A BRIEF ACCOUNT ON LAGRANGE'S ALGEBRAIC IDENTITY

MARIAN GIDEA AND CONSTANTIN P. NICULESCU

In *Indiscrete Thoughts* [18], G.-C. Rota remarked, "The mystery, as well as the glory of mathematics, lies not so much in the fact that abstract theories do turn out to be useful in solving problems, but, wonder of wonders, in the fact that a theory meant for one type of problem is often the only way of solving problems of entirely different kinds, problems for which the theory was not intended. These coincidences occur so frequently, that they must belong to the essence of mathematics." Indeed, it happens often that abstract mathematics leads to concrete applications, and real-life problems constitute a source of inspiration for sophisticated theories. The strong synergy between pure mathematics and its applications advocates for teaching methods that intertwine physical intuition with mathematical abstraction, and recognize the universality of mathematical laws throughout the sciences.

The identity

We aim to illustrate these ideas by surveying some encounters of the algebraic identity :

$$(L1)\qquad \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) = \left(\sum_{i=1}^{n} a_i b_i\right)^2 + \sum_{1\leq i<j\leq n} (a_i b_j - a_j b_i)^2,$$

for a_i, b_i, $i = 1, \ldots, n$, real or complex numbers.

Attributed to Joseph Louis Lagrange, with several fields of mathematics and mechanics, a special case of $(L1)$ is found in Fibonacci's *Book of Squares*(*Liber Quadratorum*, in the original Latin):

$$(F)\qquad \left(a_1^2 + a_2^2\right)\left(b_1^2 + b_2^2\right) = \left(a_1 b_1 + a_2 b_2\right)^2 + \left(a_1 b_2 - a_2 b_1\right)^2.$$

For integer values of the variables, this means that the product of sums of squares is again a sum of squares (see Book 13, Problem 19, in *Arithmetica* of Diophantus of Alexandria). Nowadays we can regard (F) as a consequence of complex number multiplication,

$$|a_1 + ia_2|^2 \, |b_1 + ib_2|^2 = |(a_1 + ia_2)(b_1 + ib_2)|^2 .$$

In 1773 Lagrange introduced the component form of both the dot and the cross product of vectors in $\mathbb{R}^3$ in order to study the geometry of tetrahedra and derived a special case of the identity $(L1)$:

$$\|u\| \, \|v\| = |\langle u, v\rangle|^2 + \|u \times v\|^2 \quad \text{for all } u, v \in \mathbb{R}^3.$$

See [13], page 663, lines 6-8. However Lagrange did not single out his finding and made no further comment on other, similar, results known to him.

The identity $(L1)$ appears later on in the famous *Cours d'Analyse* of Cauchy (without any mention of Lagrange: see [5], page 456, formula (31)). Cauchy used this identity to derive the inequality now bearing his name:

$$(C) \qquad \left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)$$

for every two families $a_1, ..., a_n$ and $b_1, ..., b_n$ of real numbers. Equality occurs if and only if the two families are proportional.

The beautiful book of J. Michael Steele [19] dedicated to the art of mathematical inequalities includes the above story (and many more), but leaves untouched a natural question arising from the discovery of Cauchy's inequality: Did Lagrange really know the identity $(L1)$ in that form?

A search of Lagrange's works sheds some light on this matter. In 1783, outlining an algorithm for finding the barycenters, Lagrange [14] proved an even more general identity (we call it Lagrange's second identity), as we shall see in the next section. An immediate consequence is the generalization of Cauchy's inequality in a vector-scalar melange, given by formula (sC) below. However we could not find any evidence that Lagrange made this last step to connect his findings. Thus it is very likely that the first appearance in print of the identity $(L1)$ is in Cauchy's book.

In *How to Solve It*, George Pólya [17] mentions that much can be gained by taking the time to reflect and look back at what we have done, what worked and what didn't. And last but not least, specialization is a valuable source of knowledge.

Lagrange's second identity

We may arrive at Lagrange's second identity through a problem of mass transport, a subject initiated in 1781 by Gaspard Monge [16], the inventor of descriptive geometry.

