RUSSIAN ACADEMY OF SCIENCES FEDERAL SPACE AGENCY

FIFTY YEARS OF SPACE RESEARCH

After the International Forum «Space: Science and Challenges of the XXI Century»
50th anniversary of the Sputnik launch
October, 2007
Moscow

MOSCOW
2009

УДК 520.6
ББК 39.6
П 99
ISBN 978-5-902-533-05-0

SCIENCE EDITOR *A. Zakharov*
COLLECTED BY *L. Zelenyi, A. Zakharov, O. Zakutnyaya*
EDITOR *O. Zakutnyaya, D. Odintsova*
ART DESIGN *V. Davydov*
COMPUTER PROCESSING *A. Zakharov, E. Korableva*

Collection of articles dedicated to the major results of space research in the first 50 years after the launch of the First artificial Earth satellite in 1957.

Articles based on the selected reports presented at the International Forum "Space: Science and Challenges of the XXI Century" (Moscow, October 1—5, 2007), as well as original articles are included in the book.

Subject of the book is main achievements and further prospects of studies of Earth's magnetosphere and interplanetary medium, of the Sun, the Moon and the planets of Solar system nearest to the Earth, as well as current state and future development of astrophysics and cosmic rays physics, and main results of physics and biomedical experiments, performed at the orbital stations.

Articles and illustrations are provided by authors unless otherwise stated.

Information authenticity of the articles are provided by authors.

Editorial Board may not be concordant with authors' opinion.

FOREWORD

October 4, 1957 the First artificial Earth satellite — Sputnik was launched in the Soviet Union. This event opened the Space Age and marked the beginning of a new science — space research.

To commemorate this remarkable event the International Forum «Space: Science and Challenges of the XXI Century» was held by the Russian Academy of Sciences together with the Federal Space Agency of Russia. It was held in Moscow and St.-Petersburg on the 1—5 of October 2007. The participants were greeted by S.B. Ivanov, the First Deputy Chairman of the Government of the Russian Federation; academician Yu.S. Osipov, the President of the Russian Academy of Sciences, A.N. Perminov, the head of the Federal Space Agency, academician V.A. Sadovnichy, the President of the Russian Rector's Union and the rector of the Lomonosov Moscow State University, L.M. Zelenyi, director of the Space Research Institute of the Russian Academy of Sciences.

Russian and foreign scientists and experts from various institutions of the Russian Academy of Sciences, the Federal Space Agency, industrial and educational organizations took part in the Forum. In the framework of the Forum were held twelve symposia, conferences and 'round-table' discussions, devoted to the main fields of space research and technologies, international cooperation. 225 reports were presented. At the plenary sessions of the Forum the leading scientists from different countries presented review reports on the most important topics of space research and space activities. At the solemn session held in the Russian Academy of Sciences heads of the Russian Academy of Sciences, the Federal Space Agency and foreign space agencies addressed the Forum. In the Space Research Institute the exhibition «Space Science — Past and Future» was held. The Forum was concluded with a gala concert.

Current edition includes selected reports presented at the Forum, which describe the main achievements and still unsolved problems in the field of the study of the Earth magnetosphere, interplanetary space, the Sun, the Moon, the nearest planets, astrophysics and cosmic-rays physics investigation, the results of physical and biomedical experiments carried out at orbiting space stations, issues of space technologies and international cooperation in space.

A. Zakharov

ADDRESSES TO THE PARTICIPANTS OF THE FORUM

SERGEI IVANOV
First Deputy Chairman
of the Government
of the Russian Federation

Dear Friends,

On behalf of the Government of the Russian Federation, I welcome all the participants of the International Forum dedicated to the launch of the First Artificial Earth Satellite.

This was a momentous event for all mankind and is of special significance for Russia. Fifty years ago, the Soviet Union, having survived the tragedy of the World War II, nevertheless succeeded in being the first to reach outer space. The outstanding Russian scientists and aerospace engineers S.P. Korolev and M.V. Keldysh used the ideas of K.E. Tsiolkovsky, the founder of cosmonautics, as the basis for creating and developing a new sphere of science and a new branch of industry – space engineering.

For many years the space programme has become the heart and soul of society's intellectual, industrial and technological potential, allowing our country to become one of the world leaders. A vital role in this breakthrough was played by the fundamental sciences, which not only provided the theoretical foundation for the possibility of space flights, but also took the most active part in fulfilling this possibility. Developments in the sphere of nanotechnology, information systems and other latest scientific spheres still serve as the basis for the progress in this branch.

Our task today is to rely on the potential of Russia's space industry, created over the previous decades, making effective use of its achievements in resolving the pressing problems of the 21st century. One of the keystones in this success is a consolidation of the state and society, based on the increasing activity of private institutions in space research and exploration throughout the world.

It is important that outer space is, in many ways, a sphere of effective international cooperation, as is confirmed by the holding of such a prestigious forum in Moscow.

I wish all the participants and guests productive work, interesting discussions and new discoveries.

YURI OSIPOV
The President
of the Russian Academy of Sciences

Dear Friends,

I am pleased to welcome the participants of the International Forum «Space: Science and Challenges of the 21st Century», dedicated to the 50th anniversary of the launch of the first artificial satellite of the Earth, in the Great Hall of the Academy of Sciences.

The Russian Academy of Sciences is directly associated with this anniversary. For many years (1961—1975), the Academy was headed by the «Chief Theoretician of Astronautics» — Mstislav Keldysh. He managed to establish a very close cooperation with the aerospace industry and its head — Sergei Korolev, who also had been a member of our Academy since 1953. Reading today their joint addresses to the country's leaders written in the late 1950s and early 1960s, one can see that the ideas laid at the foundation of the Russian space programme have been still evolving fifty years later and yielding fresh results.

Even before the first satellite was launched, Keldysh and Korolev had already conceived of the truly tremendous scope of the scientific objectives that might be fulfilled with the help of satellites: study of the near-Earth space and the other planets, research into the Universe using all range of electromagnetic wavelengths, remote monitoring of the Earth, experiments under microgravity conditions and much more.

It is important that Keldysh and Korolev understood right from the beginning that space research, like any other branch of science and technology, required a systematic approach and was inconceivable in isolation from other spheres of human activity.

That is why the title of the Forum held by the Academy of Sciences in this anniversary year combines two seemingly quite remote concepts: scientific research and challenges of humankind. Yet nowhere else the significance of science for the people living on our planet is demonstrated clearly as in space research.

Once, space flights were just a dream, just an idea favored by a few hundred and maybe only a few dozen people; yet today they have become a part of daily life. The importance of space flights is particularly great as we are often simply unaware there is a "space" component in some device, technology or aspect of life.

This, more than anything else, constitutes the specific nature of space research as a science. On the one hand, space research requires an adequate level of *science and technology development; on the other, it serves as a mighty stimulus to developing of new scientific ideas, hypotheses and theories – and, naturally, technologies, which can subsequently be used not only with reference to space.*

This is, however, as it were, the exterior of the interconnection between science and humankind. Today, though, one has to mention the interior aspect of their relationship.

Launch of the first satellite was primarily a scientific and technological challenge. Our great compatriot Konstantin Tsiolkovsky, whose 150^{th} anniversary has been celebrated recently, spent considerable efforts on scientific justification of space flight. Not only his technical but also his philosophical works produced a whole generation of future spacecraft designers. Tsiolkovsky's name is among the wonderful galaxy of Russian space philosophers: Nikolai Fyodorov, Vladimir Vernadsky, and Alexander Chizhevsky (whose 110^{th} anniversary has also been celebrated this year). There are the philosophers who were responsible for developing the amazing idea that humankind is inseparably linked to the Solar System and the Universe. Now, on the basis of experimental data obtained from studying solar-terrestrial links, we are becoming convinced that their predictions were right. Yet even more discoveries undoubtedly await us in the future and we can only envy those who will continue the scientific search in this area.

Hundreds of talented scientists and engineers worked to send a small aluminium sphere with a radio transmitter into orbit. Yet it was this small sphere that unexpectedly changed the perception of literally billions of people on our planet. Thus, 50 years ago Space became part of humankind life. There can be no doubt that still greater discoveries and ever more amazing achievements are in store for us on the path towards cognizing the Universe, along which humankind is currently travelling.

VIKTOR SADOVNICHY
The President of the Russian Union of Rectors
Rector of Lomonosov Moscow State University

Dear Colleagues!

In the 21st century education builds up a basis for the development of the society and directly defines quality of life and progress in all spheres of mankind's activity. The University community totally understands that the quality and the results of its activity will determine the development of the education and the prospects for progress in the 21st century. However, integration of science and education forms the basis for improvement of quality and innovativeness of the education. The University science is a powerful component of all directions of scientific research among which space research is of top-priority. Started with a launch of the First Soviet satellite, space research stimulated the development of many new scientific directions and in fact gave an impulse to the exploration of the Universe.

It's impossible to overestimate the contribution of the University scientists into the space research program. Suffice it to recollect the names of Sergey Vernov, Mstislav Keldysh, Alexander Ishlinsky, Georgy Petrov, Dmitry Okhotsimsky – the outstanding scientists and professors of Moscow University, who are inseparably associated with the University science and education. In 1957 the first scientific equipment developed by the University scientists was installed on board of the second Soviet satellite. Since then many scientific groups specialized in different fields of space research, including space physics and astronomy, applied mathematics and medicine, Earth remote sensing in behalf of geography and geology, have been set up in Lomonosov Moscow State University. Lomonosov Moscow State University is rather exclusive Russian scientific and educational center which concentrate all the most important directions of the fundamental and applied space research combined with the educational process. By all means it became possible due to the launch of the First satellite which anniversary we celebrate today.

ANATOLY PERMINOV
Head of the Federal Space Agency of Russia

Dear Ladies and Gentlemen,
Participants of the International Forum, Colleagues and Guests!

I'm glad to welcome all of you at the solemn meeting dedicated to the 50th anniversary of the launch of the First Artificial Earth's Satellite and congratulate its creators on this significant jubilee!

On October 4, 1957 the humankind celebrated the first victory over the terrestrial gravity. This outstanding engineering achievement initiated using space vehicles that were the qualitatively new and powerful implement of research activities. In these days, we, in a certain sense, summarize the half-year space activities and, at the same time, look into the future.

We are proud of the fact that our country made the road to the space. For all the fifty years, Russia has been carrying out the policy of the outer space development regularly and consistently in the interests of the nation stable development. Currently, our space potential allows to implement the full work cycle; from production of space facilities to acquisition and use of space activities' results.

Creation and further application of space rocket equipment have been the life affairs of several generations of our compatriots. Works of a great Russian scientist Konstantin Tsiolkovsky, whose 150th anniversary we celebrated in September, 2007, had had the fundamental importance to understand theoretical feasibility of flights to the outer space. The creation of the rocket equipment and the beginning of the outer space development, for us, are inseparably linked with the greatest designer of the 20th century Sergey Korolev. We have began this 'space' 2007 year with the celebration of his 100th anniversary. I want to point out particularly, in the constellation of space names, the name of the President of the Academy of Sciences Mstislav Vsevolodovich Keldysh, the author of the fundamental works on the applied celestial mechanics and rocket dynamics, who has been the inspirer and the organizer of the scientific program of the outer space exploration and is considered, by right, the Chief Theoretician of the national cosmonautics.

Scientific space explorations have cardinally changed and have constantly been widening our knowledge about the world for a short period of time by historical criteria. Research activities using space facilities are carried out from boards of automatic and manned complexes and include the whole spectrum of problems of study of the Earth, the near-Earth space, planets *and bodies of the Solar system, the Sun itself, astronomical objects of the deep space, and the astrophysical explorations, the space biology and the weightlessness physics.*

These explorations are of great importance for Russia in view of purely important discoveries of the last years, prospects of the Solar system exploration and development of fundamentally new techniques and methods of exploration, which stimulate origination and development of applied scientific tenors.

The Federal Space Program for 2006—2015 provides more than twenty scientific projects. Within it, it is planned to produce special spacecraft equipped with target complexes of scientific equipment and to install complexes of scientific equipment onboard the domestic and foreign spacecraft. For the period of up to 2015, scientific space exploration are going to be implemented for the following basic tenors:

- *extra–atmospheric astronomy;*
- *planetology;*
- *studies of the Sun, space plasma and solar-terrestrial relationships;*
- *space biology;*
- *physiology;*
- *sciences materials.*

In the planetology field, for instance, a flight to the Mars's satellite – the Phobos, to the Venus and an expedition to the Moon, and participation in a number of planetary missions implemented by NASA and the European Space Agency are planned. International cooperation in the field of research activities enlarges constantly and opens up wide feasibilities to solve global problems common to all humankind.

One of our first international programs, the INTERKOSMOS Program, commenced 40 years ago in 1967. During its realization, valuable results of the ionosphere and the Earth's magnetosphere explorations, of the study of the Sun and solar-terrestrial relationships, cosmic rays, processes and phenomena in the upper atmosphere have been acquired, various medical and biological, geophysical, technological and other researches and experiments have been implemented.

Today, I want to emphasize that the space and the science are extremely necessary for each other and practically inseparable. The space activities became feasible due to, substantially, the progress of the fundamental science in the field of mathematics, physics, mechanics, chemistry and other branches, even philosophy. On the other hand, the creation and development of the space equipment led to a great number of new scientific specialties. Now, the space *activities stimulate the fundamental and applied sciences development, build up demand for result of the research activities.*

In turn, the development of those capabilities, which are created by the space equipment for experimental equipment installed on board, transfer the scientific research of the outer space onto a new, considerably higher level.

I am sure that the up-to-date ideas and projects, which are going to be discussed at the Forum, will be widely applied in the space rocket and other science-intensive branches of economics. Taking the opportunity, I want to thank all members of the Russian Academy of Sciences and institutes dependent on it for the creative contribution to exploration and development of the outer space, which allow the Russian science to occupy fitting places including the field of fundamental research.

Esteemed scientists, we thank you for that stupendous work that you make professionally in the interests of space activities' further development!

MICHAEL GRIFFIN
Administrator
of the National Aeronautics and Space Administration,
USA

Dear Ladies and Gentlemen,

I thank you for the opportunity to be here with you to celebrate this magnificent accomplishment. Sputnik is a landmark achievement in human history. And I think it is appropriate to recognize here today that it is a Russian achievement. It is the accomplishment that fits in the history of other Russian accomplishments in science, mathematics and engineering, certainly emblematic of the work of Konstantin Tsiolkovsky, whose 150th anniversary we also celebrate, before we go along to the accomplishments of many other Russians in science and mathematics.

And for those who considered at the time that Russia was a backward nation due to the after-effects of the World War II, Sputnik served as a notice to the world that Russia was here to stay.

But Sputnik was only the start, only one of many Russian accomplishments in space work performed for the first time. First animal in space, flown shortly thereafter, the dog Laika. First lunar flyby in 1959. The first two humans to reach Earth orbit, Yuri Gagarin and German Titov. The first double space flight performed by Andriyan Nikolaev and Pavel Popovich. And Valery Bykovsky and Valentina Tereshkova — another double spaceflight, but this time with the first woman to fly in space. The first spaceflight with three people onboard: Vladimir Komarov, Boris Egorov, Konstantin Feoktistov. First space walk — a mission commanded by Pavel Belyaev and space walk executed by Alexey Leonov. The first landing of a robotic spacecraft on the Moon, by Luna-9 in 1966. The first crew transfer on orbit accomplished during Soyuz-4 and -5 missions. First triple mission — three spacecraft on orbit at one time, Soyuz-6, -7 and -8 in October, 1969. And a long, very long list of accomplishments in the Salyut and Mir programs, through the 1970s and 80s — all to the credit of the Russian people.

There were also tragedies on the way. I think it's appropriate on the 50th anniversary of the celebration of the Sputnik triumph to recognize that people have been lost during the development of space flight. Vladimir Komarov, in 1967. Viktor Patsaev, Georgy Dobrovolsky and Vladislav Volkov in 1971 on Soyuz-11. Tragedies that the Russian people above all others remember.

There is much to admire in Russian accomplishments in the space, much that I do admire. There's a lot of spaceflight within the Russian people that I believe exceeds that any other people on Earth. Russians hold on to their abilities to fly in space, with difficulties that we, Americans, have not yet encountered.

You have achieved an extraordinary amount with the much smaller investments that many other nations have put forth. And your space systems' designs, approaches and methodologies offer many lessons to those from other cultures.

I must thank you also for contributing much to our space program. The Sputnik accomplishment by the Russian people was responsible for the creation of the American space program that exists today. For without Sputnik-1, without Yuri Gagarin, there would not, I believe, have been Mercury, Gemini, Apollo, or Skylab, the key programs of the 1960s and early 70s in the United States.

Indeed, when the space race of the 1960s was over, it may be seen that America lost some of our inner momentum, some of our inner desires to accomplish things in space for their own sake.

Today we have moved beyond the competition to the theme of collaboration, working together in space. And indeed, while I believe that competition can be good, as long as it is maintained in the appropriate balance, there is more to be said in terms of cooperation with much to learn from each other. And I think we can go further together than either of us can go separately. I look forward to help, in turn, adding to that reality.

And so, I close by again thanking you for the invitation to be here to celebrate with you your accomplishments, and I look forward to many more years of working together.

Spasibo!

JEAN-JACQUES DORDAIN
Director General
of the European Space Agency

Dear Ladies and Gentlemen,

It's an honour and a privilege for me to be here today to celebrate the 50th anniversary of the launch of Sputnik, which is the 50th anniversary of the start of space exploration. And it's also an honour and privilege for ESA, for Europe, I would say, to be here next to Russia and USA, the two pioneer space powers, to celebrate this event that has changed the world and changed the life on Earth.

Speaking of Europe, we are also celebrating this year the anniversary of the Treaty of Rome. And I would like to say that there are a lot of common points between space and Europe.

First of all they are difficult. The two are difficult but the two are useful. The two are based on cooperation. The two have brought peace. The two, space and Europe, are not visible – you can not touch space, neither can you touch Europe. And this is one of the problems of the two. Consequently, they look distant from citizens.

But the two are also representing the future. The two are bringing more confidence to the future.

So Sputnik has opened a new page in the history of humanity. I must say that it is obvious after 50 years of space exploration. It was, maybe, less obvious at that time. I looked through a newspaper dated 1957 and I must say that when you are reading this paper, you feel enthusiasm, but also incredibility. A combination of hopes and fears. But it was something important for the public. And maybe still more important for the children.

Just to say that on the 1st of October, 1957 I was 10 years old, I entered the secondary school, and three days later we had Sputnik.

For me it generated a lot of dreams. And it was the same for all my friends at school. We were all dreaming of space. And I must say that there was a competition between me, getting my engineering degree, and the Apollo program. And we arrived on exactly the same date because I got my engineering degree on exactly the same day Neil Armstrong put his foot on the Moon.

Though I have never been in space, I have always dreamt of space thanks to Sputnik meaning that Sputnik has become my life.

The importance of space has never been so great as now, 50 years later. First of all it is great for the scientists and I think is a very important evidence of that is this celebration taking place at the Russian Academy of Sciences.

the world even though they don't realize that. A lot of citizens of the world use space on a daily basis. They don't know that but they use space. It is even greater for young children, since it attracts them to science and engineering.

So space today is certainly very different from what it was 50 years ago. We have moved from demonstration period to service delivered to the citizens, we have moved from the cold war time to peace, we have moved from competition to cooperation, we have moved from space to the Earth. I must say more and more people will understand that the better we understand space, the better we understand the Earth. Also, we have moved from two space powers to a lot of space powers. Today there are a lot of space powers in the world. And among these space powers ESA is still a small space power but I think, a very important space power for the world.

ESA has developed a lot of capabilities now for science, for services, for access to space. ESA has learnt a lot from its cooperation with, especially, the United States and Russia. Just to say that Europe started to work in space in '64, at the time when there were a lot of missions. Russia succeeded in a lot of "firsts", and the United States was close to the Moon.

But if there is one aspect learnt during the recent 40 years, one aspect that ESA can teach some other partners — this is certainly international cooperation. International cooperation is our daily work. 17 countries working together — I can tell you this is difficult! But this is successful. We are also cooperating with all space powers of the world. Obviously, with NASA — ESA started its career in the cooperation with NASA. We are cooperating also more and more with Russia and also with all space powers. I must say that we are the International Space Station partners, and our ESA astronauts fly only thanks to our partners NASA and Roscosmos.

The ESA laboratory Columbus will be launched in two months from now by a US Space Shuttle. A few weeks later we shall launch the cargo ATV, which will dock to the Russian part of the International Space Station. Just to say that cooperation is the daily work.

I would like to close by saying that it is always difficult to predict the future. And I shall certainly not try to tell you where we shall be in 50 years from now (at least I know where I shall be in 50 years. But for space — I don't know). I have only one strong belief — there will be more and more cooperation among the space powers, but also more and more cooperation between space powers and nonspace powers.

Although you should never regret the past, I have only one regret for the past — today the future looks much too far away. I would like to recall that there were less than four years between Sputnik and Gagarin. NASA went to the Moon in less than ten years. And I would like to get back to that type of pace because that is the only way to keep the motivation and especially the motivation of the young generation. So let's keep the motivation, let's keep the emotion.

Thank you very much!

PREAMBLE

L.M. ZELENYI
Space Research Institute
Russian Academy of Sciences,
Moscow, Russia
lzelenyi@iki.rssi.ru

FIFTY YEARS TO CHANGE OUR VIEWS ON THE WORLD

It occurs seldom in the humankind history so that the transfer from one historical epoch into another can be marked accurately to a day not to mention seconds. However, exactly fifty years is on October 4, 2007, at 22 hours 20 minutes and 8 seconds, since we have transferred into a new epoch, i.e. the Space Era, which has also become the time of space explorations.

Half a century has passed since then, and satellites have become an integral part of our life and the most important tool of research activities, which helps us to solve a number of scientific problems. Above all, these are the researches of the Universe, its origin and evolution, and our Sun and its planets. Our view on the near-earth space has changed for the fifty years passed. We could say that we have opened quite new space of the scientific search with the launch of the First Sputnik.

Why did the entry of our devices outside the Earth play the critical role in the science? A thick nitric cover wraps up our planet and, if we look at the wavelength list (Fig. 1), which are connected with various physical phenomena in the space, from gamma-

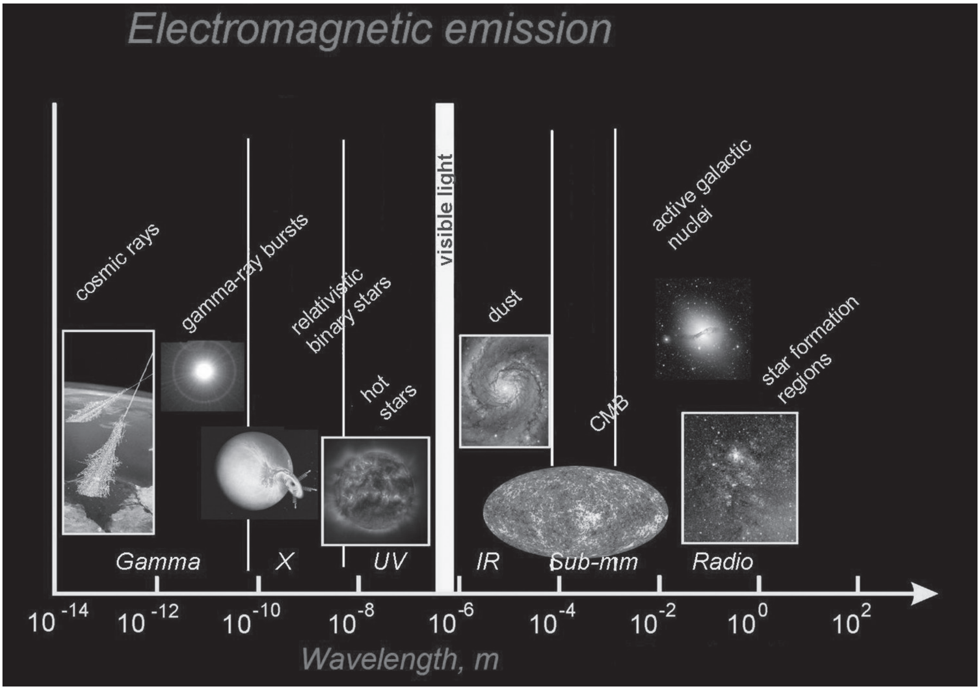


fig. 1

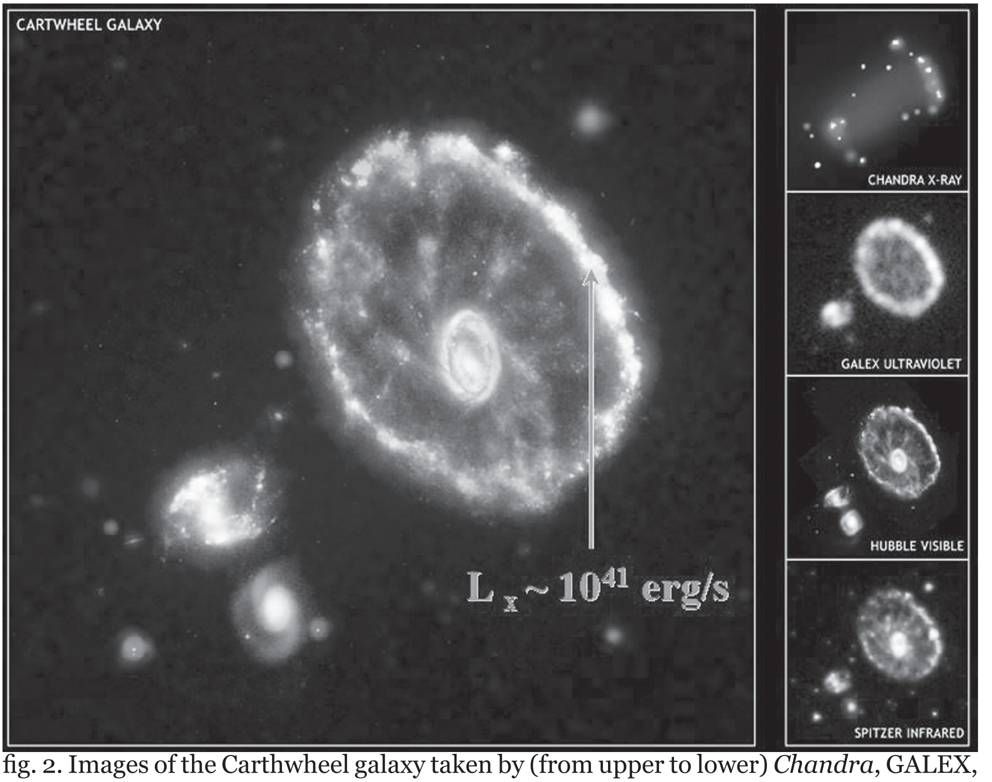


fig. 2. Images of the Carthwheel galaxy taken by (from upper to lower) *Chandra*, GALEX, Hubble Space Telescope, *Spitzer (NASA)*

radiation to radio, then we see that most of them are 'kept' from us behind a thick curtain of our atmosphere and ionosphere. We see the Universe only through two narrow spectral 'windows', i.e. the visible region and the radio-frequency region.

We can be only amazed at ingenious intuition of our predecessors working in the pre-satellite epoch, and with that how much they could see and guess looking through these two small 'windows'. If we would rather try to find an image of that happened fifty years ago, then we can remember a famous medieval gravure imaging a monk who breached the coelosphere by his head and saw an absolutely new world. The satellite played the same role. Even the well-known *beep-beeps* of the satellite, in fact, are not just sound signals; they code information about density and temperature inside the space vehicle, i.e. the small iron ball, which has changed our view on the world.

Let us compare several images of the same galaxy having a funny name the Cartwheel taken in different ranges: X-ray, ultraviolet, visible and infrared (Fig. 2).

It turns out, that we can see different processes in various pictures. In usual or visible light, we can observe no areas connected with the intensive energy yield, but the X-ray radiation allows us to see where the energy releases at the periphery of this galactic disc. Some formations being dark for us in the X-ray become visible in the ultraviolet.

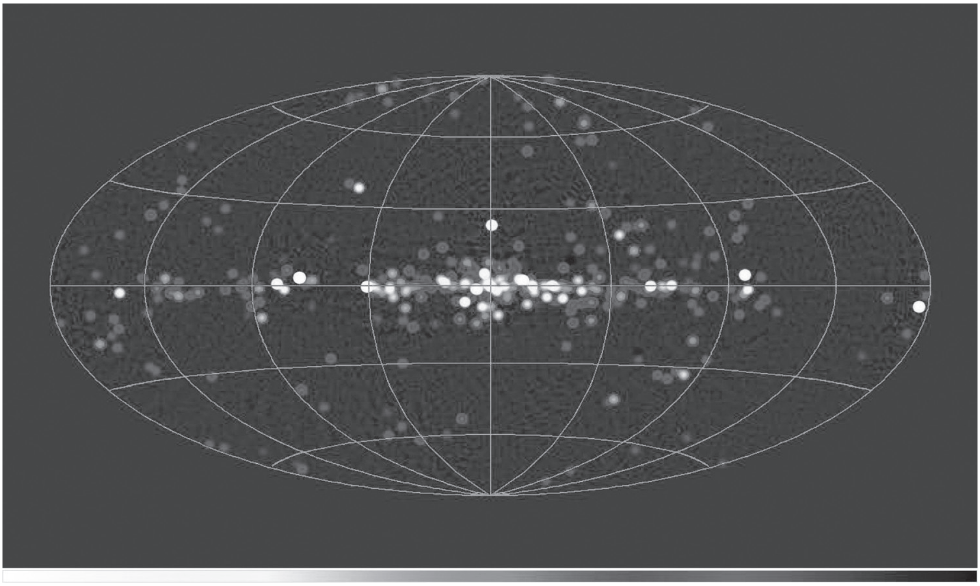


fig. 3. Image of the Milky Way by the INTEGRAL observatory (ESA)

With the all-wave astronomy, which was actually born at the very beginning of space exploration, we 'recovered our sight' and saw the world in its whole variety and beauty.

The discoveries did not delay. Probably, the most outstanding event of those years is the discovery of gamma-ray bursts, the most powerful explosions, which occur in distant areas of our Universe with energy liberation of 10^{44} W. They almost evenly cover the whole sky. As it often happens, the gamma-ray bursts were discovered accidentally by American *Vela* satellite, which had been initially designed to spy upon testing of nuclear weapons in the Soviet Union. Afterwards, similar experiments were also performed in the Soviet Union: KONUS, experiments on board Venus's interplanetary stations and other vehicles. The discovery of the gamma-ray bursts is extremely important, since it allows to look at processes of the energy release in the Universe absolutely in a new fashion. During the last years, we have succeeded to identify the gamma-ray bursts in visible range but nevertheless their nature is not clear hitherto, and the scientists continue working on this problem.

Further discoveries were in another range of lower energies. They are connected with the X-ray astronomy founded by Riccardo Giacconi, who was awarded with the Nobel Prize for his works in 2002. The main physical phenomenon discovered using X-ray telescopes was the matter accretion onto neutron stars and black holes. The matter falling onto such compact objects is accelerated, heated and starts to radiate X-rays. Thus, by means of the X-ray telescopes, we open astrophysical phenomena, which are practically invisible otherwise (Fig. 3).

Moreover, we managed to look at our Sun in a quite different way. Before the satellite era, scientists could only observe sunspots in areas of magnetic field amplification. When we looked at the Sun via the X-ray telescopes, we could see areas where the energy release occurs in fact. Comparing these data with those acquired in other ranges, i.e. the visible light and the ultraviolet light (Fig. 4), we have the whole image of what happens with our star.

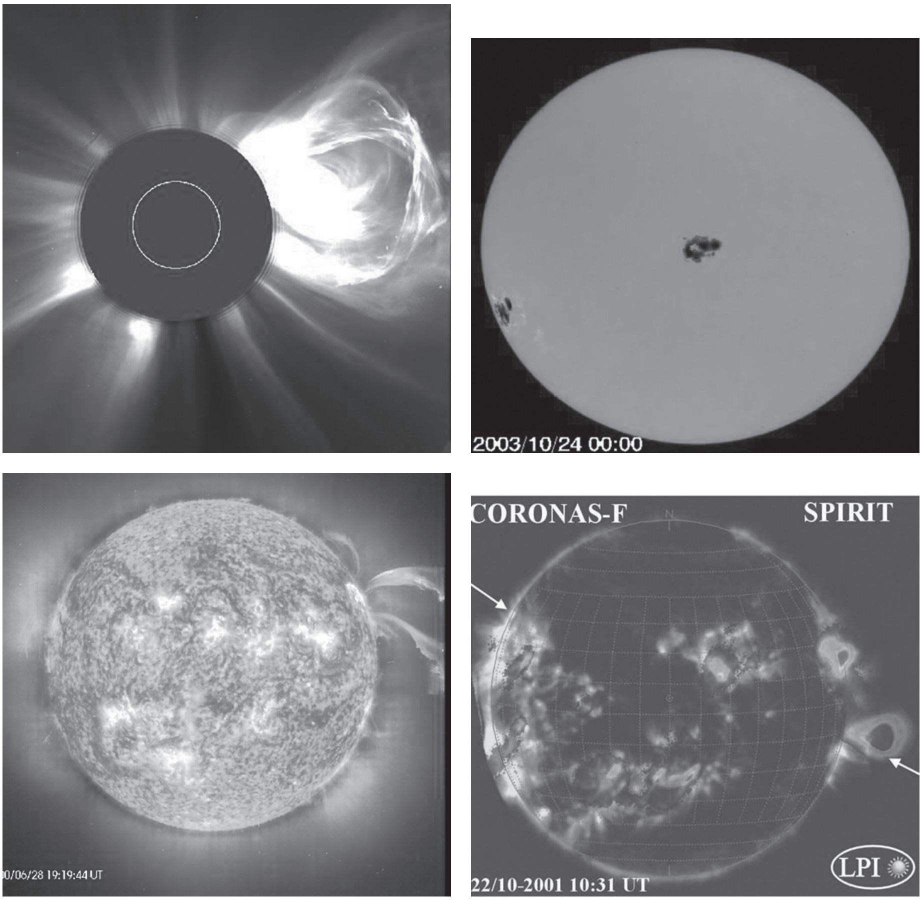


fig. 4. Images obtained with SOHO (ESA, NASA) and *Coronas-F* (Russia)

Investigating the physics of this energy release, we can understand why and how the Sun releases clouds of magnetized plasma, which spread throughout the Solar system. Some of them reach the Earth and result in quite troublesome consequences: severe magnetic storms, which can disable electric power lines, satellites in near-earth orbits and even equipment on the Earth. Understanding the physics of this chain allows us to try forecasting hazardous situations basing on observations of the Sun's activity.

