ASTROGEOMETRY: GEOMETRY EXPLAINS SHAPES OF CELESTIAL BODIES

by Andrei Finkelstein[a], Olga Kosheleva[b], and Vladik Kreinovich[c,1]

[a]*Institute of Applied Astronomy, Russian Academy of Sciences, 8 Zhdanovskaya St. 197042 St. Petersburg, Russia*

[b]*Department of Electrical and Computer Engineering and*

[c]*Department of Computer Science*

University of Texas at El Paso, El Paso, TX 79968

emails `olga@ece.utep.edu` *and* `vladik@cs.utep.edu`

Abstract. Celestial bodies such as galaxies, stellar clusters, planetary systems, etc., have different geometric shapes (e.g., galaxies can be spiral or circular, etc.). Usually, complicated physical theories are used to explain these shapes; for example, several dozen different theories explain why many galaxies are of spiral shape; see, e.g., (Toomre 1973, Strom 1979, Vorontsov-Veliaminov 1987, Binney 1989). Some rare shapes are still unexplained.

In the present paper, we show that to explain these "astroshapes", we do not need to know the details of *physical* equations: practically all the shapes of celestial bodies can be explained by simple *geometric* invariance properties. This fact explains, e.g., why so many different physical theories lead to the same spiral galaxy shape.

1. OUR MAIN IDEA: ASTROSHAPES AS GEOMETRIC OBJECTS

1.1. Physical background

To find out how shapes have been formed, let us start from the beginning of the Universe (for a detailed physical description, see,

[1] This work was partially supported by NSF grant No. EEC-9322370, and by NASA Research Grant No. 9-757. The authors want to thank A. D. Alexandrov, Yu. F. Borisov, Yu. D. Burago, V. K. Ionin, I. A. Kunin, A. V. Kuz'minyh, and A. Neumaier for valuable discussions.

e.g., (Zeldovich Novikov 1983)). The only evidence about the earliest stages of the Universe is the cosmic 3K background radiation. This radiation is highly homogeneous and isotropic; this means that initially, the distribution of matter in the Universe was highly homogeneous and isotropic. In mathematical terms, the initial distribution of matter was invariant w.r.t. arbitrary shifts and rotations.

We can also say that the *initial distribution was invariant* w.r.t. dilations if in addition to dilation in space (i.e., to changing the units of length), we accordingly change the units of mass. In the following text, we will denote the corresponding transformation group (generated by arbitrary shifts $x \to x + a$, rotations, and dilation $x \to \lambda \cdot x$) by G.

On the astronomical scale, of all fundamental forces (strong, weak, etc.) only two forces are non-negligible: gravity and electromagnetism. The equations that describe these two forces are invariant w.r.t. arbitrary shifts, rotations, and dilations in space. In other words, these interactions are *invariant* w.r.t. our group G. The *initial distribution* was *invariant* w.r.t. G; the *evolution equations are* also *invariant*; hence, we will get $G-$invariant distribution of matter for all moments of time. But our world is not homogeneous. Why?

The reason why do not see this homogeneous distribution is that this highly symmetric distribution is known to be *unstable*: If, due to a small perturbation, at some point a in space, density becomes higher than in the neighboring points, then this point a will start attracting matter from other points. As a result, its density will increase even more, while the density of the surrounding areas will decrease. So, arbitrarily small perturbations cause drastic changes in the matter distribution: matter concentrates in some areas, and shapes are formed. In physics, such symmetry violation is called *spontaneous*.

In principle, it is possible to have a perturbation that changes the initial highly symmetric state into a state with no symmetries at all, but statistical physics teaches us that it is much more probable to have a gradual symmetry violation: first, some of the symmetries are violated, while some still remain; then, some other symmetries are violated, etc.[2] At the end, we get the only stable shape: rotating ellipsoid.

Before we reach the ultimate ellipsoid stage, perturbations are invariant w.r.t. some subgroup G' of the initial group G. If a certain perturbation concentrates matter, among other points, at some point a, then, due to invariance, for every transformation $g \in G'$, we will observe a similar concentration at the point $g(a)$. Therefore, the shape of the resulting concentration contains, with every point a, the entire *orbit* $G'a = \{g(a) \mid g \in G'\}$ of the group G'. Hence, the resulting *shape consists of* one or several *orbits* of a group G'.