Suppose we have a number of sand piles located at $x_1, ..., x_n \in \mathbb{R}^2$. We seek a point x at which to collect all the sand at a minimal cost. Assuming the cost of transport of the unit mass is proportional to the square of the distance, the answer is provided by an old result about finite configurations of weighted points in the Euclidean space $\mathbb{R}^N$. (As usual, $\|\cdot\|$ denotes the natural norm, which is related to the inner product $\langle \cdot, \cdot \rangle$ by the formula $\|x\| = \sqrt{\langle x, x \rangle}$.)

Weighted least squares. *Given a family of points $x_1, ..., x_n$ in $\mathbb{R}^N$ and real weights $m_1, ..., m_n \in \mathbb{R}$ with $M = \sum_{k=1}^{n} m_k > 0$, then*

$$(WLS) \qquad \min_{x \in \mathbb{R}^N} \sum_{k=1}^{n} m_k \|x - x_k\|^2 = \frac{1}{M} \cdot \sum_{i<j} m_i m_j \|x_i - x_j\|^2.$$

The minimum is attained at one point,

$$x_G = \frac{1}{M} \sum_{k=1}^{n} m_k x_k.$$

In the above variational problem, the weights m_k do not have to be all positive. However, when all m_ks are positive, the point x_G represents the *barycenter* of the mass system $\{(m_1, x_1), \ldots, (m_n, x_n)\}$.

A special case of (WLS) is the result of Giulio Carlo Fagnano concerning the existence of a point P in the plane of a triangle ABC that minimizes the sum $PA^2 + PB^2 + PC^2$. (See [7], vol. 2.) It is worth noticing that his argument was based on calculus and can be adapted easily to cover the general case.

The solution to the weighted least squares problem was also known to Carl Friedrich Gauss, who is credited with developing the foundations of the least-squares analysis in 1795, at the age of eighteen. An early demonstration of the strength of Gauss's method came when it was used to predict the future location of the newly discovered asteroid Ceres; see [22] for a detailed account. Gauss did not publish the method until 1809, when it appeared in volume two of his work on celestial mechanics, *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium.*

We get more insight into the weighted least squares problem by considering its connection to certain polynomial identities that can be traced back to Christiaan Huygens, Gottfried Wilhelm Leibniz and Joseph Louis Lagrange. (The argument is essentially that of Lagrange; see [14], Theorem 2, page 539.)

Two polynomial identities. *For every family of points $x, x_1, ..., x_n$ in $\mathbb{R}^N$ and every family of real weights $m_1, ..., m_n$ with $M = \sum_{k=1}^n m_k \neq 0$, the following two identities hold:*

(H-Le) *the Huygens-Leibniz identity,*

$$\sum_{k=1}^{n} m_k \|x - x_k\|^2 = M \left\| x - \frac{1}{M} \sum_{k=1}^{n} m_k x_k \right\|^2 + \sum_{k=1}^{n} m_k \|x_k - \frac{1}{M} \sum_{j=1}^{n} m_j x_j\|^2;$$

($L2$) *Lagrange's second identity,*

$$\sum_{k=1}^{n} m_k \|x - x_k\|^2 = M \left\| x - \frac{1}{M} \sum_{k=1}^{n} m_k x_k \right\|^2 + \frac{1}{M} \cdot \sum_{1 \leq i < j \leq n} m_i m_j \|x_i - x_j\|^2.$$

In the language of mass transport, the Huygens-Leibniz identity says that the transport cost of the masses located at $x_1, \ldots, x_n$ to a point x equals the transport cost of the total mass of the system from the barycenter to x plus the transport cost of all masses located at $x_1, \ldots, x_n$ to the barycenter. Lagrange's second identity asserts that the same transport cost equals the transport cost of the total mass of the system from its barycenter to x plus a supplementary cost

$$\frac{1}{M} \cdot \sum_{1 \leq i < j \leq n} m_i m_j \|x_i - x_j\|^2,$$

due to the spreading of mass points around their barycenter. From Lagrange's second identity, one can immediately derive the formula (WLS).