Let us proceed. The ultraviolet astronomy has proved to be very important for the astrophysics. Young stars radiate in this very range. Our Sun is already an 'aged' star and its spectrum is set off towards the yellow side. But its temperature was much higher during the first billions years after the birth, and it radiated in a shorter waves range of the ultraviolet spectrum part. Observing the sky in the ultraviolet range, we see a 'kindergarten' of stars and we can research processes of their formation, and this allows us to understand the history of our Sun and of the Solar system.

The space researches have brought out the optical astronomy at the new level as well, though its history has already numbered thousands years. Optical observations can be carried out on the Earth but we should take into consideration its atmosphere with various aerosols and turbulences. Not without reason, large telescopes are located in mountainous regions where the sky is relatively clear. That is why the next step was important, i.e. large optical telescopes' entry to the near-earth orbits. One of the outstanding successes became launching of the Hubble Space Telescope. It has executed a great number of interesting observations but selecting from all the discoveries of 'Hubble', the most important one, in my opinion, has become the Hubble constant amendment. In the end of the 20s of the last century, American scientist Edwin Hubble, studying spectral shift of different galaxies, determined that our Universe is expanding. The farther an object is from the Earth the more redshifted its spectrum is. This law has been named 'the Hubble law', and the Hubble constant, correlating a distance to a galaxy with its velocity, in fact determines the history of our Universe and our fate. Hubble himself evaluated it as 500 km/s per megaparsec.

Since then, attempts to refine this constant have been taken for almost 70 years. Its value has fluctuated very much, and its quite exact value could be obtained only after the launch of 'Hubble'. Knowledge of this parameter is very important to understand the physics and the role of recently discovered dark energy and dark matter, which fill up our Universe.

Another greatest cosmology discovery was made in a longer wave range, the submillimetric one. It was the measurement of angular distribution of the relict radiation, which 'keeps in its memory' the very first moments of our Universe's history. After the Big Bang, the matter was almost uniform, and our metagalaxy, galaxies and stars including our Sun, i.e. the visible large-scale structure of the Universe, were gradually generated later on from its finest non-uniformities, the value of which numbered one thousandth percent.

Measurements of these non-uniformities were complicated because of the low anisotropy. Such experiments were run in our country as well; the soviet *Relict* space vehicle worked in the 80s, by means of which a group of scientists of the Space Research Institute and the Institute of Astronomy of the Academy of Sciences of the USSR executed the measurements. Unfortunately, these works were not completed for various reasons. At a later time, the COBE satellite (NASA) measured the anisotropy; and the universally recognized authors of this discovery had become American astrophysicists John Mather and George Smoot who were awarded the Nobel Prize for physics, which became the second Nobel Prize for the space explorations results (Fig. 5).

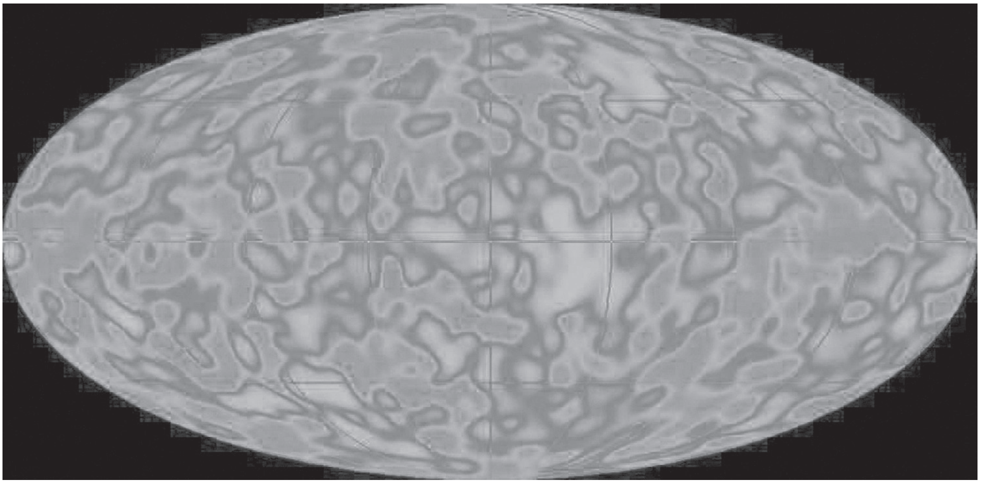


fig. 5. Image by COBE (NASA)

Let us advert to another 'classic' field — radioastronomy, which has also existed for quite long time on the Earth. For higher accuracy of radiointerferometric observations, it is necessary to locate two radiotelescopes as far from each other as possible, increasing the radiointerferometer base thereby. Being on the Earth, we are always limited by its diameter — 12,000 kilometers. This limitation can be overcome by entering of the telescopes in the outer space; and now new, much more accurate, projects are being prepared in the world. In particular, Russia plans the *Radioastron* project, where, by means of an orbital telescope, observations will be carried out on a base of tens times longer distance — 350,000 km. This will allow to research with very high accuracy the finest features of radiosource structures in our Galaxy and distant ones, quasars in particular.

A cascade of discoveries, not lesser than in the astrophysics, has followed in study of the Solar System for the passed years. We have learned a lot about almost all the planets of our Sun (Fig. 6). Probably, only the Pluto has not been researched so thoroughly, though we look forward to receiving the information from the *New Horizons*[1] spacecraft (NASA). However, now the Pluto itself is not officially considered to be a planet and, since the last year, it has turned into an asteroid[2].

Measurements of the Mercury — the first planet from the Sun have been executed as well, but it took place quite long ago; the American *Mariner-10* spacecraft transmitted the first images of the planet to the Earth in 1974—1975. However, within the next few years this planet is going to become the matter of careful researches. The American *Messenger*[3] spacecraft is flying now to the Mercury,

1 The *New Horizons* space vehicle was launched in 2006. It is expected to reach the Pluto in 2015.
2 In August 2006, The International Astronomical Union approved the resolution specifying the definition of the planet.
3 The launch was in August 2004. The space vehicle made the first flight near the Mercury in January 2008. The insertion into the orbit of Mercury's artificial satellite is expected in March 2011.

and we are waiting so much for the results. The European Space Agency also plans a mission to the Mercury, which has been named *BepiColombo*[4].

It is interesting, that perhaps the largest unexpectedness and great disappointment for the scientists became the Venus. Soviet landers, which had been sent to it in the beginning of the 70s and later, transmitted tremendous research results. The Venus has turned out to be, in a literal sense, the hot 'hell'; its atmosphere is very dense, the near-surface pressure is hundreds of the terrestrial ones and the temperature is more than 700 degrees. If the Venus was called 'the Earth's sister' before, then its example shows now how far the self-accelerating greenhouse effect can lead to.

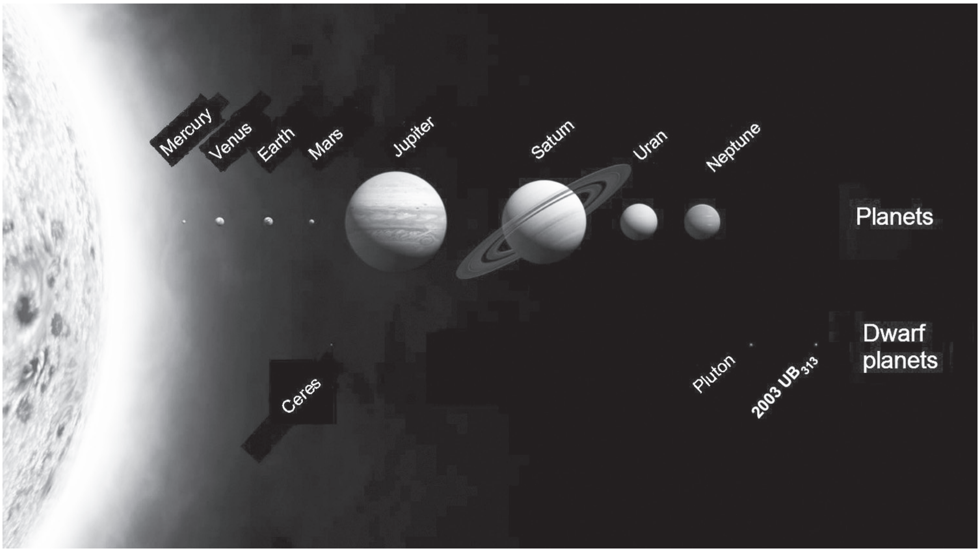
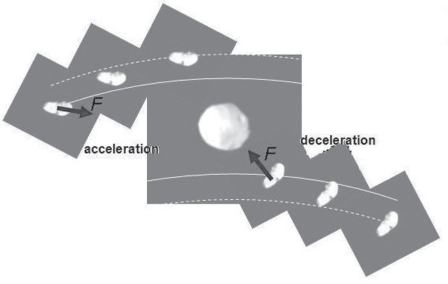


fig. 6

On the contrary, last researches of the Mars have become a nice surprise for many researchers and those who dream about colonization of the Solar system. A series of experiments, starting with the Russian HEND experiment aboard the American *Mars Odyssey* spacecraft, then radars and optical instruments of the European *Mars Express* spacecraft discovered quite vast reserves of water ice. Apparently, the Mars had spacious water areas formerly. Currently, there are disputes about how large they were but, undoubtedly, once if not oceans then rivers covered the planet, so that large scours are observed on the modern relief of the Mars.

It is impossible to tell about all the planets at once, therefore I dwell on one of them, which has particularly striked my imagination, and to a greater extent due to brilliant experiments of the last years, i.e. on the Saturn and its famous rings above all. They had been known for a very long time but the 'close sight' discovered how dynamic they were, how waves spread in them and what their fine structure was. Russian theorists of the Institute of Astronomy have developed the theory of the ring structure and predicted the so-called 'herder phenomenon', i.e. a small asteroid

4 It is planned to launch the *BepiColombo* in 2013, it is to reach the Mercury in 2019.



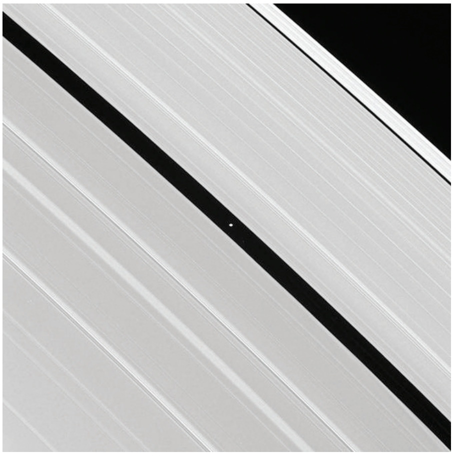


fig. 7. A. Fridman et al.

or a small satellite circling around the Saturn 'sweeping out' the materia from the ring and forming multiple fissures in it, just like a herder scattering sheep (Fig. 7).

The landing of the *Huygens* probe on the Titan — the Saturn's satellite still strikes my imagination for sure as well as yours'. This small planetary body has turned out to be similar to the young Earth in some features, and methane plays the same role here as water does on the Earth. If you look at the Titan's relief you can see that it resembles the Earth indeed, it has small mountain ridges, river scours, canyons and gullies. There is an impression that streams of liquid material existed on the Titan but it is clear that it was methane. In addition, there is very much haze consisting of some organic substances the Titan's atmosphere. Various space agencies are already planning to continue explorations of this satellite of the Saturn, and it seems to me that it will bring us a great number of surprises and discoveries.

Let us return closer to the Earth and see what new we have known about the near-earth outer space. Before, it has been a matter of optical, X-ray and ultraviolet telescopes, i.e. instruments recording electromagnetic radiations. The Satellite Era allowed experimentalists to bring into the orbit the instruments measuring the parameters of plasma, the fourth state of matter, considered to be the most widespread before the discovery of the dark energy and the dark matter.

Even the ancient Chinese, the invetors of the compass, knew the Earth had a magnetic field. Gilbert was the first to systematize these data; he showed that the Earth was a magnetic dipole, and it was quite a powerful magnet by space criteria, as we understand now. In 300 years, Kristian Birkeland – a great Norwegian scientist understood that the aurora, which was observed and systematically studied by him in Norway, was the intrusion of solar particles or corpuscles into the Earth's atmosphere. It remained unclear how systematic this phenomenon was and how often it occurred.

The Satellite Era has given the answer to many riddles at once. As in the astrophysics, discoveries run out like from the horn of plenty. In some months

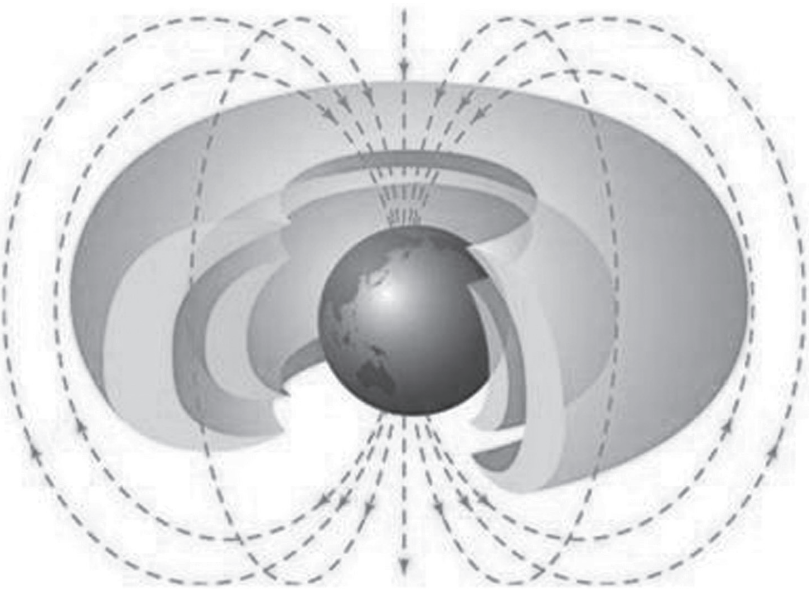


fig. 8

after the Sputnik launch, instruments for measurements of high-energetic particles ('radiations' as they were called at that time) became already working aboard the Third Soviet Sputnik. Independently and almost simultaneously James Van Allen (USA) and Sergey Vernov (USSR) published the works dealing with belts of energetic particles captured in the magnetic trap that is the Earth's magnetic field. The particles are captured in this magnetic 'bottle' (Fig. 8) and drift around the Earth.

This, by the way, as we understand now, turned out to be a very serious obstacle for space flights in turn because the radiation is very high and all research experiments, especially manned flights, must take into consideration the radiation belts, since it can be lethal to stay there.

The next train of discoveries began a little later and is connected with the discovery of the plasmasphere. It was made by a Soviet researcher Konstantin Gringauz with the help of satellite data and by an American scientist Carpenter by means of earth-based whistler observations. If you look at the plasma density, measured by Gringauz, away from the Earth, you see there is a sharp 'knee' — a jump of the particles concentration at a certain altitude, which couldn't be explained. Afterwards, the calculations allowed to understand that the Earth's ionosphere, which had been known for a long time, was not limited by altitudes of 1000...5000 km but was a huge 'ball' having the size of 25,000 km, rotating along with the Earth carrying the other particles as well.

A little later, investigation of the magnetic field allowed to determine what structure the terrestrial magnetosphere system had. Professor Norman Ness — an academician of the American Academy of Sciences made this discovery. He found an important structure — the tail of the magnetosphere, which 'stretches'

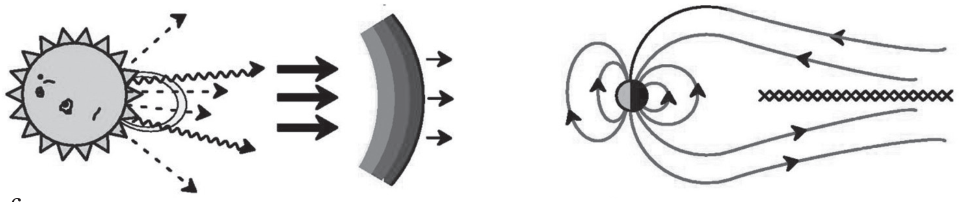


fig. 9

along the solar wind for millions kilometers. It is very important for the physics of this system since the energy is accumulated in it, and it can be released in burst-like way later on and generate magnetic storms and strong auroras on the Earth or, saying figuratively, 'shake the space weather' (Fig. 9).

A little later, the soviet *Luna-2* satellite and other spacecraft researched the solar wind. It turned out, and it was unexpected for the scientists again, that the Sun is constantly spreading the hot plasma stream, interacting with all bodies of the Solar system. There had been indirect indications of the existence of such stream before, deviations of comet tails, for instance. The researchers had suspected that there was an agent between the Earth and the Sun, carrying additional energy. I want to remember Alexander Chizhevsky — a Russian scientist, a younger contemporary of Tsiolkovsky and, in a sense, his disciple and friend, who compared the sun-spots quantity to various catastrophic occurrences on the Earth including with the quantity of murders, robberies, locust attacks, epidemics, etc. He saw the interdependency of these processes and, probably intuitively, felt that there was a complementary channel to transfer the energy from the Sun to the Earth in addition to the solar radiation itself.

Later, when other planets of the Solar system — the Mercury, the Jupiter and the Saturn were explored, it turned out to be that the Earth's magnetic field structure is universal per se. All of these celestial bodies have similar magnetosphere but of different scales; the Jupiter has the largest magnetosphere of course and the Mercury has the smallest one. A neutron star can have even a smaller magnetosphere. Radiogalaxies possess colossal magnetospheres. In other words, a new physical phenomenon — the magnetospheric interaction was discovered near the Earth.

Our Solar system is also inside a magnetosphere included in another stream — of interstellar gas. The plasma spreading from the Sun is the solar wind, which collides with the interstellar medium at a distance of about 90 astronomical units and a compound interaction originates at the border of this collision.

Finally, the following picture is beginning to emerge: magnetospheres of all planets of the Solar system are enclosed in the common magnetosphere, which the Solar system, in turn, 'cuts out' of the interstellar plasma. And its a pleasure to note that not long ago a man-made space vehicle reached these boundaries of the Solar system for the first time in the history. The American *Voyager* spacecraft crossed the above mentioned border of the termination shock in 2004 and it is going farther on. Very soon, for the firs-t time, the man-made object is going to get in not just the interplanetary space but in the interstellar one.

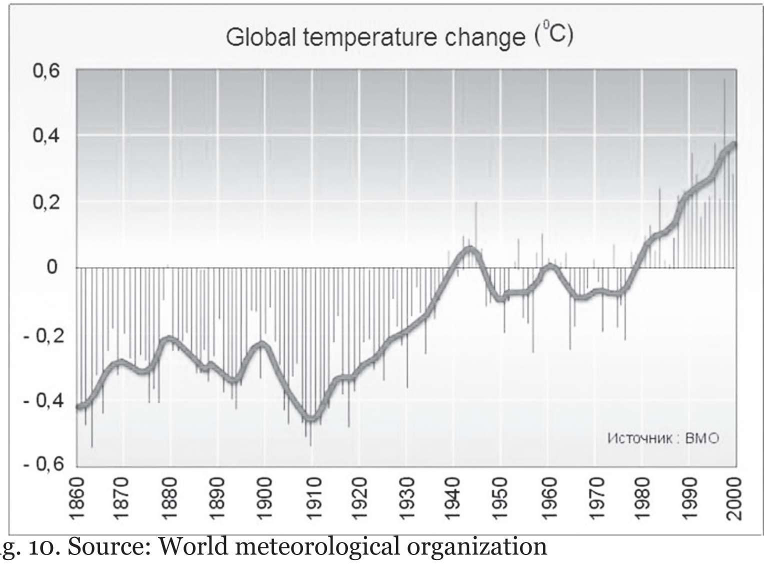


fig. 10. Source: World meteorological organization

Owing to these observations, the solar-terrestrial physics has become quite a mature science to try forecasting consequences of catastrophic occurrences on the Sun. Such forecasts of the 'space weather' become more and more essential: it is necessary to evaluate the radiation hazard for cosmonauts in the orbit and to control telecommunicational and navigational satellites. Unfortunately, there have been many examples when neglecting of such 'weather' conditions resulted in failures of some satellites. A sad example is the *Telstar* communication satellite that was lost during a magnetic storm.

Of course, we are not able to forecast the space weather since the level of our understanding is approximately the same as the level of understanding of the usual weather in the 30s of the last century. The terrestrial magnetosphere is a complex nonlinear structure, and it is very difficult to understand it per measurements in just three-four points as it has been executed so far. We need more measurement points, and consequently, more spacecraft in the near-earth space. Various space agencies including the Russian one plan such multisatellite clusters.

Let us get closer now and look at the Earth itself. Satellites allowed to see our planet as a global ecosystem and to understand better its atmosphere and climate, and observe catastrophic occurrences of typhoons, tsunamis and volcanic eruptions. There is a great number of works dealing with forecasting earthquakes per satellite data, though there is a great number of skeptical opinions in this respect as well. The satellite methods are very important for researching of the Earth's natural resources, ecology, vegetative cover and climate, and for attempts to predict what is coming in the near future.

It is no secret that the Earth's atmosphere average temperature has been increasing. The given diagram shows its history by years, and it is important that the strongest and the most significant changes of the climate is being observed in the North Eurasia (Fig. 10).

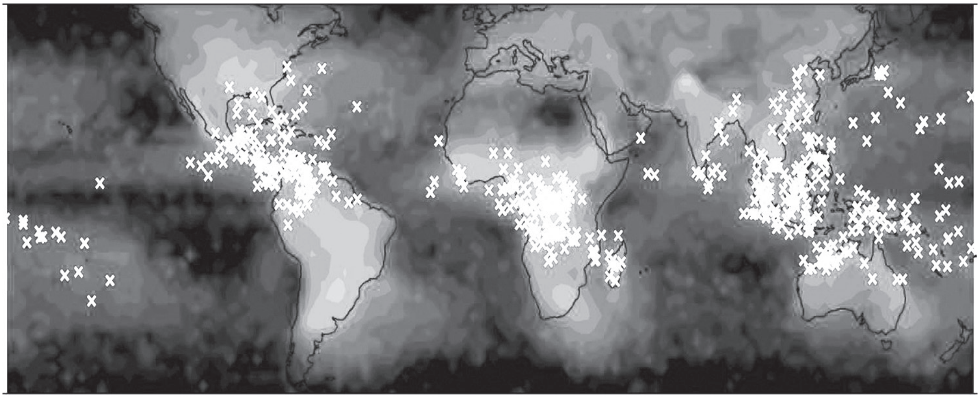


fig. 11. © NASA LIS/OTD science team

What are the consequences of these temperature changes? Are they always negative or it is possible to find positive sides? The average temperature increase results in ice melting and the ocean level rise. This, naturally, allows to use the Northern Seaway and the North-Western Passage in Canada more actively. However, the global warming has an ambiguous effect on the flora. The works on the global satellite monitoring of the Earth's vegetative cover are being already carried out. The satellites allow us to form the systematic approach to the evaluation of these occurrences.

Many other, so to speak, private discoveries, should be mentioned, which became feasible in the Space Era. Unexpectedly for many scientists, astrophysical satellite experiments gave new understanding of the lightning physics. Seemingly, it is a quite old field of research, developed in the 19^{th} century, and it is complicated to discover anything new. But unexpectedly, the American RHESSI astrophysical satellite (Fig. 11), which had not actually been designed to observe the Earth, recorded a great number of powerful flashes of the gamma-radiation taking place on our planet (they were even named 'terrestrial gamma-flashes'). Not long ago, we began to understand their nature better; they occur at unusually high acceleration of electrons in lightning discharges, which have not been predicted by any theories.

This phenomenon has generated great interest at once. A series of satellites is starting to work on its study now. Some of them are already in the flight now, i.e. the French *Taranis* satellite and the Russian *Universitetsky-Tatiana* microsatellite. Currently, the Space Research Institute of the Russian Academy of Sciences prepares the *Chibis* microsatellite to research the problems of atmospheric electricity and abnormally high acceleration of electrons when lightning discharges.

The satellites have allowed us to see one more interesting phenomenon — the magnetic pole movement (Fig. 12). The Earth's magnetic field used to change its direction periodically but it has not occurred for a long time. Probably, we are the witnesses of preparation of the Earth's magnetic field polarity change, and it is a rather hazardous tendency since the Earth, during this process, will temporarily lose the ‘magnetic shield’ protecting it from the solar plasma streams. Such global catastrophe, which is no less hazardous than meteorite blows, is also possible, and we must think about this problem as well.

Satellite geodesy based on the global navigation system has given a great number of interesting results. We managed to explore the Earth’s external gravity field, record the terrestrial crust movements and monitor the changes of the Earth’s rotation velocity with very high accuracy, i.e. 200 times higher than in the pre-satellite era.

Following the rotation of strategically selected points on the Earth surface, we can see that there is a certain tendency. The continents, very slowly, but move, and bringing all the data together makes it clear that the entire Eurasian continent is slowly rotating around a certain point located in the Tibet region. This global picture was formed with the help of satellite observations as well.

Of course, the outer space is an excellent place to carry out plasma-physical researches. The walls always bound plasma researched in laboratories, and the boundary effects are strong there, and therefore, really large-scale processes can’t be investigated herein. In the 80s and especially in the 90s, the feasibility of active experiments in the outer space, such as releases of the plasma beams and observations of their relaxation was widely examined. That was very important for understanding of common regularities of plasma behavior.

By N. Olsen and M. Mandea

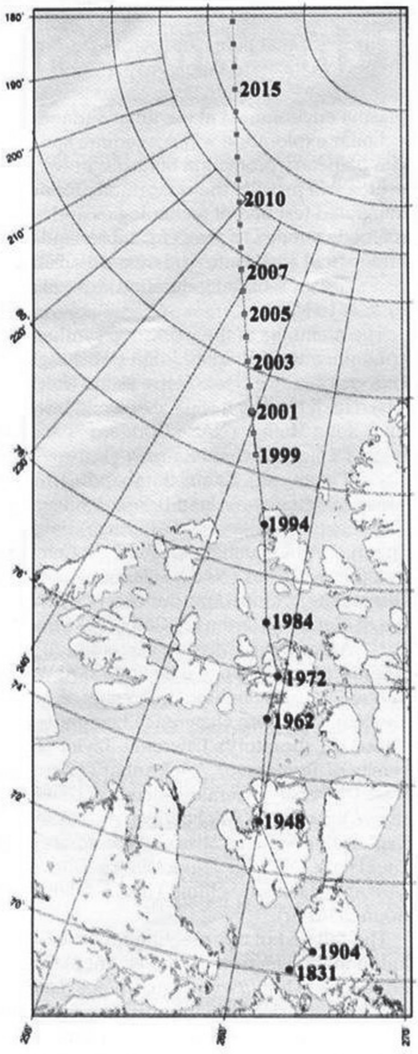


fig. 12

Analyzing these data, we found out many physical phenomena. First of all, it was the reconnection of magnetic fields — the process, which can be observed on the Earth and is responsible for magnetic storms. We have in detail researched this phenomenon it in experiments of the INTERBALL project (Russia) and the *Cluster* project (ESA). Another example is the prediction of collisionless shock waves made by theorists R.Z. Sagdeev and A.A. Galeev. Today this phenomenon is widely researched in the outer space.

During the last years the science of dust plasma has been developed. It consists not only of electrons and protons but of charged dust particles as well. The structures are formed in it much easier than in usual plasma with small weight ratio of the charged particles. Such dust plasma collapses in the Earth's gravity field, and we cannot observe its structure, therefore, experiments in weightlessness are required to study it. During such experiments, executed by Russian cosmonauts, Sergey Krikalyov in particular, a good-looking structure named the 'plasma crystal' can be seen. Such experiments allow to gain a better understanding of new physics of complex nonlinear systems and their dynamics.

Along with it, we learned much that was new about the human beings and the living organisms. Actually, new sciences have been born: the space medicine, the space biology and even the biotechnologies. However, I would like to dwell on international aspects of the Sputnik launch.

The Sputnik produced a violent reaction in the whole world. However, it is interesting that owing to it, the attitude to the Sputnik in the Soviet Union changed greatly as well. The *Pravda* newspaper, dated October 5, 1957, published a report of the News Agency of the Soviet Union only. But the situation became absolutely different in a month. The Soviet Government of that time quickly responded to the fact that the space became not only the factor of science but also the factor of prestige and politics.

Practically, the satellites determined the mode of mutual deterrence between two superpowers, the Soviet Union and the United States, and, what is the most important, transferred the rivalry between them into a different level — the level of peaceful competition. Many of political scientists emphasized the fact that the Sputnik had entirely changed the essence of the 'cold war'. Mstislav Keldysh — the President of the Academy of Sciences of the USSR spoke about the same, i.e. it is unknown, what has become more important for the state defense — a missile or the Sputnik.

The space race, in terms of the game theory, turned out a non-zero-sum game, in which the two parties win. Moreover, what is very important, the Sputnik launch rapidly resulted in international organizations' establishing. The International Committee on Space Research (COSPAR) was established almost in a year. In our country the INTERCOSMOS Council was established. Committees on legal features of space research were established in the UN. In other words, the international 'space' infrastructure existing today was originated.

The space science must necessarily be the international affair. And we are pleased that the soviet Sputnik initiated this process of internationalization of the space researches, and not only of them. I'd like to give an example of a curious incident. There was a famous 'beat generation' in the United States (such people were called 'stilyagi' in Russia), a famous writer Jack Kerouac, a poet Allen Ginsberg and many others belonged to it. After launching of the Sputnik, they renamed themselves in 'beatniks', where the Russian suffix '-nik-' meant for them something communistic, 'left' and, at the same time, that they were somewhere beyond the bounds of the usual world. Thus the Sputnik exerted such unexpected influence upon philology.

Let us get back to the science. Which main problems are to be solved by space researches? I confine myself to brief descriptions. The first one is, of course, the problem of the Universe origin, the Dark Energy and the Dark Matter — recently arizen and still unsolved. The second one is life in the Universe. The third one is the understanding of the Earth as a complex space ecosystem.

I should say that all the techniques, we apply now, are based on great discoveries of the physicists of the 19^{th} century and the beginning of the 20^{th} century. We have faced with the fact that there are such matter and energy densities, which cannot be already generated in earth-based laboratories. The astrophysics and the space research can give us the answer what the structure of the Universe is, what new laws of nature are, which we can apply for our routine tasks with time as it has always happened in the humankind history.

Looking now at the Earth through a big telescope we can see that our Earth is the third planet, which is close to the Sun — one of medium and quite ordinary stars at the periphery of the Galaxy. Now we know that the galactic cosmic rays, which get into the Solar system, the Sun’s radiation, the solar cosmic rays and the solar plasma determine, in many respects, the conditions of our subsistence as well. In other words, we live in one huge space system. And here, we should remember Konstantin Tsiolkovsky again, who intuitively felt already in the 19^{th} century that people must properly realize themselves not simply the inhabitants of the Earth but the inhabitants of the Universe, and gain the ‘space consciousness’ in a sense. This is one of those great foresights of the humankind, which prove to be correct in the long run.

I want to finish with the words of Mstislav Vsevolodovich Keldysh: ‘The humankind has entered the new epoch of acquiring concealed secrets of nature, and the knowledge, which we will get in the deep space, will be used for the improvement of life on the Earth’.

R.-M. BONNET
President
International Committee
on Space Research (COSPAR)
rmbonnet@issi.unibe.ch

SPACE FOR THE FUTURE[1]

The world in the 20th century was shaken twice by this country: the Soviet Revolution in 1917 and 40 years later the launch of Sputnik-1. More time has elapsed since the launch of Sputnik-1 than elapsed between its launch and the Soviet revolution. Difficult to believe! Today, we celebrate an event which has opened the space era, one of the most important anniversaries of our common culture, and has marked the history of humanity and the second revolution of the USSR, only 12 years after World War II, a war which has been so devastating for this country.

The dream of leaving our planet, so naively phrased by Konstantin Tsiolkovsky, the great inspirer of the development of the Soviet space program who in the late 19th century theorized many aspects of rocket propulsion and claimed that *'The Earth is the cradle of the mind, but we cannot live forever in a cradle'*, was made possible thanks to the genius of Sergei Korolev the father of Sputnik-1 and the man who placed the first human being into space, Yuri Gagarin. These people were the first to attempt realizing the dream.

By sending to us its famous *beep-beep*, the first ever signs of an extraterrestrial artificial object and the signs of one of the greatest manifestations of terrestrial intelligence, by beating the Guinness record of velocity, Sputnik-1 will remain one of the most impressive achievements of the genius of this country which together with the flight of Gagarin is a fantastic testimony of its ability to fight the extremes and to face with a legendary courage the adversity and the impossible. I am proud to be with you today and of sharing the emotion which I am sure wrings the hearts and the souls of all of you.