1.2. The result of physical analysis: description of astroshapes reformulated as a geometric problem

So, to describe all possible *shapes* of celestial bodies, it is sufficient to describe all possible *orbits* of subgroups G' of the group G (= all shifts, rotations, and dilations). In this paper, we will show that this description really describes all known astroshapes. Some preliminary results of this description were first announced in (Kreinovich 1981), (Kosheleva *et al.* 1982), and (Kosheleva Kreinovich 1989).

1.3. A word of warning: geometric shapes are only approximate

Objects of nature can only *approximately* be described by geometric figures. Correspondingly, in our physical explanation, perturbations are only *approximately* invariant w.r.t. G'. The farther away from the point a, the less similar is the point $g(a)$ to the point a. Therefore, in reality, we may observe not the *entire* orbit, but only a *part* of it.

---

[2] Similarly, a (highly organized) solid body normally goes through a (somewhat organized) liquid phase before it reaches a (completely disorganized) gas phase.

2. AS A RESULT OF APPLYING OUR MAIN IDEA, WE GET EXACTLY ALL OBSERVABLE ASTROSHAPES

2.1. Possible orbits

0−, 1−, and 2−dimensional orbits of continuous subgroups G' of the group G are easy to describe (see Appendix):

0: The only 0−dimensional orbit is a *point*.

1: A generic 1−dimensional orbit is a *conic spiral* that is described (in cylindrical coordinates) by the equations $z = k\rho$ and $\rho = R_0 \exp(c\varphi)$. Its limit cases are:

- a *logarithmic* (Archimedean) *spiral*: a planar curve ($z = 0$) that is described (in polar coordinates) by the equation $\rho = R_0 \exp(c\varphi)$.
- a *cylindrical spiral*, that is described (in appropriate coordinates) by the equations $z = k\phi$, $\rho = R_0$.
- a *circle* ($z = 0$, $\rho = R_0$);
- a *semi-line* (*ray*);
- a *straight line*.

2: Possible 2-D orbits include:

- a *plane*;
- a *semi-plane*;
- a *sphere*;
- a *semi-plane*;
- a *circular cylinder*, and
- a *logarithmic cylinder*, i.e., a cylinder based on a logarithmic spiral.

2.2. Possible orbits are exactly possible shapes

Comparing these orbits (and ellipsoids, the ultimate stable shapes) with astroshapes enumerated in (Vorontsov-Veliaminov 1987), we conclude that:

- First, our scheme describes all observed connected shapes.
- Second, all above orbits, except the logarithmic cylinder, have actually been observed as shapes of celestial bodies.

For example, according to Chapter III of (Vorontsov-Veliaminov 1987), galaxies consist of components of the following geometric shapes:

- *bars* (cylinders);
- *disks* (parts of the plane);
- *rings* (circles);
- *arcs* (parts of circles and lines);
- *radial rays*;
- *logarithmic spirals*;
- *spheres*, and
- *ellipsoids*.

It is easy to explain why logarithmic cylinder was never observed: from whatever point we view it, the logarithmic cylinder blocks all the sky, so it does not lead to any visible shape in the sky at all. With this explanation, we can conclude that we have a *perfect explanation of all observed astroshapes.*

2.3. Comment: we can also explain difficult-to-explain disconnected shapes

In the above description, we only considered connected *continuous* subgroups $G' \subseteq G$. Connected continuous subgroups explain connected shapes.