Both identities (H-Le) and ($L2$) are invariant under translation (that is, under the change of variables $x \to x + z$ and $x_k \to x_k + z$ for $k = 1, ..., n$). The presence of this symmetry allows us to reduce ourselves to the case $\sum_{k=1}^n m_k x_k = 0$. Continuing the proof of the identity (H-Le),

$$\begin{aligned}
\sum_{k=1}^{n} m_k \|x - x_k\|^2 &= \sum_{k=1}^{n} m_k \langle x - x_k, x - x_k \rangle \\
&= \sum_{k=1}^{n} m_k \langle x, x \rangle - 2 \left\langle x, \sum_{k=1}^{n} m_k x_k \right\rangle + \sum_{k=1}^{n} m_k \langle x_k, x_k \rangle \\
&= M\|x\|^2 + \sum_{k=1}^{n} m_k \|x_k\|^2.
\end{aligned}$$

As for the identity $(L2)$, it suffices to note that

$$\begin{aligned}\sum_{1\leq i<j\leq n} m_i m_j \left\|x_i - x_j\right\|^2 &= \frac{1}{2}\cdot\sum_{i,j=1}^{n} m_i m_j \left\langle x_i - x_j, x_i - x_j\right\rangle \\ &= \frac{1}{2}\sum_{i,j=1}^{n} m_i m_j \left(\langle x_i, x_i\rangle - 2\langle x_i, x_j\rangle + \langle x_j, x_j\rangle\right) \\ &= M\sum_{i=1}^{n} m_i\langle x_i, x_i\rangle - \left\langle \sum_{i=1}^{n} m_i x_i, \sum_{j=1}^{n} m_j x_j \right\rangle \\ &= M\sum_{k=1}^{n} m_k\|x_k\|^2 \\ &= M\sum_{k=1}^{n} m_k\|x - x_k\|^2 - M^2\|x\|^2.\end{aligned}$$

This concludes the proof $(H\text{-}Le)$ and $(L2)$.

Substituting in $(L2)$,

$$x = 0,\ m_k = p_k a_k^2 \text{ and } x_k = y_k/a_k,$$

for $a_k \neq 0$ $(k = 1, ..., n)$, one obtains the following stronger version of the identity $(L1)$:

A vector-scalar melange of Lagrange's identity. *Given two families $p_1, ..., p_n$ and $a_1, ..., a_n$ of nonzero real numbers such that $\sum_{k=1}^n p_k a_k^2 \neq 0$, then*

$$(sL1)\quad \left(\sum_{k=1}^{n} p_k a_k^2\right)\left(\sum_{k=1}^{n} p_k\|y_k\|^2\right) - \left\|\sum_{k=1}^{n} p_k a_k y_k\right\|^2 = \sum_{1\leq i<j\leq n} p_i p_j\|a_j y_i - a_i y_j\|^2,$$

for every family $y_1, ..., y_n$ of vectors in $\mathbb{R}^N$.

In turn, $(sL1)$ easily yields Lagrange's second identity $(L2)$. Therefore these two identities are equivalent.

An immediate consequence of $(sL1)$ is the following extension of the Cauchy inequality: *Assuming* $\sum_{k=1}^n p_k a_k^2 > 0$,

$$(sC)\qquad \left\|\sum_{k=1}^{n} p_k a_k y_k\right\|^2 \leq \left(\sum_{k=1}^{n} p_k a_k^2\right)\left(\sum_{k=1}^{n} p_k\|y_k\|^2\right),$$

with equality when $a_j y_i = a_i y_j$ for all $i, j \in \{1, ..., n\}$.

Sometimes the identity (*H-Le*) is attributed solely to Leibniz, but we could not find any concrete evidence in his works. However, there is an indirect argument that Leibniz knew at least a particular case of it. In 1672, while in Paris on a diplomatic mission, Leibniz met Huygens and persuaded him to give him lessons in mathematics. No doubt Leibniz learned some facts about the moment of inertia, a concept Huygens used in his mathematical analysis of pendulums [8].

