{e}_1\, | + \lambda_2^2\, |\, \mathbf{e}_2\, \rangle\langle\, \mathbf{e}_2\, | + \lambda_3^2\, |\, \mathbf{e}_3\, \rangle\langle\, \mathbf{e}_3\, | \ . \tag{99}$$

Omitting the terms $\lambda_i^2\, |\, \mathbf{e}_i\, \rangle\langle\, \mathbf{e}_i\, |$ with $\lambda_i = 0$ we write this

$$L^* \circ L = \sum_{i=1}^{r} \lambda_i^2\, |\, \mathbf{e}_i\, \rangle\langle\, \mathbf{e}_i\, | \ , \quad 1 \leq r \leq 3 \tag{100}$$

where now all $\lambda_i > 0$ and the $\mathbf{e}_i$ are an ortho-normal list of vectors which is uniquely determined if the λ_i are mutually different. Defining

$$\mathbf{f}_i := \frac{1}{\lambda_i}\, L\mathbf{e}_i = \text{unit vector in the direction of } \ L\mathbf{e}_i \tag{101}$$

gives us an orthonormal list $(\mathbf{f}_i)_{i\in\{1,\ldots,r\}}$ of vectors which provides the surprisingly simple representation

$$L = \sum_{i=1}^{r} \lambda_i\, |\, \mathbf{f}_i\, \rangle\langle\, \mathbf{e}_i\, | \tag{102}$$

which is known as a *singular value decomposition* of L. It allows defining the *pseudo inverse* of L as

$$L^+ = \sum_{i=1}^{r} \frac{1}{\lambda_i}\, |\, \mathbf{e}_i\, \rangle\langle\, \mathbf{f}_i\, | \ . \tag{103}$$

Further, (102) implies what is called the *polar decomposition* of L:

$$L = \hat{L} \circ |L| \,, \quad \text{where } |L| := \sum_{i=1}^{r} \lambda_i \,|\, \mathbf{e}_i \,\rangle\langle\, \mathbf{e}_i \,|\,, \quad \hat{L} := \sum_{i=1}^{r} |\, \mathbf{f}_i \,\rangle\langle\, \mathbf{e}_i \,| \tag{104}$$

which implies

$$\hat{L}^* \circ \hat{L} = \sum_{i=1}^{r} |\, \mathbf{e}_i \,\rangle\langle\, \mathbf{e}_i \,|\,, \quad \hat{L} \circ \hat{L}^* = \sum_{i=1}^{r} |\, \mathbf{f}_i \,\rangle\langle\, \mathbf{f}_i \,|\,. \tag{105}$$

Although a singular value decomposition of L is determined uniquely by L only if the λ_i all are different, the tensors L^+, $|L|$, and $\hat{L}$ are uniquely determined by L. From (104) it is clear that $(\mathbf{f}_1, \ldots, \mathbf{f}_r)$ is an orthonormal basis in $L(\mathcal{V})$, and $(\mathbf{e}_1, \ldots, \mathbf{e}_r)$ is an orthonormal basis in $\mathcal{E}(L) := \mathcal{N}(L)^\perp$, the set of all vectors which are orthogonal to all vectors in the null space $\mathcal{N}(L)$. *Essential domain* would be an expressive name for $\mathcal{E}(L)$ since all values $\mathbf{y} = L(\mathbf{x}), \mathbf{x} \in \mathcal{V}$ can already be obtained as $L(\mathbf{x}')$ for some $\mathbf{x}' \in \mathcal{E}(L)$. From (105) we see that $\hat{L}^* \circ \hat{L}$ is the projector onto $\mathcal{E}(L)$ and $\hat{L} \circ \hat{L}^*$ is the projector onto $L(\mathcal{V})$. The factor $|L|$ leaves $\mathcal{E}(L)$ invariant. $\hat{L}$, restricted to $\mathcal{E}(L)$, is a bijection onto $L(\mathcal{V})$. Restricted to the ortho-complement $\mathcal{N}(L)$ of $\mathcal{E}(L)$ $\hat{L}$ vanishes. $\hat{L}$ of such a structure is said to be *partially isometric*. Therefore, for bijective L the factor $\hat{L}$ is orthogonal. Simplifying circumstances prevail for *normal* [10] tensors, i.e for tensors L satisfying $L \circ L^* = L^* \circ L$. For them we have $\mathcal{N}(L) = \mathcal{N}(L^*)$ and therefore $L(\mathcal{V}) = \mathcal{E}(L)$. It should be noticed that in real vector spaces normality, unlike symmetry, does not guarantee the existence of a complete set of eigenvectors as it does in complex spaces. For instance, rotation with a rotation angle different from π (see next section) have only a single eigenvector, namely the rotation axis.

As an instructive example we investigate the singular value decomposition of the skew-symmetric map $A_{\mathbf{a}}$ (see (77)). For the sake of simplicity we assume $|\mathbf{a}| = 1$, to which the general case can be reduced due to (85). We form a positively oriented orthonormal basis $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ the first member of which is $\mathbf{a}$: $\mathbf{a} = \mathbf{a}_1$. First, we have to consider the expression of (99): $A^*_{\mathbf{a}} \circ A_{\mathbf{a}} = -A_{\mathbf{a}} \circ A_{\mathbf{a}} = (\text{by } (79)) = \mathbf{1} - P_{\mathbf{a}}$. By the statements following (74) this is the projection to the plane $\{\mathbf{a}\}^\perp$. More explicitly, $|A_{\mathbf{a}}| = P_{\mathbf{a}_2} + P_{\mathbf{a}_3}$. Thus, using the notation introduced so far for L: $r = 2$, $\mathbf{e}_1 = \mathbf{a}_2, \mathbf{e}_2 = \mathbf{a}_3$, $\lambda_1 = \lambda_2 = 1$, $\mathbf{f}_1 = \mathbf{a} \times \mathbf{e}_1 = \mathbf{a}_1 \times \mathbf{a}_2 = \mathbf{a}_3 = \mathbf{e}_2$, $\mathbf{f}_2 = \mathbf{a} \times \mathbf{e}_2 = \mathbf{a}_1 \times \mathbf{a}_3 = -\mathbf{a}_2 = -\mathbf{e}_1$. Since the non-zero singular values all equal 1, we see from (102) and the last equation of (104) that $\widehat{A_{\mathbf{a}}} = A_{\mathbf{a}}$.

5 Rotations

As was stated in 4.3, for each skew-symmetric map S the exponential $\exp(S)$ is orthogonal. Further, $\det(\exp(S)) = \exp(\mathrm{tr}(S)) = \exp(0) = 1$ and thus $\exp(S) \in SO(3)$. Now, each skew-symmetric $S \in \mathcal{L}$ is of the form $A_{\mathbf{a}}$ for some $\mathbf{a} \in \mathcal{V}$ (see (90)). So let us consider $\exp(A_{\mathbf{a}})$. Writing $\mathbf{a} = \varphi\, \mathbf{n}$, with the unit vector $\mathbf{n} := \hat{\mathbf{a}}$, we have $A_{\mathbf{a}} = \underline{\varphi} \circ A_{\mathbf{n}} = \varphi\, A_{\mathbf{n}}$ so

[10] $|\, \mathbf{a} \,\rangle\langle\, \mathbf{b} \,|$ fails to be normal whenever $\mathbf{a}$ and $\mathbf{b}$ are linearly independent; so there is nothing pathological about a tensor which is not normal

that we have to compute the powers of $A_{\mathbf{n}}$. This is surprisingly easy due to the following convenient relations: $A_{\mathbf{n}} \circ A_{\mathbf{n}} = P_{\mathbf{n}} - \mathbf{1}$ (see (74) for $P_{\mathbf{n}}$, and (79)), $P_{\mathbf{n}} \circ A_{\mathbf{n}} = A_{\mathbf{n}} \circ P_{\mathbf{n}} = \underline{0}$, and $P_{\mathbf{n}} \circ P_{\mathbf{n}} = P_{\mathbf{n}}$. They allow us to sum up the exponential series and to obtain

$$\exp(A_{\varphi\,\mathbf{n}}) = \exp(\varphi\, A_{\mathbf{n}}) = P_{\mathbf{n}} \;+\; \cos\varphi\,(\mathbf{1} - P_{\mathbf{n}}) \;+\; \sin\varphi\; A_{\mathbf{n}} =: R(\varphi, \mathbf{n})\;. \tag{106}$$

This represents the rotations in terms of a *rotation angle* φ and a *rotation axis* $\mathbf{n}$. The angle φ is determined uniquely, if restricted to $0 \le \varphi \le \pi$. The rotation axis is determined uniquely if $\varphi \neq \pi$, otherwise $\mathbf{n}$ and $-\mathbf{n}$ correspond to the same rotation. It is easy to see that any rotation which has an invariant vector (i.e. an eigenvector with eigenvalue 1) can be represented by the right-hand side of (106). That every rotation has in fact an invariant vector was shown in subsection 4.2. Thus

$$SO(3) = \{R(\varphi, \mathbf{n}) \;:\; \varphi \in [0, \pi],\; \mathbf{n} \in \hat{\mathcal{V}}\}\;, \tag{107}$$

where, of course, $\hat{\mathcal{V}} := \{\hat{\mathbf{v}} \;:\; \mathbf{v} \in \mathcal{V}\}$.

Rotations will play a decisive role in formulating a time-stepping algorithm for rigid bodies. In this algorithm, rotations have to act on many vectors and for many pairs of rotations R, R' one has to form the composition $R \circ R'$. It is therefore important to have fast algorithms for these operations. I use here the method of *Euler-Rodrigues parameters*, which is the most efficient one known. It seems to go back, (e.g. [13]), to the French mathematician Olinde Rodrigues, who wrote about it as early as 1840 [15]. Several rediscoveries happened in newer times, one—in 1986—by myself [16].

To arrive at this economic formulation we consider (106) when acting on a vector $\mathbf{x}$:

$$R(\varphi, \mathbf{n})\mathbf{x} = \cos\varphi\;\mathbf{x} + (1 - \cos\varphi)\,(\mathbf{n}\cdot\mathbf{x})\,\mathbf{n} + \sin\varphi\;\mathbf{n}\times\mathbf{x}\;. \tag{108}$$

As is well known, for all $\varphi \in (-\pi, +\pi)$

$$\sin\varphi = \frac{2\tan(\varphi/2)}{1 + (\tan(\varphi/2))^2}\;, \quad \cos\varphi = \frac{1 - (\tan(\varphi/2))^2}{1 + (\tan(\varphi/2))^2}\;. \tag{109}$$

Writing t for $\tan(\varphi/2)$ we get from (108)

$$R(\varphi, \mathbf{n})\,\mathbf{x} = \frac{(1 - t^2)\,\mathbf{x} + 2\,t^2\,(\mathbf{n}\cdot\mathbf{x})\,\mathbf{n} + 2\,t\,\mathbf{n}\times\mathbf{x}}{1 + t^2} \tag{110}$$

and in terms of the vector $\mathbf{r} := t\,\mathbf{n} = \tan(\varphi/2)\;\mathbf{n}$

$$R(\varphi, \mathbf{n})\,\mathbf{x} = \frac{(1 - \mathbf{r}\cdot\mathbf{r})\mathbf{x} + 2(\mathbf{r}\cdot\mathbf{x})\mathbf{r} + 2\mathbf{r}\times\mathbf{x}}{1 + \mathbf{r}\cdot\mathbf{r}} \quad =: \quad \mathcal{R}(\mathbf{r})\,\mathbf{x}\;. \tag{111}$$

What we have achieved thereby is that we have no longer two heterogeneous parameters — a rotation angle and a rotation axis — but a single parameter which is simply an element of $\mathcal{V}$. For a rotation given in terms of this new parameter $\mathbf{r}$ we write $\mathcal{R}(\mathbf{r})$ instead of $R(\mathbf{r})$ since we then can use $\mathcal{R}$ as the name of a map $\mathcal{V} \to SO(3)$ even if the context doesn't provide function arguments which would resolve ambiguities. If some $\mathbf{r} \in \mathcal{V}$ is

introduced as a means to bring about the rotation $\mathcal{R}(\mathbf{r})$ it will be called a *rotation vector*. Often it will be convenient to have for the term $\mathcal{R}(\mathbf{r})\,\mathbf{x}$ a more symmetrical notation. Therefore, we define the operation $* : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$

$$\mathbf{r} * \mathbf{x} := \frac{(1 - \mathbf{r}\cdot\mathbf{r})\mathbf{x} + 2(\mathbf{r}\cdot\mathbf{x})\mathbf{r} + 2\mathbf{r}\times\mathbf{x}}{1 + \mathbf{r}\cdot\mathbf{r}} \;; \tag{112}$$

it is linear in the second argument, non-linear in the first one, and — since the nominator never vanishes — continuous with respect to the standard topologies of the spaces involved.