It is natural to consider disconnected (in particular, discrete) subgroups as well; the orbits of these subgroups leads to disconnected shapes. Thus, we can explain these shapes, most of which modern astrophysics finds pathological and difficult to explain (see, e.g., (Vorontsov-Veliaminov 1987, Section I.3). For example, an orbit O of a discrete subgroup G'' of the 1-D group G' (whose orbit is a logarithmic spiral) consists of points whose distances r_n to the center forms a geometric progression: $r_n = r_0 \cdot k^n$. Such dependence (called Titzius-Bode law) has indeed been observed (as early as the 18th century) for planets of the Solar system and for the satellites of the planets (this law actually led to the prediction and discovery of what is now called asteroids). Thus, we get a *purely geometric explanation of the Titzius-Bode law.*

Less known examples of disconnected shapes that can be explained in this manner include:

- several parallel equidistant lines (Vorontsov-Veliaminov 1987, Section I.3);
- several circles located on the same cone, whose distances from the cone's vertex form a geometric progression (Vorontsov-Veliaminov 1987, Section III.9);
- equidistant points on a straight line (Vorontsov-Veliaminov 1987, Sections VII.3 and IX.3);
- "piecewise circles": equidistant points on a circle; an example is MCG 0-9-15 (Vorontsov-Veliaminov 1987, Section VII.3);
- "piecewise spirals": points on a logarithmic spiral whose distances from a center form a geometric progression; some galaxies of Sc type are like that (Vorontsov-Veliaminov 1987).

3. THIS IDEA ALSO EXPLAINS: EVOLUTION OF GEOMETRIC SHAPES, THEIR RELATIVE FREQUENCY, DIRECTIONS OF ROTATION AND OF MAGNETIC FIELD

Our main idea explains not only the shapes themselves, but also how they evolve, which are more frequent, etc.

3.1. Evolution of geometric shapes

Idea. In our description, we start with a homogeneous isotropic matter distribution. Then, spontaneous symmetry violation occurs, that eventually leads to geometric shapes. At first, the distinction between the points with perturbations and without them is small, so instead of seeing the geometric shape, we see an irregularly shaped object, but in course of time, a regular shape appears.

Each shape corresponds to a subgroup G' of the original group G. In course of time, further spontaneous symmetry violation can occur; the group G' will be replaced by its subgroup $G'' \subseteq G'$, and the original shape will evolve into a shape of the orbit of this subgroup G''. At the end, when no symmetries are left, we get a rotation ellipsoid.

Results. This idea leads to the *evolution tree* of different shapes. The resulting evolution tree is in good accordance with evolution trees presented in (Vorontsov-Veliaminov 1987) for different celestial objects.

Last stage of the geometric evolution (formation of an ellipsoid): where? At the last stage of our evolution, an ellipsoid starts to grow, an ellipsoid that (later on) will contain all the matter of this astronomical object. Where does this growth start? If an ellipsoid starts to grow at a certain point, then the resulting shape has, in general, fewer symmetries than before: because only transformations that keep this new point intact are symmetries of the new shape.

In line with our main idea, it is natural to assume that in most cases, the ellipsoid starts to grow at a point where this restriction of the symmetry groups is the least restrictive, i.e., where the resulting restricted symmetry group is the largest. The resulting locations are in perfect accordance with astronomical observations (Vorontsov-Veliaminov 1987):

0: For 0-D shapes:
- for two-point shapes (quasars, radiogalaxies): in the midpoint between these two points;
- for "piecewise spirals": in the vertex of the spiral (star for Titzius-Bode planets, central ellipsoid cluster for Sc galaxies, etc.)

1: For 1-D shapes:
- for a linear shape, in the middle of the line segment;
- for a radial ray, at the vertex (Section III.8);
- for a circle, in the middle (Section VI.2);
- for a spiral, in the vertex (usual type S spiral galaxies);
- for a conic spiral: at the vertex of the cone (Section III.9);

2: For 2-D shapes:
- for a disk, in the middle (example: our Galaxy);
- for a sphere, in the middle (this middle ellipsoid is called a *nucleus*);
- for cones, at the vertex (e.g., for comets);

- for a *bar* (cylinder): in the middle of the axis (Section I.3).

How evolution depends on size. The smaller the object, the shorter time it takes for each signal to go from one side to the other. Therefore, the smaller the object, the faster are the global processes that change its shape. Hence, we observe larger objects mainly in their early evolution stages, while smaller objects are mainly in their later stages.

