Mutual Distances in Celestial Mechanics

Alain Albouy, albouy@imcce.fr
UMR8028 / CNRS
IMCCE, 77, av. Denfert-Rochereau, 75014 Paris, France
Lectures at Nankai University, Tianjin, China
2004/9, oral presentation 2004/6

Index

1. Lagrange's reduced three-body problem

If we try to describe the evolution of three particles under gravitational interaction (for example the Earth, the Sun and Venus), we can introduce the "relative" motion and distinguish it from the usual or "absolute" motion. The relative motion is given when the three mutual distances are given as functions of the time. The absolute motion may be thought as less directly perceptible. To give the "absolute coordinates" (x, y, z) of each particle, we have to choose a Galilean frame. Fortunately there is a good choice for the origin of this reference frame: the center of mass of the three particles. But we should also decide what are the "fixed directions" (O, x), (O, y), (O, z). We have to decide what is a truly non-rotating frame, and this is not easy. And we are not sure that at the end we are interested in these "absolute coordinates". Is their knowledge important if we want to predict some astronomical events, such as the transit of Venus of 2004/6/8?

The idea to take the mutual distances r_{ij} between particles as variables, and to concentrate on the prediction of the future values of the r_{ij}'s, is quite old. In 1772, Lagrange obtained a very elegant system of differential equations in these variables, which happened to be the first complete "reduction" of the three-body problem. Some authors still erroneously attribute to Jacobi the first complete reduction, who called "elimination of the node" the last step of an equivalent, but less elegant, process of reduction published in 1842. Even today it is quite difficult to find in the literature the complete expression of Lagrange's system, and the easiest reference is the short exposition by Serret that follows Lagrange's paper in his "œuvres", volume 6. We will write this system explicitly after some explanation.

To present quickly the work of Lagrange, it is probably useful to explain first what a process of reduction is. Most modern presentations explain this in the Lagrangian or Hamiltonian frameworks, which is a huge pedagogical mistake. One should teach what is the reduced system obtained from a system of ordinary differential equations with symmetry to students that do not know what a Lagrangian system is (nor what a Lie group is). Once this concept has been understood, it can be applied to Hamiltonian systems, and a wonderful compatibility with the symplectic structure is observed. But it can also be applied to systems from non-holonomic dynamics with a continuous symmetry. And these are not Hamiltonian.

Reduction and separation of variables have belonged to the toolbox of theoretical mechanics since the times of Newton, Euler and Lagrange. They were used to integrate differential systems from mechanics. In the modern language of geometry, they are associated respectively to a fibration and to a product structure of the phase space. Respectively again, these structures correspond to the diagrams

$$\begin{array}{c} \mathcal{M} \\ \downarrow \\ B \end{array} \qquad\qquad \begin{array}{ccc} & \mathcal{M} & \\ \swarrow & & \searrow \\ A & & B \end{array}$$

where $\mathcal{M}$ is the phase space, i.e. the space of states of the system, or as well the space of possible initial conditions. A state or an initial condition is the data of the positions and the velocities of the particles.

Our first diagram $\mathcal{M} \to B$ (we will not discuss the second) maps the phase space on a “base” B of lower dimension (this “reduction” of the dimension explains the name of the process). Two states project on the same point of B if and only if there exists a transformation in the symmetry group of the mechanical system that sends one to the other. The base B is the “quotient space”. To any of its point is attached a whole class of states. For example, if we imagine a configuration of n particles, with n velocity vectors attached, as an “object” that we can rotate, all the objects deduced from this first object by rotation or translation will project on the same point of B. *It is enough to know the projection on B of some initial condition to predict the future evolution of the projection on B.*

There is another kind of “reduction”, of more trivial nature: the reduction corresponding to a first integral. Its geometrical counterpart is a foliation of $\mathcal{M}$ in the level hypersurfaces of this first integral. Here again the dimension of the phase space is “reduced”, the new phase space being one of these invariant hypersurfaces of $\mathcal{M}$.

The modern theory of quotient spaces and the old theory of reduction were explicitly related for the first time by S. Smale in 1970, in his paper “Topology and mechanics”. This resulted in a strong renewal of the old tool of classical mechanics*, marked at the beginning by the works of K.R. Meyer in 1973, and Marsden and Weinstein in 1974. These works were dedicated to

* Significant efforts of description of B (denoted by M_7) in the 3-body case were made by G.D. Birkhoff at a time when the word “quotient space” was not yet in use: “The ‘reduced manifold M_7 of states of motion’ corresponds to the ∞^7 set of states of motion given by sets of coördinates such as $u_1, \ldots, u_7$, which are distinct except in orientation about the axis of angular momentum.” (see [Bir], p. 285)

the "wonderful compatibility" with the symplectic structure mentioned above.

Let us come back to Lagrange's work on the 3-body problem. Our "object" is the triangle of the three particles, with three arrows attached to the vertices representing the instantaneous velocities. How can we describe this object, up to translation and rotation? Curiously it is both simpler and more satisfying to replace the word "rotation" by the word "isometry", which includes reflections. Second objection: if we add the same vector to each velocity, which is called a "boost", the future evolution of the three mutual distances is the same. Our expression "translation and isometry" is not accurate to describe the full Galilean symmetry.

Well, we are meeting some complications, which were resolved more than 200 years ago. The Jacobian way, the "dirty one", was to report everything to the center of mass of the system, which has a uniform rectilinear motion, then remark that two position vectors of the particles determine the third. Finally, Jacobi chose as well as it can be two vectors to represent the three.

But we will try to be as elegant as Lagrange was. We will object first: wait a minute, what have the masses to do with our problem of parameterizing the isometry classes of objects? Let us forget the masses and continue. We call q_1, q_2, q_3 the positions of the particles, $q_{23} = q_3 - q_2$, q_{31}, q_{12} the three separation *vectors*, from the second particle to the third, etc. As $q_{23} + q_{31} + q_{12} = 0$ our three vectors actually give the information that two vectors could give. The remark on boost symmetry suggests that the situation is similar with the velocities. So we take three more vectors $\dot{q}_{23}$, $\dot{q}_{31}$, $\dot{q}_{12}$. This is enough. The symmetries of translation and boost are treated. We will call the data we get the *six-vector object*. It is the figure formed by the six vectors. We should not forget that $q_{23}+q_{31}+q_{12} = \dot{q}_{23}+\dot{q}_{31}+\dot{q}_{12} = 0$, thus four vectors determine the whole object, which can be of dimension 1, 2, 3 or 4. Newton's equations are such that this dimension does not vary during the motion.

From the six-vector object to the 10 Lagrange's variables. To "reduce" the isometry symmetry, we replace the 6 vectors, which are actually four, by their mutual scalar products. Four vectors give 10 possible scalar products (four of which being the scalar products of the vectors with themselves). Thus we need 10 scalar products, but we need an elegant choice. The list $\|q_{23}\|^2$, $\|q_{31}\|^2$, $\|q_{12}\|^2$, $\|\dot{q}_{23}\|^2$, $\|\dot{q}_{31}\|^2$, $\|\dot{q}_{12}\|^2$, $\langle q_{23}, \dot{q}_{23}\rangle$, $\langle q_{31}, \dot{q}_{31}\rangle$, $\langle q_{12}, \dot{q}_{12}\rangle$ is quite nice. One variable is missing. Lagrange chose, and he was absolutely right, as we will see later,

$$\rho = \langle q_{23}, \dot{q}_{31}\rangle - \langle q_{31}, \dot{q}_{23}\rangle = \langle q_{23}, \dot{q}_1\rangle + \langle q_{31}, \dot{q}_2\rangle + \langle q_{12}, \dot{q}_3\rangle.$$

The differential system of the 3-body problem reduces to a differential system in these ten variables. The resulting "reduced system" will describe the *relative* motion of the particles. We did not say anything about the dimension of the space, which is naturally three. A surprising feature of Lagrange's reduced system is that it describes the motion as well if the motion is one-dimensional (the three particles remain on a fixed line, which implies that they collide in the past or/and in the future), if it is planar, spatial, or even if it is four-dimensional! In the collinear and planar cases we must admit we have too many variables. F.D. Murnaghan devised a smaller set of adapted variables.

In the 3D case we would need only 9 variables, because the four vectors have 12 coordinates, and the rotation group has 3 dimensions. It is easy to see what happens. The so-called Gram matrix of four vectors is the symmetric four by four matrix with the 16 scalar products, which are only 10 by symmetry, as entries. The condition that the four vectors generate a 3D space of dimension at most 3 is: the determinant of the Gram matrix vanishes. It is a complicated relation between the 10 Lagrange's variables. Lagrange wrote explicitly its long expansion (the notation for a determinant was not yet invented). Of course, if we compute the 10 scalar products from an initial condition of the 3D three-body problem, the relation will be satisfied. We do not always need to know the long formula.

What concepts do we need to learn in order to be able to deduce the reduced system? Shall we need to master the theory of Lie algebras, the moment map, the coadjoint actions? No, we only need to compute the first derivatives of our ten variables with respect to time. So we first write Newton's equations and deduce $\ddot{q}_{23}$, etc.

Newton's equations. Let $q_1, \dots, q_n$ represent the positions at time t of n particles with respective masses $m_1, \dots, m_n$. We have a notation for the vectorial separations $q_{ik} = q_k - q_i$, and we put $S_{ik} = S_{ki} = \|q_{ik}\|^{-3/2}$. The so-called Newton's equations are

$$\ddot{q}_i = \sum_{k \neq i} m_k S_{ik} q_{ik}. \tag{1}$$

There is a nice way to write $\ddot{q}_{ij}$. We put

$$\Sigma_{ij} = (m_i + m_j) S_{ij} + \frac{1}{2} \sum_{k \neq i,j} m_k (S_{ik} + S_{jk}). \tag{2}$$

Then

$$\ddot{q}_{ij} = -\Sigma_{ij} q_{ij} - \sum_{k \neq i,j} m_k (S_{ik} - S_{jk}) \big(q_k - \frac{1}{2}(q_i + q_j) \big). \tag{3}$$

Equations (3) form a "closed" system, involving only the vectorial separation q_{ij} and their derivatives. From (3) we deduce for example that

$$\langle q_{ij}, \ddot{q}_{ij} \rangle = -\Sigma_{ij} \|q_{ij}\|^2 - \frac{1}{2} \sum_{k \neq i,j} m_k (S_{ik} - S_{jk}) \langle q_{ik} + q_{jk}, q_{ij} \rangle. \tag{4}$$