THE LEGACY OF SPUTNIK-1

The legacy of Sputnik-1 is incommensurable. Apart from its profound political message, in opening the third dimension Sputnik-1 has offered the dazzling demonstration that space is accessible and that it represents a nearly infinite reserve of discoveries.

Following the launch of Sputnik-1, numerous rockets and artificial satellites took off from all parts of the world, from Russia again, less than a month later with Sputnik-2 and Laika, from the USA, and also from Sahara, Australia and Japan, in the course of the most stunning scientific 'treasure hunting' of history.

In the past 50 years, we have extensively traveled through the Solar system; we have landed on the Moon, on Mars, Venus, Titan, and on asteroids. Soon, thanks to ESA's *Rosetta* mission, we will land on the nucleus of a comet and plans are that we will soon explore new and potentially dangerous asteroids, the icy moons of Jupiter and return to the Moon and possibly to the Titan. With the

1 Parts of this report are referring to the book *Surviving one thousand centuries. Can we do it?*, co-authored by R.-M. Bonnet and L. Woltjer, PRAXIS-Springer, Eds., publ. in 2008.

Pioneers and the *Voyagers* we are reaching the limits of the heliosphere and explore the virgin territories of deep space. These robots represent our first and very successful attempt to make Tsiolkovsky's dream come true.

By observing the sky from above the Earth's atmosphere which lets only the tiny portion of the visible part of the electromagnetic spectrum go through, we have accessed all the hidden portions of that spectrum: the UV-, the X- and the gamma rays, the infrared, and the sub-millimetric wavelengths. We have started exploiting with an enormous amount of luck and success that nearly inexhaustible gold mine of discoveries that is space astronomy.

The revolution in knowledge and understanding of the Universe thus created has a dimension comparable if not greater than that opened by Galileo Galilei with the use of the telescope. The first discovery of an astronomical X-ray source in 1962 granted the 2002 Nobel Prize in physics to Riccardo Giacconi, and the observation of the black body radiation of the cosmic background in 1989, with the COBE satellite, that of 2006 to John Mather and George Smoot.

We have accessed the most thinkable extremes: extremes of distances, of temperature — from the several million degrees of the solar corona to the near absolute cold of the deep universe — of vacuum, of density and gravity, and of time. We have discovered black holes everywhere confirming in an unprecedented way the prediction of Einstein's theory of relativity, another gigantic revolution of the last century. We have discovered water everywhere in the Universe including in our Solar system, the subsurface of Mars, Europa and the moons of Saturn.

Maybe even more essential has been the revelation that space observations of our planet represent one of the most promising tools ever invented to serve humanity. The most mediatic picture of the 20th century will remain for a long time to come the picture of our blue Earth taken by the *Apollo* astronauts, hanging above the lunar horizon. With no concerns about political barriers and borders, the successors of Sputnik-1 have proven their indispensable role for measuring our globe and its deformations, for observing and forecasting the weather and soon the climate, the melting of ice, the rising of the sea level, the depletion of the ozone layer and the anthropogenic and natural hazards that threaten us more and more.

WHAT IS THE FUTURE OF SPACE?

Fifty years after Sputnik-1, have we discovered all that has to be discovered? The never vanishing power of our space telescopes offers the possibility of observing the faintest objects, those which are either too small or too far away to send us enough light, extending our horizon in the past to only a few thousand years after the Big Bang.

These new discoveries open themselves new questions on the evolution of the Universe, of its content, of the existence of the mysterious Dark Matter and of Dark Energy. Not less fascinating is the prospect of detecting other planets similar to the Earth orbiting other stars and probably, not so long in the future, some signs of life on some of them, responding to the anguishing question of our loneliness in the Universe.

Fifty years after Sputnik-1, we will continue to observe our Solar system with the permanent concern of detecting whether life has developed elsewhere than on the third planet. We will explore Mars extensively, and also the moons of Jupiter and Saturn, we will land again on asteroids and on comets.

Unfortunately, the naïve dream of Tsiolkovsky remains a dream. The farthest distance to Earth Man has ever been is the Moon, at one light-second. The astronauts on the ISS are circling the Earth at a distance much less than that which separates Moscow from St.-Petersburg or a little more than a thousandth of a light-second, and the nearest star is still 4 light-years away, reachable, if ever, in some 50,000 to 70,000 years with the most rapid spacecraft we are able to build presently! In the last 50 years, we have not invented a faster rocket than the *Semiorka* which sent Sputnik-1 in orbit.

In the next 50 years, we will probably go back to the Moon, we will land on Mars, but as the late Hubert Curien often said, not for science but for the pleasure of the sport or just because if one goes there the others will follow. I am personally afraid that beyond the orbit of Mars, the dream may indeed soon become a nightmare for those brave people, naïve enough to dream of venturing these deadly and inhospitable territories, where survival will never be easier than on the Earth, where resources will be finite, distances ever increasing, autonomy and detachment from the Earth the most probable fate. The only perspective offered to them is that of living in a limited vessel with no biodiversity, no culture, no family and probably no food.

Is the dream impossible? I interpret Tsiolkovsky's dream differently than just sending colonies into space. In fact where are they these colonies and these cities imagined by the visionary O'Neil? Still in the realm of fiction! At the time when Tsiolkovsky thought of leaving the cradle of the Earth, a little more than hundred years ago, Einstein had not yet invented his theory, the Universe was static and its age evaluated at a few million years only, the population of the Earth was less than 2 billion people. Hope was unlimited! Everything was possible!

Today the population of the Earth is above 6 billions and the UN forecasts a factor two increase in the next hundred years in one of its most conservative scenarios. The next 100 years will most likely see the exhaustion of oil, of precious resources and of certain metals. Humanity will have to adapt to the most rapid evolution of the climate ever faced. Clearly, the Earth requires to be managed.

Would Tsiolkovsky still dream of leaving the Earth in this context? He might! Many other famous people do presently echo his predictions. Indeed, leaving a planet where the resources are limited and which may become inhabitable if not properly managed or even dangerous, is an alternative to our future that many have envisioned. Others however like astronaut Mark Kelly do think differently. In a recent issue of *New Scientist* he said:

'We will go back to the Moon though we won't set up a permanent colony there. I don't believe we'll ever screw up the Earth so badly that living on the Moon will be easier. The Moon has no atmosphere and no water. How bad would things have to get for it to be easier to live there? Even if an asteroid hits the planet it'll be easier to live on Earth than on the Moon or Mars. That's why we have to look after our own planet — we're stuck on it'.

I do think the same!

SURVIVING BEYOND 2100

Even though we might for a moment assume that some time in the future we may possess the right technologies to colonize the Moon, Mars or other objects, we will not abandon the Earth. The population of the Earth will not be in a worse state than the new Lunar or Martian populations. And who will decide who will be those who will leave Earth and those who will stay? If we could attain that level of technological development, we would also have all the means to control the demographic, technical and industrial development on the Earth and maintain it livable.

This does not imply that we won't live on Mars or on other objects of the Solar system, but just as we inhabit the Antarctica today: for science, resource exploitation or tourism! We will do it "attached" to the Earth, "stuck" to it and not independently or autonomously. We need the mother planet, to where we will continue to return after our space trips.

The Earth is indeed the only space station which is able to feed and maintain in an acceptable level of living its 11 billion inhabitants. Our future requires that this "space station" is maintained in a situation where the limitations are neither disputable nor negotiable. Not accepting this is equivalent to adopting collective suicide as the ultimate fate of humanity. We are not accepting such a fatalistic view. We have no other choice than maintaining our "common space station" in a livable state and in the most bio-diverse environment possible. In fact this is also the necessary condition that will allow humanity to possibly envisage occupying a larger share of space and using the Moon or Mars for possible future expansion.

Colonization of the Solar system will not solve the problem of overpopulation of the Earth. Those who tend to promote such concepts are somewhat irresponsible as they refuse to face the reality of an immediate predicted disastrous future which requires drastic redirection of the global management of the Earth. The Earth deserves more attention than the Moon presently! One is already dead, the other should survive! We are bound to the Earth!

The 21st century is a unique test case for our ability to survive longer. It is the century of globalization and the first to fully confront humanity to the negative effects of blind driving a world where we are soon reaching the limits of growth. The race to growth, presently involving nearly all countries, leads to an unavoidable impasse sometimes this century or in the next one. It is a narrow path in time that all nations must succeed to go through collectively and not individually. If they manage that difficult transition they may have a better chance of surviving longer. This test case should be successful in all circumstances: there is no other choice!

Hopefully, the successors of Sputnik-1 do offer the most promising prospects to solve humanity's existential problems for the near and long-term future:

- For monitoring the solid Earth and its deformation with the most accurate global positioning systems;
- For measuring the sea surface temperature, the rising of the oceans and the sinking of our cities;
- For studying the deterioration of our atmosphere;
- For observing the cryosphere and the melting of the ice caps;
- For weather forecasting and mitigating against the related hazards;
- For avoiding tsunamis and their deadly consequences;
- For monitoring volcanoes and forecasting their eruptions;
- For managing our water resources, our land and our cities;
- For improving health, level of living and...
- ...For knowledge and education;
- For dreaming of other worlds.

CONCLUSION

The management of these measures and means requires a new governance of the world in all domains where the global future is at stake: to orchestrate and fund the development of all indispensable means, in particular the space means, to properly maintain the "common space station" whose permanent evolution and complexity requires continuous re-evaluation.

Such governance does not exist and it will be hard to implement. We have nevertheless the right of thinking that it should exist. May be in 2057! Hundred years after Sputnik-1, our civilization will be a space civilization but not as O'Neil imagined it, but very simply because space will be necessary to possibly envisage that we will be able to occupy Earth for many more centuries to come.

Thanks to Sputnik!

SELECTED RESULTS OF SPACE EXPLORATION IN THE FIRST 50 YEARS

T. E. MOORE
Heliophysics Division
Goddard Space Flight Center
National Aeronautics
and Space Administration
Greenbelt, MD 20771, USA
thomas.e.moore@gsfc.nasa.gov

50 YEARS OBSERVING PLASMAS IN SPACE

We live in a reassuringly heliocentric world these days and have arrived at the International Heliophysical Year after 50 years of the space age. It all began, in space, with the International Geophysical Year of 1957—58, when the USA and the USSR kicked off a vigorous competition to loft clever experiments into space. There we found interesting phenomena such as the radiation belts and the plasmas whose violent behavior gives rise to them, driven by a supersonic blast emanating from the Sun. We soon figured out that plasmas in motion give rise to electromagnetic fields that accelerate a few particles to enormous energies, but though we now have theories, we are still not sure exactly how this works. Here I offer a personal perspective on the discoveries and findings from the first 50 years of space plasma observations, attempting to identify outlines and trends that have emerged and suggesting how future investigations will help us understand our place in the Solar system and the Universe at large. One significant trend has been steady progress in assessing the extent of terrestrial material, from the ionosphere proper to the plasmasphere, to the auroral plasma fountain, and to a fourth (plasma) geosphere that is now known to extend to the magnetopause, as terrestrial material is ablated by the solar wind, expands to fill the magnetosphere, and is carried off downstream in the solar wind.

1. ATMOSPHERE

1.1. ASCENDING THE SPHERES

Since classical scholars identified the states of matter, it has been clear that those forms are stratified by gravitation and their physical properties, especially density, but also compression and shear strength or solidity. In his seminal history of the space age, *The Heavens and the Earth* [1], Walter MacDougall begins by noting that the leap of mankind into the atmosphere and space was indeed a “giant step for humankind” that is comparable in every sense to the leap made by Icthyostega and other aquatic species from the ocean to land, about 400 million years before the present, during the Devonian period, relatively soon after the Cambrian explosion of new species. Life may have begun in semi-solid clays, but it bloomed in the oceans and then climbed back out onto land, leading to evolution of all the mammals, and even some that returned to the sea. The one species that has made a leap into space remains a land-dweller, though one might say that amphibious tendencies are revealed by the style of returns to Earth practiced by the US space program!

This leap into space was convolved with the politics of the ‘cold war’, as we know. It stemmed from a competition among human ideas based on perceived

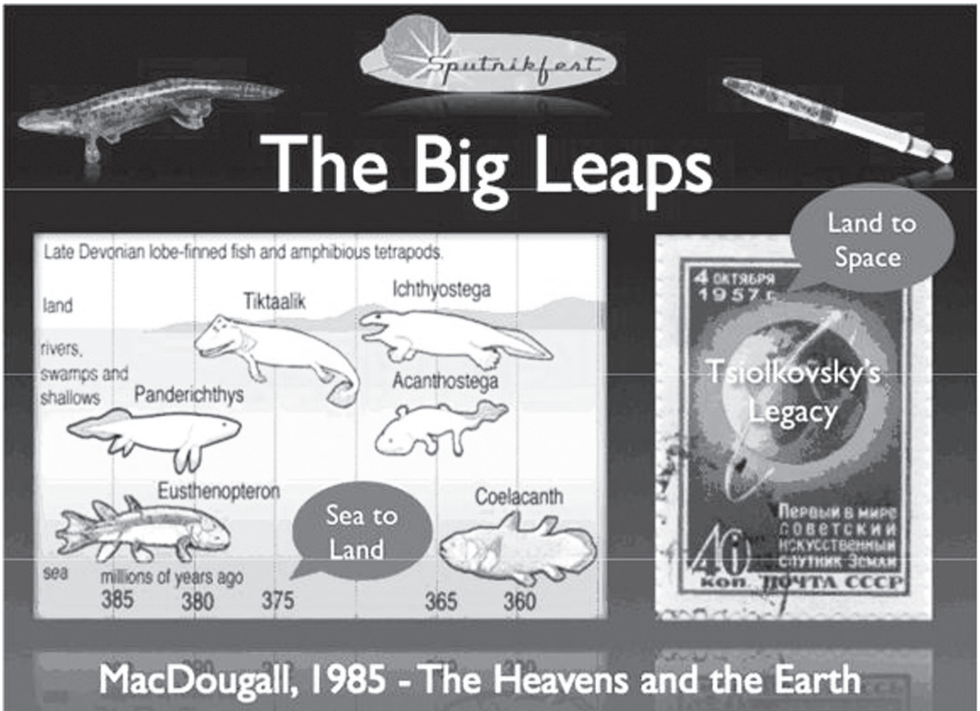


fig. 1. A perspective on the significance of the leap of humankind into space, illustrating the point of MacDougall [1]

defensive advantages. And it forced the USA to adopt centralized government science and technology development practices that were repugnant to many Americans at the time, and which continue to be controversial in the USA. We in NASA clearly owe our agency's existence to the challenge posed by Soviet and Russian space ambitions, which could be met only by a world-class competitor with an organization of requisite size and management complexity, supported by public financing. This competition spurred us onward to the greatest leap of all time, into space. It is fortunate that the early phase has now passed and we have moved on to collaboration among many nations.

1.2. REMOTE SENSING

As the space age began, humans had been thinking about Solar system astronomy for some time using a combination of theory and remote sensing, mainly using optical telescopes. The theory of the solar wind was controversial in detail but there was little doubt that some sort of a solar wind was present outside the Earth's magnetosphere; there was mainly a question about its velocity and intensity. At the same time, ideas about the Earth's ionosphere were in a state of some disarray. In 1953, a Cambridge graduate student named Owen Storey had worked out some ideas to explain the strange noises heard in radios, which appeared to be naturally generated and related to electrical storms in the atmosphere. Storey inferred that a region with a plasma density of 800 electrons/cc at 12,000 km altitude could explain these sounds, called 'whistlers'. But Storey's thesis adviser did not think this was credible, on grounds that measurements from sounding rockets had by then shown that the plasma density falls off steeply with altitude above the top layers of the atmosphere and ionosphere. Such a steep fall of density with altitude was incompatible with the high altitude densities that Storey inferred.

In what follows, I relate the story of plasma observations in space. These were initially direct measurements of density, flow, and temperature, limited in coverage, but effective in validating inferences from remote sensing. More recently, diverse methods for remotely sensing plasmas from space platforms have been developed, as exemplified by the NASA IMAGE mission (2000—2005). Despite this new age of direct space measurements, remote sensing from both the ground and in space remains an important method for gaining global perspective. Remote and *in situ* sensing are now strong partners in the study of space plasma weather.

In recounting our awareness of the extent of terrestrial plasmas, I will follow a path from their source in the ionosphere through the magnetosphere, and argue that we should identify their extent with a 'fourth (plasma) geosphere' or geogenic plasma cloud around the Earth. A final section briefly summarizes some future directions of this field of research.

2. IONOSPHERE

2.1. PLASMASPHERE

In 1959, the USSR launched two *Lunnik* rockets toward the Moon to demonstrate the power of their rocket motors. One of those rockets carried a plasma analyzer. The responsible scientist, Konstantin Gringauz, found that densities up to several hundred per cc indeed extended out beyond 15,000 km altitude, but his results were discounted in USSR science circles, on the same grounds that Storey's results had been.

In parallel with the development of space flight, a network of recording stations for whistlers had been established on the west coast of the USA by Stanford University. A graduate student there, Don Carpenter, found clear evidence that ionospheric plasma did extend out at least to a relatively sharp boundary, through the region that Storey had postulated. Then in 1963, not long after the Cuban missile crisis, Carpenter traveled to a scientific conference in Tokyo, where he encountered Gringauz. There, they compared notes and were delighted to find that their results agreed and therefore provided real substantiation of an extended ionosphere.

Scientists at the meeting, in the USSR, and in the USA all were persuaded by this agreement between a remote sensing technique and a direct measurement in space. The plasmasphere and its reasonably well-defined boundary at approximately 3 to 5 Earth radii (soon dubbed the plasmapause, or 'Carpenter's knee'), became an established feature of the geospace environment. Essentially, the plasmasphere was an upward, light ion extension of the ionosphere. The

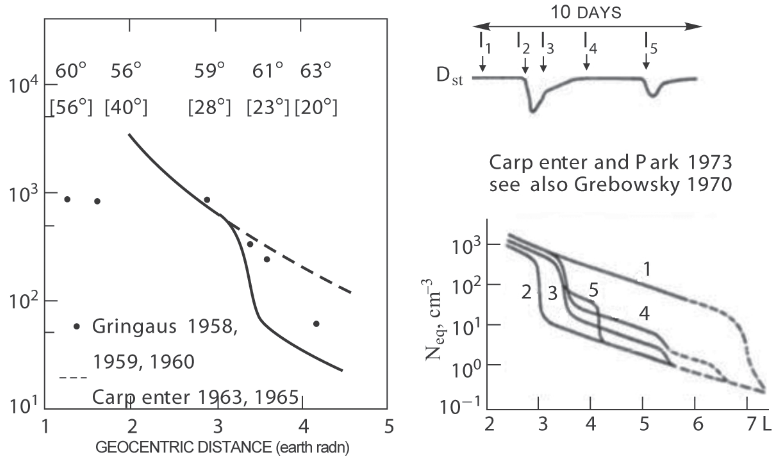


fig. 2. (left panel) Densities as a function of geocentric distance inferred by remote sensing (Carpenter) and measured directly (by Gringauz). (right panel) Variations of the plasma density profile inferred to result from geomagnetic activity, as measured by the storm disturbance (Dst)

presence of a relatively abrupt outer boundary was explained by ideas concerning magnetospheric plasma circulation that were developed by Nishida and Brice in 1967. Fig. 2 illustrates the feature and its variability during geomagnetic activity.

2.2. PLASMASPHERIC PLUMES

As thinking about the topside ionosphere and plasmasphere was developing around 1970, Grebowsky developed a model of the plasmasphere that realistically combined global magnetospheric circulation with ionospheric light ion outflows. The model computed the system response to changes in global circulation strength, with sunward flow through the inner magnetosphere, driven by the combination of dayside and nightside (Dungey cycle) reconnexion. This model showed the expected form of "plasmaspheric plumes" that were predicted to erode the outer plasmasphere, reducing it in size, and transport the removed plasma to the dayside magnetopause, where it would enter the magnetospheric boundary layers at both high and low latitudes.

This model waited for definitive observational confirmation until additional direct and remote sensing observations could be obtained. There were a number of *in-situ* observations of dayside plasmaspheric features reported by Chappell based on OGO-5 (NASA, 1968—72) data, but the global structure of the outer plasmasphere remained somewhat uncertain. The combination of spatial structure and dynamics made it difficult to determine whether there was a continuous plume, or 'detached plasma regions'. With Sandel's EUV imager on the IMAGE mission, a global view of the plasmasphere and its dynamics was achieved, at several minutes resolution. This immediately confirmed the idea that enhanced circulation strips away the outer plasmasphere and transports it to the magnetopause in a plume-like shape. It also revealed considerable structure, with the overall form subsequently winding into a spiral owing to the corotation of the source region plasmas.

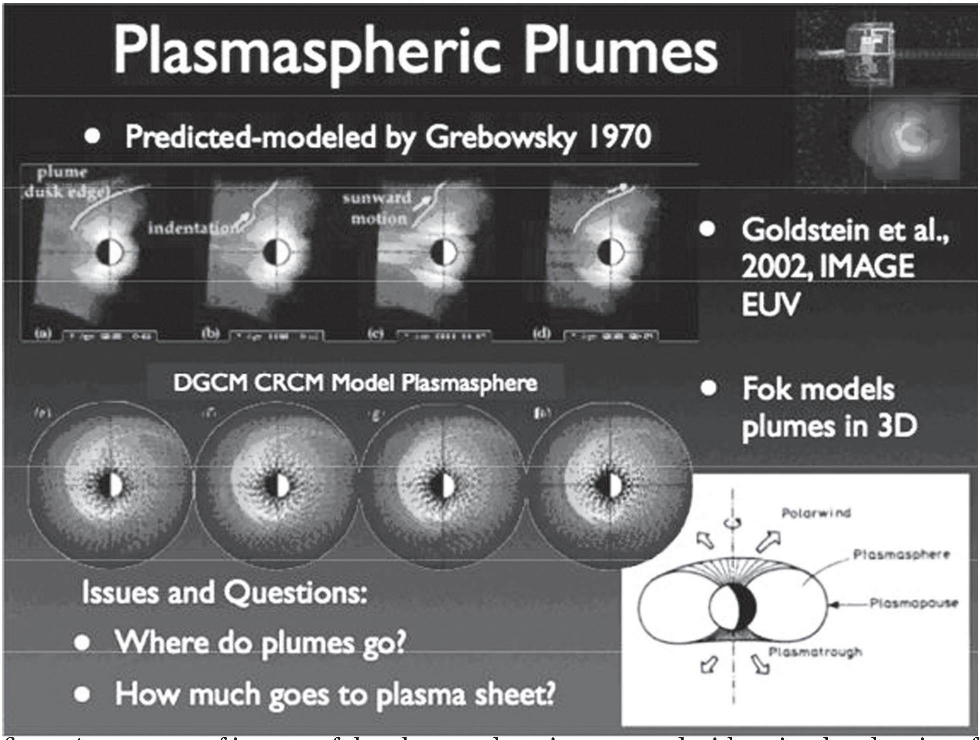


fig. 3. A sequence of images of the plasmasphere is compared with a simulated series of images from an updated plasmaspheric model. Inset shows the latitude relationship of plasmasphere and polar wind, posing the questions shown

2.3. POLAR WIND

Beginning in the late 60s, Ian Axford introduced a concept called 'polar wind' to describe the ionospheric plasma outflow along high latitude magnetic flux tubes, feeding plasma through the magnetotail lobes to the plasma sheet and downstream solar wind. The assumption was that flux tubes were drawn back and stretched out into the magnetotail by the action of the solar wind, and these would contain a negligible back pressure of solar wind plasma, in view of its supersonic antisunward motion. Banks and Holzer argued that a one-dimensional model of field-aligned flows, with a zero pressure boundary condition at infinity, would adequately describe this process. The result was qualitatively similar to the contemporaneous theory of the solar wind, predicting a steady supersonic outflow dominated by the lighter ions with a proton flux of order 10^{12} ions/m^2. The expected flow was a rapid acceleration in the topside ionosphere owing to the ambipolar field in regions where O^+ ions were dominant. As a corollary, similar flows would occur on plasmaspheric flux tubes, but would fill those flux tubes up to an equilibrium pressure. The existence of the polar wind was difficult to confirm at high altitudes, owing to the low plasma density in such regions of supersonic flow, and the consequent tendency of spacecraft to float to a positive potential, excluding the low energy

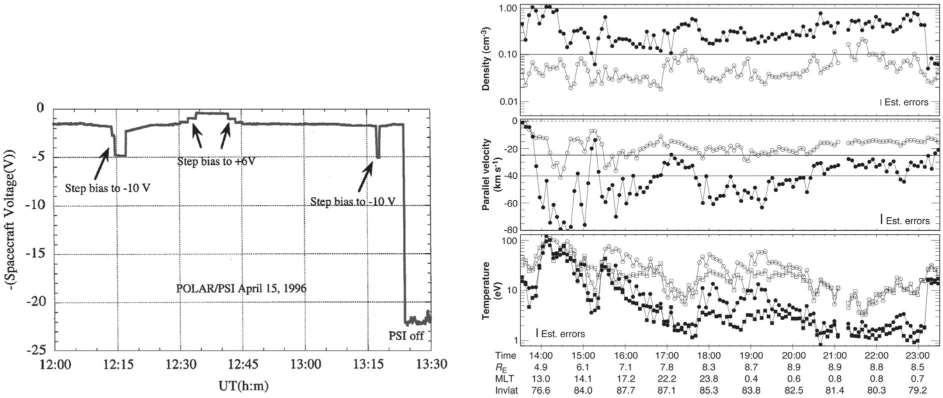


fig. 4. (left panel) The operation of a plasma or ion source on a spacecraft readily eliminates and regulates positive photoelectric charging in sunlight, shown here on the *Polar* spacecraft. (right panel) Observations of proton and O+ components of the polar wind at altitudes up to 9.3 RE was achieved for the first time by TIDE on *Polar*, with PSI potential regulation at ~+2.5 V

ion streams. However, a spacecraft with a neutralizing plasma source finally permitted this confirmation by Moore and coworkers, and the polar wind is now more routinely observed. A review of theory of the polar wind is given by Ganguli [2] and with observations, by Moore and Horwitz [4]. Fig. 4 illustrates the operation of a spacecraft neutralizer, and the resultant capability to observe polar wind characteristics at high altitude.

2.4. AURORAL WIND

Embedded within the polar wind region outside the plasmasphere, auroral processes became the focus of space plasma studies during the 1970s. In particular, auroral systems of current sheets heat and accelerate charged particles through a variety of processes. Early sounding rocket studies of McIlwain and then, Evans suggested that free acceleration of electrons downward into the atmosphere by a magnetic field-aligned electric field is responsible for enhanced auroral light emissions in thin arcs. Soon, observations revealed the corresponding upward acceleration of ions, first parallel to the magnetic field and then mainly perpendicular to it by Shelley, Sharp, and coworkers using instruments on the S3-3 spacecraft. They concluded that discrete bright auroral arcs are produced by a discharge occurring within the Birkeland current sheets coupling solar wind or magnetospheric motions with those of the ionosphere. The discharge was found to be asymmetric, with higher particle energies (and lower ion fluxes) produced in upward current regions than in downward regions. More recently, Carlson's FAST spacecraft (NASA, 1996—97) has delineated the variety of processes and their spatial relationships to magnetospheric current systems with higher time resolution. It also revealed a third type of plasma acceleration that is driven by absorption of Alfvén wave energy radiating into the ionosphere from high

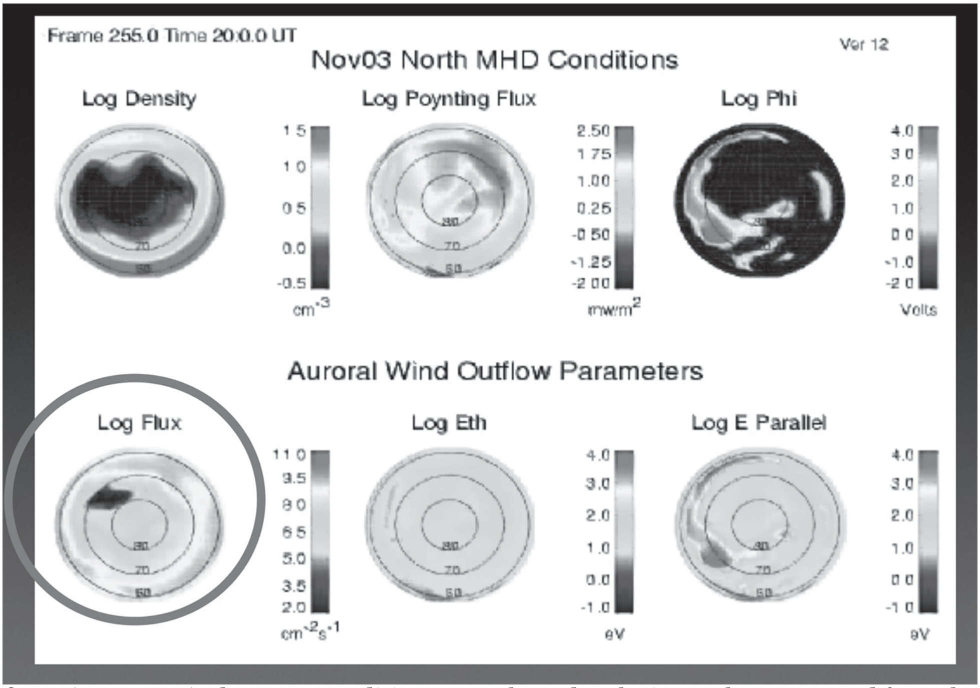


fig. 5. (upper row) The MHD conditions as evaluated at the ionosphere, mapped from the inner boundary of the MHD simulation, evaluated at 20:00 hours into a simulation of the 20 Nov., 2003 storm. From left to right are the hot plasma density, the DC Poynting flux flowing into the ionosphere, and the upward current linking the ionosphere, translated to a parallel potential drop using the Knight-Lyons relationship. (lower row) The resultant O+ outflow characteristics from the Strangeway scalings. From left to right are the local efflux, the ion thermal energy, and the parallel ion energy as given by the current-driven parallel potential drop

altitude disturbances along magnetic flux tubes. The most powerful 'Alfvénic' heating is found in the poleward part of the midnight auroral zone and in the dayside cleft regions. These regions produce the largest upward fluxes of plasma, exceeding 10^{13} $m^{-2}s^{-1}$.

Strangeway used FAST measurements to define empirical relationships between energy inputs and ionospheric outflows, and this was found to at least qualitatively reproducible using *Polar* (NASA, 1996—2008) observations by Zheng and coworkers. Using such empirical scalings, we have estimated the global distribution of outflow driven by an MHD model of the solar wind interaction, and computed the circulation of the auroral wind throughout the magnetosphere. Fig. 5 shows a frame from a simulation of a magnetospheric substorm as simulated within the framework of the LFM MHD code, showing the factors that drive the indicated outflow parameters. A review of auroral wind and atmospheric ablation by solar wind energy is given by Moore and Horwitz [4].

3. MAGNETOSPHERE

3.1. CONNECTING CELLS

As 1980 arrived, Sonnerup showed from ISEE mission data the importance of dayside magnetic reconnexion in driving magnetospheric currents and related processes. The magnetosphere acts like a magnetized cell of plasma that originates from the ionosphere, embedded within the heliosphere and its magnetized solar wind. There is perhaps a viscous interaction between the two plasma cells, independent of their magnetic linkage, but it is relatively weak compared with that which attends the creation of linked magnetic flux tubes by the reconnexion process.

The motions of conducting gases or fluids produce electrical currents that in turn produce magnetic fields. Much as surface tension acts to confine water in droplets or air in bubbles, magnetic tension and pressure act to confine plasmas in magnetic cells. The behavior of magnetic cells is analogous to the behavior of surface tension bubbles, but the magnetic field and its cohesive influence are distributed throughout a plasma cell and are not concentrated at the boundaries, as surface tension is. Magnetic field lines act like a connective tissue of fibers that thread the entire cell of plasma, rather than as a membrane at the outer boundary.

When two magnetic cells collide, magnetic field lines that were initially limited to the respective cells become linked from one cell to the other. Cohesive stresses are thus created that act to accelerate each of the cells toward the velocity of the

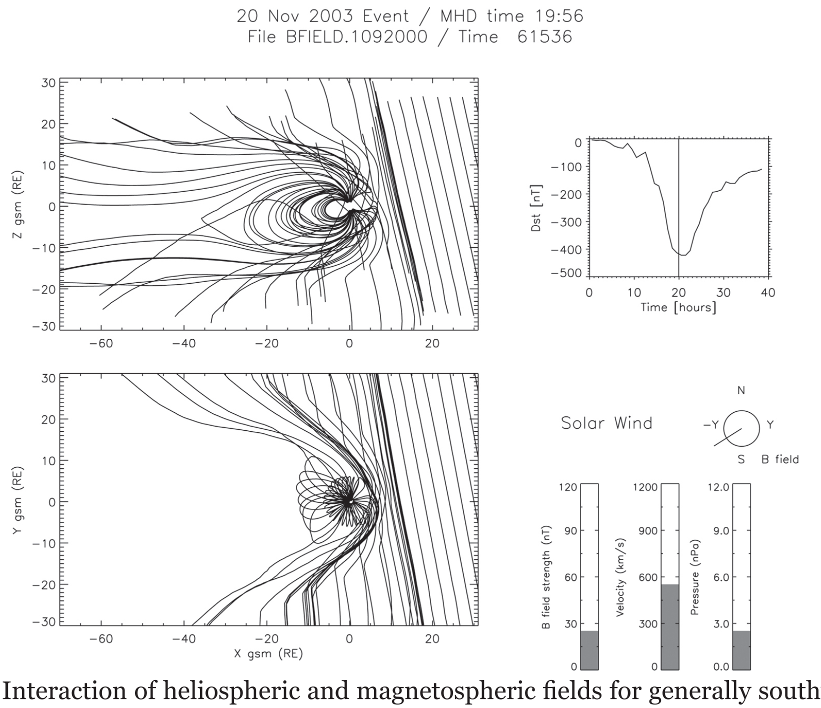


fig. 6. Interaction of heliospheric and magnetospheric fields for generally southward IMF, from the Lyon-Fedder-Mobarry global simulation

other, tending toward a merger and the formation of a single unified plasma cell with properties weighted by the relative contributions of the two merged cells.