Actual observations confirm this geometric conclusion:

- *Galaxies* are mainly observed in their intermediate stages (very frequently, of spiral shape), while those *stellar clusters* that have a definite shape are mainly ellipsoidal in shape. In our *Solar system*, practically all the mass is concentrated in the central ellipsoid (the Sun).
- In *galaxies*, 0-D forms (that correspond to the latest stages of evolution) are rare, mainly 2-D and 1-D forms are observed (that correspond to the earlier stages). In the *Solar system*, vice versa, there are many 0-D systems and few 1-D systems (rings).
- In the *Solar system*:
  - *larger planets* have rings (1-D orbits);
  - *smaller planets* have 0-D structures (satellites that follow Titzius-Bode law);
  - *the smallest planets* (Mercury and Venus) have no satellites at all (i.e., are in the final stage of ellipsoid shape).

Where are the young ones? Objects of earlier type (with more symmetries) are younger and must, tehrefore, contain mostly younger population. This conclusion is confirmed by astronomical data (Vorontsov-Veliaminov 1987):

- *Clusters of galaxies and galaxies:*
  - *irregular* (= very young) clusters consist mainly of *younger* galaxies (irregular and spiral); there are also some *elliptical* galaxies, but only very *small* ones ("dwarf galaxies"), that, due to their size, had time to evolve into the final stable state.

    - elliptical clusters (that correspond to the *last* stage of evolution) contain mainly elliptical ("*old*") galaxies.
- *Galaxies and stellar clusters:* in our Galaxy:
    - irregular (*youngest*) stellar clusters are located mostly near the disk (planar component, of *earlier* evolution type), while
    - ellipsoidal stellar clusters, that correspond to the *latest* evolution stages, are located mostly in the so-called "spherical subsystem" (ellipsoidal component, of the *latest* evolution type).
- *Galaxies and stars:* the spectral type of a galaxy (determined by its most frequent stars) is:
    - types A–F (corresponding to the *youngest* stars) in irregular (*youngest*) galaxies;
    - types F–G (*older* stars) in spiral galaxies (*next evolution step*); and
    - types G–K (*old* stars) in elliptical galaxies, that represent the *latest evolution stage*.

  In particular, in our Galaxy:
    - *younger* stars are located mainly in the *younger* (disk and spiral) components, while
    - *older* stars are located mainly in the *ellipsoidal* components (spherical subsystem and nucleus) that correspond to the *last* evolution stage.
- *Stellar clusters and stars:*
    - irregular (*younger*) clusters mostly contain *younger* stars, while
    - elliptical (*older*) clusters mostly contain *older* stars.

Where the "wild" particles are? From the physical viewpoint, the evolution of the shape means getting closer to the statistical (thermodynamic) equilibrium. In the equilibrium, particles are distributed according to the well-known Maxwell's distribution, according to which most of the particles are of approximately the same velocity, and particles that are much faster than the average are extremely rare. In non-equilibrium states, there usually are more fast particles. Therefore, the earlier (more symmetric) shapes must have more fast particles. In particular, in a multi-component system, fast particles are mainly located in the earlier-stage (more symmetric) components.

This conclusion is in agreement with the fact that in our Galaxy, fast (relativistic) particles are mainly located in the disk and in the spirals, while their density in the ellipsoid components (in nucleus and in the spherical subsystem) is very small.

3.2. Relative frequency of different shapes

Main idea in physical terms. We have already mentioned that gradual symmetry violations are more frequent than abrupt ones. Hence, the more symmetries are violated, the less frequent is the corresponding perturbation.

Main idea in geometric terms. In geometric terms, if we start with the shape that is an orbit of a group G', then the most probable evolution results are the shapes that correspond to subgroups $G'' \subseteq G'$ of the highest possible dimension.

Results. Initially, we have a group G. Out of all above-described shapes, the *plane* has the symmetry group G' of the largest possible dimension: $\dim G' = 4$ (two shifts, rotation, and dilation). Hence, out of the first shapes that emerged in the Universe the most frequent ones were planes. This conclusion is in good accordance with modern astrophysics, according to which matter first clustered into "disks" from which galaxies later evolved.