But $q_{ij} = q_{ik} + q_{kj}$ so $\langle q_{ik} + q_{jk}, q_{ij} \rangle = \|q_{ik}\|^2 - \|q_{kj}\|^2$, that we can substitute above. Doing such manipulations in the case of three bodies, we arrive to Lagrange's system. Let

$$a = \|q_{23}\|^2, \quad b = \|q_{31}\|^2, \quad c = \|q_{12}\|^2, \quad a' = \langle q_{23}, \dot{q}_{23} \rangle, \quad b' = \langle q_{31}, \dot{q}_{31} \rangle, \quad c' = \langle q_{12}, \dot{q}_{12} \rangle,$$

$$a'' = \|\dot{q}_{23}\|^2, \quad b'' = \|\dot{q}_{31}\|^2, \quad c'' = \|\dot{q}_{12}\|^2, \quad \rho = \langle q_{23}, \dot{q}_{31} \rangle - \langle q_{31}, \dot{q}_{23} \rangle.$$

be the ten variables. We need intermediate expressions that we leave as above with indices:

$$\Sigma_{23} = (m_2 + m_3) a^{-3/2} + \frac{1}{2} m_1 (b^{-3/2} + c^{-3/2}), \quad \Sigma_{31} = \cdots, \quad \Sigma_{12} = \cdots,$$

where the last are simply circular permutations of the first. Lagrange's system is

$$\dot a = 2a', \quad \dot b = 2b', \quad \dot c = 2c', \quad \dot a' = a'' - a\Sigma_{23} - m_1(c^{-3/2} - b^{-3/2})(c-b)/2, \quad \dot b' = \cdots,$$
$$\dot c' = \cdots, \quad \dot a'' = -2a'\Sigma_{23} - m_1(c^{-3/2} - b^{-3/2})(c' - b' - \rho), \quad \dot b'' = \cdots, \quad \dot c'' = \cdots,$$
$$\dot\rho = \frac{1}{2}\begin{vmatrix} 1 & 1 & 1 \\ m_1(a-b-c) & m_2(b-c-a) & m_3(c-a-b) \\ a^{-3/2} & b^{-3/2} & c^{-3/2} \end{vmatrix}. \tag{5}$$

The fundamental observation is that this system is "closed", i.e. the second members are expressed in the ten variables. One could reasonably guess that there exists a "theory of reduction" that is able to predict this observation, avoiding the computations above. This is only half true. On the open set of $\mathcal{M}$ corresponding to states of dimension 4, one could predict that the second member may be expressed uniquely as a function of the scalar products.

But this is much less than what we got. We got simple algebraic expressions that are available in any dimension. Lagrange indicated that we can actually derive a rational system. It is enough to take the $r_{ij} = \|q_{ij}\|$ and the $\dot r_{ij}$ as variables instead of the $\|q_{ij}\|^2$ and $\langle q_{ij}, \dot q_{ij}\rangle$, keeping the other four variables unchanged.

This contrast between theoretical prediction and computation is a known difficulty that is commonly met when dealing with invariants. We are working with scalar products, which are the basic invariants under the action on the isometry group. Some theoretical arguments may predict that the second members of Lagrange's system are invariant under the action of this group, and that they can be expressed in terms of the basic invariants. However, it seems more difficult to predict that these second members will have simple (algebraic or even rational) expressions, well-behaved when the dimension of the state changes.

In his presentation of the theory of invariants, Hermann Weyl discussed this point: "In those cases, and here is the point I wish to emphasize, one will find the purely functional part—asserting the value of all invariants are determined by the values of basic invariants—almost trivial; the essential difficulties lie in the algebraic part only."

In brief, the computations above, due to Lagrange, remain today the most reasonable approach to discuss the reduction in term of mutual distances. As we wish to study this reduction more deeply, we will have to understand those computations better. Our tools will not come from differential geometry, but from linear algebra.

The first challenge is to pass to the n-body problem. There are some difficulties with the variable ρ, but Betti solved them and obtained the long system generalizing Lagrange's system for $n \geq 3$ particles. But these long pages of formulas raise questions: is it possible to do anything at all with this system? Is it possible, for example, to solve with $n > 3$ the problems that Lagrange solved in the case $n = 3$? We will discuss these problems in the next chapter.

Last comments about Lagrange's reduction. The aim of the reduction is to reduce the order of the differential system. What is the order of Lagrange's system? Our definition of the order is such that an autonomous system of n first order ordinary differential equations

$$\dot x_1 = f_1(x_1, \ldots, x_n), \quad \dot x_2 = f_2(x_1, \ldots, x_n), \quad \ldots, \quad \dot x_n = f_n(x_1, \ldots, x_n)$$

is of order n. Of course there exist other conventions, but this one seems to be very natural.

Lagrange's system is thus of order 10, but it admits several first integrals. Fixing the first integrals, we can eliminate some variables and reduce the order. If the dimension of the state is three, there is also a relation expressing the three-dimensionality, which can be used to eliminate one variable. The first integrals are the energy and the Euclidean norm of the angular momentum. *The order of the system is 7.* If the dimension is four, there are two independent $O(4)$-invariants of the angular momentum instead of one. The dimension is 7 again. It can be shown that for some particular values of the angular momentum (equal diagonal terms in the diagonalization), the reduced dimension falls to 5. Of course all these statements assume that there is no other first integral or symmetry than the known ones. Several results about this have been established, starting with the famous result of Bruns in 1887. They all leave some possibilities for (improbable) new invariants.

For information, we give expressions of the energy H and the squared norm of the angular momentum $\|C\|^2$ in Lagrange's set of variables. We make first $M = m_1 + m_2 + m_3$ and

$$I = \frac{1}{M}(m_2 m_3 a + m_3 m_1 b + m_1 m_2 c), \quad J = \frac{1}{M}(m_2 m_3 a' + m_3 m_1 b' + m_1 m_2 c')$$

$$K = \frac{1}{M}(m_2 m_3 a'' + m_3 m_1 b'' + m_1 m_2 c''), \quad U = m_2 m_3 a^{-1/2} + m_3 m_1 b^{-1/2} + m_1 m_2 c^{-1/2}.$$

$$\psi = -aa'' - bb'' - cc'' + ab'' + bc'' + ca'' + a''b + b''c + c''a,$$

$$\phi = -a'^2 - b'^2 - c'^2 + 2a'b' + 2b'c' + 2c'a'.$$

Then

$$H = \frac{1}{2}K - U, \qquad \|C\|^2 = \frac{m_1 m_2 m_3}{2M}(\phi - \psi + \rho^2) + IK - J^2. \tag{6}$$

2. Special motions

Lagrange used his system to classify the self-similar motions of three bodies. Here are three definitions, including the definition of a self-similar motion (called "homographic" by Wintner).

Definitions. A motion of three bodies is called *self-similar* if the ratios of mutual distances r_{31}/r_{23} and r_{12}/r_{23} remain constant with time. A motion is called *rigid* if the three mutual distances r_{23}, r_{31}, r_{12} are constant. A motion is called a *relative equilibrium* if the ten Lagrange's scalar products are constant.

A relative equilibrium is an equilibrium of the reduced system. In such a motion, the 6-vector object $q_{23}, \dots, \dot{q}_{12}$ remains in the same isometry class. We will also say that the *relative state* is constant. In a rigid motion, the *relative configuration* is constant.

Lagrange completely classified self-similar motions. The set of configurations he obtained is well-known: there are the collinear configurations previously discovered by Euler, and the equilateral triangle. In 1843, Gascheau discovered that Lagrange's equilateral configuration was linearly stable if one of the three bodies has a much bigger mass than the other two. It was discovered later that such configuration is realized in the solar system: some asteroids have such a motion that the triangle they form with the Sun and Jupiter remains nearly equilateral.

These "Lagrange's solutions" remain the striking discovery in the 95 pages long "Essai sur le problème des trois corps". But one should not forget that this work also contains an elegant and effective reduction of the three-body problem, which is a necessary tool to prove the following non trivial result (in contrast, Jacobi's reduction would not help to deduce it):

Proposition (Lagrange). Any self-similar motion of the three-body problem is planar.

The motion is assumed to be at most three-dimensional, and the conclusion is that it is two-dimensional. To avoid any ambiguity, let us state again what we call "dimension of the motion". It is the dimension of the vector space generated by the six vectors $q_{23},\dots,\dot q_{12}$. As already asserted, this dimension is 1, 2, 3 or 4. It is what one would call "dimension of the motion", provided one correctly chooses the Galilean frame. For example, the dimension of an elliptic motion of two bodies is two. In a frame where the velocity of the center of mass is non-zero and not in the plane of the elliptic orbit, one could think that the dimension is three. This is why one has to fix a correct Galilean frame, where the center of mass is at rest, to determine the dimension. But if we wish to avoid considerations about the center of mass, it is enough to look at System (3) instead of Newton's equations.

To show the effectiveness of Lagrange's system, let us undertake the classification of *rigid motions*. We assume that a, b and c are constant. Then $a' = b' = c' = 0$. The following equations assign to a'', b'' and c'' fixed values. Then $0 = \dot a'' = m_1(c^{-3/2} - b^{-3/2})\rho = \dot b'' = \cdots$. If the triangle is not equilateral, these equations imply $\rho = 0$. If the triangle is equilateral, we can check from the last equation that $\dot\rho = 0$. Thus *the motion is a relative equilibrium.*

We will now prove that a motion of relative equilibrium is of dimension 2 or 4. This point is rather technical if we use Lagrange's system only, and Lagrange himself devoted to it one page of subtle computations (p. 277). With A. Chenciner we devised a different strategy, considering the rigid dynamics of the *state*, i.e. the 6-vector object. We know that this object is constant up to rotation. Suppose that the object is three-dimensional. Then there exists an instantaneous rotation vector Ω such that the six vector equations

$$\dot q_{23} = \Omega \times q_{23}, \quad \dot q_{31} = \Omega \times q_{31}, \quad \dots, \quad \ddot q_{31} = \Omega \times \dot q_{31}, \quad \ddot q_{12} = \Omega \times \dot q_{12},$$

are satisfied. We don't know *a priori* if Ω is constant. But we know that if we rotate the 6-vector object around Ω, these equations still coincide with Newton's equations. So the solution with fixed Ω, which is a uniform rotation of the 6-vector object around Ω, is a solution of Newton's equations. But the solution is unique: it is this solution, and consequently Ω is fixed.

But this motion is impossible if the configuration does not lie entirely in the vectorial plane orthogonal to Ω. This is a quite simple and intuitive argument. Consider that Ω is vertical, and take the "highest" particle. It is attracted by the other particles, situated "below". So its gravitational attraction points downward. The centrifugal acceleration points horizontally. It cannot balance the gravitational acceleration. Thus the particle will go down, which contradicts the hypothesis of a uniform rotation with vertical axis.

Consequently the hypothesis of a three-dimensional relative equilibrium leads to the conclusion that the motion is actually two-dimensional. For the corresponding result on self-similar motions, we suggest the reader could read Lagrange's paper, or our paper with Chenciner. There are also interesting works by Pizzetti, Banachiewitz, Müntz, Carathéodory, etc. The remaining steps, i.e. the classification of 2-dimensional self-similar motion, will motivate our definition of a central configuration.

But first we will describe some four-dimensional motions. Lagrange was reasonable so he did not treat these exotic cases! However, we will quickly discover something interesting. As $\dot{\rho} = 0$, the configuration is subject to the constraint

$$\begin{vmatrix} 1 & 1 & 1 \\ m_1(a-b-c) & m_2(b-c-a) & m_3(c-a-b) \\ a^{-3/2} & b^{-3/2} & c^{-3/2} \end{vmatrix} = 0. \tag{7}$$

It can be shown that any triangle with sides $\sqrt{a}$, $\sqrt{b}$, $\sqrt{c}$, satisfying this constraint can undertake a four-dimensional uniform rotation under Newton's laws. The intuition is not easy, but crudely there will be two orthogonal axes drawn on the triangle and crossing at the center of mass, such that the triangle rotates uniformly with angular velocity n_1 around one axis, and angular velocity n_2 around the other. If there exist two integers k_1 and k_2 such that $k_1 n_1 = k_2 n_2$, the motion is periodic. If not, it is only quasi-periodic. If the triangle is isosceles, with two equal mass particles at the base, one of the axes of rotation is the "vertical" axis, i.e. the axis of symmetry, while the other is horizontal.