Conversely, if any section of a magnetic plasma cell should acquire a velocity relative to the overall cell of which it is a part, the magnetic field may not be strong enough to maintain the integrity of the initial cell. The cell will initially stretch out as the errant plasma attempts to escape from the cell proper. If the amount of momentum possessed by the plasma exceeds the strength of the magnetic field, the cell magnetic field may become so distended by the stretching motion that a separate blob of plasma is formed and the field is pinched off between the two. Field lines connecting the cell with its errant sub-cell are then disconnected so that they no longer link the two and are confined to their respective cells. This behavior is strongly analogous to the joining and splitting of fluid cells confined by surface tension, a notable example being found in the familiar 'lava lamp' that became popular in the 1960s.

We've been discussing discrete cells of plasma and their magnetic fields. But a smaller cell may be embedded within a larger one, and the larger one may stream by at high velocity. This is the situation of Earth's magnetosphere, embedded within the solar wind. The large cell will then tend to engulf and acquire the smaller cell; to pick it up and carry it off downstream, and assimilate it. The solar and geomagnetic fields reconnect so that they are interlinked, and the magnetic forces that are created seek to accelerate Earth's ionospheric plasma up to solar wind speed while exerting drag on the solar wind plasma. Electrical currents form that link the solar wind to the roots of the interlinked field lines, in the auroral zone around each magnetic pole.

The slowed solar wind and accelerated ionospheric plasmas form an errant plasma cell in the tail of the magnetosphere, which episodically pinches off and escapes from the main cell, forming new cells of mixed plasma called 'plasmoids'. These are carried off downstream in the solar wind. The result is a continual ablation of the plasma in the small cell, which is fed with plasma from the sunlit atmosphere and auroral zones. A substantial amount of solar wind is slowed down and incorporated into the magnetosphere, with excess energy being either thermalized or transferred to the ionospheric plasmas. This increases the ionospheric plasma contribution to the cell and the rate of loss downstream. Fig. 6 illustrates the linkage of the two cells as obtained from the LFM simulation at the time of a major geospace storm, during a period of southward and eastward interplanetary magnetic field.

3.2. GEOSPACE STORM PLASMAS

As the solar wind seeks to erode away the plasma and magnetic fields of the Earth and carry them off downstream, the ionosphere responds to the energy dissipated in the auroral zones, releasing O^+ plasma into the magnetosphere. This is a result of the energy dissipation associated with the effort on the part of the solar wind to accelerate and assimilate the magnetosphere, part of which goes into heating neutral gas as well as topside ionospheric plasma. The situation is analogous to a water droplet suspended in a supersonic gas flow, for which the frictional interaction is so intense that it heats the droplet contents, turns them to vapor, and then carries them off downstream. Energy going into the ionosphere and atmosphere cause them to increase in extent and inflate upward against gravity into the magnetosphere.

The magnetosphere responds by developing a stretched out magnetotail, which then episodically relaxes or dipolarizes, at times releasing a plasmoid. The outflowing ionospheric plasmas are then further heated and compressed from the plasma sheet into the inner magnetosphere. The energy dissipation that goes into the magnetospheric plasma clouds (which then contain more material from the ionosphere), raises the pressure of those plasmas to the point that they inflate the magnetic field that confines them. In fact, the amount of pressure contained inside the magnetosphere typically exceeds the driving solar wind pressure by a factor of 10...20. An equivalent way of saying that the plasmas are confined in the magnetosphere, is to observe that they carry a net current in the same sense as the Earth's core, adding to the total dipole moment of the planet and inflating the geomagnetic field. This is the so-called 'ring current', though its ring is often asymmetric or incomplete, and linked to the ionosphere by the 'region 2 currents'.

The ring current was originally discovered by 1917 as a reduction of the magnetic field near the equator of the Earth, by as much as 1...2%. It was inferred that this was caused by a large scale current flowing around the Earth in space by Chapman and Ferraro. The advent of spacecraft measurements led to direct observations of the responsible particles, which are mainly ions in the energy range of 5 keV to 500 keV. Energetic electrons carry at most 20% of the current. Beginning in the 70s, ion composition observations showed that the ring current acquires a substantial component of oxygen ions when it is enhanced. These O^+ ions have certainly come from the ionosphere, illustrating the point made above that solar and terrestrial plasmas mix when the solar wind interaction is very strong. The largest ring currents can consist almost entirely of ionospheric O^+ plasmas, but modest ring currents consist mainly of H+ ions. Fig. 7 illustrates the simulated ring current contributions of solar, polar, auroral, and plasmaspheric winds, for a single isolated substorm event.

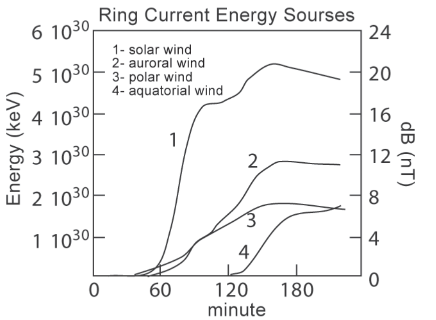


fig. 7. Ring current contributions made by a single isolated MHD substorm, courtesy of M.-C. Fok

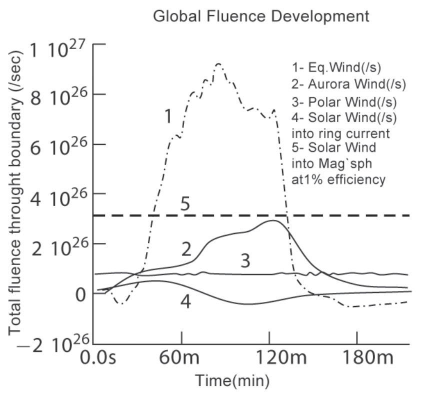


fig. 8. The total fluence of ionospheric ions from Earth, in a global ion kinetic simulation that tracks outflows of the Polar Wind (H+), Auroral Wind (O+), and Plasmasphere (H+)

3.3. PLASMASPHERIC WIND

Recently, we have learned to simulate the dynamic global circulation of plasmaspheric plumes, or what might be termed the 'plasmaspheric wind'. To do this, we have used an inner magnetosphere model of the plasmasphere that derives from the original one of Grebowski, but is updated to run within an inner magnetospheric simulation of Fok and Wolf, which in turn runs within a global LFM simulation. At the outer boundary of the plasmaspheric region near geosynchronous orbit, we initiate proton trajectories distributed along magnetic flux tubes, at all local times. These protons are then tracked until they precipitate into the atmosphere or escape from the simulation space. Each is weighted with a partial density corresponding to the value given by the plasmasphere model. By tracking millions of such particles, we can obtain a distribution of them in each cell of the simulation and compute resultant bulk plasma parameters, as described above. The fluence of particles escaping from the inner magnetosphere is plotted in Fig. 8.

We obtained surprising results from a simulation of plasmaspheric plume circulation for an isolated MHD substorm event of 3.5 hours duration. A very large release of plasmaspheric protons occurred when global convection was enhanced by southward interplanetary magnetic field for about two hours, reaching a maximum of 1×10^{27} ions/sec. As shown in Fig. 8, the fluence of plasmaspheric protons was about an order of magnitude larger than the typical fluence of polar wind protons from the Earth. Despite this large release of plasma escape, the amount that was recirculated back through the inner magnetosphere was smaller, and the amount of ring current that was developed from equatorial wind protons was rather modest and about the same as the ring current developed from polar wind protons. This result can be seen in Fig. 7. It seems that much of the plasmaspheric proton supply was either lost downstream in the solar wind or was ineffectively heated and energized to form hot storm plasmas within the inner magnetosphere.

4. FOURTH GEOSPHERE

4.1. CHARTING THE GEOPAUSE

From the above, it can perhaps be seen that the space age has revealed that the Earth's ionosphere extends much farther into space than expected before the space age began, when it was thought to be a thin layer with a scale height of no more than 100 km. The light ion plasmasphere extends out to 4—5 Earth radii, and even farther during quiet periods. This is no surprise considering the mass and ionospheric temperature of protons. The polar wind permeates the high latitude regions outside the plasmasphere with a low-density but supersonic light ion plasma. The auroral zone energizes the polar wind and adds to it a highly variable auroral wind of O^+ ions (and at times, molecular species), which can greatly exceed the flux of polar wind H^+, leading to substantial densities and pressures of ionospheric heavy ions. The auroral wind is dependent upon the amount of energy dissipated in the auroral zone from the solar wind interaction. Moreover, the auroral wind tends to remain closer to the Earth, in part because

fig. 9. This figure is intended to more comprehensively cite the many observers who have contributed so much toward charting the geosphere, as it has expanded (at least in our awareness) from the topside ionosphere to the limits of the magnetosphere, and downstream solar wind. Names of authors are loosely associated with the regions from which their observations were reported. Note that theoretical ideas have deliberately been omitted to focus on the observations

it is only marginally freed from gravity, but also because oxygen is substantially slower moving for given energy. The auroral wind tends to circulate around the magnetosphere more than once before reaching the magnetopause and being lost into the downstream solar wind. In contrast, the large fluence of ionospheric protons released from the plasmasphere during strong solar wind disturbances expands readily to the dayside magnetopause in plumes that feed such plasmas into the dayside reconnexion regions, leading to escape from the Earth and its magnetosphere.

There is in general a considerable extent of ionospheric terrestrial plasma beyond the ‘plasmapause’. To emphasize this, Dominique Delcourt and I coined the term ‘geopause’ to describe another boundary, lying beyond the plasmapause, between terrestrial and solar plasmas. Though considerable mixing of these two source plasmas clearly occurs, a boundary can always be defined where the densities or pressures of plasmas from the two sources are equal. This concept was evaluated in the first effort to develop a multifluid plasma dynamic simulation of the magnetosphere, able to track separate solar wind and ionospheric plasmas, by Winglee. He found that the geopause could move quite far out in the magnetosphere, especially in the magnetotail, but also in the dayside magnetosphere, when reconnexion drives strong circulation.

More recently, cold plasmaspheric plume plasmas have been found at the magnetopause in substantial densities. These were found by Su and coworkers at geosynchronous orbit when high solar wind pressure moved the magnetopause in that far. Later they have also been found at the more typical position of the magnetopause, by Chandler, Chen, and Moore working with *Polar* mission data and, during 2008, by McFadden and co-workers working with THEMIS mission data. We interpreted some of these observations to show that the geopause approaches the magnetopause at times of strong sunward convection of the plasmasphere, characteristic of highly active periods.

Many individuals have contributed to the charting of the geopause over the space age; too many to fully reference and cite here. To indicate their contributions, Fig. 9 crudely illustrates the regions where individual contributions have added most to our global understanding of the geosphere.

4.2. GEOSPACE STORMS

We have thus learned that terrestrial plasmas often fill the magnetosphere. In the case of O^+, this occurs mainly in connection with strong excitation of the dayside and/or nightside aurora, especially if sustained for the hours-long duration of the magnetospheric circulation cycle. In the case of H^+, this occurs mainly in connection with strong magnetospheric circulation that erodes the outer plasmasphere and carries it to the dayside magnetopause, where the cold plasma is subject to acceleration processes. As we saw earlier, the plasmaspheric wind protons may be lost downstream in the solar wind for

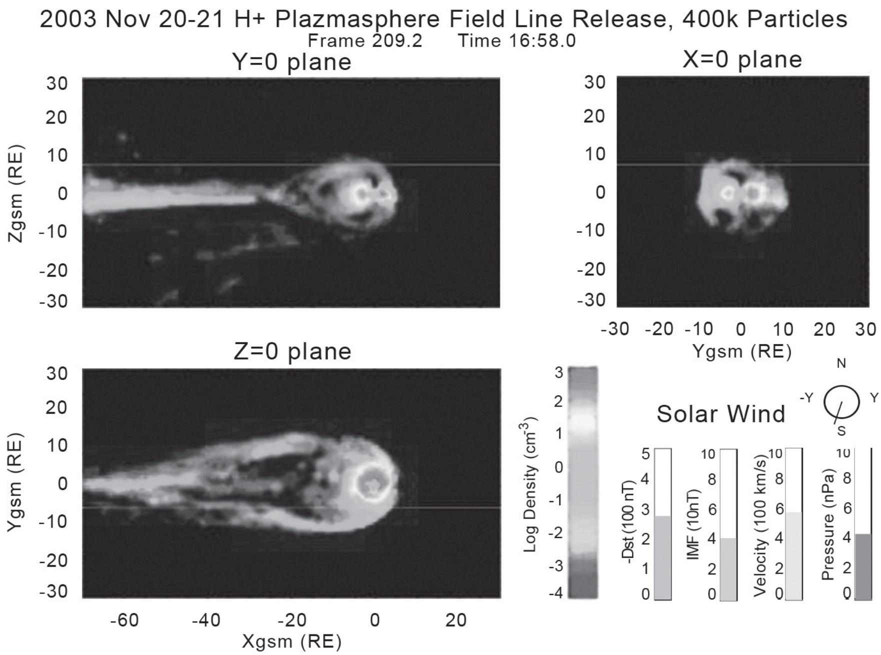


fig. 10. The distribution of plasmaspheric wind pressure in the magnetosphere as constructed using global ion kinetics to track the motion of representative H+ ions in MHD fields from the LFM simulation of the indicated storm. The protons have been released with weightings from the Ober plasmasphere model running inside the CRCM inner magnetospheric convection model, which in turn is driven by the transpolar potential from the LFM global MHD simulation

short periods of activity such as those leading to an isolated substorm. However, plasmaspheric wind plasmas may become more important contributors to inner magnetospheric pressure during full-fledged geospace storms.

As this is written we have completed an investigation of the participation of ionospheric plasmas in a major geospace storm, where processes may be quite different than for isolated substorms. We have performed an initial study of one of the largest of storms that occurred on 20—21 Nov., 2003. The plasmaspheric circulation is illustrated in Fig. 10 near the peak rate of increase of Dst. A publication in preparation will assess the impact of these plasmaspheric wind plasmas upon the ring current.

4.3. THE NEW MAGNETOSPHERE

We have come to know the magnetosphere, over the past fifty years, not as an exclusively magnetic structure, but rather as one that contains plasma pressure and dynamics as significant as that of the heliospheric (solar wind) plasma that drives it. Before the age of space exploration, we thought that the ionosphere had a scale height of <100 km and was negligible above the F-layer. Through a combination of remote sensing and in-situ observations, we learned this was not the case and that the ionosphere extends out to at least 4...5 Earth radii in the equatorial regions, creating a plasmasphere. We also learned that the ionosphere is continuously working to fill polar flux tubes but cannot do so owing to the steady loss of plasma into the downstream solar wind as high latitude flux tubes circulate through day and nightside reconnexion regions. The fluctuating strength of this circulation alternately erodes away the plasmaspheric outer reaches, then allows them to refill when it subsides.

Embedded within the high shear centers of the circulation cells are the auroral zones, which heat and accelerate the ionosphere using energy extracted from the solar wind, but at times stored within the distorted magnetic fields and hot plasmas of the magnetotail. The auroral zones emit a distinct auroral wind of heavier plasmas, mostly O^+ but at times the heating is strong enough and low enough to force molecular ions upward into the magnetosphere. Lighter ions are also heated and accelerated by auroral processes, but their fluxes are set by source strengths and relatively independent of these energetic processes. Nevertheless, their flow paths are affected by low altitude accelerations.

The classical 'magnetic storm' was discovered long before the space age, through remote sensing by means of magnetic measurements on the ground. The presence of a 'ring current' was inferred as necessary to create the observed reduction of field strength at low latitudes. Through studies of plasmas in space, we now know that the ring current associated with such storms stems from a global inflation of the geomagnetic field by the pressure of plasmas trapped within it, that those plasmas provide the charge carriers that carry the required currents, and that those currents are produced by pressure gradients of the plasmas. Since we have come to understand that a magnetic storm is produced by plasma pressure, it seems proper to update the nomenclature and refer to these as 'geospace storms' as suggested by Daglis.

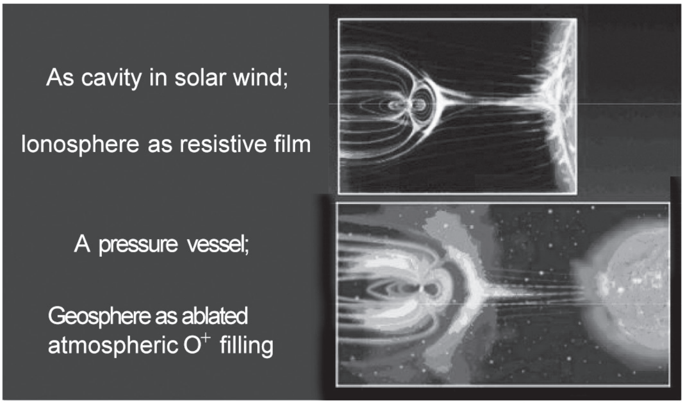


fig. 11. The new magnetosphere: from a resistively coated cavity to a pressurized vessel

The global magnetohydrodynamic perspective has provided a significant improvement over the ‘engineering’ approach to fields produced by a hypothetical current structure. It integrates the role of plasma pressures and resultant flows, which in turn are driven by energy sources. When combined with our new understanding of the role of the ionosphere, particularly in the largest events, we may now refer to a geospace storm as a transient event in which the solar wind heats and ablates the Earth’s upper atmosphere and ionosphere to the degree that it expands to not only fill but also inflate the magnetosphere, as it simultaneously leaks out into the solar wind and is carried away downstream [4]. This emerging view of the new magnetosphere is visualized schematically in Fig. 11.

5. THE FUTURE

5.1. MULTI-FLUID MAGNETOSPHERIC CIRCULATION

One fruitful direction beyond global magnetohydrodynamics is the current move toward global ion kinetic treatment of the plasma. This can be seen in the global test particle simulations described above, similar to the ‘Large Scale Kinetics’ practiced by the UCLA group, and in the bounce averaged Vlasov-Boltzmann approaches taken by CRCM and RAM codes for the inner magnetosphere. Other groups are implementing ambitious global hybrid simulations that are yielding new insights as well, but to date these have been focused on a single species treatment of the solar wind ions.

Multi-fluid circulation models of planetary magnetospheres have been around for some time owing to the importance of satellite or ring sources of plasma in the giant gas planet magnetospheres. The planets themselves are so cold and massive that they tend to hold tight their light atom atmospheres, but recently McComas and coworkers report evidence for H_3^+ ions in Jupiter’s magnetotail. A decade ago, there was only one such multi-fluid simulation of the terrestrial magnetosphere, but in the past year or so, a number of the more established

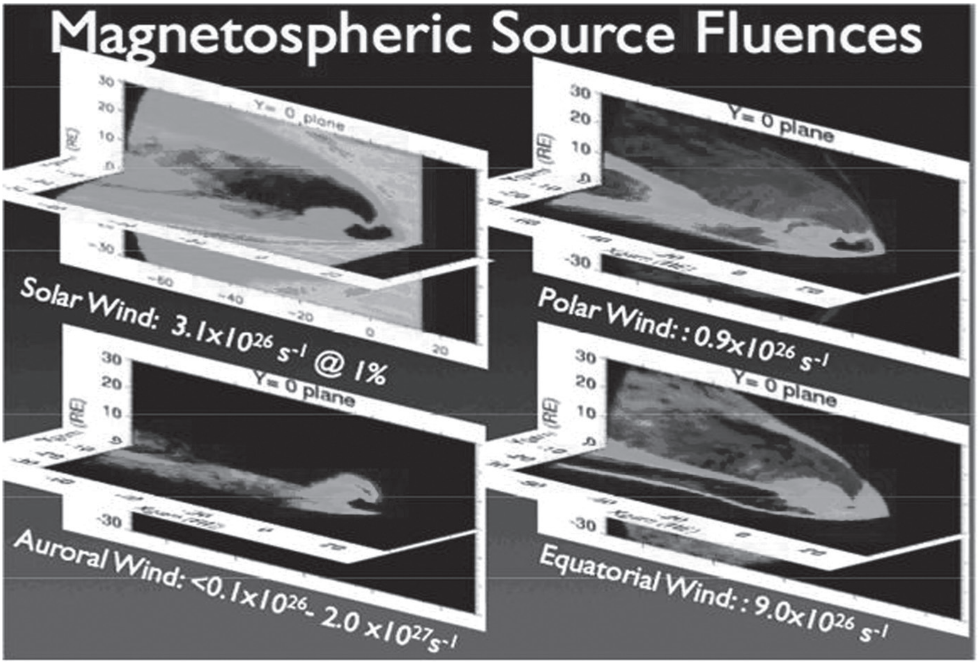


fig. 12. The importance of the various sources of storm time plasmas in the magnetosphere, expressed as fluences into the system

global simulations have been implementing a multi-fluid approach and testing their results against observations. The *Living With a Star* program has during 2007 funded a focused science team effort to study plasma redistribution during geomagnetic activity.

In view of this increased activity, we can expect interesting and significant new results to be forthcoming in the near future. The writer speculates that inclusion of internal sources of plasma will make a significant improvement in the ability of fluid simulations to generate realistic storm time plasmas and their associated ring current. It is commonly said that fluid simulations do not include important physical effects, often referred to as "energy dependent drift effects." And it has been thought for some time that the inclusion of such effects is their principal deficiency. While the full range of physical effects may be lacking in fluid simulations, it seems likely to me that the lack of internal plasma sources is a more fundamental impediment to the reproduction of realistic storm plasma pressures and global signatures such as the ring current. Solar wind energy is certainly responsible for the heating of the terrestrial ionosphere, but the expansion of the magnetosphere must result from the trapping of plasma pressure (up to 10 times the solar wind dynamic pressure) within the closed magnetic structure of the magnetosphere, and this is greatly facilitated for a source internal to that system.

Fig. 12 summarizes the four sources of magnetospheric plasmas, described above. It illustrates with selected simulation movie frames their circulation within the magnetosphere. While it should be stressed that fluence does not map directly to ring current (plasma pressure) impact, all four sources must clearly be considered as contributors to storm effects.

5.2. MULTI-SCALE COUPLING

In space physics, our systems are naturally enormous, yet they consist of the same fundamental components as systems of more familiar scale. Hence, multi-scale problems and cross-scale coupling are of increased importance for space physics. The hypothetical inverse cascade of evolution from micro-structures to meso-scale transport rates illustrates cross-scale sensitivity. Magnetic reconnexion, in particular, is thought to depend upon the coupling of the micro-physics of electron dissipation along an X line, to the global rates of reconnexion, and thus to topological evolution of a much larger-scale system.

Because the geomagnetic field provides strong coupling between regions with vastly different scales (the solar wind to the upper atmosphere), it is inevitable that relevant physical interaction processes must array themselves over a similar range of scales. Moreover, the presence of multiple ion species with widely varying masses per particle leads automatically to another embedded range of scales that must prevail even in the context of a single local structure, for example the reconnexion diffusion region at the dayside magnetopause. Recognition of the multiplicity of scales involved in magnetospheric dynamics has led to a variety of suggested new missions for their study. These include the NASA *Magnetospheric Multi-Scale* (MMS), the JAXA *Scale COupling in the Plasma universE* (SCOPE), the ESA *Cross-Scale*, and the NASA *Magnetospheric Constellation* (MC) missions. The first of these to reach fruition will likely be NASA's *Magnetospheric Multi-Scale Mission*, which has entered development as of Jan., 2008.

5.3. MAGNETOSPHERIC MULTI-SCALE MISSION

The NASA MMS mission represents a natural step beyond the ESA *Cluster II* mission, which pioneered the use of four spacecraft to provide true 3D, space and time de-aliased observations of plasmas and fields in the magnetosphere. *Cluster II* was limited in time resolution to the spin period of the spacecrafts, or about 4 sec. Reconnexion is strongly influenced by processes reflecting the shortest scales in a plasma, the electron and ion inertial or gyration lengths (<10km, and <100 km). The entire structure fluctuates in position at speeds of ~100 of km/s, as at the dayside magnetopause. Thus, a clear view of such structures as they convect past the spacecraft can only be obtained with time resolution of 0.1...1 sec for electrons and ions, respectively. Since no spacecraft could ever be spun so fast when equipped with booms and antennae for field probes, it is clear that the plasma instrumentation must view the entire sky at video frame rates of 30 per sec to resolve the electron scale structures. Moreover, such short exposures will require robust signal strengths, that is, a large aperture for plasma particles. The best way to simultaneously achieve both of these requirements is to distribute multiple sensors around the spacecraft, each of which can sample a slice of the sky.

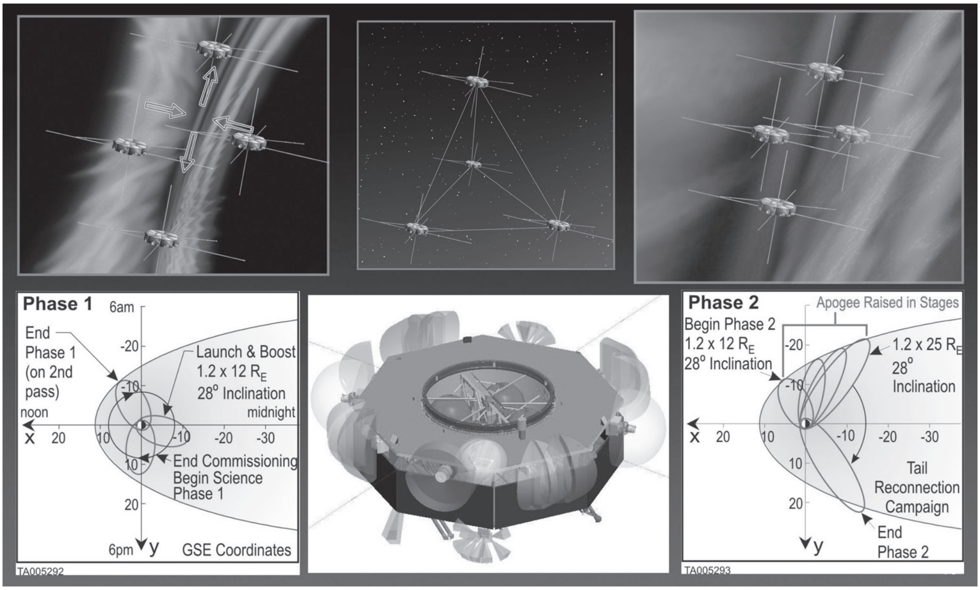


fig. 13. An MMS mission collage with, CCW from bottom left: sampling of the dayside low latitude region, spacecraft layout of plasma fields of view, sampling of the nightside current sheet, schematic mission placement relative to the reconnection X line, tetrahedral orientation and variable spacing

As shown in Fig. 13, this approach has been implemented for MMS, using four sensor boxes for each of electrons and ions, each box containing two sensors with a pole-to-pole fan of look directions, set at a 45 deg. angle to each other so that they both look nearly radially outward. Each aperture has angular response width of approximately 11 deg and is equipped for deflection by ± 5.6 and ±17 deg so that 32 independent apertures each with a fan are spread equally around the spacecraft, with some overlap of pixels near the poles at the spin axes of the spacecrafts.

The four spacecraft are maintained in a high quality (volume) tetrahedral configuration during the active science region of interest, and are varied in separation from 10 km to 400 km, allowing a diagnosis over that range of scales. Two phases of the mission provide for equatorial magnetopause skimming orbits during the season of dayside apogees. In a later phase, the apogee is raised to 25 Earth radii for a nightside passage through the magnetotail current sheet and its active reconnexion region. MMS will provide definitive resolution of the reconnexion diffusion region at electron and ion scales. It will also provide definitive plasma composition and energetic particle distributions at lower time resolution, supporting the study of characteristic ion gyro or inertial scales.

The high time resolution requirement necessitates unprecedented data rates but no downlink relief is available, so that one is essentially operating a security video camera over a dialup modem connection. This in turn requires an intelligent burst data acquisition system with a large data buffer and parallel downlink channels for low resolution survey data, in which frames will contain smeared variations, and for high resolution data transmitted in priority order. Onboard difference imaging and other analyses are used to identify periods that contain phenomena and variations of interest. Moreover, lossy data compression can optionally be used to increase the amount of high-resolution data acquired by a factor of ten or more, with negligible errors, thereby balancing the requirements for spatial and temporal resolution.

CONCLUSION

The history of plasma observations in the magnetosphere has been one of increasing appreciation of the extent and importance of the ionosphere and atmosphere, both as the sink of solar wind energy in the system, and as a responsively variable source of plasma pressure within the magnetosphere. A medium initially thought to be gravitationally confined to a thin layer has turned out to be extended throughout the magnetosphere, reaching the boundary layers and escaping outward into and flowing downstream with the solar wind, especially during times of strong solar wind interaction, when the pressure of plasma within the magnetosphere substantially exceeds the responsible dynamic pressure of the solar wind on the magnetopause, leading to a substantial inflation of the inner magnetosphere and the introduction of additional characteristic scales into a multiscale system.

ACKNOWLEDGMENTS

The author missed the first 16 years of the subject period while being schooled in physics and mathematics, but is indebted to all of his mentors and collaborators over the past 34 years for their memories and perspectives. Some of the mentors bear special mention: David Evans, Ted Speiser, Tom Holzer, Roger Arnoldy, Rick Chappell, Dave Young, Jim Burch, and Joel Fedder. Co-workers who contributed much to this perspective include Dominique Delcourt, Jim Horwitz, Barbara Giles, Mike Chandler, Craig Pollock, Mei-Ching Fok, Michael Collier, Oleg Vaisberg, and Steve Slinker. This work benefited from sustained support from many sources, notably the NASA projects for ATS-6, DE-1/RIMS, *Polar*/TIDE-PSI, IMAGE/LENA, and MMS/FPI. Sustained support from the NASA R&A programs in Heliophysics was essential to the simulation work reported here.

REFERENCES

1. MacDougall W.A. The Heavens and the Earth: A Political History of the Space Age. Johns Hopkins Univ. Press. 1985.
2. Ganguli S B. The Polar Wind // Revs. Geophys. 1996. V. 14. P. 311.
3. Hultqvist B., Øieroset M., Paschmann G., and Treumann R. Magnetospheric Plasma Sources and Losses: Final Report of the ISSI Study, Kluwer Academic Publishers. 1999.
4. Moore T.E., Horwitz J.L. Stellar ablation of planetary atmospheres // Rev. Geophys. 2007. V. 45, RG3002, doi:10.1029/2005RG000194

V.D. KUZNETSOV
Pushkov Institute of Terrestrial Magnetism, Ionosphere, and Radio Wave Propagation
Russian Academy of Sciences
Troitsk, Russia
kvd@izmiran.ru

SPACE RESEARCH OF THE SUN

INTRODUCTION

The advent of the Space Era opened a new stage in the study of the Sun and the impact of solar activity on geospace. The generally adopted concept of space weather controlled by active events in the Sun requires a detailed investigation of the Sun and forecast of solar activity. The influence of space weather on various aspects of human life increases with the development of high technologies, such as GPS, GLONASS, satellite communications, etc. In the context of future manned flights to the Moon and Mars, the forecast of space radiation becomes vital. The problems associated with the impact of solar flares and ejections on the ionosphere, radio communications, spacecraft equipment and orbits, weather and climate, technological systems, and biosphere also remain important.

Besides the applied aspect of the solar-terrestrial relationships, the study of the Sun has a great fundamental value. The Sun is an ordinary star as many in the Universe, but it is the only one we can investigate in detail by observing its atmosphere and detecting its direct radiations and particle fluxes. A lot of problems still remain unsolved, such as the physical model of the solar cycle, corona heating and solar wind acceleration, origin of coronal mass ejections, mechanisms of particle acceleration and transport, etc.

Solar observations from space gave birth to a new discipline — solar extra-atmospheric astronomy. Besides the optical observations, it involves also the UV and X-ray measurements that are impossible to make from the Earth, which increase significantly the bulk of information on processes and phenomena at the Sun.

For 50 years of the space era, the solar studies on spacecraft have made enormous progress (see Table 1, p. 61). These were usually integrated studies that combined observations of the Sun and simultaneous measurements of heliospheric disturbances in the solar wind (e.g., SOHO mission — *Solar and Heliospheric Observatory* [1]) and near-Earth space (e.g., CORONAS-F mission [2]). For a long time, the Sun could be only observed from the Earth, while the study of many solar phenomena (e.g., solar ejections propagating towards the Earth, polar regions, etc.) required simultaneous observations from at least two spacecraft or from outside the ecliptic. The idea of using heliocentric and out-of-ecliptic orbits due to their advantageous position with respect to the Sun-Earth line was suggested by NASA [3] and was realized in a number of space missions, such as ULYSSES (out-of-ecliptic orbit) [4], SOHO (orbiting around the libration point L1 of the Sun-Earth system) [1], STEREO — *Solar Terrestrial Relations Observatory* (two satellites offset from one another with respect to the Sun-Earth line, one being placed ahead of the Earth in its orbit and another behind it) [5], and a series of projects to be carried out within the frames of the *International Living With a Star* program (ILWS) [6].

This paper contains a review of some basic results of spacecraft studies of the Sun obtained for the recent years. Information is also provided on the solar space projects under preparation. All missions listed in Table 1 are divided into four parts: recently completed missions, current missions, missions under preparation, and planned missions. The launch dates (operation terms for the completed missions) given in the table are important in the context of the phases of the 11-year solar cycle (cycle 23 — 1996—2007, maximum in 2001).