Next symmetry violation leads to a planar form. Of all planar forms, the generic is the *logarithmic spiral*, and all others can be viewed as its limit (degenerate) cases. So, we can conclude that the most frequent Galaxy shape is a *logarithmic spiral*. If we do not count ellipsoids (ultimate stable forms), then spiral galaxies are indeed the most frequent ones.

So, we got an explanation of the galaxies' spiral shape without using any detailed physics. This result explains why so many different physical theories all explain the same shape: because you will get these shapes from *any* physical theory (as long as this theory has the simple geometric symmetries necessary for our geometric explanation[3].

From spirals (orbits of a 1-D symmetry group), next symmetry violation leads to a *discrete* subgroup, whose orbits describe Titzius-Bode law. Thus, we explain why this law is the most frequent organization of planetary and satellite systems.

3.3. In what directions do astroshapes rotate?

Astroshapes rotate. It is well known that the angular momentum of a body is invariant. A random cloud of particles moving in different directions has, in general, a non-zero angular momentum, so, except for the rare case when this momentum happens to be exactly 0, the shaped celestial body that emerges from these particles has the same angular momentum and therefore, rotates. Most of the celestial bodies (planets, stars, clusters, galaxies) do indeed rotate. What are the possible directions of this rotation?

Idea. From the geometric viewpoint, adding rotation means that we restrict ourselves only to symmetries from G that leave the rotation axis invariant. In view of our main idea, it is natural to assume that the most probable location of the rotation axis is where the corresponding restriction of the symmetry group is the least restrictive.

---

[3] Arnold has shown (see, e.g., (Thom 1975, Arnold 1978)) that dynamical systems theory explains why the observed shape should be *topological homeomorphic* to a spiral. We have explained even more: not only that this shape is *homeomorphic* to the spiral, but that geometricly, this shape is *exactly* a *logarithmic spiral.*

Results. The results are in perfect accordance with the observations (Vorontsov-Veliaminov 1987):

0: 0-D shapes:
- piecewise spiral rotates around the central body (example: Solar system);

1: 1-D shapes:
- linear galaxy rotates around its axis (Section IX.3);
- spiral – around its vertex (e.g., for S type galaxies);
- conic spiral – around the axis of the cone (Section III.3); in (Vorontsov-Veliaminov 1987) this direction of rotation is described as puzzling and difficult to explain;

2: 2-D shapes:
- disk – around its center;
- sphere – around a line passing through its center.

3.4. Magnetic fields: directions explained by geometry

The magnetic field vector $\vec{B}$ must satisfy the equation $\mathrm{div}\vec{B} = 0$ (one of Maxwell's equations).

Idea. If we have a shape that has a symmetry group G', and we add a magnetic field to the shape, we thus restrict ourselves only to symmetries from G' that preserve this field. The direction of the magnetic field is most probably chosen in such a way that the corresponding restriction on G' is the least restrictive.

Unlike rotation, magnetic field is not necessitated by any conservation law. So, if for some shape, there is no magnetic field whose addition can preserve its symmetry, then most probably, no magnetic fields appear in this celestial body.

Results. The directions of magnetic fields of the astrosources can be determined from the polarization of their radiation. The resulting magnetic fields (Vorontsov-Veliaminov 1987) are in accordance with this idea:

- A *point* is invariant w.r.t. arbitrary rotation. No vector can be invariant w.r.t. all possible rotations, so, most probably, no magnetic field will appear in point-shaped objects. This conclusion is consistent with the fact that the synchrotron radiation of tri-component radio galaxies (radiation that is caused by the motion of charged relativistic particles in a strong magnetic field) is mainly coming not from the central point source, but from the double endpoint source.

- For 1-D shapes, the demands that the vector $\vec{B}$ is located on these clusters, and that div$\vec{B}$ imply that the vector $\vec{B}$ is tangent to the component.
- For 2-D shapes (e.g., for a disk) there is no invariant magnetic field, so, in most disk celestial bodies, there is no magnetic field. This conclusion is consistent with our Galaxy, in which the main magnetic field is located in spirals, and there is practically no magnetic field in the disk.