The rotations $R(\pi, \mathbf{n})$ now play an exceptional role since these would correspond to rotation vectors $\mathbf{r}$ of length ∞ (of course, such vectors don't exist in $\mathcal{V}$). For $|\mathbf{r}| \gg 1$ the term $\mathbf{r} * \mathbf{x}$ can be well approximated by a very simple one:

$$\begin{aligned} \mathbf{r} * \mathbf{x} \approx \frac{(-\mathbf{r}\cdot\mathbf{r})\mathbf{x} + 2(\mathbf{r}\cdot\mathbf{x})\mathbf{r}}{\mathbf{r}\cdot\mathbf{r}} &= 2(\hat{\mathbf{r}}\cdot\mathbf{x})\hat{\mathbf{r}} - \mathbf{x} = (2P_{\hat{\mathbf{r}}} - \mathbf{1})\mathbf{x} \\ &= (P_{\hat{\mathbf{r}}} - (\mathbf{1} - P_{\hat{\mathbf{r}}}))\mathbf{x} = R(\pi, \hat{\mathbf{r}})\mathbf{x} \end{aligned} \tag{113}$$

which does no longer depend on the length of $\mathbf{r}$ and is invariant with respect to $\mathbf{r} \mapsto -\mathbf{r}$. Actually one finds

$$\| \mathcal{R}(\mathbf{r}) - R(\pi, \hat{\mathbf{r}}) \| = \frac{2}{1 + \mathbf{r}\cdot\mathbf{r}} \;, \tag{114}$$

where the norm from (51) is used. We may wonder whether π-rotations also play an exceptional role with respect to algebra. This is indeed the case: they are the only *unipotent* elements:

$$U \in SO(3) \;\wedge\; U \circ U = \mathbf{1} \quad\Leftrightarrow\quad \exists_{\mathbf{n}\in\hat{\mathcal{V}}}\; U = R(\pi, \mathbf{n}) \;. \tag{115}$$

In the algebra $\mathcal{L}$ (of which $SO(3)$ is a subset) any unipotent element U determines an idempotent element $P = \frac{1}{2}(\mathbf{1} + U)$. Since in our case U is orthogonal, it follows that P is symmetric and, therefore, a projector. Now U can be expressed in terms of P as $U = 2P - \mathbf{1} = P - (\mathbf{1} - P)$ which is the representation from (113) with $P = P_{\hat{\mathbf{r}}}$, so that we have learned that for an unipotent $U \in SO(3)$ the projector $P = \frac{1}{2}(\mathbf{1} + U)$ has a one-dimensional range. Then $P' := \mathbf{1} - P$ is a projector with two-dimensional range and $U' := 2P' - \mathbf{1}$ is a new unipotent orthogonal element of $\mathcal{L}$. One easily sees $U' = -U$ and, hence, $\det U' = -1$, $U' \in O(3)\backslash SO(3)$. Actually, U' is a reflection with respect to the plane for which $\hat{\mathbf{r}}$ is a normal vector. We thus see that the unipotent elements of $SO(3)$ form a bridge to the larger group $O(3)$ which houses rotations and reflections.

As a direct consequence of (107) we have

$$\mathcal{R}(\mathcal{V}) = \{R \in SO(3) \;:\; R \circ R \neq \mathbf{1}\} \;. \tag{116}$$

Obviously it would be canonical to get all of $SO(3)$ by such a mapping. This can be achieved by extending (and compactifying) $\mathcal{V}$ in a natural manner to

$$\mathcal{V}_c = \mathcal{V} \cup \mathbb{P}(\mathcal{V}) \;, \tag{117}$$

where $\mathbb{P}(\mathcal{V})$ is the *projective space* associated with $\mathcal{V}$: its elements are the equivalence classes [11] in $\hat{\mathcal{V}}$ with respect to the relation $\mathbf{v} \sim \mathbf{w}$ iff $\mathbf{v} = \mathbf{w}$ or $\mathbf{v} = -\mathbf{w}$. This allows writing $\pm\mathbf{v}$ (i.e. $\{\mathbf{v}, -\mathbf{v}\}$) for the generic element of $\mathbb{P}(\mathcal{V})$. Further, we will use the 'projection onto the horizon'

$$\tilde{\pi} \,:\, \mathcal{V} \,\to\, \mathbb{P}(\mathcal{V}) \,, \quad \mathbf{v} \,\mapsto\, \{\hat{\mathbf{v}}, -\hat{\mathbf{v}}\} \,. \tag{118}$$

Since $\mathcal{V}$ and $\mathbb{P}(\mathcal{V})$ are disjoint sets (no element of $\mathcal{V}$ equals an element of $\mathbb{P}(\mathcal{V})$, although elements of $\mathbb{P}(\mathcal{V})$ have elements of $\mathcal{V}$ as elements) we may extend the definition of $\mathcal{R}$ to $\mathcal{R}_c$ by defining $\mathcal{R}_c|\mathcal{V} := \mathcal{R}$ and defining on $\mathcal{R}_c$ on $\mathbb{P}(\mathcal{V})$ by

$$\mathcal{R}_c(\pm\mathbf{v}) := R(\pi, \mathbf{v}) = P_{\mathbf{v}} - (\mathbf{1} - P_{\mathbf{v}}) \,. \tag{119}$$

It is clear that $\mathcal{R}_c$ is injective and satisfies $\mathcal{R}_c(\mathcal{V}_c) = SO(3)$. It thus provides a a bijection between $\mathcal{V}_c$ and $SO(3)$. The group $SO(3)$ carries a natural topology which may be characterized among other equivalent forms as the metric topology generated by the Hilbert-Schmidt norm (see (50) et seq.). The main topological properties of $SO(3)$ are that it is *compact* and *twofold connected*. We define the topology on $\mathcal{V}_c$ by the rule that a subset A of $\mathcal{V}_c$ is *open* iff the subset $\mathcal{R}_c(A)$ of $SO(3)$ is open. Then, of course, the mapping $\mathcal{R}_c$ is an isomorphism of topological spaces (also called *homeomorphism*) and $\mathcal{V}_c$ has every topological property that $SO(3)$ is known to have. Especially, $\mathcal{V}_c$ is compact, a property which $\mathcal{V}$, taken alone, does not have. Augmenting $\mathcal{V}$ with $\mathbb{P}(\mathcal{V})$ is a process of *compactification* and $\mathcal{V}_c$ equals the *closure* $\overline{\mathcal{V}}$ of $\mathcal{V}$; $\mathbb{P}(\mathcal{V})$ is the *boundary* of $\mathcal{V}$ in $\mathcal{V}_c$ since (114) entails for all $\mathbf{v} \in \mathcal{V}$

$$\lim_{r\to\infty} \mathcal{R}(r\,\mathbf{v}) = \mathcal{R}_c(\tilde{\pi}(\mathbf{v})) \,. \tag{120}$$

Trying to eliminate the trigonometric functions in expressing the action of rotations on vectors suggests itself to anybody who strives for computational efficiency. It comes as a gift that the formula for composition of rotations becomes unbeatably simple. Let us consider the expression $\mathbf{r} * (\mathbf{r}' * \mathbf{x})$ and try it to write it in the form $\mathbf{r}'' * \mathbf{x}$. A lengthy but straightforward computation gives $\mathbf{r}'' = \frac{\mathbf{r}+\mathbf{r}'+\mathbf{r}\times\mathbf{r}'}{1-\mathbf{r}\cdot\mathbf{r}'}$. I found this relation pre-formed in lecture notes [12], Figure 1, of my academic teacher Hans Joachim Meister which I consulted in 1986 to simplify hobby-astronomical calculations, Figure 2. In just the same time I was involved in an industrial project which aimed at modifying a video tape drive in order to make it suitable as a random access storage system for high resolution images. For following computationally the tape in its helical windings around the video drum and guiding rollers I then used the formalism of rotation vectors extensively. The first document in my project folder that used rotation vectors is dated 1986-01-27; it derives the conversion between Euler angles and rotation vectors.

The combination law (product, written as $\circ$) in $SO(3)$ can be carried over via the bijection $\mathcal{R}_c$ to a combination law in $\mathcal{V}_c$ (also written as $\circ$) [12]

$$\circ \,:\, \mathcal{V}_c \times \mathcal{V}_c \,\to\, \mathcal{V}_c \,, \quad (\mathbf{a}, \mathbf{b}) \,\mapsto\, \mathcal{R}_c(\mathcal{R}_c^{-1}(\mathbf{a}) \circ \mathcal{R}_c^{-1}(\mathbf{b})) \,. \tag{121}$$

[11] there are realizations of projective spaces which don't make use of unit vectors

[12] As Figure 2 shows, I initially used the symbols $*$ and $\circ$ differently.

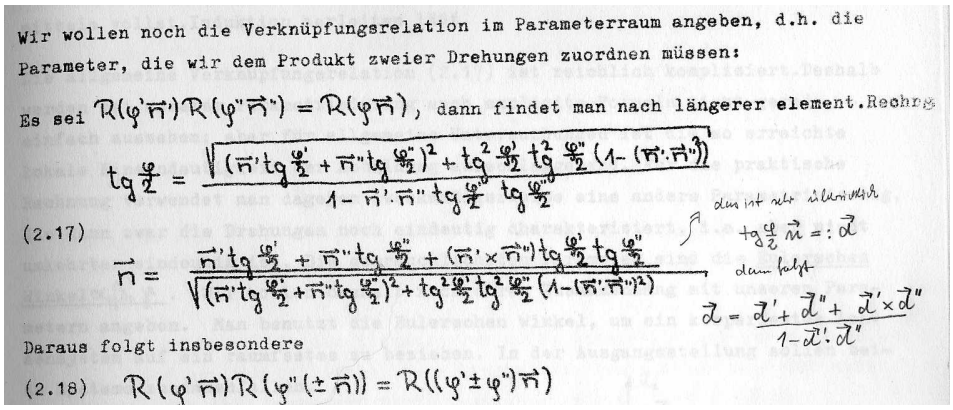


Figure 1: The composition of rotation axis and rotation angle in the lecture notes edited by Professor Meister 1965

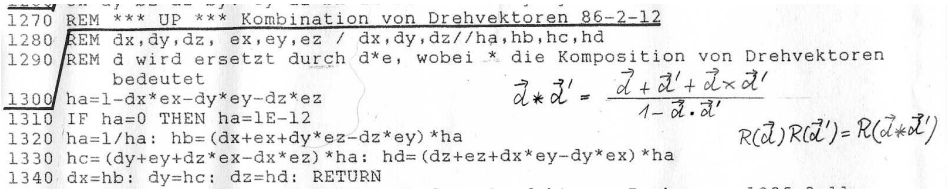


Figure 2: Combination of rotation vectors in one of my Basic programs of 1986

For the case of normal rotations (i.e. rotations with a rotation angle different from π) we have as pointed out before:

$$\mathbf{r} \circ \mathbf{r}' := \frac{\mathbf{r} + \mathbf{r}' + \mathbf{r} \times \mathbf{r}'}{1 - \mathbf{r} \cdot \mathbf{r}'} \quad \text{if} \quad \mathbf{r} \cdot \mathbf{r}' \neq 1 \,. \tag{122}$$

