4 June 1971, Volume 172, Number 3987 SCIENCE

Plasma Physics, Space Research, and the Origin of the Solar System

Hannes Alfvén

The center of gravity of the physical sciences is always moving. Every new discovery displaces the interest and the emphasis. Equally important is that new technological developments open new fields for scientific investigation. To a considerable extent the way science takes depends on the construction of new instruments, as is evident from the history of science. For example, after the development of classical mechanics and electromagnetism during the 19th century, a new era was started by the construction of highly developed spectrographs in the beginning of this century. For its time, these were very complicated and expensive instruments. They made possible the exploration of the outer regions of the atom. Similarly in the thirties the cyclotron—for its time, a very complicated and expensive instrument—was of major importance in the exploration of the nucleus. Finally, the last decade has witnessed the construction of still more complicated and expensive instruments, the space vehicles, which are launched by a highly developed rocket technology and instrumented with the most sophisticated electronic devices. We may then ask the question: What new fields of research—if any—do these open for scientific investigation? Is it true, also in this case, that the center of gravity of physics moves with the big instruments?

Scientific Aims of Space Research

The first decade of space research mainly concentrated on the exploration of space near the earth: the magnetosphere and interplanetary space. These regions earlier were supposed to be void and structureless, but we now know that they are filled with plasmas, intersected by sheathlike discontinuities, and permeated by a complicated pattern of electric currents and electric and magnetic fields. The knowledge gained in this way is fundamental to our general understanding of plasmas, especially cosmic plasmas. Indirectly it will hence be important to thermonuclear research, to the study of the structure of the galaxy and the metagalaxy, and to cosmological problems. Our advancing knowledge in cosmical electrodynamics will make it possible to approach these fields in a less speculative way than hitherto. The knowledge of plasmas is also fundamental to our understanding of the origin and evolution of the solar system because there are good reasons to believe that the matter which now forms the celestial bodies once was dispersed in a plasma state.

The second decade of space research seems to display a different character, at least to some extent. As several of the basic problems of the magnetosphere and interplanetary space are still unsolved, one can be sure that these regions will still command much interest. However, the lunar landings and also the deep-space probes to Venus and Mars have supplied us with so many new scientific facts that the emphasis in space research is moving toward the exploration of the moon, the planets, and other celestial bodies in the solar system.

The first phase of this exploration is necessarily of a character somewhat similar to the exploration of the polar regions and other regions of the earth which have been difficult to reach: a detailed mapping-out combined with geological, seismic, magnetic, and gravity surveys and an exploration of the atmospheric conditions. However, when applying this research pattern to the moon and the planets, one is confronted with another problem, namely, how these bodies were originally formed. In fact many of the recent space research reports end with speculations about the formation and evolution of the solar system. It seems that this will necessarily be one of the main problems—perhaps the main problem—on which space research will center in the near future. Already at an early date NASA stated that the main scientific goal of space research should be to clarify how the solar system was formed. This is indeed one of the fundamental problems of science. We are trying to write the scientific version of how our earth and its neighbors once were created. From a, shall we say, philosophical point of view, this is just as important as the structure of matter, which has absorbed most of the interest during the first two thirds of this century.

Plasma Physics and Its Applications

Before we concentrate on our main topic: how the solar system originated, we should make a brief summary of the state of plasma physics. As you know, plasma physics has started along two parallel lines. The first one was the hundred-years-old investigations in what was called electrical discharges in gases. This approach was to a high degree experimental and phenomenological, and only very slowly reached some degree of theoretical sophistication. Most theoretical physicists looked down on this field, which was complicated and awk-

The author is professor of applied electrophysics at the University of California, San Diego, La Jolla. This article is the lecture he delivered in Stockholm, Sweden, on 11 December 1970 when he received the Nobel Prize in Physics, a prize he shared with Professor Lois Néel. Minor corrections and additions have been made by the author. It is published here with the permission of the Nobel Foundation and will also be included in the complete volume of *Les Prix Nobel en 1970* as well as in the series Nobel Lectures (in English) published by the Elsevier Publishing Company, Amsterdam and New York. Dr. Néel's lecture will be published in a subsequent issue.

Table 1. Cosmical electrodynamics.

First approach	Second approach
Homogeneous models	Space plasmas have often a complicated inhomogeneous structure
Conductivity $\sigma = \infty$	σ depends on current and often suddenly becomes 0, $E_{\parallel}$ often $\neq 0$
Electric field $E_{\parallel} = 0$	
Magnetic field lines are "frozen in" and "move" with the plasma	Frozen-in picture often completely misleading
Electromagnetic conditions illustrated by magnetic field line picture	It is equally important to draw the current lines and discuss the electric circuit
Electrostatic double layers neglected	Electrostatic double layers are of decisive importance in low-density plasmas
Filamentary structures and current sheets neglected or treated inadequately	Currents produce filaments or flow in thin sheets
Theories mathematically elegant and very well developed	Theories still not very well developed and partly phenomenological

ward. The plasma exhibited striations and double layers, the electron distribution was non-Maxwellian, there were all sorts of oscillations and instabilities. In short, it was a field which was not at all suited for mathematically elegant theories.