1. MAIN SOLAR SPACE MISSIONS OF 1990—2000

Celebrating the 50th Anniversary of the Space Era, we can assert that we are having the 'Golden Age' of solar physics in space. The most important results obtained in heliophysics for the recent decades are mainly associated with spacecraft. A fleet of up-to-date solar space observatories, such as *Yohkoh*, SOHO, TRACE, CORONAS-F, RHESSI, STEREO, *Hinode*, etc. has observed and continues detailed observations of the Sun in different phases of solar activity. Discoveries have been made that give us a deeper insight into how the Sun works, what accounts for the variety of solar active phenomena, and how we can improve the forecast of solar activity in order to prevent its damaging effect on the Earth.

Let us describe briefly these solar space missions.

Stage	Name	Space agency
Completed	*Yohkoh* (*Solar-A*) (1991—2001) CORONAS-F (2001—2005) ULYSSES (1990—2008)	ISAS, NASA, UK Roskosmos ESA, NASA
Current	SOHO (1995) TRACE(1998) RHESSI (2002) *Hinode (Solar-B)* (2006) STEREO (2006)	ESA, NASA NASA NASA ISAS, NASA, UK NASA
Under preparation	CORONAS-PHOTON (2009) SDO (2009) SOLAR PROBE+ (2015)	Roskosmos NASA NASA
Under development	SOLAR ORBITER (>2014) INTERHELIOPROBE (>2014) SENTINELS (2017) SOLAR POLAR IMAGER (>2012) *KuaFu* STEREOSCOPE PEP SMESE (2012)	ESA Roskosmos NASA NASA China Russia Russia China, France

Table 1. Stages of realization of the recent solar space projects as for March 2008. The operation periods (for the future projects — the scheduled launch date) are given in brackets

First of all, it is necessary to mention the *Yohkoh*[2] mission, which was completed more than seven years ago and gave start to a series of outstanding solar projects. *Yohkoh* was a joint experiment of Japan, USA, and UK. Its four devices were investigating the high-energy activity associated with solar flares and the global structure of the solar corona. Two devices were the Soft X-ray Telescope (SXT) for corona imaging (cooperative Japanese-US experiment) and the Bragg Crystal Spectrometer (BCS, designed jointly by UK, USA, and Japan). The other two were Japanese instruments: the Hard X-ray Telescope (HXT, extended Fourier synthesis) ensured a significantly better sensitivity and resolution than the imaging X-ray detector flown earlier and the Wide-Band Spectrometer (WBS) realized the monitoring of high-energy X-ray and gamma-ray emissions. Many types of observations were first practiced on *Yohkoh*. The world first space-borne X-ray CCD camera controlled by computer was used in the SXT. The HXT provided the world first X-ray images of the Sun at high energies (> 30 keV). For the first time, the survey of solar X-rays was conducted continuously for ten years.

For more than ten years of operation, *Yohkoh* provided an enormous bulk of data. The images obtained demonstrate the dynamics of the solar corona with rapidly changing coronal magnetic features and corroborate the occurrence of solar flares as a result of magnetic reconnection. The reconnection effects were first observed in solar flares and active regions (AR). Brightening effects were detected in active regions with the occurrence rate depending on their energy by analogy with the similar distribution in flares. Discovered were also X-ray jets — explosive phenomena in the corona having their origins in the energy release by reconnection, X-ray manifestations of coronal mass ejections (CME), formation of giant arcades inside and outside the active regions, X-ray dimmings (darkenings), etc. The *Yohkoh* observations made it possible to ascertain the conditions of occurrence of CME-type eruptive events in AR. It was established that large active regions with S-shaped X-ray loops tend to produce such events. This result is of great importance to the space weather forecast. The *Yohkoh* images of the Sun for the period of approximately one 11-year cycle revealed that the coronal brightness was changing significantly (enhancing and decreasing) depending on the number of ARs. The wealth of data provided by *Yohkoh* was used successfully to develop the space weather forecast [7, 8].

The main objective of the CORONAS-F[3] mission was a complex study of solar activity and its manifestations in the near-Earth space [2, 9—12]. The CORONAS-F scientific payload involved both the facilities for remote observations of the Sun and a solar cosmic ray complex (SCR). The former comprised the DIFOS spectrophotometer, solar X-ray telescope (SRT-K), X-ray spectroheliograph (RES-K), DIAGENES spectrophotometer, RESIK X-ray spectrometer, solar spectropolarimeter (CPR-N), flare spectrometer (IRIS), GELIKON gamma-ray spectrometer, X-ray spectrometer (RPS-1), amplitude-time spectrometer (AVS-F), solar UV radiometer (SUFR-Sp-K), and solar UV

2 http://solar.physics.montana.edu/sxt/;
http://hesperia.gsfc.nasa.gov/sftheory/yohkoh.htm

3 http://coronas.izmiran.ru

spectrophotometer (WUSS-L). The SCR complex consisted of the cosmic ray monitor (MKR), energy and ion composition spectrometer (SKI-3), and solar neutron and gamma-ray spectrometer (SONG). These instruments were used for measuring high-energy particles of solar and magnetospheric origin along the spacecraft orbit (500 km initial altitude, 83° inclination). The operation period of the CORONAS-F mission (2001—2005) coincided with the maximum and decline of cycle 23 when a series of outstanding events were recorded in the Sun, which allowed a comprehensive study of the solar-terrestrial coupling [13].

ULYSSES[4] was orbiting the Sun in heliocentric, out-of-ecliptic trajectory with the inclination of 80° to the solar equator, perihelion distance of 1.3 AU, aphelion distance of 5.4 AU, and orbital period of 6.2 years, which was reached through the encounter with Jupiter. The spacecraft was passing over both poles of the Sun at a distance of about 2 AU. The south pole was passed over on Sept. 13, 1994; Nov. 27, 2000; and Feb. 7, 2007; and the north pole on July 31, 1995; Nov. 13, 2001; Jan. 11, 2008 (these are the dates when the highest heliolatitude of 80.2° was reached). The mission was completed in March 2008. Its main objective was to observe the high-latitude heliosphere previously unexplored from space and to study its cycle variations. The scientific payload of the ULYSSES mission comprised nine instruments: a magnetometer, the solar wind plasma and ion composition instruments, a low-energy ion and electron detector, a cosmic ray and solar particle detector, a unified radio and plasma wave instrument, a cosmic dust detector, an interstellar neutral gas detector, and a gamma-ray burst detector. The data obtained with these instruments made it possible to study a broad range of the events associated with the solar wind, heliospheric magnetic field, solar radio bursts and plasma waves, solar and interplanetary energetic particles, galactic cosmic rays, interstellar neutral gas, space dust, and gamma-ray bursts [4].

The SOHO[5] mission is the most outstanding solar space project for the whole history of spacecraft observations. It has been operating by now for 13 years and provides rich information on the Sun from the interior and hot dynamic atmosphere to the solar wind and its interaction with interstellar medium. Staying at the libration point between the Sun and the Earth, SOHO is able to observe the Sun continuously and to measure the solar wind streams and high-energy particle fluxes as they approach the Earth's magnetosphere. The studies based on the SOHO data committed a revolution in our understanding of the Sun in the spheres that form the main target of the SOHO mission: the structure and dynamics of the solar interior, heating and dynamics of the solar corona, and acceleration and composition of the solar wind.

The SOHO mission is equipped with unprecedented set of unique instruments: the GOLF, VIRGO, and MDI devices for global oscillations; the SUMER and CDS ultraviolet spectrometers; an extreme ultraviolet imaging telescope (EIT); an ultraviolet coronograph spectrometer (UVCS); a large-angle and spectrometric coronograph (LASCO), a device for measuring the solar-wind

4 http://ulysses.jpl.nasa.gov/; http://helio.esa.int/ulysses
5 http://sohowww.nascom.nasa.gov/

anisotropy (SWAN), a charge, element, and isotope analysis system (CELIAS) for measuring the solar-wind ion composition, a comprehensive suprathermal and energetic particle analyzer (COSTEP), and a device for measuring energetic and relativistic nuclei and electrons (ERNE). A number of important discoveries have been made with these instruments. Images of the subphotospheric flows and activity on the far side of the Sun were obtained for the first time, and highly dynamic solar atmosphere and transport of magnetic energy up to the corona were observed. The sources of fast solar wind were identified, and an extensive database on solar phenomena was built up. The SOHO data contributed significantly to our knowledge of solar-terrestrial coupling and improved the space weather forecast. A pleasant surprise was the discovery of great many comets passing in the circumsolar space.

The telescope satellite TRACE[6] (*Transition Region and Coronal Explorer*) placed in the solar synchronous orbit is exploring solar magnetic fields in the optical, UV, and hard UV ranges with a high spatial (1 arcsec) and time resolution. It provides information on 3D structure and time evolution of the magnetic field; its response to photospheric motions; time variations in the fine coronal structure; and topology of thermal features in the corona and transition region. The TRACE instruments are designed to take measurements in narrow UV and extreme UV spectral bands that contain the emission lines formed in the chromosphere, transition region, and corona. This ensures simultaneous observation of different temperature domains in the Sun and reveals the relationship between the fine-structure elements in the photospheric magnetic field and large-scale features in the corona. The TRACE unique high-resolution data have corroborated the permanently dynamic state (discovered earlier on *Yohkoh*) of the solar corona, which responds to impulses from the photosphere.

The RHESSI[7] mission (*Ramaty High-Energy Solar Spectroscopic Imager*) is operating in a circular orbit with the altitude of 600 km and inclination of 38°. Its objective is to study the particle acceleration and energy release in solar flares, localization of released energy, secondary effects, and ejection of beams of accelerated particles to the dense layers of the solar atmosphere. Nine onboard detectors ensure various types of observation of solar flares and related radiations and phenomena: hard X-ray and gamma-ray imaging of solar flares, high-resolution spectra of the flare-generated X-ray and gamma-ray emissions, spectroscopic flare imaging in hard X-rays, etc.

The *Hinode*[8] mission in the Sun-synchronous polar orbit has for its objective the study of the solar activity mechanisms with the aid of three advanced instruments — a solar optical telescope, an X-ray telescope, and an extreme-ultraviolet imaging spectrometer. The photosphere and the chromosphere are observed simultaneously. The main targets of the mission are the generation, transport, and decay of solar magnetic fields; transfer of energy of magnetic-field variations from the photosphere to the corona; associated variations in

6 http://trace.lmsal.com/
7 http://hesperia.gsfc.nasa.gov/hessi/
8 http://solarb.msfc.nasa.gov/

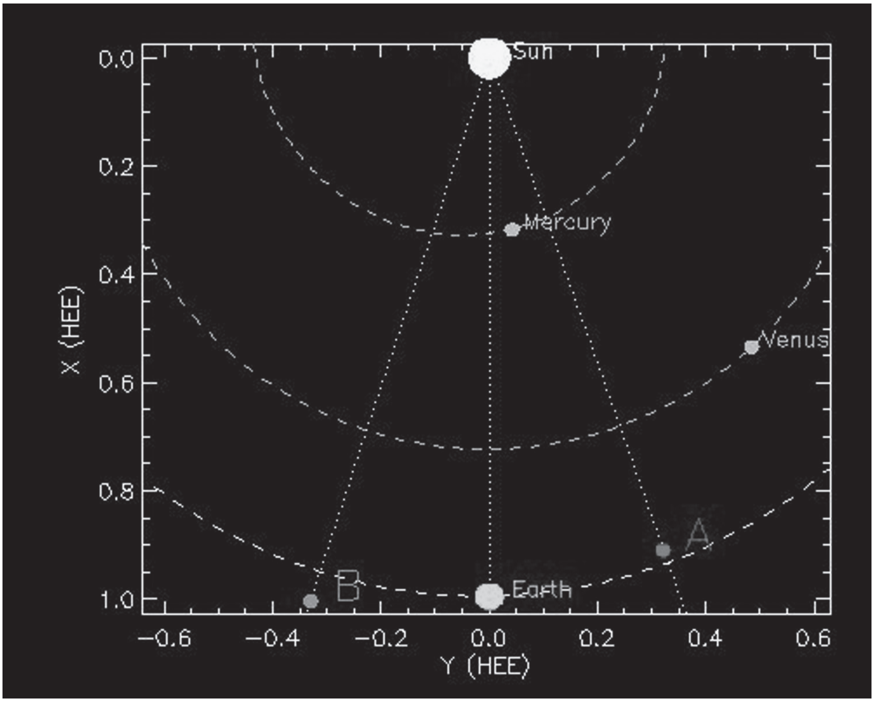


fig. 1. Orbital positions of two satellites of the STEREO mission with respect to the Earth on March 20, 2008. © NASA (http://stereo.gsfc.nasa.gov/where.shtm).

the coronal dynamics and structure; and dynamics of interplanetary medium — ejections and solar wind [14].

Two satellites of the STEREO[9] mission placed ahead of and behind the Earth in its orbit ensure observation of the Sun from outside the Sun-Earth line and, when operated simultaneously, a stereoscopic survey of the solar atmosphere and ejections propagating towards the Earth [5, 15, 16]. For the first time, the Sun and heliosphere are studied in three dimensions. Both satellites are gradually moving away from the Earth in opposite directions up to the distance of 20...30° from the Sun-Earth line. Their position on March 20, 2008 is shown in Fig. 1. Two satellites of the mission expand the heliolongitude range of vision of the solar disk, which increases the advance time of the forecast of space weather at the Earth orbit. The task is to register the commencement of the ejection and trace its propagation until encounter with the Earth's magnetosphere. This must serve to sophistication of the geomagnetic forecast techniques. The hitherto available images did not display the front of the solar disturbance all along its propagation to the Earth, so that the onset of the magnetic storm could only be estimated to within half a day. The heliospheric telescope of the STEREO mission makes it possible to trace the disturbance front from the Sun to the Earth and to predict its arrival within two hours. Alongside with these tasks, in-situ measurements of the solar wind and high-energy particles have been made, diagnostics of ejections

9 http://stereo.jhuapl.edu/; http://www.nasa.gov/mission_pages/stereo/main/index.html

has been carried out, and particle acceleration mechanisms have been studied. The scientific equipment of both satellites is identical and comprises the SECCHI complex (Sun—Earth Connection Coronal and Heliospheric Investigation) consisting of a hard UV telescope, two white-light coronographs, a heliospheric telescope; the SWAVES (STEREO/WAVES) device for tracing interplanetary radio bursts; the IMPACT complex (In-situ Measurements of Particles and CME Transients) for measuring energetic particles and local magnetic-field vector; and the PLASTIC device (PLAsma and SupraThermal Ion and Composition) for measuring the solar-wind ion composition and CME.

The data from SOHO, *Yohkoh*, TRACE, RHESSI, CORONAS-F, and other space missions were analyzed together and provided a series of important scientific results. The ground-based observation facilities were also used with the utmost efficiency. More than 1142 coordinated observations were carried out within the frames of the SOHO project alone. In 373 of them, ground-based facilities were involved; 110 observation campaigns were carried out in cooperation with *Yohkoh*, more than 442 in cooperation with TRACE, etc.

2. BASIC RESULTS OF THE SOLAR SPACE RESEARCH

2.1. HELIOSEISMOLOGY AND INTERIOR OF THE SUN

The mechanism of the 11-year periodicity of solar activity resides in the solar interior, which is the subject of helioseismic studies. MDI (Michelson Doppler Imager) is one of three helioseismic instruments onboard the SOHO mission. It provides high-quality data on global oscillations of the Sun. These data have been used to obtain profiles of the solar differential rotation (Fig. 2), to detect dramatic variations of plasma velocity in the solar interior, and to reveal a shear layer at the bottom of the convection zone (tachocline) at the latitude of about 60° [17]. The latter is of particular interest, because that is where the solar

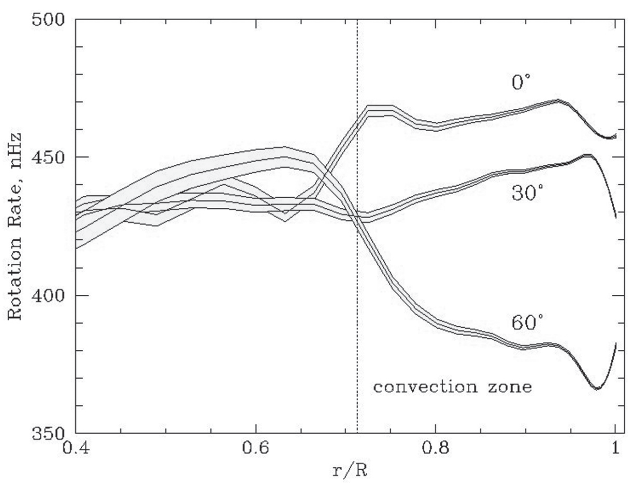


fig. 2. Radial profiles of the angular rotation rate of the Sun at different heliolatitudes (SOHO) [17]

dynamo determining the solar activity cycle is supposed to act. In this region, the rotation profile changes drastically. In the vicinity of the equator, the outer layers rotate faster than the inner ones, while at mid-latitudes and at the poles the situation is opposite.

Fig. 3 illustrates the solar rotation and flows near the poles as inferred from the SOHO/MDI data. The left-hand part of the figure shows the difference in the rotation rates of various regions at the solar surface. Such a pattern based on the SOHO data extends down to about 20,000 km. Sunspots generated by disturbances of the subphotospheric magnetic field tend to occur at the edge of these bands. The sectional view in the right-hand part of Fig. 3 represents the rotation rate in the solar interior. The large dark band under the solar equator is a massive fast flow (rotation) of hot plasma. Very weak flows have been revealed right at the poles. The lines in the right-hand part of Fig. 3 show the surface flow from the solar equator to the poles, which, according to SOHO observations, extends to the depth of at least 26,000 km (4% of the solar radius) [18] and is likely to be an important factor in the solar dynamo, although the speed of the flow (10...20 m/s) is small compared to the chaotic motions on the surface (1 km/s). The reverse flow, shown at the bottom of the convection zone, is expected to exist in accordance with the model approximation, but has not yet been observed.

fig. 3. Solar rotation and flows in the polar regions as inferred from the SOHO/MDI data. Please, for coloured image refer to [18]

Unique 3D images of the solar interior and a pattern of large-scale sub-surface dynamics with significant cycle changes have been obtained by new investigation techniques, such as the time-distance inverse methods, helioseismic holography, and analysis of ring diagrams based on the SOHO/MDI high-resolution data. The zonal flows in the convection zone play an important part in the organization of solar activity, since the active regions tend to appear in the shear layers between the fast and slow flows. New branches of the zonal plasma flows at high latitudes detected with the MDI instrument in 2002 indicate the beginning of the next activity cycle prior to the appearance of the first sunspots. As shown by observations, these new evolving zonal flows originate deep in the convection zone. Their generation mechanism is still unknown [19].

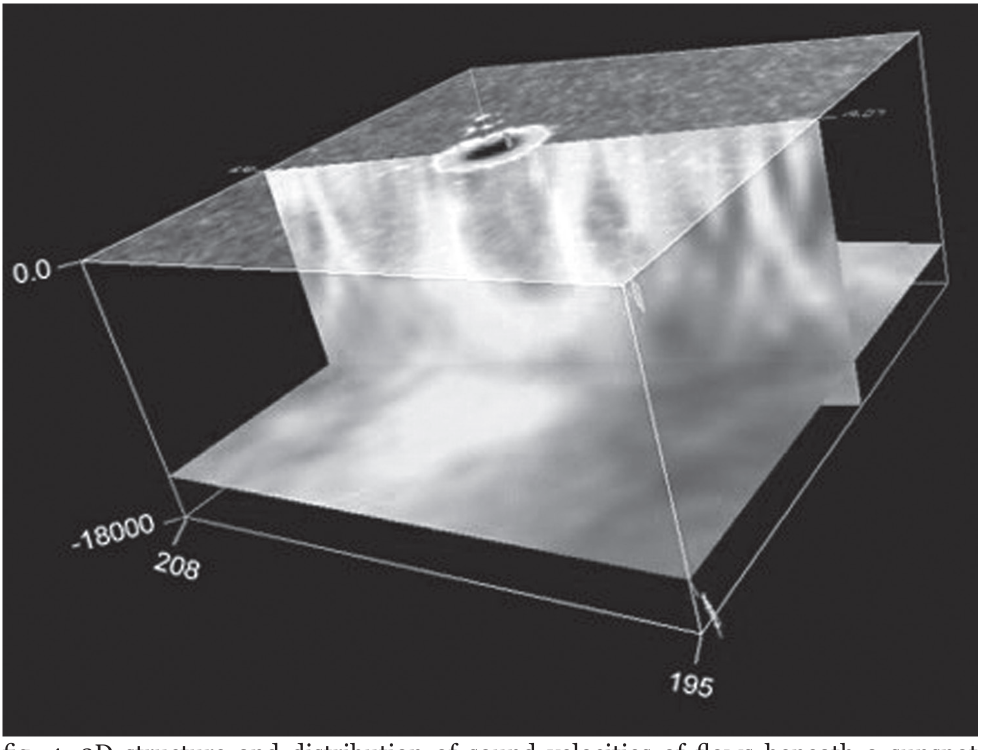


fig. 4. 3D structure and distribution of sound velocities of flows beneath a sunspot obtained by the time-distance inverse methods from the SOHO/MDI measurements. Please, for coloured image refer to [21]

The SOHO/MDI measurements of the sound velocity based on the Doppler shift of flows under a sunspot and the use of the time-distance seismology (acoustic tomography) made it possible to obtain the first 3D images of sunspots and the flow pattern underneath [20—22] (see Fig. 4). Plasma was shown to flow into the sunspot immediately under the surface as predicted by the theory. The analysis proves that sunspots are amazingly shallow features. Change from the regions cooler than the ambient atmosphere to the regions hotter than the medium occurs at mere 5,000 km under the surface. The information obtained explains the long-standing problem of heliophysics: how can sunspots exist for a few weeks without decaying? It also allows us to predict the formation of sunspots in the convection zone from on-line MDI data before they emerge at the surface.

The first images of the far side of the Sun were derived by applying seismic holography to the SOHO/MDI pressure distribution on the visible disk [20, 23]. The activity on the far side of the Sun is localized by modifying the pressure distribution (sound velocity field) on the visible side, and thus, the appearance of active regions on the limb by solar rotation is predicted. This method increases the advance time of the forecast of geoeffective phenomena in the Sun — active regions and associated flares and CME events — that are the main source of space weather in the Earth environment.

CORONAS-F/DIFOS observations of p-modes of the global oscillations of the Sun in a broad wavelength range (350...15,000 nm) made it possible to extend the observed spectrum of oscillation amplitudes to both shorter and longer wavelengths and approximate it by the law $\lambda^{-1.2}$ in agreement with earlier observations [24].

2.2. SOLAR ATMOSPHERE FROM THE PHOTOSPHERE TO THE CORONA

The photosphere is the 'face' of the Sun or, as we call it, the solar disk. The features and phenomena on the solar disk form the visible foundation, which allows us to gain an idea of the Sun as a whole.

A great number of high-resolution images of the Sun obtained with the telescopes of the SOHO, TRACE, CORONAS-F, *Hinode*, and STEREO missions in different spectral lines corresponding to different temperature layers in the solar atmosphere made it possible to localize and study the morphology of various active features in all phases of the activity cycle.

The SOHO/MDI observations revealed wave-like oscillations of the solar supergranulation with a period of 6...9 days [25]. This explains why the supergranules seem to rotate faster than the ambient plasma.

Supergranulation is a pattern of horizontal surface motions with a scale of 30 Mm and lifetime of one day or more, which is seen as a network of fine magnetic features. Earlier, when the supergranulation was believed to correspond to the preferable scale of the thermal convection cells, its dynamics was not clear. In particular, we could not explain why the supergranulation pattern appears to rotate around the Sun faster than the magnetic structures. The analysis of the SOHO/MDI data has shown that supergranules possess the properties of waves propagating preferably in the direction of solar rotation. This explains the visible faster rotation of the supergranulation pattern. The waves propagate in all directions along the solar surface, but they are stronger, i.e., have larger amplitudes, in the direction of solar rotation. This creates the illusion of faster rotation, because the waves moving in the direction of solar rotation are seen better. There is a hypothesis that supergranulation waves may be excited by interaction of the convection and rotation. It remains unclear, however, why the supergranules have a typical scale of 30 Mm and how deep they are.

On the high-resolution images from *Hinode*, one can readily see the granules and bright points corresponding to the tiny magnetic features with a strong magnetic field. The bright features in the Ca II H line (3,970 Å) indicate to the heating of the chromosphere. Magnetic fields are measured with a high accuracy, and heating around the sunspots, flares, and ejections is traced. The stages of the active region birth and evolution are clearly pronounced. Ejections over sunspots in the limb AR were observed for the first time owing to a low level of scattered light in the optical telescope. The high spatial resolution of the *Hinode* X-ray telescope made it possible to reveal a lot of X-ray bright points in the quiet Sun, which turned out to be small loop-like features existing all over the Sun, including the most quiescent regions. These

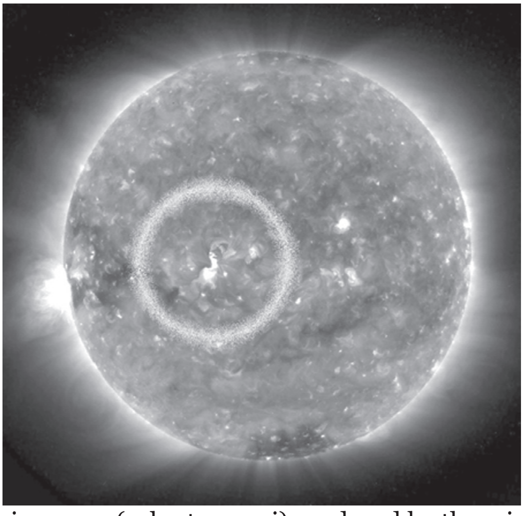


fig. 5. Huge ring wave (solar tsunami) produced by the minor flare and CME in AR 977 on Dec. 14, 2007 and propagating over the Sun at a speed of 500...1,500 km/s. STEREO image

observations made us change the definition of the 'quiet Sun'. Some results of the *Hinode* mission are described in [26, 27].

The STEREO mission started not long ago, but it has already provided some interesting images, such as the side view of a number of solar prominences displaying their twist and dynamics controlled by the magnetic fields; a huge ring wave (solar tsunami) produced by a minor flare and CME in AR 977 on Dec. 14, 2007 and propagating over the Sun at a speed of 500...1,500 km/s (see Fig. 5); a large twisted loop 400,000 km long (Fig. 6), which emerged at the surface and broke down during two days demonstrating a complex topology and dynamics of solar magnetic fields responsible for solar activity; numerous small-scale (500 km in diameter) spicules near the poles produced by the shock waves

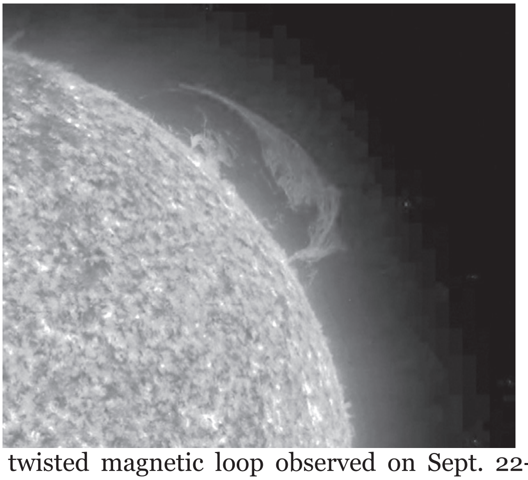


fig. 6. Long twisted magnetic loop observed on Sept. 22—23, 2007. STEREO image

that occur as the sound waves propagate from the solar surface to the atmosphere (the lifetime of the spicules is about 5 min, more than 100,000 spicules are recorded simultaneously on the solar surface); and a chain of ARs near the equator observed for 36 hours with a spatial resolution from 2 to 0.5 min, which displayed the finest details of plasma dynamics along the magnetic field lines [28].

A spectral atlas of solar coronal features (coronal holes, quiet corona, active corona, and flares) covering the wavelength range of 670…1609 Å and a similar atlas for the solar disk were compiled based on the SOHO/SUMER data [18]. The atlas of solar coronal features comprises 504 lines, 300 of which (60%) were identified. The measured line intensities provide rich information on electron densities and temperatures, opacity, and content of elements in the upper solar atmosphere. The CORONAS-F onboard spectrometers SPIRIT, DIAGENES, and RESIK also provided a wealth of spectral data [29, 30], which stimulated the development of new facilities and planning of the future missions.

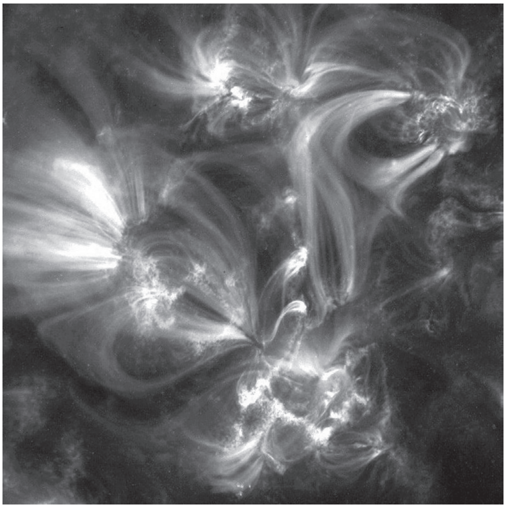


fig. 7. Multi-coupling of sunspot magnetic fields. TRACE image

High-resolution images from the TRACE mission visualized many features unobservable before, such as initiation of flares by the emerging magnetic flux, growth or rise of the coronal loops, ejection of twisted magnetic tubes from the solar surface, fine magnetic structure of the corona, sunspot rotation and associated coronal effects, transverse oscillations of the coronal loop structures as a whole, etc. [31]. The other results worth mentioning are the observation of plasma outflow from active regions; co-existence of dark and bright regions above the photosphere; nearly constant emission intensity depending on altitude; mainly inhomogeneous heating at the feet of the coronal loops in the temperature range of 0.8…1.6 MK; suddenly appearing and disappearing features; the variety of the structures most of which are controlled by the magnetic field; multi-coupling of sunspot magnetic fields (Fig. 7); the existence of different-size magnetic loops; predominance of the magnetic forces manifested in plasma ejections and motions along the field.

The SOHO/SUMER observations have revealed many characteristic features of hot (9...20 MK) magnetic loops [32]. Giant loops of very hot gas rising over the solar surface vibrate violently during the sporadic active events. The loops 350,000 km long vibrate as the whole with a period of 20 min. The hot gas moves along the line-of-sight at a speed up to 100 km/s. The gas is cooled rapidly, and the motion ceases after two or three oscillations. The vibration is presumably due to oscillations of the strong magnetic field of the loop that are excited by the bursts of high-energy particles from the solar atmosphere below. The heated high-temperature gas in the loop gets ionized and begins to emit intensively in the hot lines of iron producing oscillations inside the loop.

The temperature of the solar corona is as high as 1...2 million degrees K. This is one of the mysteries of the solar and stellar physics, since the temperature of the underlying photosphere is only 6,000° K. The questions associated with the corona heating are: where does the energy come from; how is it transported to the corona; and how does it dissipate in the rarefied corona to heat it up to that temperature?

A series of observations on SOHO, TRACE, CORONAS-F, RHESSI, and *Hinode* shed light to the mechanism of the corona heating. In particular, evidence is obtained that the heating of the solar corona may be due to numerous reconnection processes (nano- and microflares) and MHD waves from the photosphere.

Observations with the RES-K spectroheliograph (SPIRIT/CORONAS-F experiment) in the resonance line MgXII (8.42 Å) in the solar corona have revealed a whole class of the previously unknown phenomena — fast-dynamics plasma features with the temperatures up to 20 million degrees [12, 33]. Images of these high-temperature features were obtained for the first time, and their dynamics was studied. They are likely to form when the flare-generated hot plasma fills the magnetic configuration in the corona, and the plasma emission makes the latter noticeable. As a result, the features may take on a variety of the shapes (see Fig. 8), such as 'hot clouds', 'spiders', loops, propagating wave fronts, successively flaming up magnetic arcs, etc. The hot plasma features observed in the corona manifest the mechanism of corona heating at the expense of energy released in

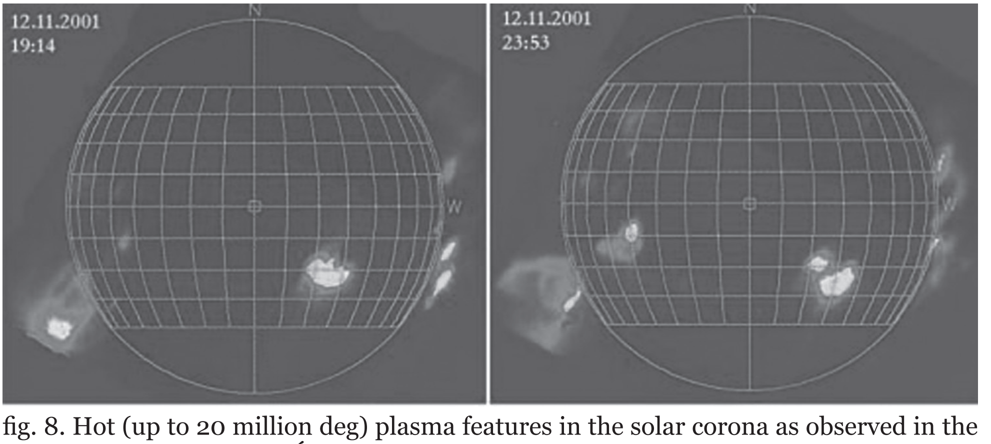


fig. 8. Hot (up to 20 million deg) plasma features in the solar corona as observed in the MgXII spectral line (8.42 Å). CORONAS-F/SPIRIT image

magnetic configurations and converted to plasma energy. It is established that mass ejections from the solar atmosphere observed with the SOHO coronograph are often associated with these hot plasma features in the corona.