What we want to emphasize is a lemma concerning only Eq. (7). *If $m_1 = m_2 = m_3$, then the solutions of* (7) *are all the isosceles triangles.* This statement is particularly nice and simple, as is its proof. If ϕ is a strictly convex* function, or a strictly concave function, then

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ \phi(a) & \phi(b) & \phi(c) \end{vmatrix} = 0 \qquad \Longrightarrow \qquad a = b \quad \text{or} \quad b = c \quad \text{or} \quad c = a.$$

The proof is evident. Indeed the determinant is the area the triangle $(a, \phi(a))$, $(b, \phi(b))$, $(c, \phi(c))$. It is zero iff the triangle is flat, i.e. the three points are collinear. But this is not possible on the graph of a strictly convex function, if a, b and c are distinct.

We apply the argument to $\phi(x) = x^{-3/2}$, which is strictly convex on the domain $x > 0$, and this proves the lemma. The choice of a, b and c as the variables is crucial in the success of this deduction.

Less exotic dimensions. Let us classify the self-similar motions of n particles in the plane. We associate to the particle j of mass m_j a complex number z_j, its position on the complex plane. The center of mass is at zero. In a self-similar motion, we have $z_j = re^{i\theta} z_j^0$, which gives $\dot{z}_j = (\dot{r} + ir\dot{\theta})e^{i\theta} z_j^0$, and $\ddot{z}_j = (\ddot{r} + i(2\dot{r}\dot{\theta} + \ddot{\theta}) - r\dot{\theta}^2) z_j$.

* A real function is called convex if any segment from one point of its graph to another is "above" the graph. It is strictly convex if furthermore the interior of the segment does not contain any point of the graph.

Let us deduce two fundamental identities from Newton's equations. As $\ddot{z}_j = \sum_{k\neq j} m_k S_{kj} z_{jk}$, we get first $\sum m_j \ddot{z}_j = 0$, then $\sum m_j \mathrm{Im}(\bar{z}_j \ddot{z}_j) = 0$. Substituting $\ddot{z}_j$ from the above equation, we obtain $(2\dot{r}\dot{\theta} + \ddot{\theta}) \sum m_j |z_j|^2 = 0$. This implies $2\dot{r}\dot{\theta} + \ddot{\theta} = 0$. There exists at any time a *real* number $\lambda = \ddot{r} - r\dot{\theta}^2$ such that $\ddot{z}_j = -\lambda z_j$.

This motivates the introduction of the main object of these lectures, the central configurations. The name was given by Wintner, but the definition goes back to Laplace.

Definition. A central configuration of n particles $q_1,\ q_2, \dots, q_n$ in a Euclidean space of finite dimension, with masses $m_1,\ m_2, \dots m_n$, with center of mass q_G, is a configuation such that there exists a real number λ with

$$\lambda q_{Gj} = \sum_{k\neq j} m_k S_{kj} q_{kj} \tag{8}$$

for any j, $1 \leq j \leq n$. We denote as usual $q_{jk} = q_k - q_j$ the vector from the point q_j to the point q_k, and $S_{ki} = \|q_{ki}\|^{-3/2}$. We also make $q_{Gj} = q_j - q_G$.

Remark. The center of mass q_G is defined by the formula $Mq_G = \sum m_j q_j$, where $M = \sum m_j$. Recently with R. Moeckel we slightly extended the definition of a central configuration in order to adapt it to the n-vortex problem in hydrodynamics. In this problem, what plays the role of the mass of a particle is the vorticity of the vortex, and it can be negative. The sum of the masses may be zero. The definition of q_G fails and the center of mass does not exist. The extended definition is not more complicated, but somewhat less intuitive (see [Alb], [Cel]).

Let us relate the different definitions. A *relative equilibrium* is a *state* of the n-body problem. How will we describe a state independently of the choice of a Galilean frame? We will simply give the q_{jk} and their derivatives $\dot{q}_{jk}$. There are obvious relations between these vectors. A state of relative equilibrium is characterized by the fact that all the scalar products $\langle q_{jk}, q_{hl}\rangle$, $\langle q_{jk}, \dot{q}_{hl}\rangle$, $\langle \dot{q}_{jk}, \dot{q}_{hl}\rangle$ are constant with time. The more general states of self-similar motion are such that the ratios $\langle q_{ij}, q_{kl}\rangle / \|q_{12}\|^2$ are constant with time.

A central configuration is a *configuration*, i.e. is given by the vectors q_{jk} only. A result of Pizzetti is: *the configuration in any state of self-similar motion in a three-dimensional space is a central configuration.* As we saw, this statement is false in dimension four, already with a relative equilibrium of three bodies. In dimension three, most of the self-similar motions are planar, but there are exceptions: three-dimensional central configurations exist if $n \geq 4$, and a state made of such a configuration and of zero velocities generates a particular self-similar motion, called *homothetic*. In a homothetic motion the configuration does not rotate; only its scale changes. The configuration is central in any homothetic motion.

In the three-body case, for any choice of the masses, the central configurations are the three collinear Euler configurations and the equilateral Lagrange configuration. In the $n > 3$ body case, the number of central configuration does depend on the choice of the masses.

Now these lectures bifurcate into two parts. Chapters 3 and 4 may be read independently after this chapter. Chapter 4 introduces a tool to handle mutual distances, which will allow us to use standard results of linear algebra to prove results about central configurations. Chapter 3 will focus on results of symmetry of central configurations, using elementary algebra, Dziobek's theory, and some elements of convex analysis.

3. Dziobek's theory and symmetry of central configurations

Dziobek published in 1900 nice identities satisfied by central configurations. A related recipe was published in 1932 by MacMillan and Bartky. Here we will simplify both presentations, picking Dziobek's variables rather than MacMillan's, but choosing a nice argument by MacMillan in the proofs, rather than the heavier Cayley determinant and variational characterization by Dziobek. These choices were inspired by Hampton's thesis [Ham].

Dziobek's approach concerns central configurations of n bodies in dimension $n-2$. It is the second highest possible dimension for a configuration of n particles. In dimension $n-1$ the theory of central configurations is trivial, and left to the reader. Dziobek was interested in planar configurations of 4 bodies, which is also our main interest.

Suppose we have n points $q_1, \dots, q_n$ in an affine space of dimension $n-2$. The affine space can of course be Euclidean but we will not need any metrical relation at the beginning. The $n-1$ vectors q_{12}, $q_{13}, \dots, q_{1n}$ cannot be independent. There is a linear combination with real coefficients relating them. To be elegant, we will write it:

$$\sum_{i=1}^{n} \Delta_i q_i = 0, \quad \text{with} \quad \sum_{i=1}^{n} \Delta_i = 0, \qquad (\Delta_1, \dots, \Delta_n) \neq (0, \dots, 0). \tag{9}$$

We can consider that the q_i are vectors. Then the second relation shows that we can choose any origin for the vectors q_i in the first. But we can also consider that the q_i are *points* and refer to a standard convention in affine geometry: a linear combination of points with real coefficients is defined in two cases: 1) the sum of the coefficients is zero; then the combination is a vector; 2) the sum of the coefficients is one; then the combination is a point. Here we are in the first case.

The following statement is clear: *if the q_i are not contained in an affine subspace of dimension $n-3$, there exists a non-zero $(\Delta_1, \dots, \Delta_n) \in \mathbb{R}^n$ satisfying* (9), *which is unique up to multiplication by a real factor.*

Let us now show a useful manipulation of Equations (8) for central configurations. In the first member we read λq_{jG}. But $M q_{jG} = \sum m_k q_{jk}$ according to the definition of the center of mass. The equations become:

$$0 = \sum_{k \neq j} m_k (S_{jk} - \lambda/M) q_{jk}, \quad \text{with} \quad M = m_1 + \cdots + m_n. \tag{10}$$

But we deduce from (9) that $\sum \Delta_k q_{jk} = 0$, and we said that this relation is the unique linear relation, up to a factor, between the $n-1$ vectors q_{jk}, $k \neq j$. Thus, for any j there exists a factor μ_j such that $m_k(S_{jk} - \lambda/M) = \mu_j \Delta_k$. As $S_{kj} = S_{jk}$ we get $\mu_j \Delta_k / m_k = \mu_k \Delta_j / m_j$. This is true for any j, k thus $(\Delta_1/m_1, \dots, \Delta_n/m_n)$ and $(\mu_1, \dots, \mu_n)$ are proportional. We obtained Dziobek's relations:

$$S_{jk} = \|q_{jk}\|^{-3} = \frac{\lambda}{M} + \mu \frac{\Delta_j \Delta_k}{m_j m_k}, \tag{11}$$

for some real number μ. As there is a mutual distance in the left hand side, it is natural to try to express Δ_i as a function of the mutual distances. However, such an expression contains square roots, which complicates a lot the computations. Much better are the following implicit relations between the Δ_i and the $s_{ij} = \|q_{ij}\|^2$. First notice that the quantity $\sum_i \Delta_i \|q - q_i\|^2$ does not depend on the point q. This result is obtained expanding $\|q - q_i\|^2$ and using both relations (9). Consequently the quantities

$$t_i = \sum_j \Delta_j s_{ij}, \quad \text{with} \quad s_{ij} = \|q_{ij}\|^2 \tag{12}$$

are equal: $t_1 = t_2 = \cdots = t_n$. These are the implicit relations.

From (12) we deduce

$$0 = t_i - t_j = s_{ij}(\Delta_j - \Delta_i) + \sum_k \Delta_k (s_{ik} - s_{jk}),$$

and

$$0 = \Big(\frac{\Delta_i}{m_i} - \frac{\Delta_j}{m_j}\Big) s_{ij}(\Delta_j - \Delta_i) + m_k \sum_k \Big(\frac{\Delta_i \Delta_k}{m_i m_k} - \frac{\Delta_j \Delta_k}{m_j m_k}\Big)(s_{ik} - s_{jk}).$$

Substituting with (11) a term of the sum in k becomes $(S_{ik} - S_{jk})(s_{ik} - s_{jk})/\mu$. But the function $s \mapsto S = s^{-3/2}$ decreases. This quantity has the sign of $-\mu$. Thus

$$\mu\Big(\frac{\Delta_i}{m_i} - \frac{\Delta_j}{m_j}\Big)(\Delta_i - \Delta_j) \leq 0.$$

Let us choose the index i corresponding to the smallest Δ_i, and j corresponding to the greatest Δ_j. We have $\Delta_i < 0 < \Delta_j$, because $\sum \Delta_k = 0$. We get

$$\mu < 0 \qquad \text{and} \qquad \Big(\frac{\Delta_i}{m_i} - \frac{\Delta_j}{m_j}\Big)(\Delta_i - \Delta_j) \geq 0 \quad \text{for any } i, j, \quad 1 \leq i < j \leq n. \tag{13}$$

It can be easily shown that the case of equality in the second inequality corresponds to a symmetric configuration with $m_i = m_j$. The fact that the Δ_i/m_i's are ordered as the Δ_i's is a strange observation, that is extended to the "non-Dziobek" case, i.e. to arbitrary central configurations in [Alb].