The other approach came from the highly developed kinetic theory of ordinary gases. It was thought that with a limited amount of work this field could be extended to include also ionized gases. The theories were mathematically elegant and when drawing the consequences of them it was found that is should be possible to produce a very hot plasma and confine it magnetically. This was the starting point of thermonuclear research.

However, these theories had initially very little contact with experimental plasma physics, and all the awkward and complicated phenomena which had been treated in the study of discharges in gases were simply neglected. The result of this was what has been called the thermonuclear crisis some 10 years ago. It taught us that plasma physics is a very difficult field, which can only be developed by a close cooperation between theory and experiments. As H. S. W. Massey once said (in a somewhat different context): "The human brain alone is not able to work out the details and understanding of the inner workings of natural processes. Without laboratory experiment there would be no physical science today."

The cosmical plasma physics of today is far less advanced than the thermonuclear research physics. It is to some extent the playground of theoreticians who have never seen a plasma in a laboratory. Many of them still believe in formulas which we know from laboratory experiments to be wrong. The astrophysical correspondence to the thermonuclear crisis has not yet come.

I think it is evident now that in certain respects the first approach to the physics of cosmical plasmas has been a failure. It turns out that in several important cases this approach has not given even a first approximation to truth but led into dead-end streets from which we now have to turn back. The reason for this is that several of the basic concepts on which the theories are founded are not applicable to the condition prevailing in the cosmos. They are "generally accepted" by most theoreticians, they are developed with the most sophisticated mathematical methods; and it is only the plasma itself which does not "understand" how beautiful the theories are and absolutely refuses to obey them. It is now obvious that we have to start a second approach from widely different starting points.

Characteristics of First and Second Approach to Cosmic Plasma Physics

The two different approaches can be summarized in Table 1.

If you ask where the border goes between the first approach and the second approach today, an approximate answer is that it is given by the reach of spacecraft. This means that in every region where it is possible to explore the state of the plasma by magnetometers, electric field probes, and article analyzers, we find that, in spite of all their elegance, the first approach theories have very little to do with reality. It seems that the change from the first approach to the second approach is the astrophysical correspondence to the thermonuclear crisis.

The Origin of the Solar System

From what has been said it is obvious that astrophysics runs the risk of getting too speculative, unless it tries very hard to keep contact with laboratory physics. Indeed it is essential to stress that astrophysics is essentially an application to cosmic phenomena of the laws of nature found in the laboratory. From this follows that a particular field of astrophysics is not ripe for a scientific approach before experimental physics has reached a certain state of development. As a well-known historic example, before the advance of nuclear physics, the attempts to understand how the stars generated their energy could not possibly be more than speculations without very much permanent value.

The problem of how the solar system originated has been the subject of a large number of highly divergent hypotheses. The reason for this has been that there was not enough basic knowledge of physics in the fields essential for the understanding of the phenomena and for a decision about which processes were possible.

However, before we discuss any details of a theory of the origin and evolution of the solar system, it is essential to define what general character such a theory should have. In the past too much attention has been concentrated on the formation of planets around the sun. One of the unfortunate results of this is that many theories of the origin of the solar system have been based on theories of the early history of the sun. This is a very shaky basis because the formation of the sun (and other stars) is a highly controversial subject. Recognizing that the satellite systems of Jupiter, Saturn, and Uranus are very similar to the planetary system, and, at least as regular as this system, it seems now more appropriate to aim at a general theory of the formation of secondary bodies around a central body, regarding the formation of the planetary system as only one of the applications of such a general theory.

The study of the sequence of processes by which the solar system originated has often been called *cosmogony*, a term which, however, is used in many other connections. As the origin of the solar system is essentially a question of the repeated formation of secondary bodies around a primary body, the term *hetegony* (from Greek *hetairos* or *hetes*, companion) has been suggested.

It seems likely (and it is fairly generally agreed) that the sequence of events leading to the formation of the solar system is likely to have been as shown in Fig. 1 (we are here following what has been called the "planetesimal" approach). A primeval plasma was concentrated in certain regions around a central body, and condensed to small solid grains. (Even the primeval plasma may have contained grains.) The grains accreted to what have been called embryos and, by further accretion, larger bodies were formed: planets if the central body was the sun, and satellites if it was a planet. The place of the asteroids in the hetegonic diagram is controversial. They have formerly been generally considered to be fragments of a broken-up planet, but there are now an increasing number of arguments for the view that they represent—or at least are similar to—an intermediate state in the formation of planets. A clarification of these two alternatives is important.