Equation (120) determines also the results of map $\circ$ for π-rotations as arguments or as results:

$$\mathbf{r} \circ \mathbf{r}' := \tilde{\pi}(\mathbf{r} + \mathbf{r}' + \mathbf{r} \times \mathbf{r}') \quad \text{if} \quad \mathbf{r} \cdot \mathbf{r}' = 1 \,, \tag{123}$$

$$(\pm\mathbf{n}) \circ \mathbf{r} := \begin{cases} -\frac{\mathbf{n}+\mathbf{n}\times\mathbf{r}}{\mathbf{n}\cdot\mathbf{r}} & \text{if } \mathbf{n} \cdot \mathbf{r} \neq 0, \\ \tilde{\pi}(\mathbf{n} + \mathbf{r} + \mathbf{n} \times \mathbf{r}) & \text{if } \mathbf{n} \cdot \mathbf{r} = 0 \,, \end{cases} \tag{124}$$

$$\mathbf{r} \circ (\pm\mathbf{n}) := \begin{cases} -\frac{\mathbf{n}+\mathbf{r}\times\mathbf{n}}{\mathbf{n}\cdot\mathbf{r}} & \text{if } \mathbf{n} \cdot \mathbf{r} \neq 0, \\ \tilde{\pi}(\mathbf{r} + \mathbf{n} + \mathbf{r} \times \mathbf{n}) & \text{if } \mathbf{n} \cdot \mathbf{r} = 0 \,, \end{cases} \tag{125}$$

$$(\pm\mathbf{n}) \circ (\pm\mathbf{n}') := \begin{cases} -\frac{\mathbf{n}\times\mathbf{n}'}{\mathbf{n}\cdot\mathbf{n}'} & \text{if } \mathbf{n} \cdot \mathbf{n}' \neq 0, \\ \tilde{\pi}(\mathbf{n} \times \mathbf{n}') & \text{if } \mathbf{n} \cdot \mathbf{n}' = 0 \,. \end{cases} \tag{126}$$

The operation $*$ extends naturally to the case that the first argument is in $\mathbb{P}(\mathcal{V})$:

$$(\pm\mathbf{n}) * \mathbf{x} \;:=\; \mathcal{R}_c(\pm\mathbf{n})\,\mathbf{x} \;=\; (2P_{\mathbf{n}} - \mathbf{1})\,\mathbf{x} \,. \tag{127}$$

Let us compile the basic properties of the operations $*$ and $\circ$, where now $\mathbf{r}, \mathbf{r}' \in \mathcal{V}_c$, $\mathbf{x} \in \mathcal{V}$, and for $\mathbf{r} \in \mathcal{V}_c \backslash \mathcal{V}$ the element $-\mathbf{r}$ is set to mean $\mathbf{r}$.

$$\mathcal{R}_c(\mathbf{r})\, \mathbf{x} = \mathbf{r} * \mathbf{x} \,, \tag{128}$$

$$(\mathbf{r} \circ \mathbf{r}') * \mathbf{x} = \mathbf{r} * (\mathbf{r}' * \mathbf{x}) \,, \quad \mathbf{r} \circ (-\mathbf{r}) = \mathbf{0} \,, \quad \mathbf{r} \circ \mathbf{0} = \mathbf{r} \,, \quad \mathbf{r} \circ \mathbf{r}' = (\mathbf{r} * \mathbf{r}') \circ \mathbf{r} \,, \tag{129}$$

$$\mathbf{r} \circ (\mathbf{r}' \circ \mathbf{r}'') = (\mathbf{r} \circ \mathbf{r}') \circ \mathbf{r}'' \,, \quad \mathbf{r} * (\mathbf{r}' * \mathbf{r}'') = (\mathbf{r} * \mathbf{r}') * \mathbf{r}'' \,, \tag{130}$$

which can be used to define $\circ$ in terms of $*$ and vice versa. From this we have the following properties of map $\mathcal{R}_c$

$$\mathcal{R}_c(\mathbf{r} \circ \mathbf{r}') = \mathcal{R}_c(\mathbf{r}) \circ \mathcal{R}_c(\mathbf{r}') \,, \tag{131}$$

which motivates the notation $\mathbf{r} \circ \mathbf{r}'$ and

$$\mathcal{R}_c(-\mathbf{r}) = \mathcal{R}_c(\mathbf{r})^{-1} \,, \quad \mathcal{R}_c(\mathbf{0}) = \mathbf{1} \,, \quad \mathcal{R}_c(\mathbf{r}) \circ \mathcal{R}_c(\mathbf{r}') \circ \mathcal{R}_c(\mathbf{r})^{-1} = \mathcal{R}_c(\mathcal{R}_c(\mathbf{r})\, \mathbf{r}') \,. \tag{132}$$

From the following two equations we encountered the first one just at the beginning of this section and its well-known twin should not remain unmentioned:

$$\det(R(\varphi, \mathbf{n})) = 1 \,, \quad \operatorname{tr}(R(\varphi, \mathbf{n})) = 2\, \cos(\varphi) \,+\, 1 \,. \tag{133}$$

Let us consider the computational realization of these expressions: As (112) is written, it defines the term but taking it literally as a recipe for computation would be not wise. What one would actually do is

$$\begin{aligned} s_0 &:= \mathbf{r} \cdot \mathbf{r} \,, \quad s_1 := 1/(1 + s_0) \,, \quad s_2 := 1 - s_0 \,, \quad \mathbf{x}' := 2\mathbf{x} \,, \quad s_3 := \mathbf{r} \cdot \mathbf{x}' \,, \\ \mathbf{r} * \mathbf{x} &:= s_1(s_2\mathbf{x} + s_3\mathbf{r} + \mathbf{r} \times \mathbf{x}') \,. \end{aligned}$$

If one needs to compute $\mathbf{r} * \mathbf{x}$ for some fixed $\mathbf{r}$ and a multitude of $\mathbf{x}$'s it pays first to compute the 3×3-matrix which represents $\mathcal{R}(\mathbf{r})$. Expressing everything in terms of components (identifying $\mathcal{V}$ with $\mathbb{R}^3$ which a program would do anyway) one can do even slightly better by extracting a common factor (c in the following): Putting $\mathbf{r} = (r_1, r_2, r_3)$ and $\mathbf{x} = (x_1, x_2, x_3)$ we define $\mathbf{y} = \mathbf{r} * \mathbf{x}$ by

$$R := \begin{pmatrix} {r_1}^2 + b & r_1 r_2 + r_3 & r_1 r_3 - r_2 \\ r_1 r_2 - r_3 & {r_2}^2 + b & r_2 r_3 + r_1 \\ r_1 r_3 + r_2 & r_2 r_3 - r_1 & {r_3}^2 + b \end{pmatrix} \,, \tag{134}$$

$$y_i := c(R_{i\,1} x_1 + R_{i\,2} x_2 + R_{i\,3} x_3) \,, \quad i = 1, 2, 3 \,, \tag{135}$$

where the numbers b and c are defined as follows:

$$\begin{aligned} \rho &:= {r_1}^2 + {r_2}^2 + {r_3}^2 \,, \\ b &:= (1 - \rho)/2 \,, \\ c &:= 2/(1 + \rho) \,. \end{aligned} \tag{136}$$

For a numerical realization of the composition law of rotations we never need to consider π-rotations since these can be approximated by rotations with an rotation angle close to π arbitrarily well. We select a small number $\epsilon \approx \sqrt{\epsilon_1/2}$, (here ϵ_1 is the spacing of the numerical lattice at number 1.0, thus $\epsilon_1 = 1.11 \cdot 10^{-16}$ for the double precision floating point numbers of C++) and define

$$\mathbf{r} \circ \mathbf{r}' := \frac{\mathbf{r} + \mathbf{r}' + \mathbf{r} \times \mathbf{r}'}{N} \,, \tag{137}$$

where the nominator N is defined

$$N := \begin{cases} 1 - \mathbf{r} \cdot \mathbf{r}' & \text{if} \quad |1 - \mathbf{r} \cdot \mathbf{r}'| > \epsilon \\ \epsilon \ \text{sign}(1 - \mathbf{r} \cdot \mathbf{r}') & \text{else (here sign}(0) := 1) \,. \end{cases} \tag{138}$$

The idea in setting ϵ is that a rotation vector $\mathbf{r}$ of length N generates a rotation which differs from a $\pi-$rotation only by numerical noise. For this difference to a $\pi-$rotation we employ (114). What kind of approximation is implied in this proposed approach to ignore the π-rotations? Since these form a part of $SO(3)$ we seem to work with a defective structure which isn't even a group [13] . Actually, there is no problem involved: within our *computational model* of $SO(3)$ every pair of elements has a well-defined product, and each element has its well-defined inverse. The group axioms, such as associativity of multiplication, is satisfied up to numerical noise just as it is satisfied for the computational model of the real numbers as provided by the implementation of floating point arithmetic on any computer. Our numerical representation of $SO(3)$ approximates the ideal to represent a group in the same spirit as computer numbers approximate the ideal to represent the field of real numbers.

According to my understanding, it would be completely adequate to introduce this computational model without the preparations presented here with the complicated-looking additions (123), (124), (125), (126) to the simple basic composition law (122). I carried out these things here, since I experienced that scientists tend to consider the Euler-Rodrigues formalism suffering from deficiencies such as 'not being able to represent π-rotations'. So I considered it potentially helpful to make completely clear that π-rotations are obtained via compactification (which is a closure with respect to a suitable topology). In all what follows, I'll make no difference between $\mathcal{R}$ and $\mathcal{R}_c$ and consider only rotations $\in \mathcal{R}(\mathcal{V})$ (again, since any rotation can be approximated arbitrarily well by those). Actually, I consider the rotation group no longer as the group $SO(3)$ but as the set $\mathcal{V}$ endowed with the operation $\circ$ as multiplication. To be precise, the multiplication is considered to be defined according to equations (137) and (138) with a parameter ϵ set according to the numerical precision of the floating point number type being employed. For clarity, we denote exactly this version of the rotation group with a special symbol, namely $\mathcal{G}$ (for *group*).

For $L \in \mathcal{L}$ and $\mathbf{r} \in \mathcal{V}$ we define the 'rotated operator' $L_{\mathbf{r}}$ by

$$L_{\mathbf{r}}\mathbf{v} \ := \ \mathbf{r} * L((-\mathbf{r}) * \mathbf{v}) \ \text{ for all } \mathbf{v} \in \mathcal{V} \,, \tag{139}$$

[13] It is a *groupoid* if we restrict the definition of the product to those cases in which the result is not a π-rotations.

and get

$$L_{\mathbf{r}} = \mathcal{R}(\mathbf{r}) \circ L \circ \mathcal{R}(-\mathbf{r}) = \mathcal{R}(\mathbf{r}) \circ L \circ \mathcal{R}(\mathbf{r})^{-1} \,. \tag{140}$$

The form of a rotated dyad is especially simple and computationally efficient:

$$|\,\mathbf{a}\,\rangle\langle\,\mathbf{b}\,|_{\mathbf{r}} = |\,\mathbf{r}*\mathbf{a}\,\rangle\langle\,\mathbf{r}*\mathbf{b}\,| \,. \tag{141}$$

Also $\mathcal{R}(\mathbf{r})$ is closely related to dyads: Let $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be an Euclidean system of reference, then also

$$(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3) := (\mathbf{r}*\mathbf{e}_1, \mathbf{r}*\mathbf{e}_2, \mathbf{r}*\mathbf{e}_3)$$

is an Euclidean system of reference and

$$\mathcal{R}(\mathbf{r}) = \sum_{i=1}^{3} |\,\mathbf{f}_i\,\rangle\langle\,\mathbf{e}_i\,| \tag{142}$$

which is just the polar decomposition of $\mathcal{R}(\mathbf{r})$ (the factor $|\mathcal{R}(\mathbf{r})|$ is the identity).

There is an interesting generalization of (106) when we consider the rotation angle φ as descriptor of the the complex number $\exp(\,\mathrm{i}\,\varphi)$ and generalize to the general complex number $z = r\,\exp(\,\mathrm{i}\,\varphi)$ in a natural manner:

$$R(z, \mathbf{n}) = P_{\mathbf{n}} \,+\, r\,\cos\varphi\,(\mathbf{1} - P_{\mathbf{n}}) \,+\, r\,\sin\varphi\,A_{\mathbf{n}} \quad \text{where} \quad r := |z|,\ \varphi := \arg z \,. \tag{143}$$

For each unit vector $\mathbf{n}$ the mapping $\mathbb{C} \to \mathcal{L}\,,\ z \mapsto R(z, \mathbf{n})$ is a homomorphism of rings. (Remember the mapping $\mathbb{R} \to \mathcal{L}\,,\ r \mapsto \underline{r}$ which also is a homomorphism of rings.) This allows, for instance, expressing the resolvent of rotations elegantly

$$(z\,\mathbf{1} - R(\varphi, \mathbf{n}))^{-1} = R\left(\frac{1}{z - \exp(\,\mathrm{i}\,\varphi)}\,,\,\mathbf{n}\right) \tag{144}$$

which helps to code the exponential function and the logarithm — topics that for the rotation group will be treated in the following subsection — for the Euclidean group. An efficient algorithm for applying $R(z, \mathbf{n})$ to a vector is:

$$R(z, \mathbf{n})\,\mathbf{x} = \mathbf{x}_{\parallel} + \mathrm{Re}(z)\,\mathbf{x}_{\perp} + \mathrm{Im}(z)\,\mathbf{n}\times\mathbf{x}_{\perp}\,, \text{ where} \quad \mathbf{x}_{\parallel} := (\mathbf{n}\cdot\mathbf{x})\,\mathbf{n},\ \ \mathbf{x}_{\perp} := \mathbf{x} - \mathbf{x}_{\parallel} \,. \tag{145}$$

5.1 Curves of rotations

We now consider curves in the 'manifold of rotations', i. e. functions [14] $\mathbb{R} \to \mathcal{G}$. As we will see in the next subsection such curves correspond to motions of rigid bodies in a rather direct way, and it is this connection that motivates looking into this matter here. In the present subsection we look to the subject from a mathematical, structural, point of view.