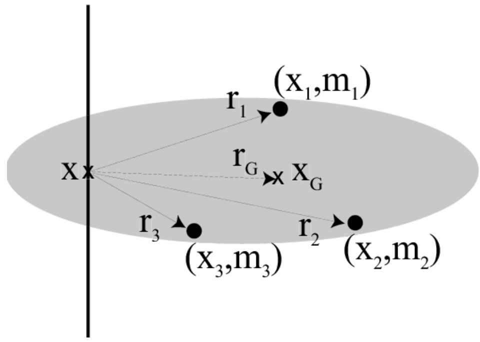

Figure 1. The Huygens-Steiner theorem

The *moment of inertia* of an object measures how easily the object can rotate about some specific axis. The moment of inertia of a system of mass points (x_k, m_k), with $k = 1, \dots, n$, about a given axis is, by definition, the scalar $I = \sum_{k=1}^{n} m_k r_k^2$, where r_k represents the perpendicular distance from x_k to the axis. Assume that we have a system of mass points lying on a light plate that rotates about a perpendicular axis which meets the plate at a point x. Let x_G be the position of the center of mass and r_G be the distance from x_G to x. The *Huygens-Steiner theorem* in mechanics (also known as the parallel-axes theorem) says that the moment of inertia $\sum_{k=1}^{n} m_k r_k^2$ about the axis through x equals the moment of inertia $(\sum_{k=1}^{n} m_k) r_G^2$ of the total mass of the system placed at x_G about the axis through x plus the moment of inertia of the system $\sum_{k=1}^{n} m_k (r_k - r_G)^2$ about a parallel axis through x_G. (See Figure 1.)

The Huygens-Steiner theorem implies that the period of a physical pendulum is the same for all locations of the axis equidistant from the center of mass. A proof can be found in [10], but this result is just a special case of the Huygens-Leibniz identity.

The identity (*H-Le*) has also a nice mechanical interpretation in terms of kinetic energy, a concept that was familiar to Leibniz. Indeed, he used the concept of *vis vita* (Latin for living force) for twice the modern kinetic energy. He realized that the total energy would be conserved in certain mechanical systems, and initiated a famous dispute in epoch concerning the "force" of a moving body. (See [15].)

For a system of n particles x_i of masses m_i and velocities v_i, the *total kinetic energy* is

$$K = \frac{1}{2} \sum_{i=1}^{n} m_i \|v_i\|^2.$$

Let v_G be the velocity of the barycenter of the system and $v_i' = v_i - v_G$ be the velocity of the particle of mass m_i relative to the barycenter. Applying Leibniz's identity with $(x, m) = (0, 0)$ we obtain

$$K = \frac{1}{2}\left(\sum_{i=1}^{n} m_i\right)\|v_G\|^2 + \frac{1}{2}\sum_{i=1}^{n} m_i\|v_i'\|^2.$$

The quantity $K' = (1/2)\sum_{i=1}^{n} m_i\|v_i'\|^2$ is called the rotational kinetic energy. If I represents the moment of inertia of the system about an axis of rotation passing through the center of mass, and ω represents the angular velocity about that axis, then $K' = (1/2)I\omega^2$. Therefore

$$K = \frac{1}{2}\left(\sum_{i=1}^{n} m_i\right)\|v_G\|^2 + \frac{1}{2}I\omega^2,$$

that is, the total kinetic energy equals the sum of the kinetic energy of the center of mass motion and of the rotational kinetic energy. This formula was established by Johann Samuel König in 1751. (See [12].)

In mechanics, there is another famous identity known as Lagrange's identity, relating the moment of inertia of a system of material points to its kinetic energy and potential energy. For a system of bodies in a homogeneous potential U of degree -1 (e.g., the Newtonian gravitational potential), that identity states that

$$\frac{d^2I}{dt^2} = 4K - 2U.$$

See [3, 11]. An immediate consequence of this second Lagrange identity is the *virial theorem*: if the time-average of d^2I/dt^2 is 0, then twice the time average of K equals the time average of U. This theorem is used by astronomers to make estimates on the total mass of galaxy clusters, as in this case d^2I/dt^2 has nearly zero average, and the velocities of the component galaxies can be measured directly.