Even if the diagram of Fig. 1 is fairly generally accepted as it stands, this does not mean that the different processes are clarified. To a high degree they are still of a hypothetical character. Up to rather recently this has necesarily been the case because the basic processes have not been known very well. To some extent we have been in the same situation as the astrophysicists trying to clarify the energy generation in stars before the advent of nuclear physics. However, the situation seems now to be changing so that there is a good hope to bring the whole field of research from the state of a discussion of more or less bright hypotheses to a systematic scientific analysis.

Basic Knowledge for Reconstruction of the Hetegonic Processes

Besides plasma physics, which we have already discussed, there are a number of other fields of research which are basic for the reconstruction of the hetegonic processes.

1) *Plasma chemistry* means the field of research concerned with chemical reactions in a plasma. These are basically different from the reactions in nonionized gases. It should also be considered to include the separation of different elements which takes place in an inhomogeneous plasma due to, for example, temperature gradients and electric currents. Furthermore, the interaction between a plasma and a solid grain condensing from it is highly dependent

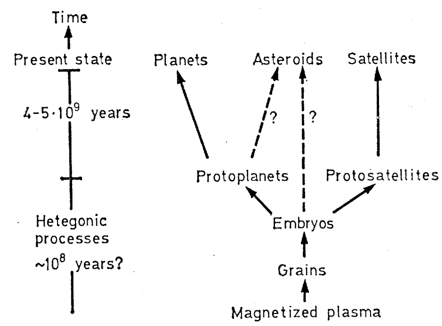

Fig. 1. Events leading to the formation of the solar system.

on the state of ionization. The laboratory results and their application to cosmic conditions are relevant for the understanding of the different chemical composition of the celestial bodies.

For the next process in our evolutionary diagram, namely, the accretion of larger bodies from the initial condensation, the following fields of research are essential.

2) *Solid-body collisions.* The grains which are the primary result of the condensation will move in Kepler orbits around the central body, but their motion will be disturbed by several effects. One of them is due to the mutual collisions. The relative velocities at these collisions may have any value from zero up to some 10 kilometers per second. This means that in many cases we are in the region of "hypervelocity" collisions. This is a field which is not yet understood very well. Available laboratory results seem to be scarce, and their application to cosmic conditions is uncertain because we know very little about the structure of the grains. Collisions between bodies with fluffy shock-absorbing surface layers are likely to differ from collisions between hard "marbles." Meteorite studies are supplying us with some information. The Apollo results about meteoroid impact on the lunar surface are another important source of knowledge. In these cases, however, we do not gain very much information about the structure of the grains in space, because the particles we recover have either passed the terrestrial atmosphere or been destroyed by impact on the lunar surface.

3) The study of *Kepler motion in a viscous medium* is essential for our understanding of the evolution of the orbits of the grains and the embryos. From a formal point of view this problem is similar to some basic problems in plasma physics, which are also concerned with a large number of interacting particles. It turns out that in the neighborhood of a central body the condensed grains have a tendency to move in similar orbits, thus forming what has been called "jet streams" in space.

4) *Celestial mechanics* serves of course as a general background for the whole hetegonic process. This field has been rejuvenated by the application of computer analysis to many of the problems which were formerly impossible to handle. Connected with this is the discovery of the importance of *resonance phenomena* in the present structure of the solar system. It seems likely that at hetegonic times also resonances played a decisive role.

5) The hetegonic processes took place 4 to 5 billion years ago. The evolution of the primary product of these processes into our present-day solar system has consisted of a number of relatively slow changes: geological forces have transformed the structure of the planets, tidal effects have braked the spins of some of the bodies (especially of the satellites), collisions have taken place in the asteroidal belt, and there have been meteor impacts on the planetary surfaces, etc. All these effects are important for the reconstruction of the state of the system immediately after the hetegonic processes ended. It is only after "correcting" for them that the solar system data we observe today are of value for the reconstruction of the hetegonic processes.

Space Observations Relevant to the Hetegonic Problem

From the analysis we have made it is evident that the background knowledge necessary for the understanding of the hetegonic processes is rapidly increasing through advances in several different fields of research. We shall now discuss the question of what sort of space missions are of special value for the study of the hetegonic problem.

Let us first state that many of the space missions which are carried out today or planned for the future give valuable contributions. Increased knowledge of the behavior of cosmic plasmas is gained by spacecraft carrying out plasma and particle measurements in the magnetosphere and interplanetary space. Further, meteor impacts on spacecraft supply us with information of the very small bodies in our environment, which are probably related to those small bodies out of which our present planets once accreted. Particularly important is the study of meteor impacts on the moon (and on Mars). Hence these and other investigations "automatically" contribute to the background knowledge necessary for the solution of the hetegonic problem. But although this is satisfactory there are a number of crucial problems which cannot be solved unless space research is purposely directed toward solving them. We shall now discuss how this could be done.