The fact that $\mathcal{G}$ defines a multiplication distinguishes curves $\phi : \mathbb{R} \to \mathcal{G}$ which satisfy

$$\phi(t_1) \circ \phi(t_2) = \phi(t_1 + t_2) \quad \text{for all} \ \ t_1, t_2 \in \mathbb{R} \tag{146}$$

[14] This is an ubiquitous theme in mathematics: sets get endowed with structure by defining real-valued functions on them (e.g. coordinates) and functions from $\mathbb{R}$ into them (e.g. curves).

and will be called *one-parameter groups in* $\mathcal{G}$ if they satisfy the mild regularity property that for each $\mathbf{x} \in \mathcal{V}$ the mapping $\mathbb{R} \to \mathcal{V}$, $t \mapsto \phi(t) * \mathbf{x}$ is continuous. It is one of the virtues of the Euler-Rodrigues formalism that it allows to represent these particular curves in a simple and explicit manner: They all arise from the function $\exp : \mathcal{V} \to \mathcal{G}$ that allows us to express the rotation vector $\mathbf{r}$ directly from the vector $\mathbf{a} := \varphi\,\mathbf{n}$ (rather than from angle φ and direction $\mathbf{n}$ separately) as $\mathbf{r} = \exp(\mathbf{a})$. Let us define this function first:

$$\exp(\mathbf{a}) := \tan\frac{|\mathbf{a}|}{2}\,\hat{\mathbf{a}}\,, \quad \exp(\mathbf{0}) := \mathbf{0}\,. \tag{147}$$

For small arguments $|\mathbf{a}| \ll 1$ this implies the approximation

$$\exp(\mathbf{a}) \approx \frac{1}{2}\mathbf{a}\,. \tag{148}$$

There are good reasons to consider the factor $1/2$ which appears here as unnatural and to define 'natural rotation vectors' which are simply 2 times the conventional rotation vectors. Some factors than would disappear (as in the formula under present consideration) and in other places new factors would appear. If one would start from zero, one probably should introduce such 'natural rotation vectors'. Given that there are scientific communities working with Euler-Rodrigues parameters (see [13],[14]), I prefer not to provoke confusion by a new convention. We now have the two related facts: For each $\mathbf{a} \in \mathcal{V}$ the mapping $t \mapsto \exp(t\mathbf{a})$ is a one-parameter group in $\mathcal{G}$ and each one-parameter group in $\mathcal{G}$ is obtained by this construction. For the action of exponentials on vectors, we have a simple relation

$$\mathcal{R}(\exp(\mathbf{a}))\,\mathbf{x} = \exp(\mathbf{a}) * \mathbf{x} = \mathbf{x} + \mathbf{a}\times\mathbf{x} + \frac{1}{2}\mathbf{a}\times(\mathbf{a}\times\mathbf{x}) + \ldots \;=\; \exp(A_{\mathbf{a}})\,\mathbf{x}\,, \tag{149}$$

which on the right-hand side is free of Euler-Rodrigues operations and uses the conventional exponential function of linear maps. This representation implies

$$\mathcal{R}(\exp(t\,\mathbf{a}))\,\mathbf{x} = \exp(t\,\mathbf{a}) * \mathbf{x} = \mathbf{x} + t\,\mathbf{a}\times\mathbf{x} + \frac{t^2}{2}\mathbf{a}\times(\mathbf{a}\times\mathbf{x}) + \ldots \;=\; \exp(t\,A_{\mathbf{a}})\,\mathbf{x}\,, \tag{150}$$

and

$$\lim_{h\to 0}\,\frac{1}{h}\,(\exp(h\,\mathbf{a}) * \mathbf{x} - \mathbf{x}\,) = \mathbf{a}\times\mathbf{x}\,. \tag{151}$$

This exponential function is obviously bijective and its inverse, log, is given by

$$\log(\mathbf{r}) = 2\arctan|\mathbf{r}|\;\hat{\mathbf{r}} \approx 2\mathbf{r}\,, \quad \log(\mathbf{0}) = \mathbf{0}\,. \tag{152}$$

The question naturally arises how the algebraic structure of $\mathcal{V}$ is related to the algebraic structure of the one-parameter groups which are made from the elements of $\mathcal{V}$ by the exponential function. Preparing the answer we rewrite what we already know: For each $\mathbf{a} \in \mathcal{V}$ we denote by $e^{\mathbf{a}}$ the one-parameter group defined by $e^{\mathbf{a}}(t) := \exp(t\,\mathbf{a})$. The vector $\mathbf{a}$ is said to be the *generator* of this one-parameter group, and it can be recovered from it by

$$\mathbf{a} = \frac{1}{t}\log(e^{\mathbf{a}}(t)) \tag{153}$$

which is valid for all $t \neq 0$. It is a natural situation that one has only an approximate formula for $e^{\mathbf{a}}(t)$ at hand which assumes $t << 1$. Then this formula will give $\mathbf{a}$ exactly if the limit for $t \to 0$ is taken in (153). Now there are the following relations between the algebraic operations in $\mathcal{V}$ and the algebraic operations in $\mathcal{G}$: The one-parameter group $e^{\mathbf{a}+\mathbf{b}}$ can be constructed out of the one-parameter groups $e^{\mathbf{a}}$ and $e^{\mathbf{b}}$ as

$$e^{\mathbf{a}+\mathbf{b}}(t) = \lim_{n\to\infty} \left(e^{\mathbf{a}}(t/n) \circ e^{\mathbf{b}}(t/n) \right)^n \ , \tag{154}$$

and the one-parameter group $e^{\mathbf{a}\times\mathbf{b}}$ can be constructed out of $e^{\mathbf{a}}$ and $e^{\mathbf{b}}$ as

$$e^{\mathbf{a}\times\mathbf{b}}(t) = \lim_{n\to\infty} \left(e^{\mathbf{a}}(\sqrt{t/n}) \circ e^{\mathbf{b}}(\sqrt{t/n}) \circ e^{\mathbf{a}}(-\sqrt{t/n}) \circ e^{\mathbf{b}}(-\sqrt{t/n}) \right)^n \ . \tag{155}$$

Finally, for any $\alpha \in \mathbb{R}$ the one-parameter group $e^{\alpha\,\mathbf{a}}$ can be constructed out of $e^{\mathbf{a}}$ without needing limits

$$e^{\alpha\,\mathbf{a}}(t) = e^{\mathbf{a}}(\alpha\, t) \ . \tag{156}$$

The three previous equations establish that $\mathcal{V}$ endowed with the vector product as multiplication is the *Lie algebra* [15] of the rotation group $\mathcal{G}$. Further for all $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ there is a one-parameter group $e^{\mathrm{Ad}(\mathbf{b})\,\mathbf{a}}$ defined as

$$e^{\mathrm{Ad}(\mathbf{b})\,\mathbf{a}}(t) = \mathbf{b} \circ e^{\mathbf{a}}(t) \circ (-\mathbf{b}) \ . \tag{157}$$

It is instructive to compute the term $\mathrm{Ad}(\mathbf{b})\,\mathbf{a}$ from this definition. First we get from the fourth equation of (129) and from the small argument approximation of the exponential function (148)

$$\mathbf{b} \circ e^{\mathbf{a}}(t) \circ (-\mathbf{b}) = \mathbf{b} * e^{\mathbf{a}}(t) \sim \mathbf{b} * \frac{t}{2}\,\mathbf{a} \ .$$

and finally making use of (153) and the approximation of the logarithm in (152)

$$\mathrm{Ad}(\mathbf{b})\,\mathbf{a} = \lim_{t\to 0} \frac{1}{t} \log(\mathbf{b} \circ e^{\mathbf{a}}(t) \circ (-\mathbf{b})) = \lim_{t\to 0} \frac{1}{t} \log(\mathbf{b} * \frac{t}{2}\,\mathbf{a}) = \lim_{t\to 0} \frac{1}{t} (\mathbf{b} * t\,\mathbf{a}) = \mathbf{b} * \mathbf{a} \ .$$

Hence

$$\mathrm{Ad}(\mathbf{b})\,\mathbf{a} = \mathbf{b} * \mathbf{a} = \mathcal{R}(\mathbf{b})\,\mathbf{a} \ . \tag{158}$$

This indicates an exceptional situation the appreciation of which requires a look into a more general scenario: The present definition of the *adjoint representation* works for any *Lie group*. It associates with any group element an isomorphism of its Lie algebra and does this in a manner that a (not necessarily faithful) *representation* of the Lie group is generated. As an illustration of the general situation we consider the Lorentz group. Its Lie algebra is 6-dimensional (generated by 3 angular displacements and 3 velocity displacements ('boosts')). The 6-dimensional adjoint representation is denoted $D^{(1,0)} \oplus D^{(0,1)}$ whereas the 'natural action' is on the 4-dimensional Minkowski space and

[15] It is the simplest Lie-algebra in the infinite set of so called *simple Lie algebras*. It belongs to what is called class A_1 and is the only *real compact form* in this class.

is denoted $D^{(\frac{1}{2},\frac{1}{2})}$. For the rotation group the 'natural representation' (i.e. $\mathcal{R}$) and the adjoint representation coincide.

For a general curve $\mathbf{r}(\bullet) : \mathbb{R} \to \mathcal{G}$ the *angular velocity* $\boldsymbol{\omega}(t)$ at instant t is defined by

$$\mathbf{r}(t+h) = \exp(2h\,\boldsymbol{\omega}(t)) \circ \mathbf{r}(t-h) + O(h^2)\ . \tag{159}$$

and it turns out that for one-parameter groups the angular velocity is constant and equals the generator of the group. This is in the same spirit as defining the derivative of real-valued functions in terms of the locally best-approximating straight line. Here straight lines are replaced by one-parameter groups. From (159) we get

$$\exp(2h\,\boldsymbol{\omega}(t)) = \mathbf{r}(t+h) \circ -\mathbf{r}(t-h) + O(h^2)\ . \tag{160}$$

and, due to (152)

$$\boldsymbol{\omega}(t) = \lim_{h\to 0} \frac{1}{2h} \log\left(\mathbf{r}(t+h) \circ -\mathbf{r}(t-h)\right) = \lim_{h\to 0} \frac{1}{h} \left(\mathbf{r}(t+h) \circ -\mathbf{r}(t-h)\right)\ . \tag{161}$$

This is a natural generalization of the definition of a derivative from a linear space to a group, where one has to take into account that the factor $\frac{1}{h}$ in the second (log-less) limit which is well defined in the present context, would not make sense if $\mathcal{G}$ would be an arbitrary group. This is even more true for the definition of a standard derivation

$$\dot{\mathbf{r}}(t) := \lim_{h\to 0} \frac{1}{2h} \left(\mathbf{r}(t+h) - \mathbf{r}(t-h)\right) \tag{162}$$

which allows us to express the angular velocity as

$$\boldsymbol{\omega}(t) = \frac{2}{1 + \mathbf{r}(t)\cdot\mathbf{r}(t)} \left(\dot{\mathbf{r}}(t) + \mathbf{r}(t) \times \dot{\mathbf{r}}(t)\right)\ . \tag{163}$$

Interestingly a similar formula holds for the 'back-rotated' angular velocity:

$$\boldsymbol{\omega}'(t) := (-\mathbf{r}(t)) * \boldsymbol{\omega}(t) = \mathcal{R}(\mathbf{r}(t))^{-1}\,\boldsymbol{\omega}(t) = \frac{2}{1 + \mathbf{r}(t)\cdot\mathbf{r}(t)} \left(\dot{\mathbf{r}}(t) - \mathbf{r}(t) \times \dot{\mathbf{r}}(t)\right)\ . \tag{164}$$

Expressing $\dot{\mathbf{r}}$ in terms of ω and $\mathbf{r}$ is also possible:

$$\dot{\mathbf{r}}(t) = \frac{1}{2} \left(\boldsymbol{\omega}(t) + \boldsymbol{\omega}(t) \times \mathbf{r}(t) + (\boldsymbol{\omega}(t)\cdot\mathbf{r}(t))\,\mathbf{r}(t)\right)\ . \tag{165}$$

We will also need the behavior of a 'constant factor' with respect to differentiation. One easily finds by going back to the definition (122) that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{r}(t) \circ \mathbf{c}) = \frac{1}{1 - \mathbf{r}(t)\cdot\mathbf{c}} \left(\dot{\mathbf{r}}(t) + \dot{\mathbf{r}}(t) \times \mathbf{c} + (\dot{\mathbf{r}}(t)\cdot\mathbf{c})\,(\mathbf{r}(t) \circ \mathbf{c})\right) \tag{166}$$

which is far from the naive expectation $\dot{\mathbf{r}}(t) \circ \mathbf{c}$. Notice that, as seen from (159) directly, the angular velocities along the curves $t \mapsto \mathbf{r}(t)$ and $t \mapsto \mathbf{r}(t) \circ \mathbf{c}$ coincide.