The moment of inertia of a mass distribution about a given axis is analogous to the variance of a probability distribution. Indeed, if X is a discrete random variable taking the values $x_1, \ldots, x_n \in \mathbb{R}^N$ with probabilities $p_1, \ldots, p_n$, its variance $\operatorname{var}(X)$ is

$$\operatorname{var}(X) = \sum_{k=1}^{n} p_k \|x_k - E(X)\|^2,$$

where $E(X) = \sum_{k=1}^{n} p_k x_k$ represents the expectation value of X. Notice that $\sum_{k=1}^{n} p_k = 1$. The variance is a measure of how far a set of points are spread out from each other. Indeed, according to Theorem 1,

$$\operatorname{var}(X) = \sum_{1\le i<j\le n} p_i p_j \|x_i - x_j\|^2.$$

In probabilistic terms, Lagrange's identity relates the variance of a random variable X to the variance of a perturbation $X - x$ of it:

$$\operatorname{var}(X - x) = \|E(X) - x\|^2 + \operatorname{var}(X).$$

This can be rephrased using standard deviation,

$$\sigma(X) = \sqrt{\operatorname{var}(X)},$$

A useful property of standard deviation is that, unlike variance, it is expressed in the same units as the data.

A connection to metric geometry

Metric geometry exhibits many interesting formulas relating the side lengths a, b, c of a triangle ΔABC to the distances between different special points, such as the centroid G , the center of the circumscribed circle O, the center of the inscribed circle I, etc.

The following result (attributed to Leibniz) gives us a formula for the radius of circumscribed circle:

$$R^2 = OG^2 + \frac{1}{9}(a^2 + b^2 + c^2).$$

As a consequence,

$$a^2 + b^2 + c^2 \leq 9R^2,$$

with equality if (and only if) the triangle is equilateral. This follows easily from Lagrange's second identity ($L2$), when applied to the family of vertices of the triangle, with equal weights $m_1 = m_2 = m_3 = 1$, and for the choice of x as the center of the circumscribed circle.

If we consider the same family of points and the same choice of x, then weights

$$m_1 = a/(a+b+c),\ m_2 = b/(a+b+c) \text{ and } m_3 = c/(a+b+c),$$

leads us to the center I of the inscribed circle as a barycenter. This is a consequence of the theorem on the angle bisectors in a triangle. In this case the Huygens-Leibniz identity (H-Le) yields the equality

$$R^2 = OI^2 + 2Rr,$$

equivalently,

$$OI^2 = R(R - 2r),$$

a formula discovered independently by W. Chapple (1746) and L. Euler (1765). As usually, r denotes the radius of the inscribed circle. A consequence of this formula is the celebrated inequality

$$2r \leq R.$$

There are many interesting geometric consequences of the Huygens-Leibniz identity in higher dimensions too. For example, if R is the radius of the smallest ball containing a finite family of points $x_1, \ldots, x_n \in \mathbb{R}^N$, then

$$\frac{1}{n}\left(\sum_{i<j} \|x_i - x_j\|^2\right)^{1/2} \leq R.$$

How general is Lagrange's second identity?

Lagrange's second identity ($L2$) and the Huygens-Leibniz identity (H-Le) are equivalent and hold in any real vector space endowed with an inner product (or just with a semi-definite symmetric bilinear form). How do these identities depend on the metric on the underlying space?

The identity ($L2$) contains the parallelogram law as a particular case,

$$\|x_1\|^2 + \|x_2\|^2 = \frac{1}{2}\|x_1 - x_2\|^2 + \frac{1}{2}\|x_1 + x_2\|^2,$$

which corresponds to a configuration $\{(x_1, 1), (x_2, 1)\}$ of two points of equal weights, and to the choice of x as the origin. Consequently, each of the identities ($L2$) and (H-Le) can be viewed as higher dimensional generalizations of this identity. A classical result of P. Jordan and J. von Neumann ([6], pp. 151-152) asserts that the parallelogram law distinguishes the Euclidean norm from all other norms on a (finite dimensional) vector space. Consequently, if the identities $(L2)$ and $(H$-$Le)$ work in a normed vector space then necessarily that space is isometric to the Euclidean space of the same dimension. The same is true for the weighted least squares formula (WLS). In fact, if a norm on $\mathbb{R}^N$ satisfies the inequality

$$\|x_1\|^2 + \|x_2\|^2 \geq \frac{1}{2}\|x_1 - x_2\|^2 + \frac{1}{2}\|x_1 + x_2\|^2$$

for all $x_1, x_2 \in \mathbb{R}^N$, then that norm derives from an inner product (see [6], p. 152).