As shown by SOHO observations, the permanently emerging and swelling up magnetic loops release their energy and, then, collapse down to the surface. This process is suggested to be one of the heating mechanisms in the solar corona based on conversion of the magnetic energy to the thermal energy of the coronal plasma [18].

The magnetic heating of the corona by active regions is also well pronounced on the records of X-rays spectra obtained with the CORONAS-F/RPS-1 spectrometer for different number of sunspots [10, 34]: the more sunspots are involved the harder is the X-ray spectrum recorded (see Fig. 9) and, correspondingly, the stronger is the impact on the Earth's atmosphere. The spectra obtained provide quantitative information on the enhancement of such impact.

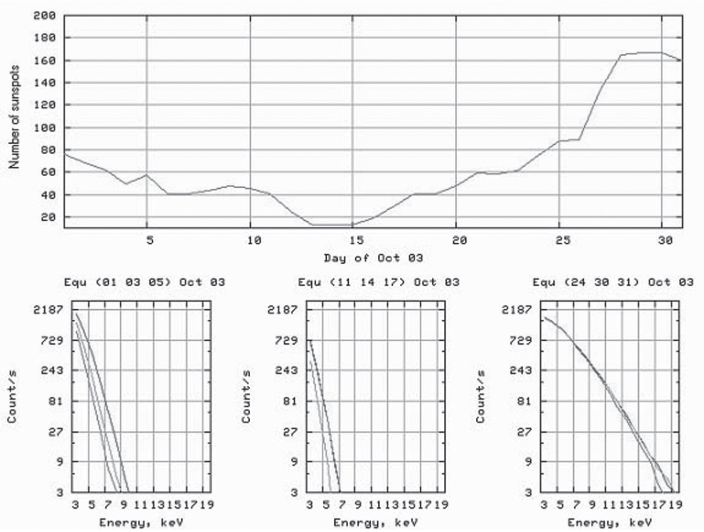


fig. 9. Dependence of the background (devoid of flares) X-ray spectra of ARs on the number of sunspots illustrating the magnetic heating of the solar atmosphere. CORONAS-F/RPS-1

Strong oscillations of the Doppler shift of spectral lines in the hot loops over active regions [32, 35] detected by the SOHO/SUMER experiment and identified as slow magneto-acoustic standing waves can only be seen in the hot flare lines (> 6 MK, i.e., Fe XVII, Fe XIX, and Fe XXI) and are invisible in the lines that form under the normal coronal temperatures (1 MK, i.e., Fe XII, Ca XIII, Ca X). The discovery of these oscillations opened a new page in the seismology of the solar corona and provided a useful tool for the study of its heating.

The dominant role of the magnetic field in the heating of the solar corona was established using the SOHO/MDI data [20]. The small-scale magnetic field emerges permanently from under the photosphere. We see it appear, break to fragments, mix up, and vanish replaced by a new field for the time interval less than 40 hours [36]. It were discovered a great many of small, closed, interlaced magnetic loops that form a dense magnetic carpet. The interaction of these loops may cause numerous processes of magnetic reconnection and formation of electric fields. The dynamic patterns obtained with the SOHO/EIT telescope reveal closed magnetic loops of various spatial (from small bright points to large loop systems) and temporal scales (from minutes to days). This pattern fits the Parker hypothesis of 1988 [37], according to which the corona heating may be due to the energy released in numerous nanoflares equal to 10^{24} erg or more, i.e., 10^{-9} of energy of the ordinary solar flares. The efficiency of such heating is explained by a huge number of nanoflares, particularly, in the epoch of solar minimum. Thus, it can be assumed that the magnetic field energy dissipates on small spatial scales, and the observed hot features, such as bright points, can be considered a superposition of elementary events similar to nanoflares. The theoretically estimated dissipation scale (about 100 m) is much smaller than the spatial resolution of any telescopes available, so that this suggestion cannot be tested by observations. The statistical analysis of spacecraft observations (TRACE, SOHO, *Yohkoh*, SMM, etc.) has established that solar events with different amounts of energy released (nanoflares, microflares, and ordinary flares) possess a common feature — power-law distribution of the measured intensities [38]. This result suggests that the solar atmosphere can be regarded as a system with fully developed turbulence down to the smallest scales, on which the dissipation occurs.

Super-high resolution of the *Hinode* data (spatial resolution of the solar features of 150 km and time resolution of 5 s) also allowed us to learn a great deal of the key part of the magnetic field and Alfven waves in the energy transfer, heating of the solar corona, and acceleration of the solar wind [26]. An important finding of the *Hinode* mission is the long-sought observational evidence of ubiquitous presence of MHD waves as the main carrier of energy from its reservoir on the solar surface (convection) to the atmosphere. The convection motions of plasma in the photosphere possess sufficient energy to serve the source and reservoir of the corona heating. Two well-known and widely used groups of theories were proposed to account for energy transfer to the corona. One of them [37] mentioned above claims that the magnetic field lines rooted with one end deep in the photosphere are continuously disturbed by random photospheric motions and, as a result, get braided high in the atmosphere. This leads to numerous reconnections of oppositely directed field lines. At the reconnection points, there arise two oppositely directed jets that are accelerated and pushed apart by the reconnecting and strongly curved field lines. This is a form of the release of magnetic energy and its conversion to the energy of plasma motion.

The unexcelled *Hinode* observations revealed a great number of such jets in active regions [39] almost everywhere in the chromosphere [40], including the penumbral chromosphere [41], as well as in large-scale features (up to 2 Mm wide and 100 Mm long) [42], which is a new evidence of the reconnection. Whether the energy of these ejections (often referred to as nanoflares) is sufficient to heat the corona will be clear when we study their statistical characteristics, such as the occurrence rate, energy distribution, total number, etc.

The theories of the second group suggest that the heating of the solar corona and acceleration of solar wind may be due to MHD waves. The fast and slow magnetosonic waves, to judge from their amplitudes, do not possess energy enough to heat the corona and can not go far to the corona [43], while the field-aligned Alfven waves are ubiquitous in the solar atmosphere, and their energy is sufficient to heat the corona and accelerate the solar wind [40, 42]. They have amplitudes of 10…25 km/s and periods of 100…500 s (B = 10 G, V_A = 45…200 km/s). With allowance for the wave transport coefficient (about 3%), this ensures the minimal energy flux to the corona equal to about 100 W/m^2. The mechanism of dissipation of these waves in the rarefied corona still remains unclear. Various models have been proposed, including the resonance absorption of Alfven waves when their frequency becomes equal to the gyrofrequency of different plasma ions (high-frequency models); dissipation of low-frequency waves due to such effects as self-interference at reflection, medium compressibility, parametric decay of waves, etc. [40]. The first results of the *Hinode* mission are of great importance to further development of the theory of the solar and stellar corona heating.

2.3. FLARES AND CORONAL MASS EJECTIONS

The SOHO, CORONAS-F, TRACE, RHESSI, *Hinode*, and other space missions observed a great number of flares and coronal mass ejections (CME) varying in shape and evolution. These observations supplied rich information on the major phenomena of solar activity and powerful geoeffective disturbances in the Sun-Earth system. During cycle 23, the spacecraft facilities recorded outstanding activity in the Sun characterized by extremely intensive flares and ejections [44].

Observation of solar flares in the X-ray and gamma-ray frequency ranges provided us with the spectral, dynamic, and energy characteristics of the flare processes and allowed the diagnostics of flare-generated plasma and the study of acceleration and nuclear processes in flares.

Fig. 10 represents a reconstruction of the magnetic field in the solar corona during the major flare of Sept. 7, 2005 as observed in the CORONAS-F/SPIRIT experiment: the closed loops of the magnetic field become open after the flare.

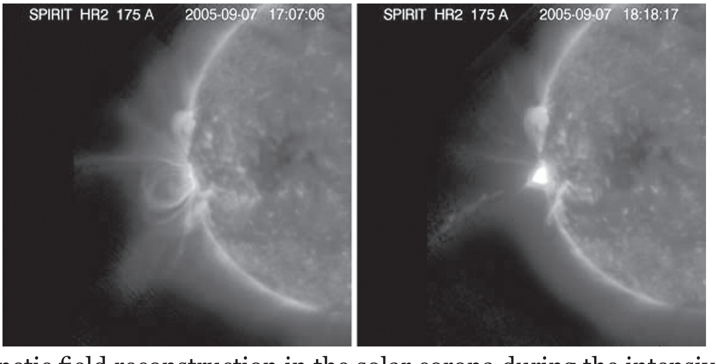


fig. 10. Magnetic field reconstruction in the solar corona during the intensive limb flare of importance X17+ on Sept. 7, 2007. CORONAS-F/SPIRIT data

A detailed spectroscopic diagnostics of the flare-generated and coronal plasma was performed with the CORONAS-F, RHESSI, and SOHO X-ray instruments. The full-profile spectral lines from the most outstanding flares were obtained for the first time with the CORONAS-G/RESIK and DIAGENES instruments. This promoted the study of energy release and dissipation in solar flares: non-maxwellian and nonequilibrium processes, excitation of inner atomic shells and dielectronic recombination, spectral-line width variations due to a noticeable plasma turbulence, etc. The other new results are as follows: the absolute Doppler shifts of X-ray spectral lines in solar flares and radial velocities of the emitting plasma were measured throughout the explosive phase of the flare; lines and effects associated with electron transitions in the Ar XVIII (Lyβ) and Si XIV ions, satellite lines of atomic transitions of the type 1s2-1snp (n = 3, 4, 5) (Si XII), and a number of new spectral lines were discovered and identified [29, 30, 44]; the absolute content of rare elements (potassium, chlorine, argon, and sulfur) was determined in solar flares; and the temperature dependence of the spectra of the solar plasma ions corresponding to different activity levels was investigated. The X-ray ion spectral lines of solar plasma for high values of the quantum number *n* (Fig. 11) were observed for the first time in astrophysical plasma. These data open up new opportunities for temperature diagnostics of the coronal plasma.

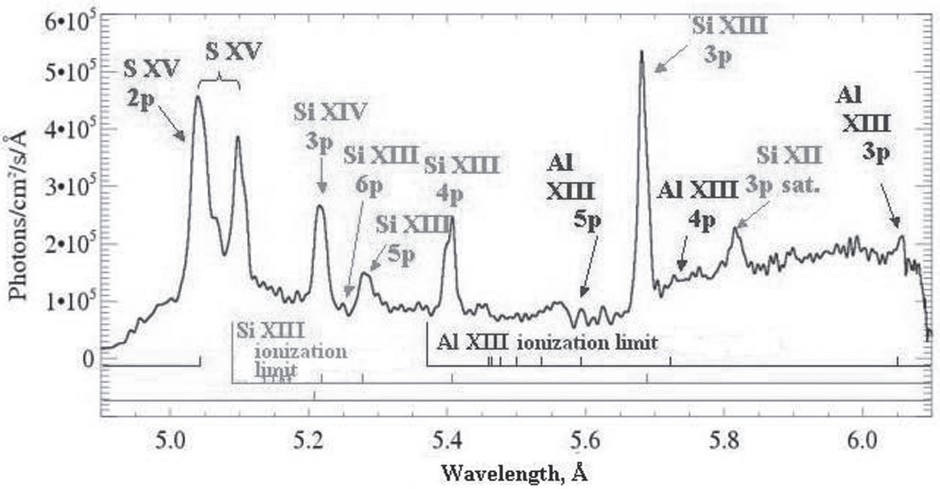


fig. 11. Ion spectral lines of Si XV, Si XIV, and Al XIII corresponding to the atomic transitions with a high quantum number n: from n = 3 up to the ionization limit. CORONAS-F/RESIK data

The CORONAS-F/SPR-N device measured a high (up to 85%) linear polarization of hard X-rays (20...100 keV) in the impulsive phase of the major flares of October 29, 2003 [45], which could not be detected earlier and which was indicative of generation and acceleration of charged particle beams by the impulsive flare mechanism.

RHESSI observations have provided answers to many questions associated with solar flares [46]. For example, combined with the TRACE data, they revealed that energy release in flares usually occurred in the form of plasma loops and helmet streamers heated to tens of millions of degrees and high-energy electrons propagating downward from the upper corona and heating the plasma in lower layers. High-resolution observations of nonthermal X-rays (30...80 keV) made it possible to identify the emission source as the site of intrusion of accelerated particles to the dense atmosphere at the feet of the magnetic loop. The sources at different feet of the same loop have different time profiles and different brightening times. The flare energy release in the corona was localized by superimposing the TRACE and RHESSI images and was related to the magnetic reconnection that occurred in the stretched helmet feature in accordance with the standard model applied to the observations.

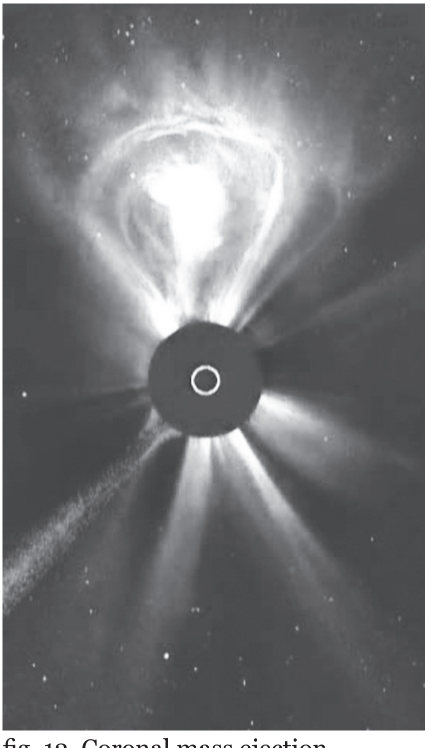


fig. 12. Coronal mass ejection observed on February 27, 2000. SOHO/LASCO image

Numerous microflares (a probable source of the corona heating) that occur permanently in the solar active regions, last for a few minutes, and are accompanied by particle acceleration were detected for the first time in the X-ray energy range. As shown by observations, the small-scale features are the key to explaining the solar activity and its effects. For the first time, SOHO observed the flare-generated global oscillations. They had the form of the waves propagating from the flare origin throughout the solar atmosphere [31] similar to those shown on the STEREO image in Fig. 5 (p. 70).

The SOHO/LASCO coronograph has provided a wealth of remarkable CME images and movies (see Fig. 12) that give an idea of the most outstanding active events in the Sun and the release of huge amounts of energy and mass ejection to the heliosphere. The mass ejected to the solar wind by CME proved to be greater than believed earlier, as well as the CME occurrence rate. The SOHO observations contribute significantly to the study of CME effect on geospace and technological systems (communication and navigation satellites, etc.), as well as on the solar system as a whole. An important result of the SOHO mission is the opportunity to find out which solar ejections are dangerous [18].

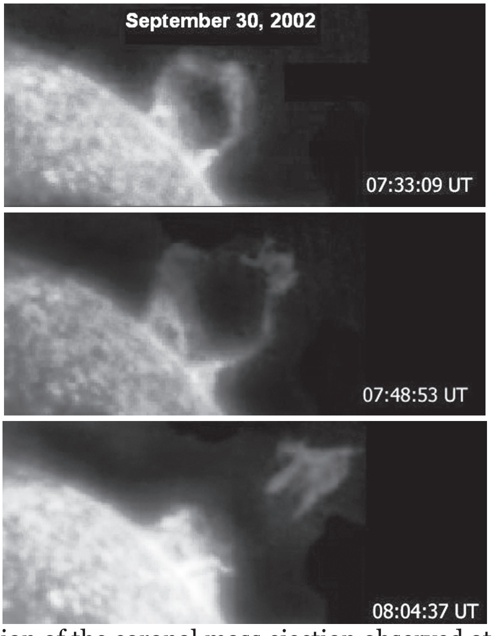


fig. 13. Evolution of the coronal mass ejection observed at low altitudes in the solar atmosphere. CORONAS-F/SPIRIT data

The CORONAS-F solar X-ray telescope (SPIRIT) operated in the coronograph regime recorded low-altitude ejections (Fig. 13), which supplemented the high-altitude pattern obtained with the SOHO/LASCO coronograph ($R > 1.1\ R_{\odot}$).

The first 3D reconstructions of CME were obtained by polarization analysis of the SOHO/LASCO images [18]. For a few CME events, their 3D structure, location, and propagation velocities were calculated from the brightness ratio of the polarized and nonpolarized emission.

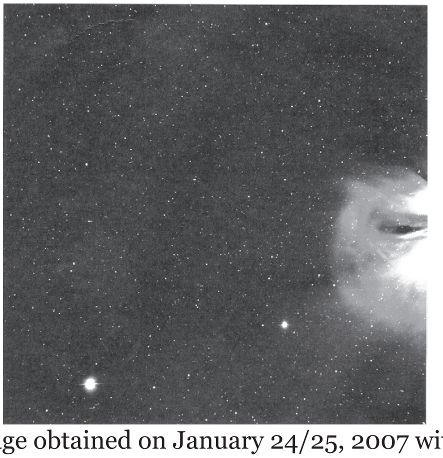


fig. 14. CME image obtained on January 24/25, 2007 with the STEREO heliospheric telescope (borrowed from [47]). The large bright points are Mercury and Venus

One of the first CME images obtained with the STEREO heliospheric telescope at a large distance from the Sun is shown in Fig. 14 [47]. The positions of the ejection, Mercury, and Venus give an idea of the potentialities of the instrument.

2.4. SOLAR WIND AND HELIOSPHERE

An important problem of the solar physics still remains the origin and acceleration of the solar wind. Where does the solar wind originate and how is it generated and accelerated? Observations on SOHO, ULYSSES, and other spacecraft contributed largely to the solution of this problem.

ULYSSES measurements have been used to determine the heliolatitude dependence of the solar wind velocity in the heliosphere (Fig. 15), which revealed a pronounced difference between the solar wind characteristics at the poles and at the equator. At the minimum of the activity cycle, fast (about 800 km/s), stable solar-wind streams were detected to come from the poles, and slow (about 400 km/s), variable streams from the equatorial region. In the maximum epoch, the difference between the polar and equatorial solar wind is not as clear. The maximum of cycle 23 (1995—2007) was observed in 2000—2001, so that ULYSSES passed over the poles (northern and southern) twice at the maximum of the cycle and three times (twice over the northern pole and once over the southern one) at the minimum.

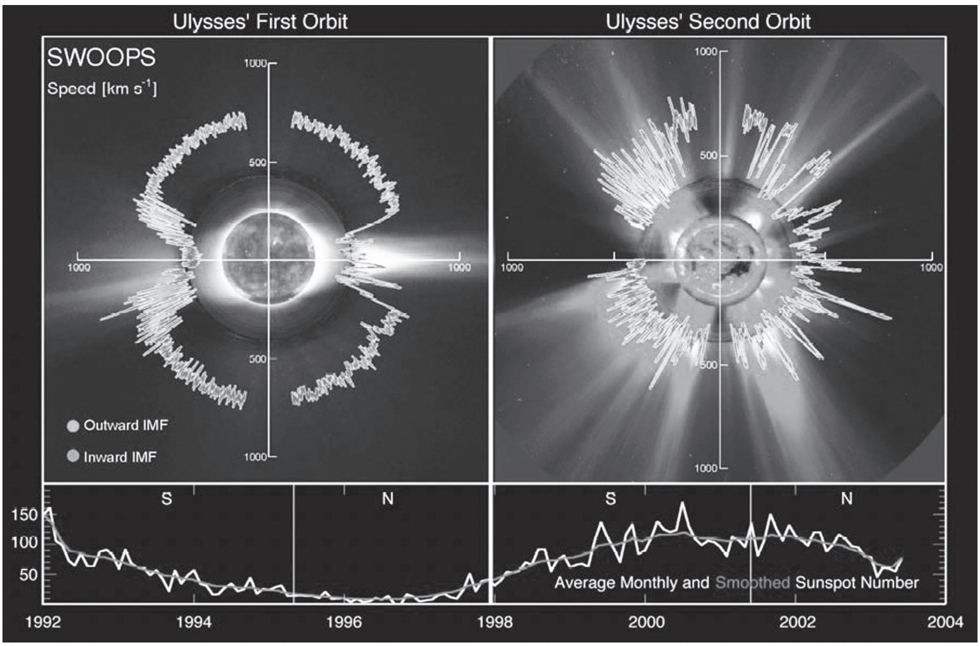


fig. 15. Heliolongitudinal dependence of the solar wind velocity obtained during the first (1994—1995; left) and the second (2000—2001; right) pole passes of the ULYSSES mission. The sunspot number curve below shows that the first pass took place at the decline of activity, near the minimum of the cycle, and the second one, in the vicinity of the maximum. The difference between the solar wind characteristics in the polar and equatorial regions at the maximum is poorly pronounced. © NASA/JPL-Caltech

The fast and slow solar-wind streams in the outer solar atmosphere could be mapped using the SOHO data, and the wind outflow could be related to the magnetic topology by tracing its acceleration from the solar surface to a few solar radii [20]. The SOHO/SUMER observations corroborated the earlier suggestion that the fast solar wind originates in the polar coronal holes — dark features with open magnetic field [48]. The Doppler shift of the Ne VIII line formed at the base of the corona was measured to establish the relation between the speed of the plasma outflow and the structure of the chromospheric magnetic network identified by the Si II line. This relation suggests that solar wind is formed along the boundaries of the magnetic network (where the plasma velocity is 5…12 km/s) and at their intersections (where the plasma moves at 10…20 km/s). The SOHO/SUMER and MDI data (vertical magnetic field component) made it possible to localize the source of the fast solar wind in the magnetic funnel (open field configuration) between the altitudes of 5,000 and 20,000 km above the photosphere [18]. According to a recently proposed model, the closed magnetic loops are carried away by convection to the funnel regions, where they reconnect with the open field lines, and the plasma contained within the loops is released and accelerated to form the solar wind. *Hinode* observations have revealed significant amounts of material (1/2 of the solar wind mass loss rate) to emanate from active regions to the upper corona [49] at velocities insufficient for its escape to the solar wind (approximately 140 km/sec). Future observations must show whether these outflows may indicate a probable source of formation of the solar wind.

Thus, the solar wind is much more effectively accelerated at the poles and in other open-field regions than at the equator. Here, the solar wind velocity reaches 800 km/s at a distance of about three solar radii. The acceleration is less effective at the equator, where closed field lines form the base of the streamers that are, probably, the source of the slow solar wind. The SOHO/UVCS measurements have revealed [50] that the speed of the solar wind stream emanating from equatorial coronal holes in the epoch of solar maximum is 3 to 5 times lower than a similar speed for the polar coronal holes at the solar minimum. Since the solar wind velocity at the Earth's orbit in the epochs of solar minimum and maximum is determined, respectively, by the polar and equatorial CH and is equal in both cases, the acceleration must occur beyond the UVCS range of observation, i.e., farther than three solar radii.

Observational evidence (SOHO/UVCS and NASA's *Spartan 201* mission) has been obtained that Alfven waves generated in the solar atmosphere accelerate solar wind acting as the sea surf and transmitting energy to the solar wind particles [51]. This information is available for several atoms of the solar plasma, for heavy atoms, such as oxygen and, provisionally, for hydrogen, which is the main constituent of the solar wind plasma. However, many other atoms, e.g., helium (second abundant element), have never been detected in the acceleration region in the corona. Therefore, new observations are required for better understanding of the wave-solar wind interaction [52].

The solar-wind plasma (ULYSSES/SWOOPS) and ion composition (ULYSSES/SWICS) have been measured to study the solar wind streams in high-latitude coronal holes [35]. In the epoch of solar minimum, the solar wind emanating from coronal holes was shown to have the ion composition similar to that of the chromosphere and corona. The high-speed solar wind was observed to precede both from small and from large coronal holes, though the highest velocities were recorded at the center of the major CH. While moving along the orbit to lower latitudes, ULYSSES crossed CH boundaries and took measurements in the transition layer at their periphery, where plasma was decelerated and was passing gradually to the ambient solar wind.

The properties of the transient solar wind were studied with the aid of two spaced satellites in the joint ULYSSES-SOHO experiment [18]. In November 2002, when ULYSSES was passing at a distance of 4.3 AU from the Sun at the latitude of 27° north of its western limb, one and the same very hot plasma feature could be identified, first, in the vicinity of the Sun and, then, in the vicinity of the spacecraft. At the end of November, SOHO detected four major CME events leaving the Sun for a few days and propagating towards ULYSSES. They reached ULYSSES 15 days later. In interplanetary space, these ejections, apparently, merged to form a single feature in the solar wind and produce a strong interplanetary shock wave. The plasma of this merged feature contained an unusually large amount of highly ionized iron ions indicating to its high-temperature source. Such highly ionized features are frequently observed in the solar wind and are identified as interplanetary CMEs. The SOHO/UVCS observations also detected highly ionized Fe in plasma, particularly, after the CME of November 26, 2002. It that case, very hot plasma (of (6…10)·106 K) was, apparently, generated high in the solar atmosphere — above 1.5 solar radii — by reconnection of post-flare loops.

When passing over the solar poles and crossing the ecliptic plane, ULYSSES recorded a lot of heliospheric features and events that gave us an idea of solar-heliospheric relationships and proved useful to the study of solar-terrestrial coupling. These are, in particular, the corotating stream interaction regions (CIRs) formed at low and intermediate latitudes when the high-speed solar wind from coronal holes encounters the slower ambient wind [53]. The interaction region is confined by the front and rear shock waves and rotates with the Sun. The shock waves can accelerate the charged particles giving rise to periodic enhancements of the high-energy particle fluxes recorded by spacecraft. Above the streamer belt (source of the slow solar wind), CIRs are weakened significantly. As a result, the front shock waves disappear, and the intensity of charged particles decreases.

An interesting phenomenon was the change of the magnetic polarity of the Sun, which took place in 2004 as ULYSSES was passing over the pole [54]. The polarity reversal was found out to be a complex process lasting for months. During this time, the corona is evolving and reflecting the changes that occur at the photospheric level. Therefore, the magnetic polarity in the heliosphere measured at the south pole by ULYSSES corresponded to the 'old' magnetic cycle, while the polarity measured less than a year later at the north pole was opposite in sign and reflected the characteristics of the 'new' cycle. This pattern agreed perfectly well with the behaviour of the solar wind in that period.

An important discovery of the ULYSSES mission is that charged particles in the heliosphere are transported relatively easily in heliolatitude [4, 55]. If the Sun did not rotate, the solar wind would carry away the open magnetic field lines from the solar surface radially in all directions. Since the charged particles are linked to the magnetic field, there should be no noticeable transport across the field. Owing to the rotation, the magnetic field is twisted and forms the Archimedean spiral, in which the field lines lie on the cones of heliolatitude. Under these conditions, the charged particles cannot move easily in heliolatitude. Actually, however, such an ideal pattern does not exist. ULYSSES recorded a large number of high-energy particles over the solar poles, far from their parent flares. For example, significant transient enhancements of energetic particle fluxes (comparable to those observed in the ecliptic plane) recorded up to the highest heliolatitudes in 2000—2001 as ULYSSES was passing over the poles were due to the flares and CME events that occurred at lower latitudes or even in the opposite hemisphere. A particular mechanism that would allow the charged particles cross easily the magnetic field lines still remains unknown. Either the particles may 'jump' over the field or the filed lines themselves may distort strongly enough to connect the low-latitude sources to higher latitudes. Anyhow, ULYSSES has demonstrated that the particle propagation pattern in the heliosphere needs revision. This result is not only of pure scientific interest. Astronauts in far space can be exposed to radiation from the sources, which, by their position in the Sun, were believed earlier not dangerous. When the event lasts 10 days or longer, a considerable range of heliolongitudes becomes involved owing to the solar rotation, and thus, solar flares and/or CME events are able to fill the entire volume of the inner heliosphere with charged particles in a short time and create a reservoir of radiation.

High-precision measurements of the anisotropy of proton fluxes with the SOHO/ERNE device have revealed an interplanetary 'channel', in which the solar high-energy particles can propagate freely [18]. By observing high-energy particles from the solar event of May 2, 1998 during four hours it was established that the intensity of the proton flux along the magnetic field was a factor of 1,000 higher than across the field. As shown by observations, the CME magnetic ropes form a 'channel' for transport of solar protons along the field with a path length of at least 10 AU.

2.5. SOLAR-TERRESTRIAL AND SOLAR-HELIOSPHERIC COUPLING

The spacecraft remote observations and in-situ measurements of the Sun have provided a wealth of valuable information. It was used to study the solar active phenomena, solar-terrestrial and solar heliospheric coupling, and to develop methods of forecast of the space weather and protection from its damaging effect on the Earth. ULYSSES gave us a global idea of the heliosphere and its response to solar flares and ejections. Staying at the libration point on the Sun-Earth line, SOHO was continuously monitoring the Sun and recording heliospheric disturbances before they reached the Earth's magnetosphere. CORONAS-F was orbiting the Earth inside the magnetosphere and was monitoring solar active events simultaneously with their effects in near-Earth space.

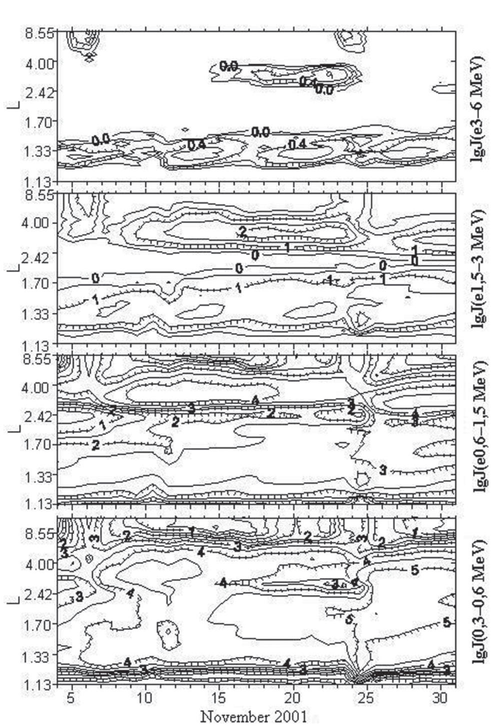


fig. 16. An example of the intensity-time-L-shell map (lines of equal intensity of electron flux) for the period of November 4-30, 2001. Structural changes in the radiation belt of electrons are demonstrated. The strokes are directed towards decreasing intensity. CORONAS-F/MKL data

A number of outstanding solar events were recorded in cycle 23 (November, 2001, October-November, 2003, November, 2004, and January and September, 2005), which provided a good opportunity to study the solar-terrestrial coupling [13, 56]. During these events, CORONAS-F recorded flare-generated energetic gamma-rays and neutrons [57]. The spectrum of relativistic protons with energies up to 14 GeV was measured for the first time by the magnetic cut-off effect [58]. Measurements of high-energy particle fluxes along the CORONAS-F orbit with SCR complex provided information about the deformation and restructuring of the magnetosphere, dynamics of the Earth's radiation belts and change of their boundary positions and maxima during strong magnetic storms, shift of the polar cap boundary and the boundary of particle penetration to small heights, etc. [59, 60]. All these processes can be traced on three-dimensional (intensity-time-L-shell) maps. Such a map is represented in Fig. 16 as an example. Variations of charged particle fluxes in the Earth's radiation belts were studied, and the distribution of radiation at the height of 450 km was plotted for electrons in the energy range from 300 keV to 3 MeV and protons with the energies of 1...26 MeV. A new effect was discovered in the main phase of the storm. The outer radiation belt of electrons was shown to vanish at energies above 1.5 MeV. This effect is associated with restructuring of the magnetosphere – destruction of the adiabatic orbits of trapped particles.

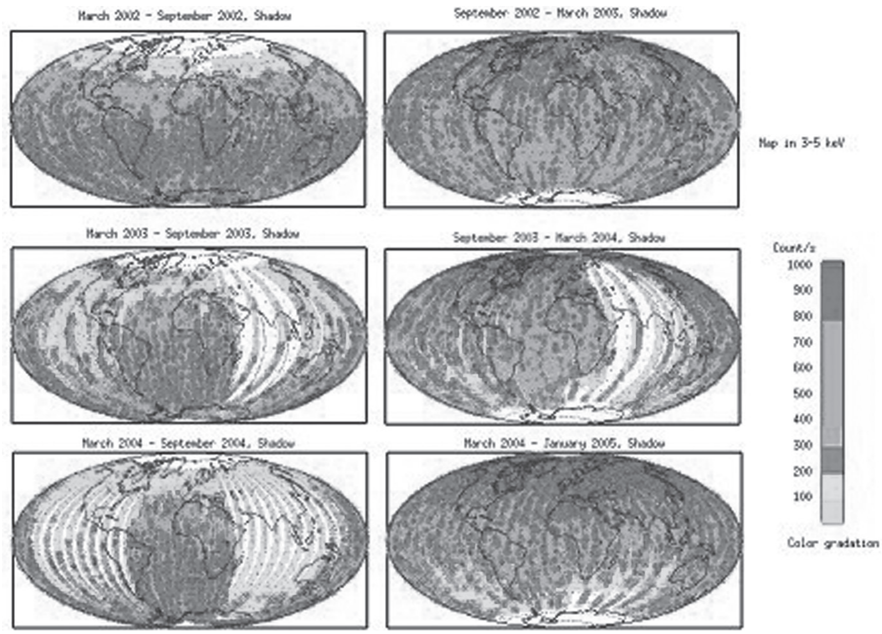


fig. 17. Airglow dynamics in the nighttime upper atmosphere in the X-ray energy range of 3…5 keV. CORONAS-F/RPS-1 data. For coloured image, please, refer to [34]

X-ray observations of the nighttime atmosphere on CORONAS-F made it possible to plot global maps of the atmospheric glow (Fig. 17) representing its seasonal variation and dependence on solar activity. Information on energetic precipitating electrons was obtained, and the energy balance of the Earth's atmosphere in the range of 3…31.5 keV was investigated [34].