From (12) we can extract weaker identities. Why they are useful is something we cannot fully explain. We write

$$Q_{ijk} = \begin{vmatrix} 1 & 1 & 1 \\ t_i & t_j & t_k \\ \Delta_i & \Delta_j & \Delta_k \end{vmatrix} = 0. \tag{14}$$

Of course $Q_{ijk} = 0$ if $t_i = t_j = t_k$. But if all the Q_{ijk} are zero, we can only deduce that $t_i = \alpha \Delta_i + \beta$ for some $(\alpha, \beta) \in \mathbb{R}^2$, and for all i. We should also remark that if we add the same constant to all the s_{lm}, $1 \leq l < m \leq n$, the value of Q_{ijk} is unchanged. This property simplifies some arguments in what follows.

Convex and concave configurations. Symmetry of central configurations.

The case where one of the Δ_i is zero is special. The point q_i is distinguished. Eq. (9) becomes an affine relation between the other points. It means that they are in a $(n-3)$-dimensional affine subspace. On another side, Eq. (11) shows that all these points are equidistant from q_i.

In the case of four bodies this is impossible. If three points are on a line, and also on a circle with center q_i, two of them must coincide. This is not allowed. Thus the case $\Delta_i = 0$ is a good "frontier" to define classes of central configurations.

Definition. A configuration of n particles, of dimension $n-2$ is called *concave* if all the Δ_i's have same sign except one. It is called *convex* in the remaining cases.

Discussion. Indeed this is not the general definition; it is rather a useful characterization. A convex configuration is more generally a configuration of n particles in any dimension such that the boundary of the convex hull contains all the points. The word "concave" is used for non-convex. Let us check the agreement between both definitions. If the configuration has dimension $n-2$, at most one particle is not on the boundary of the convex hull. The remaining particles form a simplex. The barycentric coordinates of the interior particle are positive. We have a relation as $q_n = (\alpha_1 q_1 + \cdots + \alpha_{n-1} q_{n-1})/(\alpha_1 + \cdots + \alpha_{n-1})$, with $\alpha_i > 0$. We make $\Delta_i = \alpha_i > 0$ and $\Delta_n = -\alpha_1 - \cdots - \alpha_{n-1} < 0$. This is the concave case in the above definition.

Symmetry. In 1995, I managed to prove that *any central configuration of four particles of equal masses possesses some symmetry.* This result was completed later by some polynomial calculations with a computer, giving the complete classification of these central configurations. Two results of symmetry were indeed necessary, one in the convex case, one in the concave case. Long and Sun published interesting extensions of both results, where only some of the masses are supposed to be equal. The study of the concave case below includes the extension. I will not present the convex case. The extensions are less easy and up to now incomplete. The reader should consult the paper by Long and Sun. I have also an unpublished paper in French, that needs to be improved and corrected before publication. In the 5-body case, the concave case is not easy. Almeida Santos [San] obtained several interesting results.

One of the results of symmetry. It corresponds to the concave case if we have four bodies, and to some convex cases in higher dimensions, e.g. 5 bodies in dimension 3.

Proposition. A central configuration of n bodies in dimension $n-2$, such that (i) three of the Δ_i's in Eq. (9) are negative, let us say $\Delta_1 < 0$, $\Delta_2 < 0$ and $\Delta_3 < 0$, (ii) $m_1 = m_2 = m_3$, (iii) the remaining Δ_i's are non-negative, possesses some symmetry. It satisfies $\Delta_1 = \Delta_2$ or $\Delta_2 = \Delta_3$ or $\Delta_3 = \Delta_1$. If for example $\Delta_1 = \Delta_2$, by Equation (11) the perpendicular bisector hyperplane of particles q_1 and q_2 contains all the remaining particles.

Proof. We use Equation (14) $Q_{123} = 0$. To expand it, we expand t_1, t_2 and t_3 in the determinant. We write first the terms in Δ_1, Δ_2, Δ_3.

$$Q_{123} = \begin{vmatrix} 1 & 1 & 1 \\ \Delta_2 s_{12} + \Delta_3 s_{13} & \Delta_1 s_{12} + \Delta_3 s_{23} & \Delta_1 s_{13} + \Delta_2 s_{23} \\ \Delta_1 & \Delta_2 & \Delta_3 \end{vmatrix} + \sum_l \begin{vmatrix} 1 & 1 & 1 \\ s_{1l} & s_{2l} & s_{3l} \\ \Delta_1 \Delta_l & \Delta_2 \Delta_l & \Delta_3 \Delta_l \end{vmatrix}.$$

Suppose $\Delta_1 < \Delta_2 < \Delta_3 < 0$ and all the remaining Δ_i are non-negative. Then

$$0 < \Delta_2\Delta_3 < \Delta_1\Delta_3 < \Delta_1\Delta_2, \qquad \text{and} \qquad \Delta_1\Delta_l < \Delta_2\Delta_l < \Delta_3\Delta_l < 0, \tag{15}$$

for any $\Delta_l > 0$. If $\Delta_l = 0$ the corresponding term vanishes. By (11) and $\mu < 0$ the determinant in Δ_l has the sign of

$$-\begin{vmatrix} 1 & 1 & 1 \\ s_{1l} & s_{2l} & s_{3l} \\ S_{1l} & S_{2l} & S_{3l} \end{vmatrix}.$$

This reminds us the determinant used to study the four-dimensional 3-body relative equilibria with equal masses. The determinant is the oriented area of the triangle (s_{1l}, S_{1l}), (s_{2l}, S_{2l}), (s_{3l}, S_{3l}). As the function $s \mapsto S = s^{-3/2}$ is convex, and $S_{1l} > S_{2l} > S_{3l}$, the determinant is positive and the term is negative.

Now the first determinant in the expression of Q_{123} above must be studied separately. We will use a technique, the simplex method, in the simplest case (of a 1-dimensional simplex). The general case is explained and used in [AlL] and [San].

Let us consider the three points in $\mathbb{R}^2$ with coordinates (s_{12}, S_{12}), (s_{13}, S_{13}), (s_{23}, S_{23}). The ordinates satisfy $S_{12} < S_{13} < S_{23}$ by (15) and (11) with $\mu < 0$. The three abscissa are ordered in the other way $s_{23} < s_{13} < s_{12}$ because $s \mapsto S = s^{-3/2}$ is a decreasing function. Furthermore, the convexity of this function implies an inequality $s_{13} < s_{13}^0$, where s_{13}^0 is defined by the collinearity condition

$$\begin{vmatrix} 1 & 1 & 1 \\ s_{12} & s_{13}^0 & s_{23} \\ S_{12} & S_{13} & S_{23} \end{vmatrix} = 0. \tag{16}$$

We claim that

$$\begin{vmatrix} 1 & 1 & 1 \\ \Delta_2 s_{12} + \Delta_3 s & \Delta_1 s_{12} + \Delta_3 s_{23} & \Delta_1 s + \Delta_2 s_{23} \\ \Delta_1 & \Delta_2 & \Delta_3 \end{vmatrix}$$

is indeed negative for any s such that $s_{23} \le s \le s_{13}^0$, so it is negative for $s = s_{13}$. As it is an affine function of s, it is sufficient to prove it is negative for $s = s_{23}$ and for $s = s_{13}^0$. At $s = s_{23}$, the above determinant has the same sign it has on $(s_{23}, s, s_{12}) = (0, 0, 1)$, because we can change (s_{23}, s, s_{12}) in $(s_{23}+x, s+x, s_{12}+x)$ without changing the value of the determinant. It is the sign of $\Delta_2^2 - \Delta_1^2 + \Delta_3(\Delta_1 - \Delta_2) = (\Delta_1 - \Delta_2)(\Delta_3 - \Delta_1 - \Delta_2) < 0$. To estimate the sign on $s = s_{13}^0$, we can replace (s_{12}, s, s_{23}) by $(-S_{12}, -S_{13}, -S_{23})$ using the same invariance and (16), and then replace by $(\Delta_1\Delta_2, \Delta_1\Delta_3, \Delta_2\Delta_3)$ using (11). Subtracting the third line multiplied by $\Delta_1^2 + \Delta_2^2 + \Delta_3^2$ we get

$$-\begin{vmatrix} 1 & 1 & 1 \\ \Delta_1^3 & \Delta_2^3 & \Delta_3^3 \\ \Delta_1 & \Delta_2 & \Delta_3 \end{vmatrix}.$$

The function $\Delta \mapsto \Delta^3$ being concave for $\Delta < 0$, this quantity is negative. Finally, all the terms in the expansion of Q_{123} are non-positive, and several of them are non-zero. Then $Q_{123} = 0$ is impossible. The condition $\Delta_1 < \Delta_2 < \Delta_3$ is impossible.

4. A tensor of mutual distances

From n points $q_1, \dots, q_n$ in a finite-dimensional Euclidean space we obtain the list s_{12}, s_{13}, $\dots, s_{n-1,n}$ of the squared mutual distances. A very natural question is: given such a list of $(n-1)n/2$ real numbers, is there a configuration of n points in a Euclidean space such that it is the list of the squared mutual distances? Obviously, the answer is not always yes. For example, the s_{ij}'s must be non-negative. There are also the triangular inequalities, etc.

Some conditions must be satisfied. There is a good way and a bad way to write them. Curiously, the bad way is the most widely known: *the Cayley determinant of the configuration, and of all the subconfigurations, must be non-negative.* We recall that the Cayley determinant of a tetrahedron $q_1, \dots, q_4$ is

$$\begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & s_{12} & s_{13} & s_{14} \\ 1 & s_{12} & 0 & s_{23} & s_{24} \\ 1 & s_{13} & s_{23} & 0 & s_{34} \\ 1 & s_{14} & s_{24} & s_{34} & 0 \end{vmatrix}. \tag{17}$$

Why do we call this the bad way to express the conditions? Because in most natural questions it leads to intractable computations. The good way is the following.

Lemma 1 (Borchardt). The real numbers $s_{12}, \dots, s_{n-1,n}$ are the squared mutual distances of a configuration in a Euclidean space if and only if the quadratic form

$$-\sum_{1 \le i < j \le n} s_{ij} \xi_i \xi_j$$

restricted to the hyperplane $H \subset \mathbb{R}^n$ of equation $\xi_1 + \ldots + \xi_n = 0$ is non-negative. The rank of this restricted quadratic form is the dimension of the configuration.

To illustrate Lemma 1, we compute the quadratic form on $(\xi_1, \dots, \xi_n) = (1, -1, 0, \dots, 0) \in H$. If a configuration exists, this number is non-negative according to Lemma 1. It is s_{12}. For a triangle, we can also take $(\xi_1, \xi_2, \xi_3) = (2, -1, -1)$, which gives an example of funny condition: $2s_{12} + 2s_{13} \ge s_{23}$.

Lemma 1 is similar to a well-known "vectorial" statement, associated to the Gram matrix.