It is worth mentioning that Lagrange's second identity has an analogue in the space-time $\mathbb{R} \times \mathbb{R}^N$, endowed with the Minkowski inner product

$$\langle (t,x),(s,y)\rangle_M = -c^2 ts + \langle x, y\rangle,$$

where c is the speed of light. Indeed, by applying $(L2)$ to $a = (ict, x)$ and $b = (ics, y)$, we obtain:

$$\|(t,x)\|_M \|(s,y)\|_M - \langle (t,x),(s,y)\rangle_M^2 = -c^2 \sum_{i=1}^{N} (ty_i - sx_i)^2 + \sum_{1\leq i<j\leq N} (x_i y_j - x_j y_i)^2.$$

The case of spaces with a curved geometry

Since Lagrange's identity and the notion of a barycenter are closely tied to the metric, it is natural to wonder: what is the effect of a curved metric on this identity? And can one define a natural barycenter for a system of mass points in a curved space?

A remarkable class of curved metric spaces are spaces with global nonpositive curvature (global NPC spaces), which we discuss below. They have important applications to the study of groups from a geometrical viewpoint, and to certain rigidity phenomena in geometry. Informally, a global NPC space is characterized by the fact that its triangles are not "fatter" than the corresponding triangles in the Euclidean plane.

A starting point for the formal definition is the formula for the length of a median in a triangle in $\mathbb{R}^2$. For a triangle with vertices $x_0, x_1, z \in \mathbb{R}^2$, the length of the median from z is given by

$$\left\| z - \frac{x_0 + x_1}{2} \right\|^2 = \frac{1}{2}\|z - x_0\|^2 + \frac{1}{2}\|z - x_1\|^2 - \frac{1}{4}\|x_0 - x_1\|^2 .$$

This formula follows easily from the parallelogram law (and is actually equivalent to it).

A *global NPC space* is a complete metric space $E = (E, d)$ with the property that for each pair of points $x_0, x_1 \in E$ there exists a point $y \in E$ such that for all points $z \in E$,

$$(NPC) \qquad d^2(z,y) \leq \frac{1}{2}d^2(z,x_0) + \frac{1}{2}d^2(z,x_1) - \frac{1}{4}d^2(x_0,x_1).$$

In a global NPC space E, each pair of points x_0 and x_1 can be connected by a unique geodesic. The point y in (NPC) is the unique midpoint of the geodesic segment $[x_0, x_1]$.

Every Hilbert space is a global NPC space, and the midpoint of $[x_0, x_1]$ is given by the usual formula

$$y = \frac{x_0 + x_1}{2}.$$

The upper half-plane $\mathbf{H}$= $\{z \in \mathbb{C} : \operatorname{Im} z > 0\}$, endowed with the Poincaré metric,

$$ds^2 = \frac{dx^2 + dy^2}{y^2},$$

is another example of a global NPC space. In this case the geodesics are the semicircles in $\mathbf{H}$ perpendicular to the real axis and the straight vertical lines ending on the real axis. The Gauss curvature of $\mathbf{H}$ is -1.

A Riemannian manifold (M, g) is a global NPC space if and only if it is complete, simply connected and of nonpositive sectional curvature. Other important examples of global NPC spaces are the Bruhat-Tits buildings (in particular, Bruhat-Tits trees). See [2], [4], [9].

To measure the curvature of a global NPC space E, we compare triangles in E to triangles in the space $M^2_{-\kappa}$ of constant curvature $-\kappa$, defined as the upper-half plane $\mathbf{H}$ with the distance function scaled by a factor of $1/\sqrt{\kappa}$, $\kappa > 0$. Given a triangle Δ in E, a comparison triangle Δ' in $M^2_{-\kappa}$ is a geodesic triangle such that the lengths of the edges of Δ are equal to the lengths of corresponding edges of the triangle Δ'. Given a point p on an edge $[x, y]$ of Δ, a point p' on the corresponding edge $[x', y']$ is a comparison point if $d(x, p) = d(x', p')$. If we locally have that for all pairs of points p, q on an edge of Δ, their comparison points p', q' on Δ' satisfy $d(p, q) \leq d(p', q')$ we say that E has curvature $\leq -\kappa$. If instead we have $d(p, q) \geq d(p', q')$, then we say that E has curvature $\geq -\kappa$. Thus the concept of curvature of global NPC spaces is defined up to an inequality.