5.2 Rotational motion of a rigid body

Let us speak about how a rigid body is placed in space as its *attitude* [16] . Letting shrink the body in size, lets attitude degenerate to position. The information that gets lost in going from attitude to position is often called orientation, but this is a usage of the word which is quite different from the meaning used in the present article. So we will speak of *angular attitude* in this case. There are many ways to give a computationally useful description of attitude. A particularly convenient one is to mark four indexed points on the body (as *body-fixed* points [17]) which define a frame $\in \mathcal{F}_{\mathcal{P}}$. As the body moves, the 'engraved' frame moves with it (this is what 'body-fixed' means) and always represents the changing attitude of the body. As a result, we get a function of time (formalized as $\mathbb{R}$) with values in $\mathcal{F}_{\mathcal{P}}$. We write the time-dependent frame in the mixed form with one point and three vectors (see (41)) as

$$F(t) = (p_0(t), \mathbf{e}_1(t), \mathbf{e}_2(t), \mathbf{e}_3(t)) = (F_0(t), \vec{F}(t)) \ , \quad t \in \mathbb{R} \ , \quad F_0(t) \in \mathcal{P} \ , \quad \vec{F}(t) \in \mathcal{F}_{\mathcal{V}} \ . \tag{167}$$

This is our description of the general motion of a rigid body; it implies that the *configuration space of the rigid body* is here represented by the set $\mathcal{F}_{\mathcal{P}}$ of frames in affine space $\mathcal{P}$.

We will be concerned first with a special case of this general notion in which there is a body-fixed point which also is held fixed in space. Let us call such a motion a *rotational motion* although most writers seem to associate with this term not only a fixed point but an invariant rotational axis. The body engaged in rotational motion is also referred to as a (the) *spinning top*. Without loosing generality we assume that in marking our body-fixed frame we chose this fixed point as origin. Than the component $F_0(t)$ of our frame is constant in time and only $\vec{F}(t)$ depends on t. This implies that the *configuration space of the spinning top* is represented by the set $\mathcal{F}_{\mathcal{V}}$ of frames in the linear space $\mathcal{V}$.

From (17) we know that $\mathcal{F}_{\mathcal{V}}$ is closely related to $SO(3)$ and thus with $\mathcal{G}$. The corresponding statement (45) relates $\mathcal{F}_{\mathcal{P}}$ with $ISO(3)$. Corresponding to the transition $SO(3) \mapsto \mathcal{G}$ we will later construct a computational version $\mathcal{E}$ of $ISO(3)$. The spaces $\mathcal{F}_{\mathcal{P}}$ and $\mathcal{F}_{\mathcal{V}}$ are closely related to the groups $ISO(3)$ and $SO(3)$ in so far as these groups act transitively and effectively on these spaces.

In addition to the body-fixed frame $(\mathbf{e}_1(t), \mathbf{e}_2(t), \mathbf{e}_3(t))$ let us agree on a space-fixed frame $\underline{F} = (\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3)$. This determines a time-dependent rotation $g(t) \in \mathcal{G}$ via

$$g(t) := F(t)/\underline{F} = \sum_{i=1}^{3} |\, \mathbf{e}_i(t) \,\rangle\langle\, \underline{\mathbf{e}}_i \,| \ , \tag{168}$$

Recall here the division of frames as defined in (15). The motion of the body thus is described by a curve in space $\mathcal{G}$. The group $\mathcal{G}$ thus plays the role of a *configuration space* of the spinning top. As is obvious, this description depends on the selection of $\underline{F}$.

[16] The intended meaning is as in *attitude control in satellites*

[17] the physical constitution of a rigid body makes this an unproblematic idealization

Making a different selection $\underline{F}' = (\underline{\mathbf{e}}'_1, \underline{\mathbf{e}}'_2, \underline{\mathbf{e}}'_3)$ will give the description [18]

$$g'(t) := F(t)/\underline{F}' = (F(t)/\underline{F})\,(\underline{F}/\underline{F}') = g(t)\,(\underline{F}/\underline{F}')$$

in which all group elements g are multiplied by the fixed element $\underline{F}/\underline{F}' \in \mathcal{G}$.

From this it easily follows that associating with the pair (t_1, t_2) of times the group element

$$g(t_1, t_2) \;:=\; g(t_1)\, g(t_2)^{-1} \;=\; g'(t_1)\, g'(t_2)^{-1} \;=\; F(t_1)/F(t_2)$$

defines a mapping $\mathbb{R} \times \mathbb{R} \to \mathcal{G}$ which does not depend on selection of a *master frame* (be it $\underline{F}$ or $\underline{F}'$). For any pair (t_1, t_2) we have $g(t_1, t_2)^{-1} = g(t_2, t_1)$ and for any triplet (t_1, t_2, t_3) we have $g(t_1, t_2)\, g(t_2, t_3) = g(t_1, t_3)$. Notice that no ordering of the t-values (such as the one following the indexing) is assumed here. Finally we have

$$g(t_1, t_2)\, F(t_2) = F(t_1)\;, \quad F(t_2) = g(t_2, t_1)\, F(t_1)\;. \tag{169}$$

This shows that $g(\cdot, \cdot)$ is an 'evolution operator' for the moving frames. In groups, the one-parameter subgroups are known to be of great organizing value. These are functions $\phi : \mathbb{R} \to \mathcal{G}$ such that $\phi(t_1)\, \phi(t_2) = \phi(t_1 + t_2)$ for all $t_1, t_2 \in \mathbb{R}$. As we have seen in subsection 5.1 these are just the rotations around a fixed axis with constant angular velocity. In our present context a 'shifted version' $g(t) = \phi(t) g_0$ of a one-parameter subgroup ϕ is as natural as a genuine one-parameter subgroup. Its evolution operator satisfies $g(t_1, t_2) = \phi(t_1)\, g_0\, g_0^{-1}\, \phi(t_2)^{-1} = \phi(t_1 - t_2)$ and therefore

$$g(t_1 + h, t_2 + h) \;=\; g(t_1, t_2) \;\text{ for all } t_1, t_2, h \in \mathbb{R}\;. \tag{170}$$

This says that the evolution operator is invariant with respect to time-shift. Just as for one-parameter subgroups one can follow the curve by successive multiplications

$$\phi(n\, h) = \phi(h)^n \;\text{ for all } h \in \mathbb{R}, n \in \mathbb{N} \tag{171}$$

the frame evolution satisfying (170) can be followed by iterated multiplication:

$$g(t_1 + h, t_2 + h)\, g(t_2 + h, t_2) = g(t_1 + h, t_2) \;\Rightarrow\; g(t_1, t_2)\, g(t_2 + h, t_2) = g(t_1 + h, t_2)$$

$$\Rightarrow\; g(t_1, t_2)\, g(t_2 + h, t_2)^n \;=\; g(t_1 + n\, h, t_2)\;.$$

The kinematic role of one-parameter subgroups (and their shifted versions) for rotational motion is the same as the role of linear motion (at constant velocity) for motion of freely movable point masses: they are suitable approximations to the true motion, valid for a sufficiently small span of time. In this way they allow to define the notion of velocity, here that of angular velocity.

So far, we were dealing with the kinematic aspects of rotational motion. We now turn to the dynamics of rotational motion. We will see that the Euler-Rodrigues parameters used here for the kinematic description are well suited to formulate the equations

[18] Here, and in the following, we write multiplication in $\mathcal{G}$ without the operation symbol $\circ$ (if the context makes clear that the factors belong to $\mathcal{G}$).

of motion of the rotating rigid body, both in the Lagrangian and, even more, in the Hamiltonian framework. We will also see that this formulation allows a computational treatment of rigid body motion without constrains. Any integrator for first order differential equations can be used to evolve the system. It is to be noted, however, that any trajectory of the system reaches infinity (corresponding to rotations of rotation angle π) on a regular basis and comes back from there to finity. Natural methods will be given for handling this situation.

Let us consider the body as made of a huge number N of mass points (m_i, p_i) and let, again, be p the point held fixed in rotational motion of the body. Then the inertia-related properties of the body are described by the *total mass* m, the *center of mass* $p_{\mathrm{CM}} := p + \mathbf{x}$, and the *tensor of inertia* Θ, which are defined as follows:

$$\begin{aligned} m &:= \sum_{i=1}^{N} m_i \ , \quad \mathbf{x} := \frac{\sum_{i=1}^{N} m_i \, \mathbf{x}_i}{m} \ , \\ \Theta &:= \sum_{i=1}^{N} m_i \left(|\mathbf{x}_i|^2 \underline{1} - |\, \mathbf{x}_i \,\rangle\langle\, \mathbf{x}_i \,| \right) \ , \quad \text{where} \quad \mathbf{x}_i := p_i - p \ . \end{aligned} \tag{172}$$

Since this Θ as a sum of N symmetric tensors is a symmetric tensor, and thus has a spectral representation of the form

$$\Theta = I_1 \,|\, \mathbf{e}_1 \,\rangle\langle\, \mathbf{e}_1 \,| + I_2 \,|\, \mathbf{e}_2 \,\rangle\langle\, \mathbf{e}_2 \,| + I_3 \,|\, \mathbf{e}_3 \,\rangle\langle\, \mathbf{e}_3 \,| \ , \quad I_1 \geq I_2 \geq I_3 > 0 \ , \tag{173}$$

where the ordering of the eigenvalues is by convention and the positivity of all eigenvalues is evident from (174) since otherwise rotation could bring about negative kinetic energy.

The tensor of inertia represents the connection between the kinematic motion descriptor *angular velocity* ω and the dynamical motion descriptors *angular momentum* $\mathbf{L}$ and *kinetic energy* T in the form

$$\mathbf{L} = \Theta \, \boldsymbol{\omega} \ , \quad T = \frac{1}{2} \boldsymbol{\omega} \cdot \Theta \boldsymbol{\omega} \ . \tag{174}$$

The mutual orthogonal unit vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are along the *principal axes* of Θ and the directions are assumed to be fixed such that $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$ (this is always possible, since we may any eigenvector multiply by -1). The $\mathbf{e}_i$ are *body-fixed* vectors which we use to characterize the attitude of the body.

Using also the space-fixed vectors $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$ of the previous subsection we have a rotation vector $\mathbf{r}$ which satisfies

$$\mathbf{e}_1 = \mathbf{r} * \underline{\mathbf{e}}_1 \ , \quad \mathbf{e}_2 = \mathbf{r} * \underline{\mathbf{e}}_2 \ , \quad \mathbf{e}_3 = \mathbf{r} * \underline{\mathbf{e}}_3 \ . \tag{175}$$

Assume the body placed in space in a way that its principal axes are parallel to these vectors $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$ (the position of the fixed point is understood as unchanged, of course). Then we say that the body 'is in reference position' and its tensor of inertia is 'the reference version' of the bodies tensor of inertia in generic position. This reference version is then

$$I := I_1 \,|\, \underline{\mathbf{e}}_1 \,\rangle\langle\, \underline{\mathbf{e}}_1 \,| + I_2 \,|\, \underline{\mathbf{e}}_2 \,\rangle\langle\, \underline{\mathbf{e}}_2 \,| + I_3 \,|\, \underline{\mathbf{e}}_3 \,\rangle\langle\, \underline{\mathbf{e}}_3 \,| \tag{176}$$

and we have from (141)

$$\Theta = I_{\mathbf{r}} \ . \tag{177}$$

Equations (141), (176) make the computation of $I_{\mathbf{r}}$ particularly efficient. Also the position vector of the center-of-mass has a specific value, let us call it $\mathbf{x}_0$, in this reference position. Its connection with $\mathbf{x}$ in (172) is given by

$$\mathbf{x} = \mathbf{r} * \mathbf{x}_0 \ . \tag{178}$$

Suppose that space $\mathcal{P}$ carries a scalar field Φ such that our body has a potential energy $\Phi(p_{CM})$, where p_{CM} is the location of the center-of-mass of the body. Then the potential energy becomes a function of the bodies rotation vector $\mathbf{r}$ of the form

$$V(\mathbf{r}) = \Phi(p + \mathbf{r} * \mathbf{x}_0) \ . \tag{179}$$

In a homogeneous gravity field of strength $\mathbf{g}$ we have $\Phi(q) = -(q-p)\cdot\mathbf{g}$ and thus

$$V(\mathbf{r}) = (\mathbf{r} * \mathbf{x}_0)\cdot\mathbf{g} \ . \tag{180}$$

This term has to be added to the free Lagrangian and to the free Hamiltonian of the spinning top in order to get a theory of the top spinning under the influence of gravity or more general conservative forces if a different field Φ is chosen.