Big Bodies versus Small Bodies

It is usually thought that after the lunar landings the most important missions will be those to Venus, Mars, and the other planets. This is not necessarily true, because missions to asteroids and comets would be at least as interesting from a scientific point of view. As some asteroids are the closest neighbors of the earth-moon system, this would also be the easiest from a technical point of view.

Our analysis has indicated which fields of research will contribute to the clarification of different phases of the hetegonic processes. Plasma physics and plasma chemistry are important for the first phase, including the condensation of small grains. The study of meteorite- and asteroid-sized bodies will have bearing on the accretion. We can state as a general rule that the smaller the body, the further will the study of it bring us back in time. Thus small bodies will be relevant to earlier periods more than large bodies. This means that it is essentially through studying the properties of small bodies in space that we can hope to understand the crucial phase in the formation of the solar system when most of the matter which later formed the planets and satelilites was still dispersed.

There is evidence that during the formation of the planets and satellites a great deal of information about the formation processes was stored in them. However, to a large extent this information is either obliterated or inaccessible. The planets are likely to have accreted from "planetesimals." The earliest phase of this accretion produced a small body, the matter of which may today be in the core of the planet, which means that it is inaccessible even if a manned spacecraft should land on the surface of the planet. There is also a possibility that, for example, convection in the interior of the planet has more or less completely obliterated the information once stored there. Concerning the surface layers, geological processes—including atmospheric effects—have mostly wiped out the surface traces of the hetegonic processes in the earth and probably also in Venus. In other bodies like the moon and Mars, and probably also Mercury, there seems to be considerable information left, but only referring to the very last phase of the hetegonic processes.

Hence our conclusion is that studies of large bodies like the planets has only a limited value for the study of the origin of the solar system.

Asteroids, comets, and meteoroids are different in this respect. Even if some of these bodies are fragments produced at collisions in space, it is very likely that also these fragments contain considerable information about the condensation and the accretion processes. Because of the smallness of the bodies, there is no heating or convection in their interior which can obliterate the information stored from the time when they were produced, and at least in the very small bodies and by the fragmented bodies ther "interior" is accessible. Furthermore, a study of them will give us knowledge of the behavior of small bodies in space which will be valuable for the clarification of the hetegonic processes in general. We study in them intermediate products in the manufacturing of planets. They give us, so to say, snapshots showing the sequence of events when a planet like the earth once was created.

Old and New Fields of Science

We shall now return from our odyssey in both space and time to our starting point—how new technologies displace the center of gravity of the physical sciences. The great revolution in physics which took place in the beginning of this century meant that classical mechanics and classical electrodynamics were considered to be more or less obsolete as fields of research. The new fields that attracted the interest were the theory of relativity and quantum mechanics, and the experimental work was largely concentrated on the exploration of the electron shells of the atom. The advance of nuclear physics marked another step in a similar direction.

The new trend which is introduced by the rise of plasma physics and space research is to some extent opposite. In these fields quantum mechanics and the theory of relativity are not very important. Instead classical mechanics has become rejuvenated and is essential not only for calculating the trajectories of spacecraft but also for the study of the motion of the natural celestial bodies during their evolutionary history. Also classical electromagnetism is of decisive importance to the theory of magnetized plasmas, which is basic both for thermonuclear research and for astrophysics in general. This does not mean that we should make a mistake—similar to what was made 50 years ago—of declaring the atomic and nuclear physics to be obsolete. They are not. They have an enormous inertia which will keep them moving, and they will produce many new and interesting results. But they have very serious competitors, and, remarkably enough, these are the fields which earlier were declared dead that are now being resurrected.

It is possible that this new era also means a partial return to more understandable physics. For the nonspecialists, four-dimensional relativity theory and the indeterminism of atom structure have always been mystic and difficult to understand. I believe that it is easier to explain the 33 instabilities in plasma physics or the resonance structure of the solar system. The increased emphasis on the new fields means a certain demystification of physics. In the spiral or trochoidal motion which science makes during the centuries, its guiding center has returned to these regions from where it started. It was the wonders of the night sky, observed by Indians, Sumerians, or Egyptians, that started science several thousand years ago. It was the question why the wanderers—the planets—moved as they did that triggered off the scientific avalanche several hundred years ago. The same objects are now again in the center of science—only the questions we ask are different. We now ask how to go there, and we also ask how these bodies once were formed. And if the night sky on which we observe them is at a high latitude, outside this lecture hall—perhaps over a small island in the archipelago of Stockholm—we may also see in the sky an aurora, which is a cosmic plasma, reminding us of the time when our world was born out of plasma. Because in the beginning was the plasma.