Lemma 2. The $m(m+1)/2$ real numbers $t_{11}, t_{12}, \dots, t_{1m}, t_{22}, t_{23}, \dots, t_{mm}$ are such that there exist m vectors $v_1, \dots, v_m$ in a Euclidean vector space with $\langle v_i, v_j \rangle = t_{ij}$, $1 \le i \le j \le m$, if and only if the quadratic form $\sum_{1 \le i,j \le m} t_{ij} \eta_i \eta_j$ is non-negative on $\mathbb{R}^m \ni (\eta_1, \dots, \eta_m)$. Note that $j < i$ is allowed in the sum: we make $t_{ij} = t_{ji}$. The rank of this quadratic form is the dimension of the vectorial subspace spanned by the v_i's.

To deduce Lemma 1 from Lemma 2, make $m = n - 1$, $v_i = q_i - q_n$ (denoted by q_{ni}). We get $2t_{ij} = 2\langle q_{ni}, q_{nj} \rangle = -\|q_{ij}\|^2 + \|q_{ni}\|^2 + \|q_{nj}\|^2 = -s_{ij} + s_{in} + s_{jn}$ if $i < j$, and $t_{ii} = s_{in}$. The relations between the ξ's and the η's are: $\eta_i = \xi_i$ if $i < n$, $\xi_n = -\eta_1 - \cdots - \eta_m$. This substitution transforms $-\sum_{i<j} s_{ij} \xi_i \xi_j$ into $\sum_{i,j} t_{ij} \eta_i \eta_j$.

Some attributions. Lemma 1 was stated by Borchardt in 1866 in the case of maximal rank. He omitted to prove that the condition of positive-definiteness is sufficient. Darboux gave some geometrical interpretations of Borchardt's Lemma. Schoenberg independently published Lemma 1 with complete proofs in 1935. But instead of Borchardt's quadratic form, he used the less elegant and more obvious quadratic form $\sum_{i,j}(s_{in}+s_{jn}-s_{ij})\eta_i\eta_j$. He corrected this defect in 1938, thus establishing Lemma 1 completely. On another hand, Cayley's first mathematical publication, in 1841, contains the determinant that expresses the squared volume of a simplex as a function of the mutual distances. Actually, Cayley equalled to zero such a determinant in order to relate the ten mutual distances between five points in the three-dimensional space. The interpretation as a squared volume and the generalization to any number of points are quite trivial, so we can attribute everything to Cayley. However, different expansions of Cayley's determinant were known to be the squared volume of a tetrahedron much before Cayley. Euler and Lagrange are often quoted, but Tartaglia wrote such an expansion in his "General trattato di numeri et misure", p. 35, in 1560.

Lemma 1 shows that a quadratic form can replace the list of squared mutual distances. This will allow us to compute with mutual distances as we compute with quadratic forms. But we missed some intermediate step, that will allow us to prove Lemma 1 without using Lemma 2, and to understand Lemmas 1 and 2 as particular cases of the following Lemma 3.

Notation. Let E be a Euclidean space. We denote by Q the Euclidean "identification" $Q : E \to E^*$ from E on its dual vector space E^*. The linear map Q is symmetric: ${}^tQ = Q$. We now use the bracket notation for the duality product of E with E^*, and $\langle Qu, v\rangle = \langle u, {}^tQv\rangle$ is the Euclidean scalar product of $u \in E$ with $v \in E$. We also denote the Euclidean vector space E by the pair (E, Q).

Lemma 3. Let F be a vector space and $T : F \to F^*$ be a symmetric (${}^tT = T$) linear map from F to its dual F^*. The map T is such that there exist a Euclidean space (E, Q) and a linear map $B : F \to E$ with $T = {}^tBQB$ if and only if the quadratic form on $F : v \mapsto \langle Tv, v\rangle$ is non-negative. The rank of T is the rank of B.

Proof. The part "only if" is easy: $\langle Tv, v\rangle = \langle QBv, Bv\rangle \geq 0$ because Q is positive definite (i.e. Euclidean). We prove the part "if" taking any basis of F. We endow F with the standard Euclidean form, such that this basis is orthonormal. The matrix t of T is symmetric and non-negative: it possesses a non-negative square root, i.e. a matrix b such that $b^2 = t$ and ${}^tb = b$. The matrix b has the rank of t. It defines a linear application from F to F. We take as E the image of this application, and as Q the restriction to this image of the standard Euclidean form. We have $T = {}^tBQB$, as required. To obtain the assertion on the rank of T, it is sufficient to remark that $\mathrm{Ker}B = \mathrm{Ker}\,T$. One inclusion comes from the composition, the other is as follows. If $Tv = 0$, then $\langle Tv, v\rangle = 0$, and $\langle QBv, Bv\rangle = 0$. Thus $Bv = 0$.

Let us quickly analyze what are F, E, B and Q in Lemma 2. Clearly T is the Gram matrix $(\langle v_i, v_j\rangle)_{i,j}$, a symmetric form on $\mathbb{R}^m = F$. The v_j are elements of E. The linear map $B : \mathbb{R}^m \to E$ associate to $(\eta_1, \ldots, \eta_m)$ the linear combination $\sum \eta_i v_i$. Finally, Q is the Euclidean form on E. We have $T = {}^tBQB$.

In Lemma 1, $B : H \to E$ is the natural "affine" analogue of the above B. It associates to $(\xi_1, \ldots, \xi_n) \in H \subset \mathbb{R}^n$ the affine combination $\sum \xi_i q_i \in E$ of the points q_i. As in Chapter 3 we meet this nice distinction: the q_i are points in an affine space, but the linear combinations

$\sum \xi_i q_i$, with $\sum \xi_i = 0$, are vectors. Lemma 1 is now completely natural and as easy as Lemma 2. The reader will check easily the identity $T = {}^tBQB$ in this case.

Remark. It seems that the fundamental concepts of affine geometry are not taught everywhere in the world. This difficulty is doubled by a certain confusion on the use of the word "affine". One tends to use this word negatively, i.e. to express the lack of a certain structure. This is probably a bad habit. At the origin, it was used negatively to denote the lack of Euclidean structure. But later the finite-dimensional vector space became a more familiar object for the mathematician than the old Euclidean space. So "affine" began to be used negatively to express that the space does not possess a "base point", an element called zero. There is nothing to object. If we start with a Euclidean vector space, and want to arrive to an affine space, we must "forget" successively the Euclidean structure and that a special element is called zero. It seems difficult to speak clearly about these things today. We try to use the word positively, i.e. to express what is the structure, instead of what it is not. But we know that this raises a lot of questions from most readers. A good way would have been to follow the wise Schouten, who called "affine", "centered affine", "Euclidean", "centered Euclidean", what we call respectively "affine", "vector", "Euclidean" and "Euclidean vector" spaces.

The application $B : H \to E$ will be our configuration. The "three-vector figure" in Chapter 1 was precisely such an object. The respective images by B of the three elements $(0, -1, 1)$, $(1, 0, -1)$ and $(-1, 1, 0)$ of H are q_{23}, q_{31} and q_{12}.

We should write $B \in \mathrm{Hom}(H, E)$ but we will call $\mathcal{D}$ the dual of H and denote H by $\mathcal{D}^*$. We will also use tensorial notations. We will choose to write $B \in \mathcal{D} \otimes E$. The next chapter comes to explain this.

5. Two topics in linear and multilinear algebra

First topic. The disposition space. When we study Euclidean vector spaces, it is not necessary to explain what is the dual space of a vector space. The space and its dual are identified, they are the same space. Here our main vector spaces arrive with a natural Euclidean structure, but for computational reasons it is better to distinguish $\mathcal{D}$ and its dual $H = \mathcal{D}^*$.

A general principle is: *the dual of a subspace is a quotient space.* More precisely, if F is a finite-dimensional vector space, and $H \subset F$ is a subspace, there is a subspace $H^0 \subset F^*$ of the dual space called the annulator. In the Euclidean intuition it corresponds to the orthogonal of H. The space H^0 is the space of linear forms that vanish on H.

Our principle becomes the statement: *the dual of the subspace H of F is the quotient space F^*/H^0.* When we say "H^* is F^*/H^0" we mean "there is a canonical identification between H^* and F^*/H^0" and "we will identify these two spaces". When there is a canonical identification between two spaces, to insist in distinguishing them is often "purely pedantic", to quote Hermann Weyl, but in some cases there is a good reason to make the distinction. For example, if A and B are two sets, the Cartesian product $A \times B$ is canonically isomorphic to $B \times A$, but if you take the bad habit to identify (x, y) with (y, x), you will get into troubles when $A = B$...

To see the canonical identification $H^* \equiv F^*/H^0$, observe simply that a linear form on F is a linear form on the subspace $H \subset F$, and that two linear forms on F coincide on H if and only if their difference is in H^0.

The subspace we met is the hyperplane $H \subset \mathbb{R}^n$ with equation $\xi_1 + \cdots + \xi_n = 0$. The dual of $\mathbb{R}^n$ "is" $\mathbb{R}^n$. The annulator of H is the line $[L]$ generated by the vector $L = (1, \ldots, 1) \in \mathbb{R}^n$. Thus $H^* = \mathbb{R}^n/[L]$. Our preferred notation is $\mathcal{D} = \mathbb{R}^n/[L]$. We have $H = \mathcal{D}^*$ and $\mathcal{D}^*$ is the new name for H.

It is the second time in this lecture that we meet a quotient space. The first time was in connection with reduction. But look at $\mathcal{D}$; an element of this quotient space is a class of elements of $\mathbb{R}^n$. Two elements $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ are in the same class iff $x_1 - y_1 = \cdots = x_n - y_n$. If the numbers are the coordinates of particles on the line, two configurations of n particles are in the same class if they are deduced one from the other by a translation. *The space $\mathcal{D}$ is the space of collinear configurations "reduced" by the translations.* With Chenciner we called an element of $\mathcal{D}$ a "disposition".

Dispositions often appear in the theory of central configurations. They are not always directly related to true particles disposed on a line. In Chapter 3, we met the n-uple $(t_1, \ldots, t_n)$, and wrote the condition $t_1 = \cdots = t_n$. This condition characterizes the null disposition.

Second topic. Tensorial notations. Linear algebra is important. The theory of matrices is important. Linear algebra becomes the theory of matrices if we choose bases of the vector spaces. In our problem a good choice of a basis of $\mathcal{D}^* = H$ is induced by the so-called "Jacobi coordinates". But any specific choice of a basis leads to unnatural computations. We could also choose "any basis", an unspecified basis, and compute with matrices. But it would be difficult to distinguish $\mathcal{D}$ and $\mathcal{D}^*$. We would need to tell this old story of contravariant and covariant coordinates.

Let us call (M) a mathematician who likes the theory of matrices and finds that everything else is useless abstraction. A second mathematician (L) learned abstract linear algebra at the university. Finally, a third mathematician (T) learned tensor calculus (which is common), and decided not to be (L) anymore, i.e. to translate any question of linear algebra in term of tensor calculus (which is less common).

We met the space $\mathrm{Hom}(\mathcal{D}^*, E)$ of configurations of $n = m + 1$ points in the 3D vector space E (or in an affine space with direction E) up to translation. We ask our three mathematicians the question: *what is the dual of this space?*

(M) probably answers: the space is the space of $3 \times m$ matrices. To dualize I exchange lines and columns as I do with line vectors and column vectors. The dual is the space of $m \times 3$ matrices.