At first we consider the free case since it is only the kinetic energy term for which our decision to use Euler-Rodrigues parameters as generalized coordinates poses non-trivial problems — which, however, can be solved in a rather elegant manner, as will be shown now. Using Euler-Rodrigues parameters as generalized coordinates means to write down the Lagrangian in terms of $\mathbf{r}$ and $\dot{\mathbf{r}}$. As a physical quantity (as opposed to a function of specified variables) the free Lagrangian is the kinetic energy and thus given by

$$L = T = \frac{1}{2}\boldsymbol{\omega}\cdot I_{\mathbf{r}}\,\boldsymbol{\omega} = \frac{1}{2}\mathcal{R}(\mathbf{r})^{-1}\boldsymbol{\omega}\cdot I\,\mathcal{R}(\mathbf{r})^{-1}\boldsymbol{\omega} = \frac{1}{2}\boldsymbol{\omega}'\cdot I\,\boldsymbol{\omega}' \ . \tag{181}$$

To express this in terms of $\mathbf{r}$ and $\dot{\mathbf{r}}$ we simply replace $\boldsymbol{\omega}'$ by the expression from (164) and have

$$\boxed{L(\mathbf{r},\dot{\mathbf{r}}) = \frac{2(\mathbf{s}\cdot I\,\mathbf{s})}{(1+\mathbf{r}\cdot\mathbf{r})^2} \ , \quad \text{where} \quad \mathbf{s} := \dot{\mathbf{r}} - \mathbf{r}\times\dot{\mathbf{r}} \ .} \tag{182}$$

Using coordinates with respect to basis $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$ this Lagrangian is given as

$$L(\mathbf{r},\dot{\mathbf{r}}) = \frac{2I_1\left(\dot{r}_1 - r_2\dot{r}_3 + r_3\dot{r}_2\right)^2 + 2I_2\left(\dot{r}_2 - r_3\dot{r}_1 + r_1\dot{r}_3\right)^2 + 2I_3\left(\dot{r}_3 - r_1\dot{r}_2 + r_2\dot{r}_1\right)^2}{(1+r_1^2+r_2^2+r_3^2)^2} \ . \tag{183}$$

This expression is not too complex. But due to the nominator, the Lagrangian equations (of the second kind) become very complicated and I saw no sense so far in using them in a numerical simulation. It came as a real surprise to me that passing to the *Hamiltonian* formulation turned out to be possible and to give canonical equations which are much simpler than the Lagrangian equations. Let us see how this works. To make the

introduction of canonical momenta more transparent we use a simplified notation, which does not always mention $\mathbf{r}$:

$$M := \underline{1} + A_{\mathbf{r}} \ , \quad \rho := \frac{1}{1 + \mathbf{r} \cdot \mathbf{r}} \ , \tag{184}$$

which implies together with (84)

$$M^{-1} = \rho(\underline{1} - A_{\mathbf{r}} + |\, \mathbf{r}\,\rangle\langle\, \mathbf{r}\,|) \ , \tag{185}$$

Then

$$L = T = 2\rho^2(\dot{\mathbf{r}} \cdot M\, I\, M^* \dot{\mathbf{r}}) = \frac{1}{2}\dot{\mathbf{r}} \cdot \mathbf{p} \ , \quad \text{where} \quad \mathbf{p} := \frac{\partial L}{\partial \dot{\mathbf{r}}} = 4\rho^2 M\, I\, M^* \dot{\mathbf{r}} \ . \tag{186}$$

Since the canonical momentum is an important quantity, we express it in terms of various quantities which we encountered already:

$$\mathbf{p} = 4\,\frac{I(\dot{\mathbf{r}} - \mathbf{r} \times \dot{\mathbf{r}}) + \mathbf{r} \times I(\dot{\mathbf{r}} - \mathbf{r} \times \dot{\mathbf{r}})}{(1 + \mathbf{r} \cdot \mathbf{r})^2} = 2\,\frac{\mathbf{L}' + \mathbf{r} \times \mathbf{L}'}{1 + \mathbf{r} \cdot \mathbf{r}} = \mathbf{L} + \mathbf{L}' - 2\,\frac{\mathbf{r} \cdot \mathbf{L}'}{1 + \mathbf{r} \cdot \mathbf{r}}\,\mathbf{r} \ , \tag{187}$$

where $\mathbf{L}' := I\,\boldsymbol{\omega}' = \mathcal{R}(\mathbf{r})^{-1}\,\mathbf{L}$. This shows that the canonical momentum $\mathbf{p}$ differs from the angular momentum $\mathbf{L}$. We are now in a position to form the Hamiltonian:

$$H := \dot{\mathbf{r}} \cdot \mathbf{p} - L = \frac{1}{2}\dot{\mathbf{r}} \cdot \mathbf{p} \tag{188}$$

and are left with the task to express $\dot{\mathbf{r}}$ in terms of $\mathbf{p}$. Due to the last equation in (186) this only needs the inversion of simple matrices:

$$\dot{\mathbf{r}} = \frac{1}{4\rho^2} M^{*-1}\, I^{-1}\, M^{-1}\, \mathbf{p} \tag{189}$$

thus, due to (185) and (60)

$$H = \frac{1}{8\rho^2}(M^{-1}\,\mathbf{p} \cdot I^{-1}\, M^{-1}\,\mathbf{p}) \ . \tag{190}$$

Finally,

$$\boxed{H(\mathbf{p}, \mathbf{r}) = \frac{\mathbf{q} \cdot I^{-1}\,\mathbf{q}}{8} \ , \quad \text{where} \quad \mathbf{q} := \mathbf{p} + \mathbf{p} \times \mathbf{r} + (\mathbf{p} \cdot \mathbf{r})\,\mathbf{r} \ .} \tag{191}$$

Again, using coordinates with respect to basis $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$ this Hamiltonian is given as

$$\begin{aligned} H(\mathbf{p}, \mathbf{r}) = & \frac{(p_1 + p_2\, r_3 - p_3\, r_2 + s\, r_1)^2}{8I_1} + \frac{(p_2 + p_3\, r_1 - p_1\, r_3 + s\, r_2)^2}{8I_2} \\ & + \frac{(p_3 + p_1\, r_2 - p_2\, r_1 + s\, r_3)^2}{8I_3} \ , \quad \text{where} \quad s := p_1\, r_1 + p_2\, r_2 + p_3\, r_3 \ . \end{aligned} \tag{192}$$

This makes available all the theory and practical integration methods for Hamiltonian systems [21], [22],[23]. Especially the far-reaching results of [29] on approximate energy

conservation and conservation of 'Noether charges' can be applied. Also for the transition to a quantum mechanical treatment the conventional starting point is reached. One should note, however, that the Hamiltonian is not *separable* in the sense of [21], p. 76, so that the largest part of this standard text says nothing about our system.

It turns out that not only the Hamiltonian but also the canonical equations can be expressed in a simple manner:

$$\boxed{\begin{aligned} \dot{\mathbf{r}} &= +\frac{\partial}{\partial \mathbf{p}} H(\mathbf{p},\mathbf{r}) = \mathbf{s} + \mathbf{r}\times\mathbf{s} + (\mathbf{r}\cdot\mathbf{s})\,\mathbf{r} \\ \dot{\mathbf{p}} &= -\frac{\partial}{\partial \mathbf{r}} H(\mathbf{p},\mathbf{r}) = -\mathbf{r}\times\mathbf{s} - (\mathbf{r}\cdot\mathbf{s})\,\mathbf{p} - (\mathbf{p}\cdot\mathbf{r})\,\mathbf{s}\;, \quad \text{where} \\ \mathbf{s} &:= \frac{1}{4}\, I^{-1}\,\left(\mathbf{p} + \mathbf{p}\times\mathbf{r} + (\mathbf{p}\cdot\mathbf{r})\,\mathbf{r}\right)\;. \end{aligned}} \tag{193}$$

As a physical quantity the term $\mathbf{s}$ equals $\boldsymbol{\omega}'/2$. Here it is important, however, that it is given as a function of $\mathbf{r}$ and $\mathbf{p}$. We thus have a highly non-linear, though very regular (all functions are polynomials — at most of fifth order — in the dynamic variables) system of ordinary differential equations for the time-dependent quantities $\mathbf{r}$ and $\mathbf{p}$. In these equations the linear map I is constant in time. It is, therefore, a tensor-valued parameter. If $\mathbf{r}(t)$ has been obtained from these equations, the tensor of inertia of the moving body results from (177) for which we can give the more explicit form

$$\Theta(t) = \mathcal{R}(\mathbf{r}(t))\, I\, \mathcal{R}(\mathbf{r}(t))^{-1} = I_{\mathbf{r}(t)}\;. \tag{194}$$

As an introductory example let us consider the special case that I is a homothety.

$$I = \theta\mathbf{1}\;, \quad \theta > 0\;. \tag{195}$$

From physics it is clear that in his case a rotation with constant angular velocity is an allowed motion. We expect to get this solution by writing all the vectors $\mathbf{r}, \mathbf{p}, \mathbf{s}$ as real multiples of a constant unit vector $\mathbf{e}$:

$$\mathbf{r} = r\,\mathbf{e}\;, \quad \mathbf{p} = p\,\mathbf{e}\;, \quad \mathbf{s} = s\,\mathbf{e}\;, \quad \dot{\mathbf{r}} = \dot{r}\,\mathbf{e}\;, \quad \dot{\mathbf{p}} = \dot{p}\,\mathbf{e}\;. \tag{196}$$

For these quantities the system (193) implies

$$\begin{aligned} \dot{r} &= s\,(1+r^2) = \frac{p(1+r^2)^2}{4\,\theta} \\ \dot{p} &= -2\,s\,r\,p = -\frac{r\,p^2\,(1+r^2)}{2\,\theta} \\ s &= \frac{p\,(1+r^2)}{4\,\theta}\;. \end{aligned} \tag{197}$$

For the Hamiltonian (191) we get

$$H(p,r) = 2\,\theta\,s^2 = \frac{p^2\,(1+r^2)^2}{8\,\theta} \tag{198}$$

and easily verify that the canonical equations resulting from it are in fact (197). The key to solving these equations is to observe $\dot{s} = 0$. The most elegant way is to deduce this from Poisson-brackets: $\dot{s} = \{H, s\} = 2\,\theta\,\{s^2, s\} = 0$. Writing $r = \tan\varphi/2$ we get $\dot{r} = (\tan\varphi/2)'\,\dot{\varphi}/2 = (1 + (\tan\varphi/2)^2)\,\dot{\varphi}/2 = (1 + r^2)\,\dot{\varphi}/2$. From the first equation of (197) we now have $s = \dot{\varphi}/2$. This implies $\ddot{\varphi} = 0$ and $\varphi = \omega\,t + \varphi_0$ and $r(t) = \tan\frac{\omega\,t+\varphi_0}{2}$. The third equation of (197) says $p(t) = \frac{4\,\theta\,s}{1+r(t)^2} = \frac{2\,\theta\,\omega}{1+r(t)^2}$. This represents the general solution of (197) $r(t)$ and $p(t)$ in terms of two integration constants ω and φ_0 (and the system parameter θ). It is interesting to look at the numerical initial value problem associated with (197). Since, as we have seen, s is a constant, the third equation of (197) expresses $p(t)$ in terms of $r(t)$ so that it is sufficient to consider r. For simplicity we put $s = 1$ and set the initial condition $r(0) = 0$. Then we are left with the first equation of (197) in the form