The outstanding solar events and their effects in geospace observed on CORONAS-F in cycle 23 are described in [13, 56, 61, 62].

The heliospheric effects of the major, importance X28 flare of November 4, 2004 in the western part of the disk were recorded and investigated owing to a favorable position of ULYSSES [35]. Passing far from the Sun (at 5.3 AU), but within the line-of-sight range, ULYSSES recorded a strong interplanetary shock produced by the fast, flare-generated CME. During the entire period of high activity associated with this flare, ULYSSES was measuring high-energy particle fluxes modulated by the CME-related transient and high-speed solar wind streams emanating from a large, long-lived transequatorial coronal hole. The effects of this flare in geospace are described in detail in [62].

The AR NOAA 10720 proved to be one of the most flare-productive active regions. From January, 14 to January 22, 2005, it produced 15 flares of importance M and 5 flares of importance X. The proton event associated with the importance X1.7 flare of November, 20 was most intensive (>100 MeV) since October, 1989. The SOHO facilities allowed virtually real-time observation of this event with the information appearing simultaneously on the SOHO Internet site, which attracted increased attention of mass media [18]. CORONAS-F measured high-energy gamma-ray emission in lines and relativistic proton fluxes from this flare [13, 63, 64].

One of the most intensive solar flares in cycle 23 — flare of importance X17+ — occurred on September 7, 2005 in AR10808 on the eastern limb (see Fig. 10 p. 75). Associated with this flare was a very large and fast CME event. At that time, ULYSSES was at the distance of 4.8 AU almost exactly behind the Sun if viewed from the Earth at the heliolatitude of 30° S. This position ensured a good view of the source of activity a few days before it appeared on the disk visible from the Earth. As inferred from onboard radio observations, the active region AR10808 produced at least four major flares on the far side of the Sun. The flare of importance X17+ generated an unusually intensive radio burst detected by ULYSSES and a shock wave that was recorded on September 14 and moved at a speed of ~1,210 km/s almost throughout the distance of 5 AU. The radio bursts associated with some other flares of importance X that occurred later had surprisingly low intensities. This agrees with what was observed in 2003 during the Halloween event. A probable explanation is that the inner heliosphere was filled with energetic electrons from the X17+ flare, and their flux was sufficient to suppress the plasma instability, which had to trigger radio emission [18]. In spite of the position on the limb decreasing its geophysical efficiency, this flare produced accelerated particles, a shock wave, and a small geomagnetic storm that were detected by CORONAS-F and other spacecraft in the near-Earth space [13].

In the context of the global warming and the necessity of working out an adequate response, an important task is to study long-term variations in the solar irradiance. Measurements of the 'solar constant' taken for 26 years on SOHO (VIRGO device) and other space missions were analyzed using complex calibration models, and no evidence of a significant long-term trend in the total solar irradiance was found; i.e., the brightness of the solar radiation did not increase or decrease systematically over the time scales observed. However, while the total solar irradiance changes by less that 0.1% during a solar cycle, the hard ultraviolet radiation changes by 30% over the time scale of a week and by a factor of 2 to 100 (depending on the wavelength) during a cycle. A detailed knowledge of the solar radiation spectrum is critical to understanding the climate variations and discriminating between the natural and man-caused effects [65]. Observations with the CORONAS-F/VUSS and SUFR instruments allowed us to measure weak variations in the total solar UV flux and to determine the contribution from the most intensive flares, which, as a rule, did not exceed a few percent in the observation band (about 120 nm) [66].

The absorption of solar hard X-rays (at the wavelengths of 8.42; 175, and 304 Å) [67] was studied using the CORONAS-F/SPIRIT X-ray telescope. As a result, the altitude dependence of the absorption coefficient was obtained with about a factor 100 height resolution. The density and composition of the Earth atmosphere as a function of solar activity was studied up to 500 km, and the content of the molecular nitrogen and atomic oxygen was determined. The experimental data obtained can serve as a basis for up-to-date models of the Earth atmosphere.

3. SOLAR SPACE PROJECTS UNDER PREPARATION

The CORONAS-PHOTON[9] mission, launched in 2009, is the third spacecraft in the CORONAS series. It will measure hard flare radiations and study the processes of energy build-up and transformation in solar flares, acceleration mechanisms, propagation and interaction of high-energy particles. In-situ measurements along the orbit (600 km initial altitude and ~82° inclination) and monitoring of solar X-rays penetrating the Earth atmosphere will contribute to the study of correlation between the physical and chemical processes in the upper atmosphere and solar activity. Helioseismic observations started by CORONAS-I (1996-2001) and CORONAS-F (2001—2005) will be continued. The scientific payload of the mission will comprise a high-energy emission spectrometer, a low-energy gamma-ray telescope, a polarimeter-spectrometer of hard X-rays and gamma-ray spectrometer, a high-speed X-ray monitor, a multi-channel UV monitor, a solar imaging telescope/spectrometer, an analyzer of charged particles, an electron and proton telescope, and an optical photometer.

The launch of the SDO[10] spacecraft to geosynchronous orbit with 28.5° inclination is scheduled for 2009. The mission is designed for continuous coordinated observations of the Sun from subphotospheric layers to the outer atmosphere. The main scientific task of the mission is to gain a better understanding of the impact of solar events on the Earth's space environment and improve its forecast by tracing solar magnetic fields (their generation, propagation, and vanishing) from the solar interior to 18 solar radii; studying the magnetic topology responsible for fast release of large amounts of energy on the scales from thousands to hundreds of thousands of kilometers; monitoring the dynamic processes that control space weather; and studying variations in the solar emission and structure both on small time scales and on the scale of a solar cycle. The onboard scientific complex will comprise three instruments: a combined solar telescope, a helioseismic and magnetographic telescope, and a hard UV radiometer. The complex is designed for high-resolution (0.6") observations of the solar atmosphere, for tracing magnetic fields from subphotospheric layers to the corona, for magnetic mapping, and helioseismic observations of the Sun.

In the *Solar Probe*[11] project the ballistic concept of the mission has been changed. Instead of gravitational maneuvers (GM) near the Jupiter which were assumed earlier for the approaching to the Sun the multiple GMs near the Venus are chosen (the new name of the project is *Solar Probe+*). The spacecraft will approach to the Sun in perihelion up to 9.5 solar radius, an inclination of the orbit with respect to the ecliptic plane will be only 3.4°. Using the advantage of close flyby, the mission is supposed to determine the structure and dynamics of the magnetic fields at the sources of the fast and slow solar wind; to trace the flow of energy that heats the corona and accelerates the solar wind; to determine what mechanisms accelerate and transport energetic particles; to explore dusty

9 http://iaf.mephi.ru/coronos-photon main.htm
10 http://sdo.gsfc.nasa.gov/
11 http://solarprobe.gsfc.nasa.gov/

plasma phenomena in the near-Sun environment and their influence on the solar wind and energetic particle formation. The scientific payload will comprise the instruments for in-situ measurements (Fast Ion Analyzer, Two Fast Electron Analyzers, Ion Composition Analyzer, Energetic Particle Instrument, Magnetometer, Plasma-Wave Instrument, Neutron/Gamma-Ray Spectrometer, Coronal Dust Detector) and White-Light Hemispheric Imager.

4. SOLAR SPACE MISSIONS UNDER DEVELOPMENT

The *Interhelioprobe*[12] and *Solar Orbiter*[13] missions are under development, respectively, in Russia and in the European Space Agency (ESA) [68—70]. By means of multiple gravity-assisted maneuvers at Venus (and at the Earth for the *Solar Orbiter*), the spacecraft in heliocentric orbit will gradually approach the Sun up to 40...45 solar radii and will reach the corotation region. The orbit can be inclined to the ecliptic plane by 38° through the same GMs. The principal scientific objective of the mission is the study of the fine structure and dynamics of the solar atmosphere, corona heating and solar-wind acceleration mechanisms, characteristics of the previously unattended inner heliosphere, polar and equatorial regions as viewed from high heliolatitudes. The spacecraft orbit will allow a number of unique observations, such as high-resolution observations of the solar atmosphere at short distance, observations and measurements in the corotation regime, long-lasting in-situ measurements in the immediate proximity to the Sun, observation of the polar regions, in-situ out-of-ecliptic measurements, stereo observations in cooperation with other spacecraft, observation of the far side of the Sun, etc.

The SENTINELS[14] mission will consist of four segments. Four spacecraft will be placed in heliocentric orbits in the inner heliosphere with the perihelion distance of 0.25-0.75 AU in the vicinity of the Sun-Earth line; one spacecraft in solar-synchronous trajectory will be orbiting the Earth; and one spacecraft in heliocentric orbit at 1 AU will be moving gradually away from the Earth. The mission will study the relationship between the solar events and interplanetary disturbances, particularly, those affecting the geospace (sources, acceleration and transport of particles; origin, evolution, and interaction of ejections, shock waves, and other geoeffective features). SENTINELS will link the solar segment of the NASA/LWS project to the geospace segment. The facilities and advantages of the LWS SENTINELS mission and *Solar Orbiter* will be combined in the HELEX mission (*Heliospheric Explorers*).

The *Solar Polar Imager*[15] will be launched to heliocentric orbit with the inclination of 75° at 0.5 AU. The solar sail will be used to reach the working orbit. For the first time, the solar activity will be observed from the poles. The study of the polar regions of the Sun will involve polar magnetic fields and their relationships with the heliospheric fields, dynamo and solar cycle, photospheric

12 http://www.izmiran.ru/projects/INTERHELIOPROBE/
13 http://www.orbiter.rl.ac.uk/; http://www.esa.int/esaSC/120384_index_0_m.html
14 http://sentinels.gsfc.nasa.gov
15 http://www.lws.nasa.gov/

fig. 18. Ballistic scheme of the PEP mission

and subphotospheric motions, helioseismology of the polar regions, polar corona, azimuthal and 3D structure of the corona and mass ejections. We hope to reveal differences, if any, between the polar and equatorial radiation and to find out how the characteristics of the polar solar wind and energetic particles correlate with the coronal features. The European Space Agency investigates the possibility of realizing the solar polar orbit at less than 0.5 AU based on the solar sail technology for in-situ measurements in circumsolar space.

China (in cooperation with Germany, France, Belgium, Austria, Canada, etc.) is developing the *KuaFu*[16] mission for the study of space weather, magnetic storms and auroras [71, 72]. Two satellites in polar orbits with identical sets of instruments and a spacecraft at the libration point ahead of the magnetosphere will survey the Sun and measure the solar wind streams and disturbances arriving at the Earth much in the same way as the SOHO mission did.

Within the STEREOSCOPE[17] project, the Institute of Solar-Terrestrial Physics in Irkutsk in cooperation with other institutions is planning to create a long-term interplanetary stereoscopic observatory with space vehicles at the libration points L4 and L4 on the Earth orbit aimed at studying the 3D structure of the solar atmosphere and reducing the damaging effect of solar events on geospace [73].

The *Polar-Ecliptic Patrol* (PEP) mission is under development at IZMIRAN [74]. Two spacecraft in oppositely inclined heliocentric orbits spaced by 0,5 AU will be shifted by a quarter of a period about one another, so that the Sun-Earth line will be constantly surveyed from out-of-ecliptic position (Fig. 18). The ballistic scheme will be similar to that of the *Interhelioprobe* mission. An important task of the mission will be out-of-ecliptic mapping of the heliosphere and solar disturbances and ejections with a wide-angle optical heliospheric telescope.

16 http://sess.pku.edu.cn/research/kuafu/

17 http://www.gao.spb.ru/russian/cosm/stereo/

Mission	Telescope	Wavelength, Å	Resolution on the Sun (km on the Sun)	Field of view of the Sun
SOHO	EIT	171, 195, 284, 304	1750 (2.5 arcsec per 1 AU)	Full
TRACE	Trace	171, 195, 284, UV range	360...725 (0.5...1 arcsec per 1 AU)	Partial
Hinode	SOT, XRT	2...60	145...725 (0.2...1 arcsec per 1 AU)	Full
STEREO	EUVI	171, 195, 284, 304	1120 (1.6 arcsec per 1 AU)	Full
SDO	AIA	EUV 7 channels UV 3 channels	420 (0.6 arcsec per 1 AU)	Full
Solar Orbiter	EUI	171, 195, 284, 304	70 (0.5 arcsec per 0.2 AU)	Full and partial
CORONAS-PHOTON	TESIS	8.42, 134, 304	725 (1 arcsec per 1 AU)	Full Corona 2...5 $R_{\odot}$
Interhelioprobe	TREK	185...195, 290...310	140 (1 arcsec per 0,2 AU)	Full and partial

Table 2. Spatial resolution of solar telescopes

SMESE (*Small Explorer for Solar Eruptions*) is a joint French-Chinese mission. A spacecraft in the solar-synchronous orbit will carry out continuous monitoring of the Sun and observe mass ejections and flares [75, 76].

The strategy of the future solar space missions involves further improvement of the spatial and temporal resolution of solar observations; observations outside the Sun-Earth line; spectroscopic imaging of the Sun and solar corona from high heliolatitudes; 3D imaging of the solar atmosphere, disturbances, and ejections, and in-situ measurements in the inner heliosphere.

Table 2 provides for comparison the resolving capacity of different missions. The best resolution (70...140 km in the Sun) will be achieved in the *Solar Orbiter* and *Interhelioprobe* missions, which will observe the Sun at short distances.

The solar and solar-terrestrial projects are implemented within the frames of the international program — *International Living With a Star* (ILWS) [6] based on the NASA LWS[18] program and aimed at a comprehensive study of the Sun and its impact on geospace.

18 http://lws.gsfc.nasa.gov/

CONCLUSION

Thus, for 50 years of the space era, the solar studies on spacecraft have passed a long way from the first Earth artificial satellites to complex space observatories. These studies have provided impressive results, which enlarged significantly our knowledge of the Sun, solar activity, and its impact on the Earth. Nowadays, still more valuable results are obtained from the currently operating solar space observatories *Hinode*, STEREO, RHESSI, SOHO, and TRACE. Besides, we are in expectation of the CORONAS-PHOTON, SDO, and other space missions. Particular emphasis in these missions will be placed on sophisticating the model of the solar interior and gaining a deeper insight into the cycle dynamics of the solar magnetic field and the origin of the cycle itself. There is still no definitive verdict as to the mechanism of heating of the solar corona and solar wind acceleration, although a significant progress has been made in the recent years. A better knowledge of the nature of most active solar events — flares and mass ejections — is required to develop sound physical methods for predicting space weather and its effect on the Earth. An important role in the solar studies belongs to the *International Living With a Star* program (ILWS), which joins the effort of scientists of many countries in solving the vital problems of solar and solar-terrestrial physics.

ACKNOWLEDGEMENT

I thank E.I. Prutenskaya for the help in translation of the paper.

REFERENCES

1. Fleck B. SOHO: The Solar Heliospheric Observatory // Newsletter of the European Astronomical Society 11—13.
2. Kuznetsov V.D. CORONAS-F Mission. Observations of the Sun and Manifestations of Solar Activity // Newsletter of the European Astronomical Society (EAS) / December 2006. Issue 32. P. 9—10.
3. Space Physics Strategy — Implementation Study. Vol. 1: Goals, Objectives, Strategy. Report of Workshop 1 January 22-26, 1990, Baltimore, Maryland. NASA. April 1991.
4. Marsden R.G. ULYSSES: Exploring the 3-D Heliosphere // Newsletter of the European Astronomical Society (EAS). December 2006. Issue 32. P. 7—8.
5. Vourlidas S. Seeing Double: The STEREO Mission // Newsletter of the European Astronomical Society (EAS). December 2006. Issue 32. P. 13—15.
6. http://ilws.gsfc.nasa.gov/
7. COSPAR Information Bulletin, № 153, April 2002, p. 18. Published for COSPAR by Elsevier Science, Ltd.
8. COSPAR Information Bulletin, № 160, August 2004, p. 42. Published for COSPAR by Elsevier Science, Ltd.
9. Kuznetsov V.D., Zhitnik I.A., Sobelman I.I. CORONAS-F: Contribution to Solar-Terrestrial Physics. // Vestnik RAN. 2005. V. 75. № 8. P.704—714.
10. Kuznetsov V.D. A Review of Solar and Solar-Terrestrial Studies on Board the CORONAS-F Mission. // Astron. Vestnik. 2005. V. 39. № 6. P. 485—494.

11. Kuznetsov V.D. Results of Solar Observations on Board the CORONAS-F Mission. // Izv. RAN. Physical Series. 2006. V.70. № 10. P. 1449—1452.
12. Zhitnik I.A., Kuzin S.V., Sobelman I.I., Bugaenko O.I., Ignatyev A.P., Mitrofanov A.V., Oparin S.N., Pertsov A.A., Slemzin V.A., Sukhodrev N.K., and Urnov A.M. Principal Results of the CORONAS-F/SPIRIT Experiment. // Astron. Vestnik. 2005. V. 39. № 6. P. 495—506.
13. Kuznetsov V.D. Observations of the Sun-Earth System within the CORONAS-F mission // Journal of Atmospheric and Solar-Terrestrial Physics. 2008. V. 70, Issue 2—4. P. 234—240.
14. Watanabe T. Solar-B scientific operation and its observing modes // Adv. Space Res. 2005. V. 36. Issue 8. P. 1478—1482.
15. Matthews S.A, Culhane J.L. Magnetic coupling of the Sun–Earth system — The view from STEREO //Adv. Space Res. 2007. V. 39. № 12. P. 1791—1803.
16. Kaiser M.L. The STEREO mission: an overview // Adv. Space Res. 2005. V. 36. Issue 8. P. 1483—1488.
17. Kosovichev A.G. et al. Sounding Solar and Stellar Interiors. / Eds. Provost J., Schmider F.-X. Proc. IAU Symp. № 181. 1998. P. 203.
18. Redfearn J., Bond P., Wilson A. The Space Dimension. ESA. BR-190. 2003. P.13—14.
19. ESA's Report to the 36th COSPAR Meeting, Beijing, China, July 2006. ESA, SP-1296. P. 15—26.
20. Bocchialini K. SOHO: a unique opportunity to understand the physical conditions on the Sun // Space Research in France. Report to COSPAR. CNES, Paris, France. 2004. P. 71.
21. Kosovichev A.G., Duvall T.L., Scherrer P. Time-Distance Inverse Methods and Results // Solar Phys. 2000. V. 192. P. 159.
22. COSPAR Information Bulletin, N 159, April 2004. P. 48.
23. Lindsey C., Braun D.C. Seismic Images of the Far Side of the Sun // Science. 2000. V. 287. P. 1799.
24. Lebedev N.I., Kuznetsov V.D., Oraevsky V.N., Staude J., and Kostyk R.I. CORONAS-F/DIFOS Helioseismic Experiment. // Astron. Zh. 2004. V. 81. № 10. P. 956—960.
25. Gizon L., Duvall JR T.L., Schou J.. Wave-like properties of solar supergranulation // Nature. 2003. V. 421. P. 43—44.
26. Science. 7 December. 2007. V. 318. Issue 5856. P. 1571—1597.
27. Astronomy and Astrophysics. 2008. V. 481. № 1.
28. http://www.nasa.gov/mission_pages/stereo/main/index.html
29. Beigman I.L., Bozhenkov S.A., Zhitnik I.A., Kuzin S.V., Tolstikhina I.Yu., Urnov A.M. Solar extreme vacuum UV spectra obtained in the CORONAS-F/SPIRIT experiment. I. Catalog of lines in the 280-330 Å range. // Pisma v Astron. Zh. 2005. V. 31. № 1. P. 39—58.
30. Sylwester J., Sylwester B., Kordylewski Z., Fillips K.J., Kuznetsov V.D., Boldyrev S.I.. Observations of solar X-ray spectra with the CORONAS-F/RESIK and DIAGENES spectrometers // Astron. Vestnik. 2005. V. 39. № 6. P. 537—548.
31. COSPAR Information Bulletin, № 160. August 2004. P. 41.

32. COSPAR Information Bulletin, N 154. August 2002.
33. Zhitnik I.A., Bugaenko O.I., Ignat`ev A.P., Krutov V.V., Kuzin S.V., Mitrofanov A.V., Oparin S.N., Pertsov A.A., Slemzin V.A., Stepanov A.I., Urnov A.M. Dynamic 10 MK plasma structures observed in monochromatic full-Sun images by the SPIRIT spectroheliograph on the CORONAS-F mission // Monthly Notice of Royal Astronomical Society. V. 338. Issue 1. 2003. P. 67—71.
34. Pankov V.M., Prokhin V.L., Havenson N.G.. X-ray CdTe-pectrometer RPS-1 in the CORONAS-F experiment. // Astron. Vestnik. 2006. V. 40. № 4. P. 344—349.
35. ESA's Report to the 35th COSPAR Meeting. Paris, France, 2004. ESA. SP-1276.
36. Schrijver C.J., Title A.M., van Ballegooijen A.A., Hagenaar H.J., Shine R.A. Sustaining the Quite Photospheric Network: The Balance of Flux Emergence, fragmentation, Merging, and Cancelation // Astrophys. J. 1997. V. 487. P. 424.
37. Parker E.N // Astrophys. J. 1988. V. 330. P. 474.
38. Aschwanden M.J., Tarbell T.D., Nightingale R.W., Schrijver C.J., Title A., Kankelborg C.C., Martens P., Warren H.P. Time Variability of the "Quiet" Sun Observed with TRACE. II. Physical Parameters, Temperature Evolution, and Energetics of Extreme-Ultraviolet Nanoflares // Astrophys. J. 2000. V. 553. P. 1047.
39. K.Shibata et al // Science. 7 December, 2007. V. 318. Issue 5856. P. 1591.
40. De Pontieu et al // Science. 7 December, 2007. V. 318. Issue 5856. P. 1574.
41. Katsukaw Y. et al // Science. 7 December, 2007. V. 318. Issue 5856. P. 1594.
42. Cirtain J.W. et al // Science. 7 December, 2007. V. 318. Issue 5856. P. 1580.
43. Banerjee D., Erdélyi R., Oliver R., O'Shea E // Solar phys. 2007. V. 246. P. 136.
44. Sylwester J., Gaicki I., Kordylewski Z., Kowalinski M., Novak S., Plocieniak S., Siarkowski M., Sylwester B., Trzebinski W., Culhane J.L., Whyndham M., Bentley R.D., Guttridge P.R., Lang J., H.Phillips K.J., Brown C.M., Doschek G.A., Oraevsky V.N., Kuznetsov V.D., Stepanov A.I., Lisin D.V.. RESIK — Bent crystal solar X-ray spectrometer for studies of Coronal Plasma Composition // Solar Physics. 2005. V. 226. № 1. P. 45—72.
45. Zhitnik I.A., Logachev Yu.I., Bogomolov A.V., Denisov Yu.I., Kavanosyan S.S., Kuznetsov S.N., Morozov O.V., Myagkova I.N., Svertilov S.I., Ignatyev A.P., Oparin S.N., Pertsov A.A., Tindo I.P. Results of measurements of polarization, temporal, and spectral characteristics of flare-generated hard X-rays as inferred from the CORONAS-F/SPR-N experiment. // Astron. Vestnik. 2006. V. 40. № 2. P. 108—119.
46. Solar physics. 2002. V. 210. № 1—2. P. 3—32. Special Issue: The Reuven Ramaty High-Energy Solar Spectroscopic Imager (RHESSI): Mission Description and Early Results.
47. Harrisson R., Davis C., Eyles C., Halain J.-P., Howard R.. Searching for Solar Clouds in Interplanetary Space // Space Research Today — COSPAR's Information Bulletin. 2007. № 168. P. 25—31.
48. Hassler D.M., Dammasch I.E., Lemaire P., Brekke P., Curdt W., Mason H.E., Vial J.-C., Wilhelm K. Solar wind outflow and chromospheric magnetic network // Science. 1999. V. 283. P. 810.
49. T.Sakao et al // Science. 7 December, 2007. V. 318. Issue 5856. P. 1585.

50. COSPAR Information Bulletin. № 159. April 2004. P. 48.
51. COSPAR Information Bulletin. № 146. December 1999. P. 21.
52. Cranmer S.R. Coronal Heating versus Solar Wind Acceleration // Proc. SOHO-15: Coronal Heating. 6— Sept. 2004. St. Andrews, Scotland, ESA SP-575.
53. COSPAR Information Bulletin. № 129. April 1994. P. 36.
54. COSPAR Information Bulletin. № 159. April 2004. P. 49.
55. Inputs to Highlights in Space 2001. COSPAR for UN (Comission D). SO/01160/KPW/cn. Part of the report "Space Plasmas in the Solar System, Including Planetary Magnetospheres".
56. Veselovsky I.S., Panasyuk M.I., Avdyushin S.I. et al. Solar and Heliospheric Phenomena in October—November 2003: Causes and Effects // Cosmic Research. 2004. V. 42. №. 5. P. 435—488. Translated from Kosmicheskie Issledovaniya. 2004. V. 42. №. 5. P. 453—508.
57. Kuznetsov S.N., Kurt V.G., Myagkova I.N., Yushkov B.Yu., Kudela K. Flare-generated gamma-ray emission and neutrons recorded with the SONG device in 2001-2004 // Astron Vestnik. 2006. V.40. № 2. P. 120—126.
58. Kuznetsov S., Kurt V., Yushkov B., Myagkova I., Kudela K., Belov A., Mavromichalaki H., Caroubalos C., Hilaris A., Preka-Papadema G., Mussas X. October 28 2003 X10 Flare: High Energy Gamma Emission, Type II Radio Emission and Solar Particle Observations // Intern. J. Mod. Phys. 2005. V.20. №. 29. P. 6705—6708.
59. Kuznetsov S.N., Myagkova I.N., Yushkov B.Yu., Denisov Yu.I., Muravyeva E.A., Kudela K. Dynamics of the Earth's radiation belts during large magnetic storms as inferred from CORONAS-F data // Astron. Vestnik. 2007. V. 41. № 4. P. 369—378.
60. Kuznetsov S.N., Yushkov B.Yu., Denisov Yu.I., Kudela K., Myagkova I.N. Dynamics of the magnetosphere penetration boundary of solar cosmic rays as inferred from CORONAS-F data // Astron Vestnik. 2007. V. 41. № 4. P. 379—384.
61. Yermolaev Yu.I., Zelenyi L.M., Zastenker G.N. et al. Solar and Heliospheric Disturbances that Resulted in the Strongest Magnetic Storm of November 20, 2003. // Geomagnetism i aeronomiya. 2005. V. 45. № 1. P. 23—50. (Translated in Geomagnetism and Aeronomy. 2005. V. 45. P. 20—46)
62. Yermolaev Yu.I., Zelenyi L.M., Zastenker G.N. et al. A Year Later: Solar, Heliospheric, and Magnetospheric Disturbances in November 2004 // Geomagnetism i aeronomiya. 2005. V. 45. № 6. P. 723—763. (Translated in Geomagnetism and Aeronomy. 2005. V. 45. №. 6. P. 681—719)
63. Kuznetsov S.N., Kurt V.G., Yushkov B.Yu., Myagkova I.N., Kudela K., Kaššovicová J., Slivka M. Proton acceleration during 20 January 2005 solar flare: CORONAS-F observations of high-energy gamma emission and GLE // 29th International Cosmic Ray Conference. Pune (2005). V. 1. P. 49—52.
64. Arkhangelskaya I.V., Arkhangelsky A.I., Glyanenko A.S., Kotov Yu.D., Kuznetsov S.N. Behavior of the time profiles of gamma-ray bursts generated in the January 2005 solar flares in different energy ranges as inferred from the CORONAS-F/AVS-F data // Proc. of the National Conference "Experimental and Theoretical Studies of Foundations of Heliogeophysical Forecast", Troitsk, Moscow Region, 10-15 October 2005. P. 15—20.
65. http://sohowww.nascom.nasa.gov/

66. Nusinov A.A., Kazachevskaya T.V. CORONAS-F spacecraft observations of high-frequency ultraviolet and X-ray emissions from solar flares in 2001—2003 // Astron. Vestnik. 2006. V. 40. № 2. P. 127—133.
67. Slemzin V., Bugaenko O., Ignatiev A., Krutov V., Kuzin S., Mitrofanov A., Oparin S., Pertsov A., Stepanov A., Zhitnik I.. Investigation of absorption of solar EUV-radiation in the Earth's atmosphere at altitudes of 100—500 km using solar images in the experiments TEREK-C (CORONAS-I) and SPIRIT (CORONAS-F). Solar Variability as an input to the Earth's Environment // Proc. ISCS 2003 Symp. ESA SP-535. 2003. P. 389-392.
68. Oraevsky V.N., Galeev A.A., Kuznetsov V.D., Zelenyi L.M. Solar Orbiter and Russian Aviation and Space Agency Interhelioprobe // Solar Encounter: The First Solar Orbiter Workshop. Proc. ESA SP-493. 2001. P. 95—108.
69. Marsch E. et al. Solar Orbiter — A High Resolution Mission to the Sun and Inner Heliosphere // Solar Encounter: The First Solar Orbiter Workshop. Proc. ESA SP-493. 2001. P. XI —XXVI.
70. Marsch E., Marsden R., Harrison R., Wimmer-Schweingruber R., Fleck B.. Solar Orbiter — mission profile, main goals and present status // Adv. Space Res. 2005. V. 36. Issue 8. P. 1360—1366.
71. Wang J., Zhang J. KuaFu and the studies of CME initiation // Adv. Space Res. 2007. V. 40. № 12. P. 1770—1779.
72. Tu C.-Y., Schwenn R., Donovan E., Marsch E., Wang J.-S., Xia L.-D., Zhang Y.-W. Space weather explorer — The KuaFu mission // Adv. Space Res. 2008. V. 41. № 1. P. 190—20.
73. Grigoryev V.M. Space-born Solar Stereoscope Experiment in Solar Physics // Solar Phys. 1993. V. 148. P. 386—391.
74. Kuznetsov V.D., Oraevsky V.N. "Polar Ecliptic Patrol (PEP)" for Solar Studies and Monitoring of Space Weather // J. British Interplanetary Society. 2002. V. 55. № 11/12. P. 398—403.
75. Vial J.-C., Auchère F., Chang J., Fang C., Gan W.Q., Klein K.-L., Prado J.-Y., Trottet G., Wang C., Yan Y.H. SMESE: A SMall Explorer for Solar Eruptions // Adv. Space Res. 2007. V. 40. № 12. P. 1787—1801.
76. Vial J.-C., Auchère F., Chang J., Fang C., Gan W.Q., Klein K.-L., Prado J.-Y., Rouesnel F., Sémery A., Trottet G., Wang C.. SMESE (SMall Explorer for Solar Eruptions): A Microsatellite Mission with Combined Solar Payload // Adv. Space Res. 2008. V. 41. № 1. P. 183—189.

A. NISHIDA
2-9-25 Tamagawa-Gakuen, Machida,
Tokyo 194-0041, Japan
hirosoph@dj.mbn.or.jp

THE EARTH'S PLASMASPHERE: 40 YEARS OF THE CONVECTION MODEL

Discovery of the plasmasphere was one of the highlights of the early days of space research. This structure was soon interpreted in the framework of the growing paradigm of the magnetospheric convection in combination with the idea of the supersonic outflow of ionospheric plasma, the polar wind. This paper gives an updated overview of our understanding of the plasmasphere with emphasis on the advances made in recent 15 years. It has been seen that the convection model offers adequate interpretation to a wealth of observations made with advanced techniques, but new problems have also emerged. An outstanding problem is the nature of the notches, that is, deep indentations of the plasmasphere that can last several tens of hours while lagging behind the Earth's rotation. The energetics of the polar wind need also be clarified by incorporating the relaxation of photoelectrons in thermal electrons that seems to be much more effective than is expected from the Coulomb collisions alone.

INTRODUCTION

The International Geophysical Year (IGY) in 1957—1958 gave a great impetus to research in space surrounding the Earth. Extensive observations highlighted by the launch of Sputnik revealed that the upper atmosphere which used to be described with only a few key words such as the ionosphere and the exosphere was a complex world rich in internal structures. One of them was the plasmasphere. This is a domain of cold plasma that is bounded sharply by magnetic field lines that cross the equatorial plane at the distance of about 4 R_e. The sharp decrease in the plasma density at the plasmapause was deduced from the analysis of the whistler atmospherics by Carpenter [11; see also 12], and it was detected by one of the earliest satellite missions (*Lunik-1*) in 1959 [Gringauz, 25; see also 26].