4. OPEN PROBLEMS

In this paper, we explained the *basic* shapes of celestial objects, their evolution, etc., in simply geometric terms. In addition to the *basic* geometric facts, there are also lots of *minor* geometric details. The success of our explanations make it plausible that the details of geometric shapes can also be explained in this manner.

REFERENCES

Arnold, V. I., *Mathematical methods of classical mechanics*, N.Y., Springer, 1978.

Binney, J., "Stellar dynamics", in: I. Appenzeller, H. J. Habing, and P. Léna (eds.), *Evolution of galaxies: astronomical observations*, Springer Lecture Notes in Physics, Vol. 333, Berlin, Heidelberg, 1989, pp. 95–146.

Kosheleva, O. M., and Kreinovich, V., *Astrogeometry, or geometrical investigation of forms of celestial bodies*, Technical Report, Center for New Information Technology "Informatika", Leningrad, 1989.

Kosheleva, O. M., Kreinovich, V., and Finkelstein, A. M., "Group-theoretic approach to foundations of space-time theory," *Proceedings of the Symposium on Global Geometry and Foundations of Relativity*, Novosibirsk, 1982, pp. 76–78 (in Russian).

Kreinovich, V., Referee's comments in a review of V. A. Dubrovin, S. P. Novikov, and A. T. Fomenko, *Modern Geometry*, Moscow, Nauka, 1980, *Zentralblatt fur Mathematik*, 1981, Vol. 433, pp. 295–297.

Strom, S. E., and Strom, K. M., "The evolution of disk galaxies", *Scientific American*, April 1979; reprinted in P. W. Hodge (ed.), *The Universe of galaxies*, Freeman and Co., N.Y., 1984, pp. 44–54.

Toomre, A., and Toomre, J., "Violent tides between galaxies", *Scientific American*, December 1973; reprinted in P. W. Hodge (ed.), *The Universe of galaxies*, Freeman and Co., N.Y., 1984, pp. 55–65.

Thom, R., *Structural stability and morphogenesis*, Benjamin Cummings, Reading, MA, 1975.

Vorontsov-Veliaminov, B. A., *Extragalactic astronomy*, Harwood Academic Publishers, Chur, Switzerland, London, 1987.

Zeldovich, Ya. B., and Novikov, I. D., *Relativistic Astrophysics. Part 2. The structure and evolution of the Universe*, The University of Chicago Press, Chicago and London, 1983.

APPENDIX: HOW TO DESCRIBE ORBITS OF SUBGROUPS OF G?

A 1-D orbit is an orbit of a 1-D subgroup. This subgroup is uniquely determined by its "infinitesimal" element, i.e., by the corresponding element of the Lie algebra of the group G. This Lie algebra if easy to describe. For each of its elements, the corresponding differential equation (that describes the orbit) is reasonably easy to solve.

2-D forms are orbits of $\geq$ 2-D subgroups, so, they can be enumerated by combining two 1-D subgroups.

An alternative (slightly more geometric) way of describing 1-D orbits is to take into consideration that an orbit, just like any other curve in a 3-D space, is uniquely determined by its curvature $\kappa_1(s)$ and torsion $\kappa_2(s)$, where s is the arc length measured from some fixed point. The fact that this curve is an orbit of a 1-D group means that for every two points x and x' on this curve, there exists a transformation $g \in G$ that maps x into x'. Shifts and rotations do not change κ_i, they may only shift s (to $s + s_0$); dilations also change s to $s \to \lambda \cdot s$ and change the numerical values of κ_i. So, for every s, there exist $\lambda(s)$ and $s_0(s)$ such that the corresponding transformation turns a point corresponding. to $s = 0$ into a point corresponding to s. As a result, we get functional equations that combine the two functions $\kappa_i(s)$ and these two functions $\lambda(s)$ and $s_0(s)$. Taking an infinitesimal value s in these functional equations, we get differential equations, whose solution leads to the desired 1-D orbits.