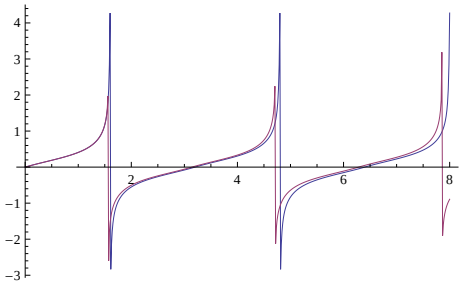
$$\dot{r} = 1 + r^2 \tag{199}$$

and the solution consistent with our initial condition is $r(t) = \tan t$. Of course, this trajectory reaches infinity in finite time $t = \pi/2$. From the most basic facts concerning the initial value problem of ordinary differential equations (e.g. [20], Theorem 2.1.3, p. 102) this is plausible since the function $r \mapsto 1 + r^2$ is not globally Lipschitz on $\mathbb{R}$. Would r denote a position in space, reaching infinity would mean that the system (or one of its components) would no longer belong to the 'real world' and following its trajectory till infinity is the most one can reasonably do. How such a situation may arise in the non-relativistic theory of gravitating point masses is described in [30]. In our case, however, the singularity on the trajectory is only an artifact of the coordinate system. Had we described the attitude of the body by a rotation angle instead of its tangent, no singularity would occur. So, in order for our description of uniform rotation to make sense from a practical point of view one should be able to continue time discrete trajectories across the singularities (the exact trajectory has singularities for all $t_n = (n+\frac{1}{2})\pi$). This is in fact the case and is explained with Figure 3. As the graphics there indicates the method renders well the qualitative features of the tangent curve (notice that a modified logarithmic y-axis is used) but is not very accurate, even for the rather short time step which is employed in this example. So in the following treatment of the non-simplified problem (193) a more accurate method is developed. Here one does not integrate over the singularity but shifts it repeatedly to other places so that it does no influence the computation.

Computing trajectories of (193) by means of the explicit *asynchronous leapfrog method* [27] or the implicit Störmer - Verlet method ([22], equation (2.10) and [23], last triplet of equations on p. 85) shows very good conservation of energy and angular moment in regions where $|\mathbf{r}|$ is not too large, say lower than 10. (For larger $|\mathbf{r}|$ one could hold the quality level of conservation by reducing the time step, which is not a favorable option.) Large $|\mathbf{r}|$ means that the rotation angle is close to 180°. For $|\mathbf{r}| = 10$ we are off by an angle of 11°.5. This angle is not an absolute property of the attitude of our body. Instead it reflects the relative attitude of our body and the basis $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$. A simple change of this basis gives our body a position $\mathbf{r}$ which is completely normal. For the computational implementation I chose a particularly simple and efficient such change: As soon as $|\mathbf{r}|$

```
(* functions *)
f[r_] := 1 + r^2
iniALF[r_, t_] := stateALF[r, f[r], t] (*ALF: asynchronous leap-frog *)
stepALF[dt_, rCrit_, stateALF[r_, v_, t_]] := Module[
  {τ = dt / 2, ra = r, va = v, ta = t},
  If[Abs[ra] > rCrit && ra * va > 0, ra = -ra]; (* using compactness *)
  ra += τ * va; ta += τ;
  va += 2 (f[ra] - va);
  ra += τ * va; ta += τ;
  stateALF[ra, va, ta]
 ]
gett[stateALF[r_, v_, t_]] := t
display[r_] := Sign[r] * Log10[Abs[r] + 1] (* for data with large dynamic range *)
getr[stateALF[r_, v_, t_]] := display[r]
(* input data *)
rCrit = 1000; tMax = 8.; nSteps = 600; r0 = 0; t0 = 0;
(* work *)
dt = tMax / nSteps; s0 = iniALF[r0, t0];
nl = NestList[stepALF[dt, rCrit, #] &, s0, nSteps];
alf = {gett[#], getr[#]} & /@ nl;
exact = {gett[#], display[Tan[gett[#]]]} & /@ nl;
ListLinePlot[{alf, exact}, PlotRange → All]
```

(a) Mathematica Code



(b) Mathematica Graphics

Figure 3: Numerical and exact solution of (199). The numerical method (blue curve) considers r and $-r$ close together for $|r| > r_{Crit}$.

is found to exceed some pre-selected limit (e. g. 10 as before) the basis $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$ gets reset to the actual values of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Then the new value of $|\mathbf{r}|$ is 0 which is quite a remarkable size reduction. A simulation of the motion of our spinning top should allow to monitor the bodies attitude in a constant frame; so we have to record the attitude of the 'temporary basis' $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$ with respect to a 'master basis' $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$ which is held constant over the whole simulation. It is convenient to think of $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$ as being selected (as a positively oriented orthonormal basis in $\mathcal{V}$) independent of the attitude of our body at any time. This then allows us to treat many bodies simultaneously, the situation treated in [17] and [18]. We thus introduce a rotation vector ρ, which describes $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \underline{\mathbf{e}}_3$ in terms of $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$:

$$\underline{\mathbf{e}}_1 = \rho * \mathbf{m}_1 \ , \quad \underline{\mathbf{e}}_2 = \rho * \mathbf{m}_2 \ , \quad \underline{\mathbf{e}}_3 = \rho * \mathbf{m}_3 \tag{200}$$

and define

$$J := I_1 \, |\, \mathbf{m}_1 \,\rangle\langle\, \mathbf{m}_1 \,| + I_2 \, |\, \mathbf{m}_2 \,\rangle\langle\, \mathbf{m}_2 \,| + I_3 \, |\, \mathbf{m}_3 \,\rangle\langle\, \mathbf{m}_3 \,| \tag{201}$$

which allows representing I as

$$I = J_\rho \ . \tag{202}$$

We are now in a position to formulate the *basis resetting step*:

$$\rho \mapsto \mathbf{r} \circ \rho \ , \quad I \mapsto J_\rho \ , \quad \dot{\mathbf{r}} \mapsto \frac{\dot{\mathbf{r}} - \mathbf{r} \times \dot{\mathbf{r}}}{1 + \mathbf{r} \cdot \mathbf{r}} \ , \quad \mathbf{r} \mapsto \mathbf{0} \ , \quad \mathbf{p} \mapsto 4\, I\, \dot{\mathbf{r}} \ . \tag{203}$$

The point is that $\mathbf{r} \circ \rho$ has the same value before and after the transformation and thus, due to the constancy of the $\mathbf{m}_i$, the basis $\mathbf{e}_i = \mathbf{r} * \underline{\mathbf{e}}_i = \mathbf{r} * \rho * \mathbf{m}_i$ remains constant in space during reset (not during motion). The transformation property of the velocity $\dot{\mathbf{r}}$ is a direct consequence of (166). Notice that the original value of $\dot{\mathbf{r}}$ results from the first equation in (193) and has to be computed in a time-stepping algorithm anyway. The transformation property of $\mathbf{p}$ comes from inserting the transformed value of $\dot{\mathbf{r}}$ into

the first equation of (187) using there that now $\mathbf{r} = \mathbf{0}$. The computational work of such a resetting step is similar to one dynamical step. Typically we have a few thousand dynamical steps between two resetting steps, so that the computational effort of resetting is ignorable. In (203) the particular form of the transformation in $\mathbf{r}$-space was exploited, having computational efficiency in mind (which is not very important as just mentioned). This obscures the general nature of the problem. There is a general way to extend an arbitrary transformation $\mathbf{f} : \mathcal{V} \to \mathcal{V}$ in $\mathbf{r}$-space (configuration space) to a *canonical transformation* $\tilde{\mathbf{f}}$ in $\mathbf{r}, \mathbf{p}$-space (phase space):

$$\begin{aligned} \tilde{\mathbf{f}} &: \mathcal{V} \times \mathcal{V} \to \mathcal{V} \times \mathcal{V} \\ (\mathbf{r}, \mathbf{p}) &\mapsto \left(\mathbf{f}(\mathbf{r}), \left(\frac{\partial \mathbf{f}(\mathbf{r})}{\partial \mathbf{r}} \right)^{*-1} \mathbf{p} \right) \end{aligned} \tag{204}$$

which, by taking the contragredient of the Jacobian, reflects the property of canonical momenta to be co-vectors. Since Jacobians can efficiently be computed numerically, this is a very practical method. It gives the same result as equation (203) in which the transformation of $\mathbf{p}$ is ad hoc.

Figure 4 shows the evolution over a time span of 10 seconds for a free top with data treated in [28]: $I_1 = 0.345, I_2 = 0.653, I_3 = 1.0$ and $\mathbf{L}' = (1.8, 0.4, -0.9)$.

Finally we treat the case that on the center of mass there acts a force which can be derived from a potential of the form (179). Let the total Hamiltonian thus be of the form

$$H(\mathbf{p}, \mathbf{r}) = H_0(\mathbf{p}, \mathbf{r}) + V(\mathbf{r}) \tag{205}$$

where H_0 is what in (192) was denoted H, and $\mathbf{x}$ is what in (178) was denoted $\mathbf{x}_0$. In the canonical equations this interaction potential adds a term $-\frac{\partial V}{\partial \mathbf{r}}$ to the right-hand side of the second equation of (193). For this term a straight-forward calculation gives

$$-\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}} = \frac{2}{1 + \mathbf{r} \cdot \mathbf{r}} \left(-\frac{|\,\mathbf{r}\,\rangle\langle\,\mathbf{t}\,|}{1 + \mathbf{r} \cdot \mathbf{r}} + |\,\mathbf{x}\,\rangle\langle\,\mathbf{r}\,| - |\,\mathbf{r}\,\rangle\langle\,\mathbf{x}\,| + (\mathbf{x} \cdot \mathbf{r})\,\mathbf{1} - A_{\mathbf{x}} \right) (\nabla\Phi)(p + \mathbf{r} * \mathbf{x}) \tag{206}$$

where

$$\mathbf{t} := (1 - \mathbf{r} \cdot \mathbf{r})\,\mathbf{x} + 2\,(\mathbf{r} \cdot \mathbf{x})\,\mathbf{r} + 2\,\mathbf{r} \times \mathbf{x} \tag{207}$$

In the canonical equations we now have in addition to I the parameter $\mathbf{x}$ (recall that this is the 'center-of-mass position vector in reference position' and that the dynamically evolving center-of-mass position vector is $\mathbf{r} * \mathbf{x}$). and just as I it is affected by the basis resetting step which we introduced in order to avoid stress for the time stepping algorithm resulting from $|\mathbf{r}|$ becoming large. In complete analogy to the representation (201) and (202) we introduce a master representation for $\mathbf{x}$:

$$\mathbf{y} = y_1\,\mathbf{m}_1 + y_2\,\mathbf{m}_2 + y_3\,\mathbf{m}_3\ , \quad \mathbf{x} = \rho * \mathbf{y}\ . \tag{208}$$

If $\mathbf{x}$ is used always in this form, i.e. expressed in terms of ρ and an invariable vector, the basis resetting step (203) needs no extension: resetting ρ resets $\mathbf{x}$ automatically.