To define the barycenter of a mass point system we need a few preliminaries.

A subset $C \subset E$ is said to be convex if $\gamma([0, 1]) \subset C$ for each geodesic $\gamma : [0, 1] \to E$ joining two points in C. A function $f : C \to \mathbb{R}$ is called convex if the function $f \circ \gamma : [0, 1] \to \mathbb{R}$ is convex whenever $\gamma : [0, 1] \to C$, $\gamma(t) = \gamma_t$, is a geodesic, that is,

$$f(\gamma_t) \leq (1 - t)f(\gamma_0) + tf(\gamma_1)$$

for all $t \in [0, 1]$.

All closed convex subsets of a global NPC space are themselves global NPC spaces. The distance from a point z,

$$d_z(x) = d(x, z),$$

provides a basic example of a convex function. Moreover, its square is *uniformly convex* in the sense that

$$d^2(\gamma_t, z) \leq (1 - t)d^2(\gamma_0, z) + td^2(\gamma_1, z) - t(1 - t)d^2(x_0, x_1)$$

for all geodesics $\gamma : [0, 1] \to C$, $\gamma(t) = \gamma_t$, all points $z \in E$ and all numbers $t \in [0, 1]$. Technically this represents the extension of the inequality (NPC) from the case of midpoints to that of an arbitrary convex combinations. (See [20]). As a consequence, the balls in a global NPC space are convex sets (in the sense defined above).

The concept of a barycenter can now be naturally defined for any probability measure μ on E that admits finite moments of first order (i.e., $\int_E d(x, y)d\mu(y) < \infty$, for all $x \in E$). Think of μ as a mass distribution over the space. By analogy with the case of weighted least squares (WLS), the barycenter of μ is defined as the unique minimizer of a uniformly convex function, more precisely, of

$$F_y(z) = \int_E \left[d^2(z,x) - d^2(y,x)\right] d\mu(x).$$

This point is independent of the parameter $y \in E$ and is usually denoted b_μ.

If the support of μ is included in a convex closed set K, then $b_\mu \in K$.

This definition of a barycenter is justified by the fact it satisfies relation analogous to the Huygens-Leibniz identity (H-Le). The difference is that one only obtains an inequality relation, known as the *variance inequality*:

$$d^2(z,b_\mu) \leq \int_E \left[d^2(z,x) - d^2(b_\mu,x)\right] d\mu(x),$$

for all $z \in E$.

It is remarkable that the 'defect' of this relation from being an identity gives a measure of the curvature of the space. Assume that E is a global NPC space whose curvature is bounded from below by $-\kappa$. Then the following *reverse variance inequality* holds:

$$\int_E \left[d^2(z,x) - d^2(z,b_\mu) - d^2(b_\mu,x)\right] d\mu(x) \leq \frac{2\kappa}{3} \int_E \left[d^4(z,b_\mu) + d^4(b_\mu,x)\right] d\mu(x).$$

See [20, 21] for details.

It is worth noticing that defining the barycenter through a variational problem, as above, works satisfactorily outside the context of global NPC spaces as well, for example in the case of Wasserstein spaces (see [1]). These spaces provide a natural framework for the solution of the Monge-Kantorovich transport problem (see [23]).

It may seem quite surprising that the simple algebraic identity that we surveyed in this paper appears in such a multitude of forms and levels of abstraction, across the centuries, and in a wide range of areas of science. At a closer look, Lagrange's identity represents just a variant of the least action principle of classical mechanics. Thus, one possible explanation for the versatility of this algebraic identity is that it is deeply rooted in our physical reality. When one considers a more general model for the physical universe, such as a space-time continuum, or a curved space, the identity no longer survives, but is replaced by an inequality that seems to reflect in a precise way the geometric characteristics of the model.

Acknowledgments
The authors would like to thank the reviewers and the editor for many suggestions that improved the manuscript.