(L) objects: dear (M), when you write ${}^tXQX$, where X is a column vector and Q a quadratic form, X and ${}^tX$ are twice the same vector. They cannot live in two different spaces, as $\mathcal{D}$ and $\mathcal{D}^*$. I don't agree that duality is related to the presentation of a vector as a line or a column.

(L) thinks one minute and answers: it is $\mathrm{Hom}(\mathcal{D}, E^*)$. He thinks one more minute and says: it may be $\mathrm{Hom}(E, \mathcal{D}^*)$ as well, I cannot really decide. It is funny: I can put stars to each space, and use $(\mathcal{D}^*)^* = \mathcal{D}$ or I can simply switch the order of the spaces.

(T) answers immediately: I don't use Hom, I am tired with heavy notation. Your question is:

what is the dual of $\mathcal{D} \otimes E$? The answer is: $\mathcal{D}^* \otimes E^*$.

And the next question comes: *if* $B \in \mathrm{Hom}(\mathcal{D}^*, E)$, *where lives* ${}^tB$*?*

(M) answers: it is a $m \times 3$ matrix, obviously!

(L) continues: oh, I was trained a lot at the university: $B \in \mathrm{Hom}(E^*, \mathcal{D})$. I know the rule by heart: you exchange the order and put stars. It is the map which pulls-back a form.

(T) answers: what I know is that if $B \in \mathcal{D} \otimes E$, then ${}^tB \in E \otimes \mathcal{D}$. My answer is much simpler than (L)'s, and as simple as (M)'s. But (M) will never answer these questions correctly, because he has only two possibilities, $3 \times m$ and $m \times 3$, while we have eight distinctions: $\mathcal{D} \otimes E$, $E^* \otimes \mathcal{D}$, $\mathcal{D}^* \otimes E$, etc.

Another question: *prove that* $\mathrm{rank}\, B = \mathrm{rank}\, {}^tB$.

(M) answers: the rank of a matrix is computed from the determinants of square submatrices, so this is obvious.

(L) answers: the rank of B is the dimension of the image of B. The image of ${}^tB$ is called the coimage. I learned this later at the university. I have $\dim \mathcal{D}^* = m$ and $\dim E = \nu$. I know that $\dim \ker B + \mathrm{rank}\, B = m$ (oh, I confess that I sometimes hesitate between ν and m...) and $\dim \ker {}^tB + \mathrm{rank}\, {}^tB = \nu$. Now $\ker {}^tB$ is called the co-kernel. It is the annulator of the image of B so its dimension is $\nu - \mathrm{rank}\, B$. Thus $\mathrm{rank}\, {}^tB = \nu - (\nu - \mathrm{rank}\, B)$.

(T) answers: what (L) calls image and coimage are for me "image at the left" and "image at the right". In the language of tensors one would rather speak about "support" than about "image"; but indeed "image" is convenient also. For symmetric or antisymmetric tensors the left support and the right support are the same, it is simply called the support. *The basic proposition is that the dimensions of the left support and the right support are equal.* This is the first statement in my theory and it is very easy to remember. The statement $\dim \ker B + \mathrm{rank}\, B = m$ is an easy corollary.

(L) objects: but you will not deal easily with the fundamental operation, the composition of linear maps.

(T) answers: the fundamental operations of tensor calculus are tensor product and contraction. The composition of linear maps in a "contracted product", a secondary operation, obtained by tensor product and contraction. I use the same symbol as yours, $\circ$, for a composition, and use it exactly as you do. But I call it contracted product. If $\alpha \in E \otimes F$ and $\beta \in F^* \otimes G$, then $\alpha \circ \beta \in E \otimes G$. I allow this "contraction" in $\alpha \otimes \beta \in E \otimes F \otimes F^* \otimes G$ because F and F^* "touch" each other in the ordered list E, F, F^*, G.

(L) answers: for me $\alpha \in \mathrm{Hom}(F^*, E)$ and $\beta \in \mathrm{Hom}(G^*, F^*)$ so I can form $\alpha \circ \beta \in \mathrm{Hom}(G^*, E)$.

(T) answers: this is a possible translation of my relation. You contract an element $\xi \in G^*$ *at the right hand side*. You contract first with β: it gives $\beta \circ \xi \in F^*$, that you contract with α. You have got an element of E. Your "composition" is for me an easy game!

(T) continues: but there is another possible interpretation of the contracted product $\alpha \circ \beta$ as a composition. Contracting $\eta \in E^*$ at the left hand side is equally good, and gives in your language: $\hat\alpha \in \mathrm{Hom}(E^*, F)$ and $\hat\beta \in \mathrm{Hom}(F, G)$ so $\hat\beta \circ \hat\alpha \in \mathrm{Hom}(E^*, G)$. I put a $\hat{}$ for you, it looks like your transposition. It is not my transposition. I don't transpose here. I use the possibility I have to read the tensorial relations as composition rules from the left to the right!

(M): Should I recognize the $\alpha_{ij}\beta^j{}_k$ of my colleagues using indices and Einstein's convention?

(T): Yes! Index notation constitutes even today the unique "official" way to deal with tensor calculus. Schouten in his "Ricci calculus" improved it and discussed it quite a lot. Here I consider that with my small tensors (at most two indices) it would be too "expensive" to write indices. But you can write them if you wish. For compatibility, I add a rule to Einstein's convention: the summation is allowed on repeated indices, one up, one down, *that are adjacent in the ordered list of indices*. In our formula $\alpha_{ij}\beta^j{}_k$, this list is $ijjk$ so the summation in j is allowed. You should use the transposition to bring together the indices as needed. In compensation, you will quickly notice that it is no longer necessary to write the indices.

Last question: *How do you write the evaluation of a quadratic form Q on a vector X?*

(M) answers ${}^tXQX$.

(L) answers: well, rather than a notation for the quadratic form, I prefer to give the name Q to the associated linear map $Q \in \mathrm{Hom}(F, F^*)$. So I use $\langle Q(X), X\rangle$. I can also do something similar to (M). I can see X as a linear map from $\mathbb{R}$ to the vector space F that contains X. It is the map $\lambda \mapsto \lambda X$. So I can use ${}^tX \circ Q \circ X$. It a "1×1 matrix", a real number.

(T) answers: $X \circ Q \circ X$. I have $X \in F$, $Q \in F^* \otimes F^*$, but Q is a symmetric form and I have a notation for this: $Q \in F^* \vee F^*$. I took this notation in Deheuvels, but many authors use $F^* \odot F^*$. There is no official notation in the language of (L).

(T) continues: *I don't transpose X.* The transposition on vectors is undefined: I cannot exchange the spaces as I did with E and $\mathcal{D}$ because there is only one space F. If we have another opportunity I will draw some pictures for you to explain why your transposition and mine are different things, even if we agree to write $Q = {}^tQ$.

6. Generalized Lagrange's system and equations for central configurations

As you noticed (L) wrote Chapter 4 and he also wrote [AlC]. But who is writing now is (T). If your way of thinking looks like (L)'s, you can translate easily. Just interpret $\circ$ as a composition. If you rather think matricially as (M), you will recognize your familiar formulas but you need to understand that we wish to distinguish clearly $\mathcal{D}$ and $\mathcal{D}^*$.

We recall that our main space is $\mathcal{D} = \mathbb{R}^n/[L]$, where $L = (1, \dots, 1) \in \mathbb{R}^n$ and $[L]$ means the vectorial line generated by L. We will often write $X = (x_1, \dots, x_n) \in \mathcal{D}$. Indeed $(x_1, \dots, x_n)$ is an element of $\mathbb{R}^n$ whose class is $X \in \mathcal{D}$. Also $\xi = (\xi_1, \dots, \xi_n) \in \mathcal{D}^*$ means $\xi_1 + \cdots + \xi_n = 0$. The duality bracket is $\langle \xi, X\rangle = \sum_i \xi_i x_i$. Of course the value depends only on the class X.

We introduce some notation. The "position" configuration is $x \in E \otimes \mathcal{D}$. Its time derivative is $\dot x = y \in E \otimes \mathcal{D}$. The mutual distances tensor is $\beta = {}^tx \circ x \in \mathcal{D} \vee \mathcal{D}$. We denote by $\mathcal{D} \vee \mathcal{D}$ the subspace of symmetric tensors in $\mathcal{D} \otimes \mathcal{D}$. The subspace of antisymmetric tensors is $\mathcal{D} \wedge \mathcal{D}$. *We make $E = E^*$, identifying with the Euclidean structure.* The formula $\beta = {}^tx \circ x$ is the same as $T = {}^tB \circ Q \circ B$ in the previous two chapters, but the convention of identification means $Q = \mathrm{Id}$, so Q disappears. This is why the convention is decided: to simplify the formulas.

If we give masses $(m_1, \dots, m_n)$ to the n particles, there is a Euclidean structure on $\mathcal{D}$. The Euclidean form is the "mass form" $\mu \in \mathcal{D}^* \vee \mathcal{D}^*$. To obtain its value on $X = (x_1, \dots, x_n) \in \mathcal{D}$ we can use one of the formulas:

$$X \circ \mu \circ X = \frac{1}{M} \sum_{1 \leq i < j \leq n} m_i m_j (x_i - x_j)^2 = \sum_{i=1}^{n} m_i (x_i - x_G)^2,$$

where

$$M = m_1 + \cdots + m_n \qquad \text{and} \qquad x_G = \frac{1}{M}(m_1 x_1 + \cdots + m_n x_n).$$

The form μ is positive definite: if $X \circ \mu \circ X = 0$, then X is the null disposition $x_1 = \cdots = x_n$.

In purely geometrical problems, there are no masses and thus no preferred identification between $\mathcal{D}$ and $\mathcal{D}^*$. But after the introduction of the masses the convention $\mathcal{D} = \mathcal{D}^*$ becomes a possibility. It simplifies the tensorial relations we will establish as much as $E = E^*$ does. But it happens that most of the formulas we meet are linear in the masses, and that we can observe this important property looking at the tensorial relations only if we don't make $\mathcal{D} = \mathcal{D}^*$.

Force function and factorization of Newton's equations. Let us deduce a useful factorization of Newton's equation from simple computations using the force function. It is well-known, and was first observed by Lagrange [La2], that Newton's equations

$$\ddot{q}_i = -\sum_{j \neq i} m_j S_{ij} q_{ji} \tag{18}$$

can be also written using the force function $U(q_1, \dots, q_n) = \sum_{i<j} m_i m_j \|q_{ij}\|^{-1}$,

$$m_i \ddot{q}_i = \frac{\partial U}{\partial q_i}. \tag{19}$$

But $x \in E \otimes \mathcal{D}$ is the configuration. We consider now U as a function of x, and write $U(x)$. The equation becomes

$$\ddot{x} \circ \mu = dU|_x. \tag{20}$$

Here $dU|_x$ means the differential of U at the point $x \in E \otimes \mathcal{D}$. It is element of the dual space $E^* \otimes \mathcal{D}^* = E \otimes \mathcal{D}^*$. Fortunately, both sides of the equation are in the same space.