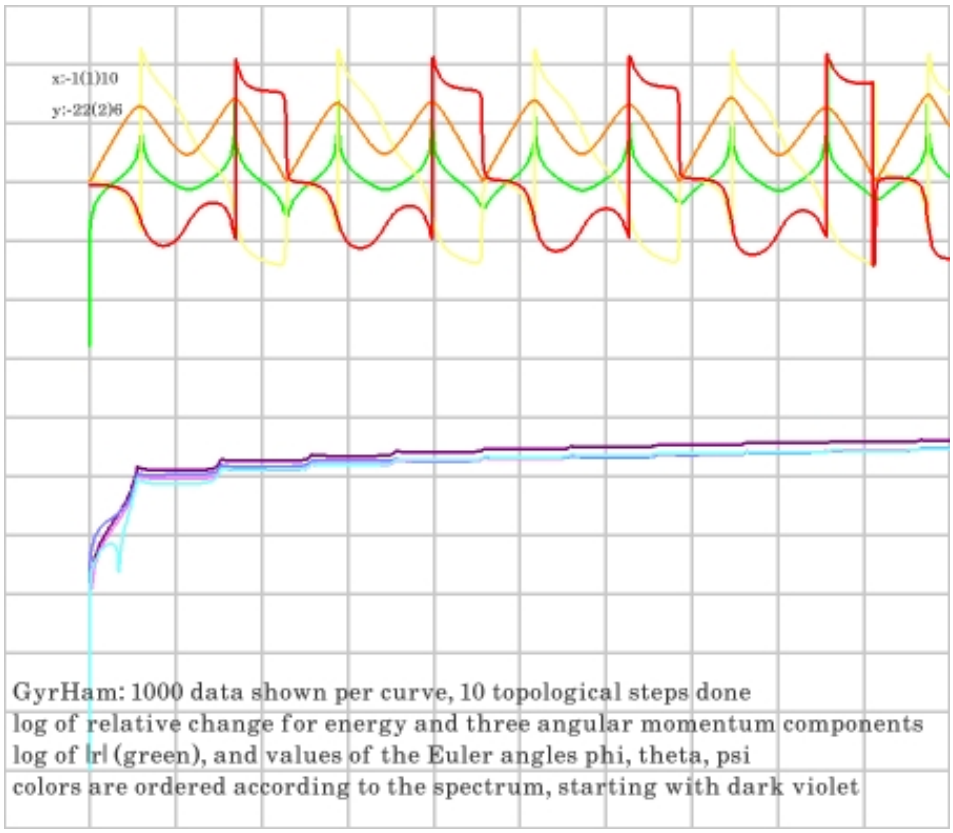


Figure 4: Evolution of state descriptors of a spinning top as obtained using the asynchronous leapfrog integrator

That this can be expressed in short and simple code can be seen from the listing of class `GyrHam` in file spinningtop.cpp in [6].

Simulation runs with such an interaction behave very similar to those without interaction but for angular momentum, there is only one of the three components constant.

I am not aware that this elegant Hamiltonian formulation of the spinning top with Euler-Rodrigues parameters as position coordinates has been obtained earlier. Actually it is hard to believe that this is a truly new result: These days, where it is common place that the configuration space of the spinning top is the Lie group SO(3), there is nothing more natural than trying whether, what in Lie group theory is known as *canonical coordinates* (see [24] p.104 Proposition 1.6), could be used as generalized coordinates in a Lagrangian and Hamiltonian approach to dynamics of the spinning top. Up to an unhappy factor 2 the Euler-Rodrigues parameters approximate and extend canonical coordinates and thus are simply the computationally adequate realization of the theoretical concept of canonical coordinates in the case of the rotation group. Carrying through his program can't end in equations essentially different from (193).

Note that the numerical integrator in [17], [18] is based on a predicting version of the

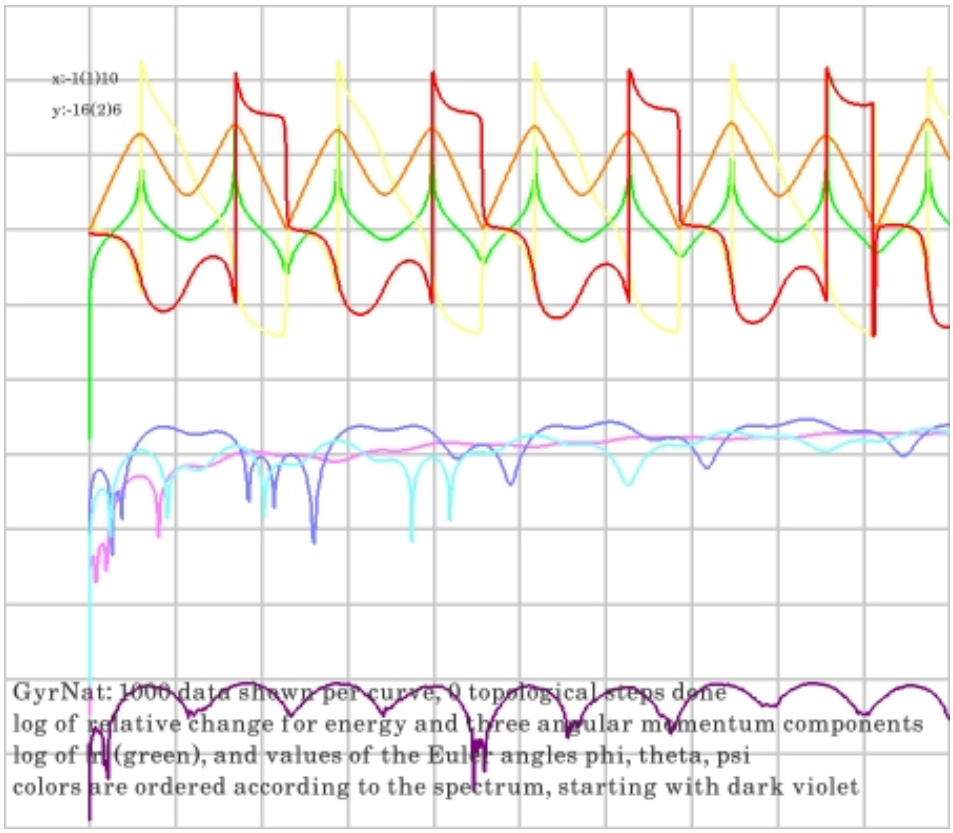


Figure 5: Evolution of state descriptors of a spinning top as obtained using the direct midpoint integrator

the action principle and does not rely on differential equations formulated in terms of some system of coordinates. For the same input data used for creating Figure 4 this *direct midpoint integrator* creates Figure 5. This diagram shows that now the points where $|\mathbf{r}|$ becomes large play no particular role. As a result of this, the conservation of energy and angular momentum holds more accurately. What makes this possible is that the group multiplication $\circ$ is used for bringing about the state change associated with an integration step instead of additive increments of coordinates. This 'natural' integration method is reversible. For the Hamiltonian method I did not investigate reversibility thoroughly. For simulation runs over a few rotations no deviations from reversibility can be found on the graphical resolution. With the Störmer - Verlet integrator I found the relative errors of energy and angular momenta always going up, whereas the asynchronous leapfrog method can be tuned to go some of the steps down again. The phenomenon is however very sensitive to numerical errors and so a suitable setting in the discretization length for computing the Jacobian numerically is needed so see the effect. Investigations with multiple precision arithmetic will have to be done to study reversibility. Again, the 'natural method' is free of these problems.

6 Motion of rigid bodies in space

We now suspend the assumption that a point of our body remains fixed in space. Further, we have in mind not only one body and allow for the possibility that bodies collide and also interact with external fields. As we have seen in subsection 5.2 the generic attitude of a rigid body can be represented by a frame $F(t) \in \mathcal{F}_{\mathcal{P}}$ and, therefore, by a group element $g(t) \in ISO(3)$ which moves the space-fixed 'master frame' $\underline{F}$ into $F(t)$. It is tempting to base the notion of velocity (which here comprises translational velocity and angular velocity) on the analysis of $t \mapsto g(t) \in ISO(3)$ just as we have based the notion of angular velocity alone on $t \mapsto g(t) \in SO(3)$, see (159) and (161). We thus will need the exponential function and the logarithm also for the group $ISO(3)$. Actually we have given these functions for the 'computational version' $\mathcal{G}$ of $SO(3)$. It suggests itself to make use of these results by replacing $ISO(3)$ by a computational version in which rotations are represented by elements of $\mathcal{G}$. This computational version of the Euclidean group will be designated $\mathcal{E}$. Defining this group, its Lie algebra, and the exponential function shall be the first topic of this section.

6.1 The Euclidean group $\mathcal{E}$

We define ...

References

[1] Ulrich Mutze: homepage
http://www.ulrichmutze.de/softwaredescriptions/pala_listing.pdf
http://www.ulrichmutze.de/sourcecode/pala.zip
http://www.ulrichmutze.de/winExe/pala_distribution.zip

[2] Cornell University, Center for Advanced Computing: Engineering breakthrough: a research collaboration pays off for Kodak.
http://www.cac.cornell.edu/about/studies/Kodak.pdf

[3] Ulrich Mutze: Predicting Classical Motion Directly from the Action Principle II
Mathematical Physics Preprint 1999
http://www.ma.utexas.edu/mp_arc/c/99/99-271.pdf

[4] Ulrich Mutze: Separated quantum dynamics
Mathematical Physics Preprint 2008
http://www.ma.utexas.edu/mp_arc/c/08/08-69.pdf

[5] Ulrich Mutze: homepage
http://www.ulrichmutze.de/softwaredescriptions/cpmProject.pdf
http://www.ulrichmutze.de/softwaredescriptions/cpm0listing.pdf
http://www.ulrichmutze.de/softwaredescriptions/cpmlisting.pdf

[6] Ulrich Mutze: homepage
http://www.ulrichmutze.de/softwaredescriptions/cpmlisting.pdf

[7] Saunders Mac Lane, Garrett Birkhoff: Algebra, MACMILLAN, 1968, Chapter XII

[8] Jeremy M. Berg, John L. Tymoczko, Lubert Stryer: Biochemistry, Freeman, Sixth edition 2008

[9] Werner Graeub: Lineare Algebra, Springer, 1958

[10] Rudolf Haag: Local Quantum Physics, Springer, Second edition 1996

[11] D.P.L. Castrigiano, U. Mutze: Covariant description of particle position, Phys. Rev. D vol. 26 pp. 3499-3505 1982

[12] H. J. Meister: Darstellungstheoretische Behandlung von Problemen der Elementarteilchenphysik I. Grupentheoretische Grundlagen. Universität München, Wintersemester 1965/66, p. 14 (2.17)

[13] Thomas Haslwanter: Mathematics of Three-dimensional Eye Rotations, Vision Res. Vol 35, No 12, pp 1727-1739, 1995

[14] Klaus Hepp: On Listing's Law, Commun. Math. Phys. 132, 285-292 (1990)

[15] Olinde Rodrigues: Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ses déplacements considérés indépendamment des causes qui peuvent les produire, Journal de Mathématique Pures et Appliquées, 5, 380-440, 1840

[16] Ulrich Mutze: Quaternions - Redundancy + Efficiency = Ternions,
Mathematical Physics Preprint 05–53, February 2005
http://www.ma.utexas.edu/mp_arc/c/05/05-53.pdf

[17] Ulrich Mutze: Rigidly connected overlapping spherical particles: a versatile grain model, Granular Matter 2006

[18] Ulrich Mutze: Polyspherical grains and their dynamics
Mathematical Physics Preprint 07–252, July 2007
http://www.ma.utexas.edu/mp_arc/c/07/07-252.pdf

[19] Ulrich Mutze: Precision-dependent symmetry breaking in simulated motion of polyspherical grains
Mathematical Physics Preprint 09–113, July 2009
http://www.ma.utexas.edu/mp_arc/c/09/09-113.pdf

[20] A.M. Stuart and A.R. Humphries: Dynamical Systems and Numerical Analysis, Cambridge University Press, 1996

[21] J.M. Sanz-Serna and M.P. Calvo: Numerical Hamiltonian Problems, Applied Mathematics and Mathematical Computation 7, Chapman & Hall 1994

[22] Ernst Hairer, Christian Lubich, Gerhard Wanner: Geometric numerical integration illustrated by the Störmer/Verlet method, Acta Numerica (2003) pp. 1-51 Cambridge University Press, 2003

[23] Benedict Leimkuhler and Sebastian Reich: Simulating Hamiltonian Dynamics, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press 2004

[24] Sigurdur Helgason: Differential Geometry, Lie Groups, and Symmetric Spaces, Academic Press, 1978

[25] Yvonne Choquet-Bruhat, CÃ©cile DeWitt-Morette, Margret Dillard-Bleick: Analysis, Manifolds, and Physics, North-Holland, Revised Edition 1982

[26] John Baez: Torsors Made Easy, June 12, 2006
http://www.math.ucr.edu/home/baez/torsors.html

[27] Ulrich Mutze: An asynchronous leap-frog method (2008)
http://www.ma.utexas.edu/mp_arc/c/08/08-197.pdf

[28] Ernst Hairer, Gilles Vilmart: Preprocessed Discrete Moser-Veselov algorithm for the full dynamics of the rigid body J. Phys. A: Math. Gen. 39 13225-13235 2006
see also: `ftp://ftp.irisa.fr/techreports/theses/2008/vilmart.pdf`

[29] J. David Brown: The Midpoint Rule as a Variational-Symplectic Integrator. I. Hamiltonian Systems
arXiv: gr-qc/0511018v1 3 Nov 2005

[30] Donald G. Saari and Zhihong (Jeff) Xia: Off to Infinity in Finite Time
http://www.ams.org/notices/199505/saari-2.pdf