References

[1] M. Agueh and G. Carlier, Barycenters in the Wasserstein space, *SIAM J. Math. Anal.* **43** (2011), No. 2, 904-924.
[2] W. Ballmann, *Lectures on spaces with nonpositive curvature*, DMV Seminar Band **25**, Birkhäuser Verlag, Basel, 2005.
[3] G. D. Birkhoff, *Dynamical systems*, AMS, Chelsea, 1927.
[4] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature,* Grundlehren der Mathematischen Wissenschaften vol. **319**, Springer-Verlag, 1999.
[5] A.-L. Cauchy, *Cours d'Analyse de l'École Royale Polytechnique*, I ère partie, Analyse Algébrique, Paris, 1821. Reprinted by Ed. Jacques Gabay, Paris, 1989.
[6] M. M. Day, *Normed linear spaces*, 3rd Edition, Springer-Verlag, Berlin, 1973.

[7] Giulio Carlo di Fagnano, *Opere Matematiche.* 3 Vols. Pubblicate sotto gli Auspici della Societa Italiana per il Progresso delle Scienze. Per cura dei Professori Senatore Vito Volterra, Gino Loria, e Donisio Gambioli. Rome, 1912. These are based on his *Produzioni matematiche*, Pesaro, Stamperia 1750.

[8] C. Huygens, *Horologium oscillatorium sive de motu pendulorum ad horologia aptato demonstrationes geometricae*, Ed. F. Muguet, Paris, 1673.

[9] J. Jost, *Nonpositive curvature: geometric and analytic aspects*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1997.

[10] T. W. B. Kibble and F. H. Berkshire, *Classical Mechanics*, Addison Wesley Longman, Harlow, 4-th ed., 1996.

[11] V. V. Kozlov, Lagrange's identity and its generalizations, *Nelin. Dinam.*, **4** (2008), No. 2, 157–168.

[12] S. König (Sam. Kœnigio), De universali principio æquilibrii & motus, in vi viva reperto, deque nexu inter vim vivam & actionem, utriusque minimo, dissertatio, *Nova acta eruditorum* (1751), 125-135 and 162-176.

[13] J. L. Lagrange, Solutions analytiques de quelques problémes sur les pyramides triangulaires. Nouveaux Mémoirs de l'Académie Royale de Berlin, 1773; see *Oeuvres* de Lagrange, vol. 3, pp. 661-692, Gauthier-Villars, Paris, 1867.

[14] J. L. Lagrange: *Sur une nouvelle proprieté du centre de gravité*, Nouveaux Mémoirs de l'Académie Royale de Berlin, 1783; see *Oeuvres* de Lagrange, vol. 5, pp. 535-540, Gauthier-Villars, Paris, 1870.

[15] G. W. Leibniz, *Specimen Dynamicum*, 1695.

[16] G. Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année* (1781), pages 666-704.

[17] G. Pólya, *How To Solve It, A New Aspect of Mathematical Method*, 2nd ed., Princeton Univ. Press, 1973.

[18] G.-C. Rota, *Indiscrete Thoughts*, Birkhäuser, Boston, 1977.

[19] J. Michael Steele, *The Cauchy-Schwarz master class: an introduction to the art of mathematical inequalities*, Cambridge University Press, 2004.

[20] K. T. Sturm, Probability measures on metric spaces of nonpositive curvature. In vol.: *Heat kernels and analysis on manifolds, graphs, and metric spaces* (Pascal Auscher et al. editors). *Lecture notes from a quarter program on heat kernels, random walks, and analysis on manifolds and graphs*, April 16–July 13, 2002, Paris, France. Contemp. Math. **338** (2003), 357-390.

[21] K. T. Sturm, Coupling, regularity and curvature. In vol.: *Interacting Stochastic Systems* (J.-D. Deuschel and A. Greven editors), Springer 2004.

[22] D. Teets and K. Whitehead, The Discovery of Ceres: How Gauss Became Famous, *Mathematics Magazine* **72** (1999), 83-91.

[23] C. Villani, *Optimal transport: old and new*, Grundlehren der mathematischen Wissenschaften vol. 338, Springer-Verlag, 2009.

Northeastern Illinois University, Chicago, IL 60625, and Institute for Advanced Study, Princeton, NJ 08540, USA
E-mail address: mgidea@neiu.edu

University of Craiova, Department of Mathematics, Craiova 200585, Romania
E-mail address: cniculescu47@yahoo.com