Observe now that U is function of the mutual distances only. There exists a function $\hat{U}$: $\mathcal{D} \vee \mathcal{D} \to \mathbb{R}$ such that $U(x) = \hat{U}(\beta)$, where $\beta = {}^t x \circ x$ is the tensor of mutual distances. Differentiating this equation, and computing the differential on a "small" variation x' of x, we obtain

$$\langle dU|_x, x' \rangle = \langle d\hat{U}|_\beta, {}^t x' \circ x + {}^t x \circ x' \rangle = 2 \langle d\hat{U}|_\beta, {}^t x \circ x' \rangle$$

because $d\hat{U} \in (\mathcal{D} \vee \mathcal{D})^* = \mathcal{D}^* \vee \mathcal{D}^*$ is a symmetric tensor. A suitable expression for the duality bracket $\langle \phi, \psi \rangle$ of $\phi \in A \otimes B$ and $\psi \in A^* \otimes B^*$ is $\langle \phi, \psi \rangle = \mathrm{trace}({}^t\phi \circ \psi) = \mathrm{trace}(\psi \circ {}^t\phi)$. It is an easy exercise to deduce from it the familiar rule $\langle a \circ b, c \rangle = \langle b, {}^t a \circ c \rangle$. Applying this rule we get $\langle dU|_x, x' \rangle = 2\langle x \circ d\hat{U}|_\beta, x' \rangle$. This is true for any x' and gives the factorization we announced

$$dU|_x = -x \circ \alpha, \qquad \text{making} \qquad \alpha = -2d\hat{U}|_\beta \in \mathcal{D}^* \vee \mathcal{D}^*, \tag{21}$$

to be substituted in Newton's equations (20).

Factorization of Newton's equations again. Direct approach. The projection of (18) on any axis of coordinates is $\ddot{x}_i = -\sum_{j \neq i} m_j S_{ij}(x_i - x_j)$. The second member depends on the *disposition* $X = (x_1, \ldots, x_n)$, and depends on it linearly. From the first member we retain only the disposition $\ddot{X} = (\ddot{x}_1, \ldots, \ddot{x}_n)$. We can write this equation $\ddot{X} = -X \circ \mathcal{Z}$, where $\mathcal{Z} \in \mathcal{D}^* \otimes \mathcal{D}$. It gives $\ddot{X} \circ \mu = -X \circ \mathcal{Z} \circ \mu$. The comparison with (21), projected on the same axis, suggests

$$\alpha = \mathcal{Z} \circ \mu. \tag{22}$$

To check this relation, let us choose any disposition $T = (t_1, \ldots, t_n)$. We have $(T \circ \mathcal{Z})_i = \sum_{j \neq i} m_j S_{ij}(t_i - t_j)$. As $\sum_i m_i (T \circ \mathcal{Z})_i = 0$, the multiplication by μ is simply the multiplication of each coordinate $(T \circ \mathcal{Z})_i$ by m_i. Consequently

$$T \circ \mathcal{Z} \circ \mu \circ T = \sum_i \Big(\sum_{j \neq i} m_i m_j S_{ij}(t_i - t_j) \Big) t_i = \sum_{i<j} m_i m_j S_{ij}(t_i - t_j)^2.$$

The last identity is obtained joining both terms in $m_i m_j S_{ij}$. To compare $\mathcal{Z} \circ \mu$ with α, we have to compute $d\hat{U}$. Let β', with "standard coordinates" s'_{ij}, i.e. such that $\xi \circ \beta' \circ \xi = -\sum_{i<j} s'_{ij} \xi_i \xi_j$, be a "small" variation of the mutual distances tensor. We have

$$U = \sum_{i<j} m_i m_j s_{ij}^{-1/2}, \qquad \langle d\hat{U}, \beta' \rangle = -\frac{1}{2} \sum m_i m_j s_{ij}^{-3/2} s'_{ij}.$$

We conclude that $\alpha = -2d\hat{U} = \mathcal{Z} \circ \mu$, using $T \circ \mathcal{Z} \circ \mu \circ T = \langle \mathcal{Z} \circ \mu, T \otimes T \rangle$, and noticing that if $\beta' = T \otimes T$, then $s'_{ij} = (t_i - t_j)^2$.

Remark. As μ and α are both positive definite on $\mathcal{D}$, the "operator" $\mathcal{Z}$ has $n-1$ positive eigenvalues.

Generalized Lagrange's system. We can now compute efficiently with the mutual distances. The absolute state is (x, y), with $y = \dot{x}$. The basic invariants under the action of the isometries are

$$\beta = {}^t x \circ x, \quad \gamma = \frac{1}{2}({}^t y \circ x + {}^t x \circ y), \quad \rho = \frac{1}{2}({}^t y \circ x - {}^t x \circ y), \quad \delta = {}^t y \circ y.$$

We decomposed ${}^t y \circ x$ in symmetric part and antisymmetric part. In the three-body case ρ has only one coordinate, the scalar ρ used in the first chapter. Lagrange's system is:

$$\dot{\beta} = {}^t \dot{x} \circ x + {}^t x \circ \dot{x} = 2\gamma,$$

$$\dot\gamma = \delta - \frac{1}{2}({}^t\mathcal{Z}\circ{}^tx\circ x + {}^tx\circ x\circ\mathcal{Z}) = \delta - \frac{1}{2}({}^t\mathcal{Z}\circ\beta + \beta\circ\mathcal{Z}),$$

$$\dot\delta = -{}^t\mathcal{Z}\circ{}^tx\circ y - {}^ty\circ x\circ\mathcal{Z} = -{}^t\mathcal{Z}\circ\gamma - \gamma\circ\mathcal{Z} + {}^t\mathcal{Z}\circ\rho - \rho\circ\mathcal{Z},$$

$$\dot\rho = -\frac{1}{2}({}^t\mathcal{Z}\circ\beta - \beta\circ\mathcal{Z}).$$

We observe again that the system is closed, i.e. that the second member is expressed in β, γ, ρ, δ and $\mathcal{Z}$ only. Note that $\mathcal{Z}$ is a function of β and the masses. In the three-body case, we proved as an exercise that a rigid motion is a relative equilibrium. Let us do this here again.

Suppose that β is constant. Then $\dot\beta = 2\gamma = 0$. By the second equation above $2\delta = {}^t\mathcal{Z}\circ\beta + \beta\circ\mathcal{Z}$. So δ is constant and $0 = {}^t\mathcal{Z}\circ\rho - \rho\circ\mathcal{Z}$ by the third relation. Let $\rho_{\mu\nu}$ and $\beta_{\mu\nu}$ be the matrices of ρ and μ in a base where $\mathcal{Z}$ is diagonal with diagonal entries $\zeta_1,\dots,\zeta_{n-1}$. This equation reads $(\zeta_\nu - \zeta_\mu)\rho_{\mu\nu} = 0$, while the last equation of the system is $\dot\rho_{\mu\nu} = -(\zeta_\mu - \zeta_\nu)\beta_{\mu\nu}/2$. For (μ,ν) such that $\zeta_\nu = \zeta_\mu$ we have $\dot\rho_{\mu\nu} = 0$. For the other values we have $\rho_{\mu\nu} = 0$. So $\dot\rho = 0 = \dot\beta = \dot\gamma = \dot\delta$. The state is of relative equilibrium.

Balanced configurations. For a motion of relative equilibrium we have $\dot\rho = 0$, thus

$${}^t\mathcal{Z}\circ\beta - \beta\circ\mathcal{Z} = 0.$$

This equation in the unknown $\beta = {}^tx\circ x$ (the relative configuration) defines what we call the *balanced* configurations. One can prove that there exist motions of relative equilibrium with any such configuration. If the configuration is not central such a motion will be at least four-dimensional, so it could be concluded that the balanced configurations are useless in usual 3D mechanics. However, Moeckel showed in 1997 that the clusters of small masses in a central configuration are asymptotically balanced configurations.

The main known result on balanced configuration has been established in Chapter 2: any balanced configuration of three particles with equal masses is an isosceles triangle. Some results in the case of three particles with different masses may be found in [AlC]. We conclude with the following characterization.

Lemma. For any balanced configuration $x \in E\otimes\mathcal{D}$ of "full" dimension $\dim E$, there exists a positive $S \in E\vee E$ such that $S\circ x = x\circ\mathcal{Z}$.

Proof. Let $F\subset\mathcal{D}$ be the right image (or coimage) of x. A necessary condition for the existence of S is: the right image G of $x\circ\mathcal{Z}$ is included in F. But F is the image of β by the "isotropy implies degeneracy" argument (used in the proof of Lemma 3, Chapter 4) and consequently G is the right image of $\beta\circ\mathcal{Z}$. Then ${}^t\mathcal{Z}\circ\beta = \beta\circ\mathcal{Z}$ implies $G\subset F$, as required. Actually as $\mathcal{Z}$ is invertible $F = G$. But we don't want to use this property of $\mathcal{Z}$ here. Now we restrict everything to $F\subset\mathcal{D}$. In particular $x\in E\otimes F$ is invertible, so we simply make $S = x\circ\mathcal{Z}\circ x^{-1} = {}^tx^{-1}\circ(\beta\circ\mathcal{Z})\circ x^{-1}$.

Central configurations. A configuration $x\in E\otimes\mathcal{D}$ is central iff there exists a $\lambda\in\mathbb{R}$ such that $\lambda x = x\circ\mathcal{Z}$. Let us denote $\check{\mathcal{Z}} = \mathcal{Z} - \lambda\mathrm{Id}$. *A configuration is central with multiplier λ iff $x\circ\check{\mathcal{Z}} = 0$.* It means that the right image $F\subset\mathcal{D}$ of x is in the kernel of $\check{\mathcal{Z}}$.

Mutual distances. We can also characterize the relative configuration β. As F is also the image of β, *a relative configuration* $\beta \in \mathcal{D} \vee \mathcal{D}$ *is central with multiplier* λ *iff* $\beta \circ \check{\mathcal{Z}} = 0$.

In the Dziobek case treated in Chapter 3, the dimension of F is $n-2$, so the rank $\hat{Z}$ is one. Let $\Delta = (\Delta_1, \dots, \Delta_n) \in \mathcal{D}^*$ and $d = (\Delta_1/m_1, \dots, \Delta_n/m_n) \in \mathcal{D}$ be the quantities met in this chapter. The reader will easily check the identity $\check{\mathcal{Z}} = \Delta \otimes d$.

Dizzy configurations. A dizzy configuration x with multiplier λ is a configuration such that $\beta = {}^t x \circ x$ satisfies

$$ {}^t\check{\mathcal{Z}} \circ \beta + \beta \circ \check{\mathcal{Z}} = 0, \qquad \text{with} \qquad \check{\mathcal{Z}} = \mathcal{Z} - \lambda \mathrm{Id}. $$

For a dizzy configuration $x \in E \otimes \mathcal{D}$ with rank $\dim E$ there exists an $\Omega \in E \wedge E$ such that $\Omega \circ x = x \circ \check{\mathcal{Z}}$. The argument is the same as in the balanced case, but now $\beta \circ \hat{\mathcal{Z}}$ is antisymmetric.

Dizzy implies central. Multiplying $\Omega \circ x = x \circ \check{\mathcal{Z}}$ by $\mu \circ {}^t x$ at the right hand side, we observe that $\Omega \circ x \circ \mu \circ {}^t x$ is symmetric. Let us choose an orthonormal base of E where $b = x \circ \mu \circ {}^t x$, which is the inertia tensor of the configuration, is diagonal with diagonal entries $b_1, \dots, b_\nu$. If in this base $\Omega = (\Omega_{ij})_{1 \le i,j \le \nu}$, then $\Omega_{ij}(b_j + b_i) = 0$. Then $\Omega_{ij} = 0$ for any $i < j$, i.e. $\Omega = 0$. Finally $x \circ \check{\mathcal{Z}} = 0$ and the configuration is central.

The dizzy configurations and the central configurations are the same thing, so the reader may ask why we introduced a new funny name. It happens that central configurations are important and defined in the same way in the problem of n Helmholtz' vortices in the plane. But what plays the role of the masses are the vorticities, and they can be negative. And there exist dizzy configurations in the plane that are not central if the masses are allowed to be negative. Taking again the argument above, we find a necessary restriction: $b_1 + b_2 = 0$. But $b_1 + b_2$ is the trace of the inertia tensor, and is also the moment of inertia I with respect to the center of mass. So we should have $I = 0$ for these configurations. Dizzy configurations give self-similar motions in the n-vortex problem, while central configurations give relative equilibria.

Expanded equations. The mutual distances tensor β has standard coordinates s_{ij}. By this assertion we mean simply that $\xi \circ \beta \circ \xi = -\sum_{i<j} s_{ij} \xi_i \xi_j$ for any $(\xi_1, \dots, \xi_n) \in \mathcal{D}^*$. To get the s_{ij} explicitly from β, we make first $e^*_{ij} = (0, \dots, 0, -1, 0, \dots, 0, 1, 0, \dots, 0) \in \mathcal{D}^*$, where the -1 is at the position i and the 1 at position j; then $s_{ij} = e^*_{ij} \circ \beta \circ e^*_{ij}$. It is a good opportunity to give the answer to an exercise left to the reader in Chapter 4. As $\beta = {}^t x \circ x$ and $x \circ e^*_{ij} = q_{ij}$, we get $s_{ij} = \langle q_{ij}, q_{ij} \rangle$.

To get expanded equations for ${}^t\check{\mathcal{Z}} \circ \beta + \beta \circ \check{\mathcal{Z}} = 0$, we could use representations of our tensors by $n \times n$ matrices. But it is faster to compute as above. The $n(n-1)/2$ standard coordinates of ${}^t\check{\mathcal{Z}} \circ \beta + \beta \circ \check{\mathcal{Z}}$ are the $e^*_{ij} \circ ({}^t\check{\mathcal{Z}} \circ \beta + \beta \circ \check{\mathcal{Z}}) \circ e^*_{ij} = 2 e^*_{ij} \circ \beta \circ \check{\mathcal{Z}} \circ e^*_{ij}$. We write

$$ e^*_{ij} \circ \beta \circ \check{\mathcal{Z}} \circ e^*_{ij} = (e^*_{ij} \circ {}^t x) \circ (x \circ \check{\mathcal{Z}} \circ e^*_{ij}) = -\langle q_{ij}, \ddot{q}_{ij} + \lambda q_{ij} \rangle $$

It happens that we already computed $\langle q_{ij}, \ddot{q}_{ij} \rangle$ in Chapter 1. We introduced

$$ \Sigma_{ij} = (m_i + m_j) S_{ij} + \frac{1}{2} \sum_{k \neq i,j} m_k (S_{ik} + S_{jk}). $$

and found

$$\langle q_{ij}, \ddot{q}_{ij} \rangle = -\Sigma_{ij} s_{ij} - \frac{1}{2} \sum_{k \neq i,j} m_k (S_{ik} - S_{jk})(s_{ik} - s_{kj}).$$

The expanded equations for dizzy configurations with multiplier $\lambda > 0$ are finally

$$0 = (\Sigma_{ij} - \lambda) s_{ij} + \frac{1}{2} \sum_{k \neq i,j} m_k (S_{ik} - S_{jk})(s_{ik} - s_{kj}).$$

These equations, associated to some equations due to Dziobek, were recently used by Hampton and Moeckel to solve an outstanding conjecture: there exists a finite number of central configurations of four particles with positive masses m_1, m_2, m_3 and m_4.

Some other configurations generalizing the central configurations. The balanced configurations generalize the central configurations. Let us shortly describe other interesting possibilities of generalization.

In a lecture at Herman's seminar in 1986, Yoccoz introduced what can be called the "minimal configurations". Given a configuration $x \in E \otimes \mathcal{D}$, we look for a "minimal configuration" in the class of x' with same right image (or coimage) $F \subset \mathcal{D}$ as x. They are the configurations of the form $R \circ x$, where R is some invertible linear transformation acting on E. The relative configurations corresponding to the class are the β's with image F. They form a connected open set in some linear subspace of $\mathcal{D} \vee \mathcal{D}$.

The function which Yoccoz minimized on such a class is $U + \lambda I/M$. As above

$$U = \sum_{i<j} m_i m_i s_{ij}^{-1/2}, \quad I = \frac{1}{M} \sum_{i<j} m_i m_j s_{ij}, \quad M = \sum_i m_i.$$

Using the property of convexity of the function U when expressed as the function of the mutual distances tensor, it is easy to prove that there is at most one critical point of $U + \lambda I/M$ in a class, that must be the minimum. Furthermore, *a central configuration with multiplier λ is a minimal configuration with multiplier λ.*

More recently Straume introduced some special configurations, defined as the configurations such that the Newtonian acceleration $\ddot{q}_i$ on the particle i is proportional to $q_i - q_G$, where q_G is the center of mass. In contrast with the central configurations, the proportionality factor may depend on the index i. The problem of finding the masses for which a given geometrical configuration is a Straume configuration is surprisingly interesting.

Acknowledgments to Martin Celli, Alain Chenciner, Richard Cushman, Jacques Féjoz, Yanning Fu, Yiming Long, Daniel Meyer, Ernesto Perez-Chavela and Shanzhong Sun.

References

[Alb] *On a paper of Moeckel on Central Configurations*, Regular and Chaotic Dynamics, 8 (2003) pp. 133–142

[AlC] A. Albouy, A. Chenciner, *Le problème des N corps et les distances mutuelles*, Inventiones Mathematicae 131 (1998) pp. 151–184

[AlL] A. Albouy, J. Llibre, *Spatial central configurations for the 1+4 body problem*, Proceedings of an International Conference dedicated to D. Saari, Contemporary Mathematics 292 (2002) pp. 1–16

[Ban] T. Banachiewitz, *Sur un cas particulier du problème des trois corps*, C. R. Acad. Sci. Paris 142 (1906) pp. 510–512

[Bet] E. Betti, *Sopra il moto di un sistema di un numero qualunque di punti che si attraggono o si respingono tra di loro*, Annali di Matematica s.2 t.8 (1877) pp. 301–311

[Bir] G.D. Birkhoff, *Dynamical Systems*, Providence, RI (1927), American Math. Soc. (1966)

[Bor] C.W. Borchardt, *Ueber die Aufgabe des Maximum, welche der Bestimmung des Tetraeders von grösstem Volumen bei gegebenem Flächeninhalt der Seitenflächen für mehr als drei Dimensionen entspricht*, Mathematische Abhandlungen der Akademie der Wissenschaften zu Berlin (1866) pp. 121–155, Werke pp. 201–232

[Car] C. Carathéodory, Sitzungsberichte der Bayerischen Akademie der Wissenschaften 1933 pp. 257–267

[Cay] A. Cayley, *On a theorem in the geometry of position*, œuvres v.1 (1841) pp. 1–4

[Cel] M. Celli, *Sur les distances mutuelles d'une choréographie à masses distinctes*, C. R. Acad. Sci. Paris, Ser. I 337 (2003) pp. 715–720

[Dar] G. Darboux, *Sur les relations entre les groupes de points, de cercles et de sphères dans le plan et dans l'espace*, Annales scientifiques de l'Ecole Normale Supérieure, s.2, t.1 (1872) pp. 323–392

[Dzi] O. Dziobek, *Ueber einen merkwürdigen Fall des Vielkörperproblems*, Astron. Nachr. 152 (1900) pp. 33–46

[Deh] R. Deheuvels, *Tenseurs et Spineurs*, Presses Universitaires de France (1993)

[Ham] M. Hampton, *Concave central configurations in the four-body problem*, thesis (2002)

[HaM] M. Hampton, R. Moeckel, *Finiteness of relative equilibria of the four-body problem*, (2004) preprint

[Jac] C.G.J. Jacobi, *Sur l'élimination des nœuds dans le problème des trois corps*, (1842), gesammelte Werke 4, Chelsea, pp. 297–314

[Lag] J.L. Lagrange, *Essai sur le problème des trois corps*, œuvres v.6 (1772) pp. 229–324

[La2] J.L. Lagrange, *Remarques générales sur le mouvement de plusieurs corps*, œuvres v.4 (1777) pp. 401–418

[LoS] Y. Long, S. Sun, *Four-Body Central Configurations with some Equal Masses*, Arch. Rational Mech. Anal. 162 (2002) pp. 25–44

[MaB] W.D. MacMillan, W. Bartky, *Permanent Configurations in the Problem of Four Bodies*, Trans. Amer. Math. Soc. 34 (1932) pp. 838–875

[MaW] J. Marsden, A. Weinstein, *Reduction of symplectic manifolds with symmetry*, Rep. Math. Phys. 5 (1974) pp. 121–130

[Mey] K. Meyer, *Symmetries and integrals in mechanics*, in: "Dynamical Systems", ed. M. Piexoto, Academic Press, New York (1973) pp. 259–272

[Moe] R. Moeckel, *Relative equilibria with clusters of small masses*, J. Dynam. Differential Equations 9 (1997) pp. 507–533

[Mun] Ch.H. Müntz, *Die Ähnlichkeitsbewegungen beim allgemeinen n-Körperproblem*, Math. Zeitschrift 15 (1922) pp. 169–187

[Mur] F.D. Murnaghan, *A symmetric reduction of the planar three-body problem*, Amer. J. Math. 58 (1936) pp. 829–832

[Piz] P. Pizzetti, *Casi particolari del problema dei tre corpi*, Rendiconti della Reale Accademia dei Lincei s.5 v.13 (1904) pp. 17–26

[San] A. Almeida Santos, *Dziobek's configurations in restricted problems and bifurcation*, to appear in Celestial Mechanics and Dynamical Astronomy

[Sc1] I.J. Schoenberg, *Remarks to Maurice Fréchet's article: Sur la définition d'une classe d'espaces vectoriels distanciés applicables vectoriellement sur l'espace de Hilbert*, Annals of Mathematics 36 (1935) pp. 724–732

[Sc2] I.J. Schoenberg, *Metric spaces and positive definite functions*, Transactions of AMS, 44 (1938) pp. 522–536

[Sch] J.A. Schouten, *Ricci-Calculus, An introduction to tensor analysis and its geometrical applications*, Second edition, Springer-Verlag, Berlin (1954)

[Sma] S. Smale, *Topology and mechanics II*, Invent. Math. 11 (1970) pp. 45–64

[Str] E. Straume, preprint (2003)

[Wey] H. Weyl, *The classical groups, Their invariants and representations*, Princeton University Press (1953) p. 30

[Win] A. Wintner, *The Analytical Foundations of Celestial Mechanics*, Princeton Math. Series 5, Princeton University Press, Princeton, NJ (1941)