1. CONVECTION MODEL OF PLASMAPAUSE FORMATION

In parallel to these findings from observations, Axford and Hines [6] and Dungey [15] proposed a canonical theory of the dynamics of the magnetosphere. According to this theory the interaction with the solar wind produces a global convection in the magnetosphere which governs its internal structure. Axford and Hines suggested that the magnetosphere was divided into two regimes by convection; the solar wind induced convection dominates in the outer regime while the corotation with the Earth dominates in the inner regime. In their convection profile the boundary between these regimes was depicted at about

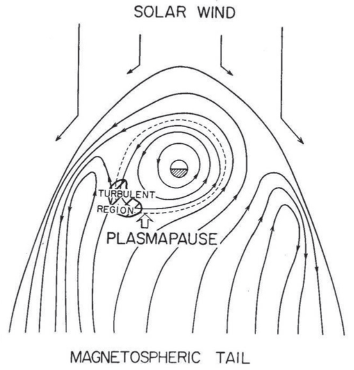


fig. 1. Convection model of the plasmapause formation as originally proposed [38]

the distance where the plasmapause was found, but they did not tell what determined this distance. The information they could use to infer the profile of the solar wind induced convection was the equivalent current system of polar magnetic disturbances (DS). Since the global distribution of these disturbances is controlled not only by the applied electric field but also by enhancement of the ionospheric conductivity due to precipitating electrons, it was difficult to derive the electric field from their ionospheric equivalent current system. The low latitude extent of the disturbance electric field was particularly difficult to determine since the fields of the geomagnetic daily variation and from the field-aligned currents had to be separated. It was only in the early 1980s that an algorithm was developed that could make such a separation objectively.

Around this time, we found a new type of geomagnetic variations having a global scale, DP2. DP2 variations typically had timescales of about one hour and appeared coherently from inside the polar cap to the equator. They do not originate from intensifications of the auroral electrojet, and the extension of the ionospheric current to the equatorial region was demonstrated by the occurrence of the equatorial enhancement. The system of the electric field deduced from DP2 consisted of twin vortices that projected to the dawnside and duskside convection cells [43; 39]. We interpreted DP2 fluctuations as the manifestation of time variations of the solar wind induced convection, and suggested further that the convection was generated by the magnetic reconnection on the dayside magnetopause because we found that DP2 fluctuations showed peak-to-peak correspondence with the southward component of the interplanetary magnetic field (IMF). The time delays between the estimated bow shock crossing time of IMF structure and DP2 on the ground were 7—9 minutes only, demonstrating a very rapid response of the convection to the driving agent [40].

Global scale of the DP2 fluctuations extending to the equator means the convection electric field induced by the solar wind penetrates deep inside the magnetosphere down to the equatorial ionosphere. It follows from this that the solar wind induced convection (SWIC) is to be added to the corotation to obtain the streamlines over the entire magnetosphere. The resulting profile of the convection is shown in Fig. 1. The streamlines are naturally divided into the inner and outer regimes, and their bounding surface shows all the observed characteristics of the plasmapause such as the average distance, local time dependence, and Kp dependence [38]. In this diagram the convection was supposed to be generated at the flanks of the magnetosphere following the picture by Axford and Hines [6]. A version based on the open model of the magnetosphere was presented by Brice [9].

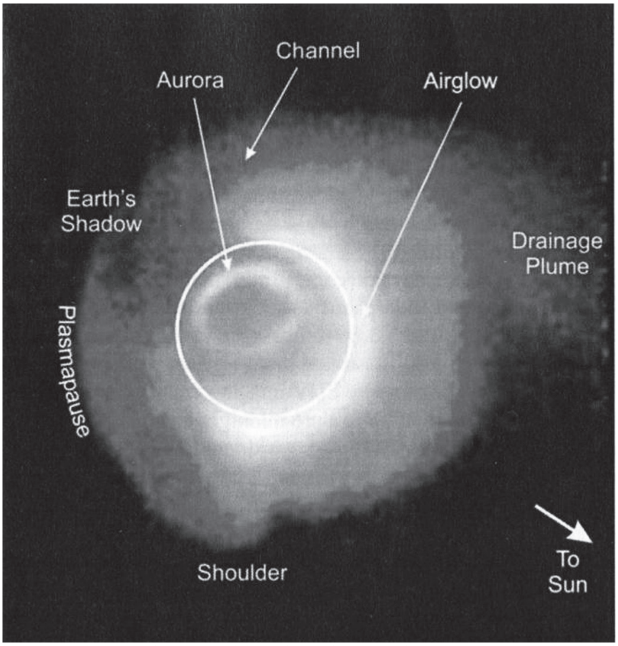


fig. 2. Plasmasphere observed by the ultraviolet imager on board IMAGE during a magnetic storm. The while circle shows the approximate size and position of the Earth [47]

Sharp drop in density observed at the plasmapause means that the ionospheric plasma that is supplied along field lines outside the bounding surface of streamlines is dumped to the solar wind when these field lines are open. Originally we estimated the escape time by assuming the loss due to diffusion and found that the plasma is evacuated in several hours [38]. But it was proposed later that the plasma could be lost supersonically as the polar wind [7; 36]. From the ground, the ionospheric counterpart of the plasmasphere was identified as the density trough over high latitudes [e.g., 41].

In some 40 years since this model was published a great wealth of information has been obtained on the plasmapause by observations both from the ground and onboard satellites. The most comprehensive among the satellite observations was the imaging of the plasmasphere by the *Image for Magnetosphere-to-Aurora Global Exploration* (IMAGE, NASA, 2000—2005) satellite. Using global images formed by lights scattered by He^+ ions, EUV (Extreme Ultraviolet Imager) observations by IMAGE have confirmed the dynamic character of the plasmasphere and made it possible to follow the global responses of the plasmasphere to the solar wind conditions continuously. An example of plasmasphere images taken by EUV is shown in Fig. 2.

Our simple model has served as the basis for interpreting these features throughout with incorporating of two elements. One is the effect of the time variation in the SWIC. The most obvious of these is the erosion of the duskside plasmapause in response to the enhancement of SWIC [24; 42]. The other is the effect of polarization of the warm plasma that flows earthward from the plasma sheet on the nightside. This warm plasma (about 1 keV) is subject not only to the cross **E** drift but also to the gradient/curvature drift that is charge dependent. Azimuthal separation of ions and electrons produces dusk-to-dawn electric field that opposes SWIC electric field [8], whereas their radial separation produces the anti-earthward electric field that enhances the westward flow in the dusk-midnight sector of the plasmapause [19]. The strength of the polarization electric field is determined by the balance between the pressure drift current $\nabla V \times \nabla P$ in the magnetosphere and the closure current in the ionosphere, where V represents volume of magnetic flux tubes [51; 28]. Simulation schemes have been built to follow the time development of the coupled system [e.g., 16; 33; 52].

2. EFFECTS OF SHIELDING

Fig. 3 shows the SWIC equipotential contours, that is, streamlines, obtained by simulations using the RCM scheme [52]. The corotation is not included. The panel (b) is for the time shortly after the electric potential across the polar cap is increased. It represents the unshielded penetration of the externally applied electric field; the departure from the exactly parallel contours is due to sharp jump in the ionospheric conductivity at the dawn and dusk terminators. The panel (a) is for the time when the shielding has fully developed just before the above increase in the potential is applied. The shielding electric field stands out in the panel (c) where the polar cap potential has been reduced by a factor of two from the time of (b). The shielding electric field is represented by two vortices centered in dusk and dawn sectors and is directed dusk to dawn in the inner magnetosphere while the applied electric field is directed essentially from dawn

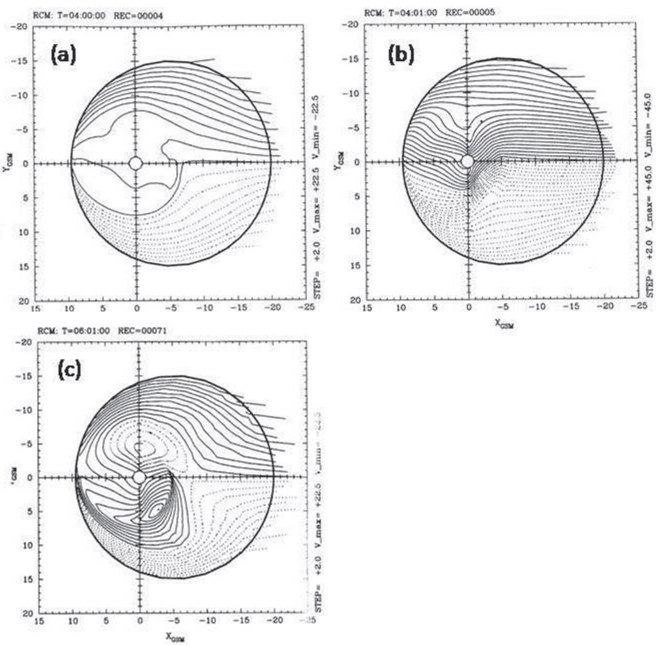


fig. 3. Equipotential contours in the equatorial plane, not including the corotation electric field, computed by RCM (a) for a time of good shielding under the polar-cap potential of 45 kV, (b) just after a sudden increase in the polar cap potential to 90 kV, and (c) immediately after the factor 2 reduction of the polar-cap potential [52]

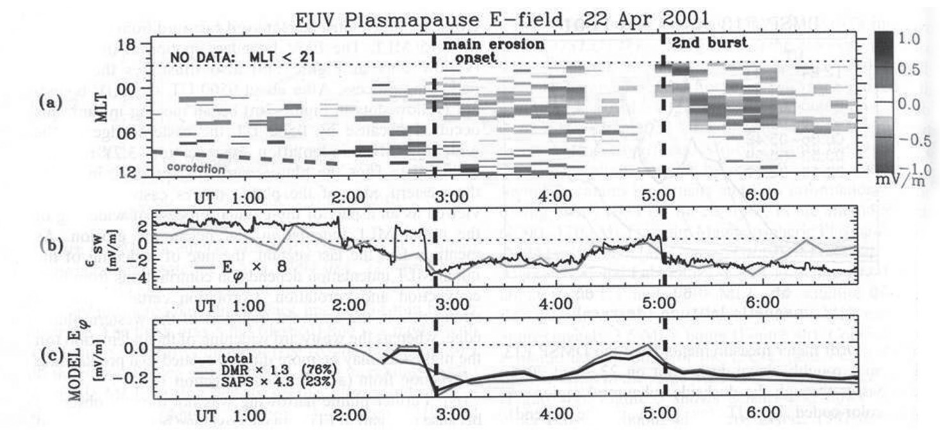


fig. 4. Electric field at the plasmapause inferred from IMAGE EUV on April, 22, 2001 during the first few hours of the plasmapause erosion. (a) Eπ(component of electric field locally tangent to the plasmapause); red(blue) represents inward (outward) radial motion at a fixed MLT, the dashed line indicates strict corotation with the Earth in this UT-MLT frame. (b) Dawn-dusk solar wind electric field (black) and the plasmapause electric field x 8 (light) in the range of 0000-0300 MLT. (c) Model plasmapause electric field in the 0000-0300 MLT range. Blue: scaled DMR (dayside magnetopause reconnection) contribution which is about 10% of solar wind field, red: scaled SAPS contribution (about 3% of the solar wind field). Black: total. For coloured image, please, refer to [22]

to dusk. The time taken for the shielding to be established is 10—20 minutes according to this simulation [50].

The electric field in the inner magnetosphere has been derived from the observations of the plasmapause motion during an interval of several hours which includes a few sharp increases in the southward component of the IMF [22]. The middle panel of Fig. 4 compares the azimuthal component of the electric field derived from the motion of the plasmapause in the 0000-0300 MLT range (light line) with the dusk-to-dawn component of the electric field in the solar wind (black line). The plasmapause electric field tracks the solar wind electric field rather well, and notably the onsets of the plasmasphere erosion indicated twice by vertical dashed lines agree with increases in the electric field. The top panel shows the intensity of the electric field parallel to the plasmapause with respect to UT and MLT in the night sector; Red bars after the southward turning of IMF indicate occurrence of strong inward flows. The bottom panel is a comparison of the models where DMR means the electric field due to reconnection on the dayside magnetopause and SAPS will be discussed later.

The electric field at the plasmapause shown in Fig. 4 is supposed to include the effect of the shielding. If the shielding takes more than a few tens of minutes to develop the ratio between the plasmapause electric field to the solar wind electric field is expected to increase first to the level of the full penetration before becoming reduced due to shielding. But such a trend is not visible in this example or in other case studies of the IMAGE data. Observations of the low-latitude electric field by the Jicamarca incoherent scatter radar have shown that the electric field stayed at an enhanced level when the IMF turned southward and remained continually southward for 2—3 hours [29]. This suggests the shielding is either established so fast that the shielded electric field is observed as soon as SWIC is intensified or it takes so long that the external conditions have changed before it becomes appreciable. It is noted in this regard that the partial ring current, which signifies the earthward entry of the warm particles from the plasma sheet and is thus related to the development of the shielding electric field, starts to develop with a short time lag (5—11 min) from the growth of the ionospheric convection in the polar cap [27].

The electric field derived from the plasmapause observations is 10...20% of the solar wind electric field which is assumed to be given simply by $V_{sw}B_{IMF, z}$ [22]. On the other hand, the polar cap potential has been found to be about 35% of the solar wind electric field that is estimated by the component reconnection model [23]. The latter estimate of the solar wind electric field includes the contribution from B_y component of IMF and is roughly twice as large as the electric field derived from B_z alone. Combining the above we find that the fraction of the unshielded electric field is about 15...30% of the polar cap potential. This estimate, however, is subject to uncertainty because it is not clear when the shielding is really established. In order to derive the degree of the shielding and its time development in a clearer manner we hope that future analyses will compare the electric field at the plasmapause with the polar cap potential rather than with the solar wind electric field.

The shoulder structure that is seen in Fig. 1 (p. 96) in the postmidnight sector of the plasmapause has been interpreted to be a manifestation of the shielding electric field [20]. When the SWIC convection is suddenly weakened at the northward turning of the interplanetary magnetic field, the shielding electric field illustrated in Fig. 3c (p. 98) remains for some time making the case of the 'overshielding'. In the postmidnight sector of the plasmapause the convection due to the shielding electric field is directed dusk to dawn and imposes an antisunward flow. In the dawn sector this convection under the shielding electric field is directed opposite to the corotation and slows the eastward motion. Hence an asymmetric bulge called shoulder can be expected to result.

The overshielding effect has been observed in the current flow in the equatorial ionosphere on the dayside. During the recovery phase of a substorm, the equatorial electrojet changed direction from eastward to westward, that is, it reversed from the direction of DP2 current at the dayside equator, when IMF suddenly turned northward [30]. In another event, the equatorial DP2 started to decrease one hour after the onset of the ring current development when IMF was still southward but decreasing in magnitude, indicating shielding effects becoming dominant at the equator [31]. Growth and decay of the shielding electric field seem to depend on several conditions including the form of IMF variation, pressure of the warm plasma injected from the tail, ionospheric conductivity, and locations of R1 and R2 field-aligned currents.

3. SUB AURORAL POLARIZATION STREAMS (SAPS)

In the basic convection profile of Fig. 1 (p. 96), the sunward flow from the nightside is deflected clockwise by the effect of the corotation and envelopes the plasmapause. The streamlines depicted here are the equipotential contours of the electric field which represent the convection path of the low energy plasma, but the plasma from the tail is in fact warm and the grad-B drift has to be added on the **E**x**B** drift. Since the grad-B drift for positive ions is directed opposite to the corotation the ions flowing from the tail can come closer to the Earth than the plasmapause. For the warm electrons from the tail this drift is directed eastward like the corotation so that they are kept further away than the plasmapause. The resulting polarization produces the electric field that is directed outward in the dusk-to-midnight sector.

This is shown by the RCM simulation in Fig. 5 [50]; see strong concentration of equipotentials in this sector around L ~ 4. The polarization produces field-aligned currents that are closed by the ionospheric Pedersen current in the subauroral region. Since the polarization is produced on the earthward side of the earthward boundary of the warm electrons from the tail, the electron precipitation is not available for enhancing the ionospheric conductivity. Hence the electric field is not much reduced and generates a strong westward drift in the ionosphere. This phenomenon has been called Polarization Jets, Sub Auroral Ion Drifts, or Sub Auroral Polarization Streams where its different aspects are emphasized.

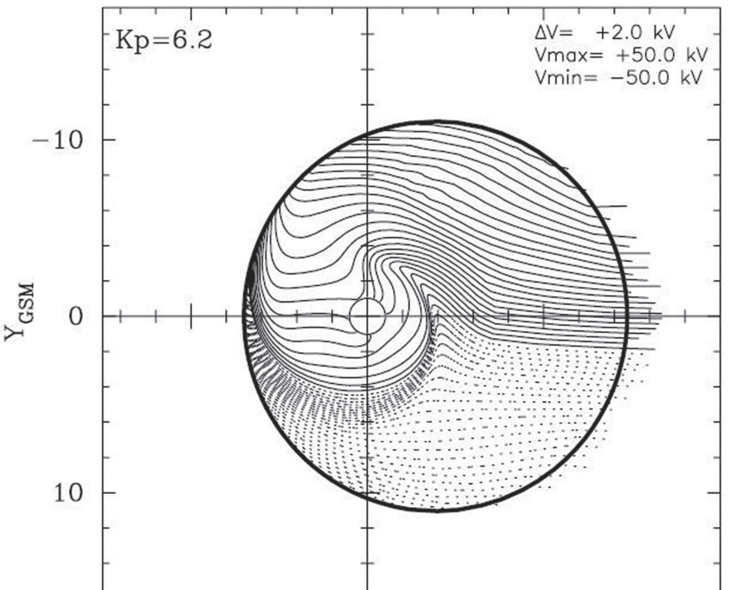


fig. 5. Equatorial equipotential contours in the corotating frame showing a narrow jet of fast flow at L=3...4 mainly in the dusk-midnight sector where SAPS is observed [50]

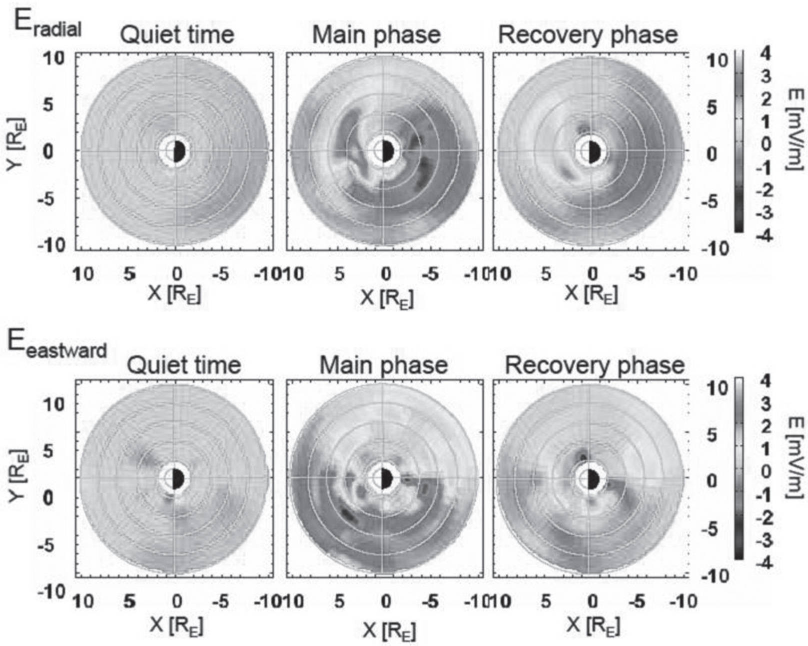


fig. 6. Electric fields observed in quiet times (left) and main and recovery phases of storms (center and right) by *Akebono* are projected to the equatorial plane. Top: radial component, bottom: azimuthal component [44; reformatted]

A statistical summary of the electric field observations by the *Akebono* satellite projected to the equatorial plane is shown in Fig. 6. The corotation field has been subtracted. Radial component is in the upper row and eastward component is in the lower row, and main and recovery phase distributions are shown in middle and right columns, respectively. The cases when the intensity of the symmetric ring current (SYM-H) was lower than -50 nT at minimum have been selected as storms. During the main phase, intensification of the dawn-to-dusk electric field is seen both in radial and azimuthal components. The maximum strength of the radial component is about 1.5 times stronger at dusk than at dawn, and this difference suggests the contribution from the polarization electric field of SAPS. The duskside peak extends to the afternoon sector, and it remains during the recovery phase in the dusk-afternoon sector [44]. The electric field observation by CRRES (USA, 1990—91) has shown that the increase in the electric field in the premidnight inner magnetosphere starts within 30 sec of the substorm onset [45].

Signatures of SAPS have been observed in the ionosphere at times of high magnetic activity. Radar observations at Millstone Hill have shown that SAPS is seen as a band of westward ion velocity in the subauroral region that extends from higher latitude at dusk to lower latitude in post midnight. The average peak amplitude is greater than 900 km/s in the dusk sector and about 400 km/s in the predawn sector when Kp is 5+ or 6o [17].

4. EFFECTS OF THE ENHANCED CONVECTION

When SWIC is intensified, position of the plasmapause changes in two ways according to the convection model. First, it moves toward the Earth on the nightside but away from the Earth on the dayside as the sunward component of the flow velocity is increased. The plasmapause becomes sharp on the nightside while a broad bulge develops on the dayside. Second, the outermost layer of the plasmasphere is peeled in the dusk region where the sunward velocity of SWIC has exceeded the anti-sunward velocity of the corotation. The resulting structure is a plume extending from the duskside plasmapause toward the dayside magnetopause.

This is illustrated by observations by the IMAGE satellite in Fig. 7 and 8 [48]. In each panel of Fig. 7 plasmapause configurations before and after a sudden

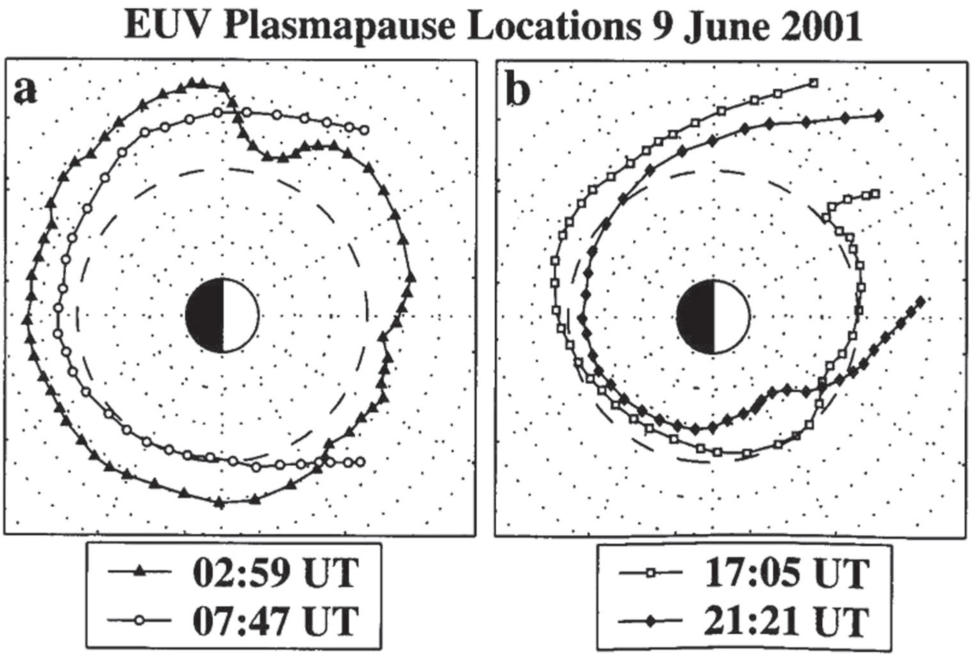


fig. 7. Changes in the plasmapause location. (a) Initial nightside erosion and dayside plume formation, and (b) secondary erosion and sunward surge [48]

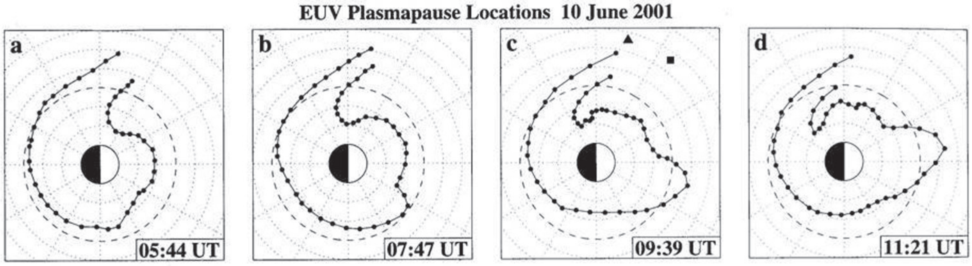


fig. 8. Changes in the plasmapause location: wrapping of the plume [48]

southward turning of IMF are compared. In panel (a) the filled triangles show the plasmapause position during an interval of northward IMF, and the open circles show the plasmapause location during the disturbance; the plasmapause became closer to the Earth by ~$1R_e$ over the entire nightside, and a broad bulge was formed on the dayside. In panel (b) the open squares show the pre-disturbance location where the plume formed in the earlier activity still remained on the dusk side, and the filled diamonds show the position during the disturbance where a broad bulge was formed on the dayside.

Fig. 8 shows development of the plume. The plume plasma was carried sunward by SWIC but its root attached to the plasmasphere moved eastward in the direction of the Earth's rotation when the intensity of SWIC became lower after the initial enhancement. Hence the plume wrapped around the plasmapause in the dusk-night sectors as seen in panel (d). The outward polarization field of the SAPS makes the plume plasma to flow faster and extend further [21].

The response time of the plasmapause to the southward turning of IMF has been estimated to be 10...30 min from the cross correlation between the solar wind electric field and the plasmapause motion for 16 such events [37]. In interpreting the response times given here and elsewhere from similar observations, we note that SWIC is driven not only directly by the reconnection with IMF but also by the release of solar wind energy stored in the tail. The former corresponds to DP2 and the latter to substorms (DP1), and the onset of the substorm expansion phase tends to be delayed by 40...60 min from the southward turning of IMF. The response time derived from these analyses could be for the sum of these components of SWIC.

When the bulge or plume reaches the magnetopause and the magnetic field lines which contain the plasma of the plasmaspheric origin become open, cold and dense plasma is mixed with the solar wind plasma in the magnetosheath. Such a mixture has been observed by the *Interball-Aurora* (Russia, 1996—2001) satellite in high latitudes [49]. On the basis of the EUV imagery of He^+ line by *Nozomi* (JAXA, launched in 1998) when this spacecraft was departing for Mars, it has been suggested that peeling of plasma from the dusk sector of the plasmasphere can occur even in a quiet/moderate geomagnetic condition due probably fluctuations in the convection field [55].

5. NOTCH AND SUB-COROTATION

Imaging of the plasmasphere by the IMAGE satellite has shown that the configuration of the plasmasphere changes dynamically in response to IMF variations and magnetic activities. Many of these changes can be interpreted by, or have been anticipated from, the convection model of the plasmapause formation. However, it has also revealed a new type of the deformation that cannot readily be interpreted by this model. This is the notch.

A notch (marked by an arrow) is followed for 55 hours in five panels of Fig. 9. Notches are characterized by deep density depletions that extend mostly radially inward to L=2 or even less. The sizes in local time range from very narrow (~0.1 hours MLT) to very broad (~3 hours MLT). Notches appear to share their origin with low-density channels, which is the gap between the plume and the plasmapause proper formed on the duskside upon recovery from the enhanced convection. Notches can remain a prominent feature for several tens of hours before refilling. Most of the notches drift at a rate between 85% and 97% of the corotation and often at the same rate across a wide range of inner L shells. This is seen in Fig. 9 where the notch lags behind the red line which corotates with the Earth [10; 18; 47].

fig. 9. EUV images of the plasmasphere where the plasmasphere density notch is recorded for 55 hours. The radial line, which bisects a notch in the first panel, tracks the Earth's rotation in the subsequent panels. For coloured image, please, refer to [10]

If notches are to be explained by the convection model it has to be assumed that the outer streamlines originating from the tail can sometimes come so close to the Earth as 1.6 R_E. For notches to be maintained in an almost radial configuration for hours it has to be assumed that the sum of SWIC and the corotation yields the same angular velocity over the entire radial extent of a notch. These seem to be difficult assumptions. Burch et al. [10] have suggested that the corotation lag is due to the flow of the heated neutral gas from higher to lower latitudes. Because of conservation of the angular momentum the azimuthal component of velocity of such gases is lower than the local velocity of rotation of the Earth. The effect is transmitted along magnetic field lines and makes the plasmasphere rotate slower than the corotation. This is an interesting idea, but the long lifetime of notches still poses a difficult problem since the equatorward atmospheric motion may not easily be expected to last so long.

6. FILLING OF THE PLASMASPHERE

In the convection model of the plasmasphere formation, distribution of plasma along the field lines outside the plasmapause is not in the static equilibrium but is being replenished by the upward flow from the ionosphere. Supply of the ionospheric plasma also occurs on the field lines which are convected from the tail and entrained in the plasmasphere due to weakening of the convection. Extensive observations have shown that the outflow (the polar wind) is supersonic and suggested that photoelectrons play a vital role in driving this flow.

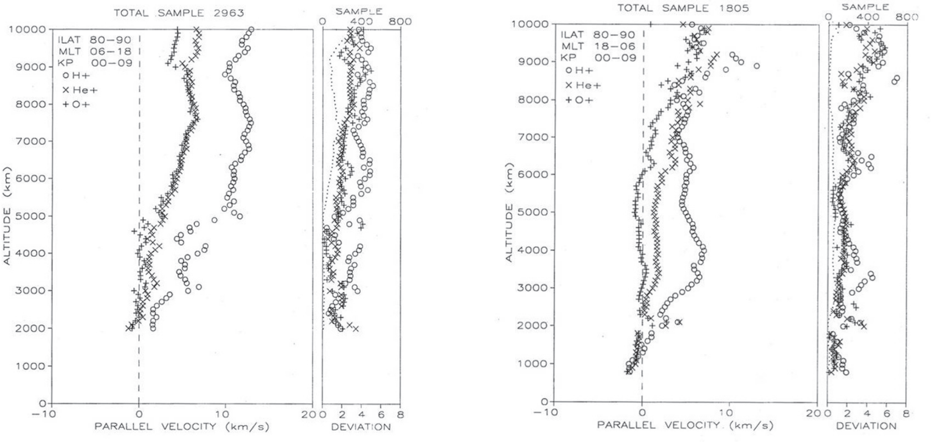


fig. 10. Altitude profiles of the H+ polar wind velocity (left panel) and standard deviation (right panel) above 80°invariant latitude observed on *Akebono* near sunspot maximum as a function of altitude, in (a) dayside and (b) nightside, respectively [1]

The polar wind ions consist primarily of H^+, He^+, O^+ ions. Fig. 10 shows the outflow velocities of these ions along magnetic field lines that were observed by the *Akebono* satellite (JAXA, launched in 1989) above the invariant latitude of 80° in the altitude range of 1,000 to 10,000 km. The daytime data are in panel (a) and the nighttime data are in panel (b). Averages are to the left and standard deviations are to the right in each panel. Velocities of all three species (H^+, He^+, O^+) increase almost monotonically with altitudes. Significant accelerations begin above 2,000 km; the H^+, He^+, O^+ velocities reach 1 km/s near 2,000 km, 3,000 km and 4,000 km respectively. The velocity ratio between the ion species lies on average between unity and the inverse square root mass ratio of the species, that is, the kinetic energy of lighter ions is lower than that of heavier ions [1]. The transition of the ions velocities from subsonic to supersonic is estimated to occur near 1,500; 3,000 and 6,000 km for H^+, He^+, O^+ ions, respectively, based on average ion temperature of 0.2 eV above 7,000 km, and the Mach numbers at 7,000 km are about ~1.5, 1.5 and 5 for each of them [54]. The average upward H^+ ion flux observed at 6,000…9,000 km in the noon sector and projected to 2,000 km altitude is in the range of $1…10·10^7$ $cm^{-2}s^{-1}$ and, surprisingly, the corresponding O^+ flux is found to be comparable to the H^+ ion flux [3]. The H^+ flux integrated over the polar ionosphere (ILAT ≥ 75°) correlates inversely with the Kp index and is 2~3 times lower in active times than in quiet times [4].

As seen in Fig. 10 velocities of all ions are significantly higher on the dayside than on the nightside. Fig. 11 shows that the outflow velocity of H^+ ions increases with electron temperature, being more pronounced at higher altitudes [2]. The temperature of the polar wind ions is generally lower than that of electrons and is in the range of 0.05…0.3 eV between 7,000 km and 10,000 km [14]. The ratio of the ion temperature in the plasmasphere to the ionospheric temperature near the footprint is higher in the daytime than in nighttime [32].

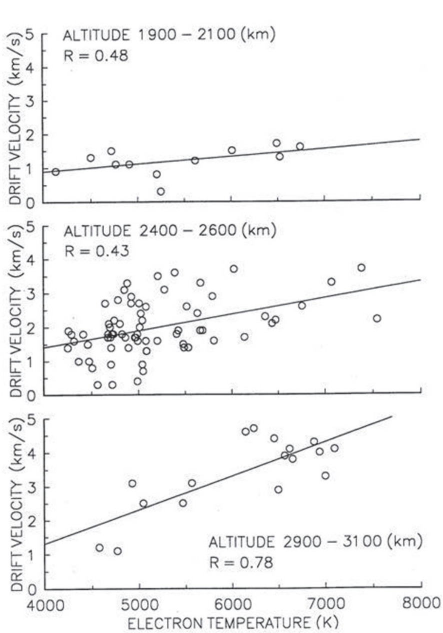


fig. 11. Scatter plots of the H+ polar wind velocity versus electron temperature in three altitude ranges [2]

Filling of the flux tubes by the plasma of the ionospheric origin has been observed in the equatorial region of the plasmasphere by satellites at the geosynchronous orbit. The filling rate estimated on the flux tubes that have arrived from the nightside and are taking the first pass on the dayside is ~0.21...0.54 cm^{-3}hour^{-1} [49], which is consistent with the upward H$^+$ flux observed in the topside ionosphere. At later passes in the dayside the filling rate is several times higher than the above.

The electron temperature increases from about 3,000 K at 1,000 km altitude to about 6,500 K at 8,000 km in daytime, but it is in the range of 2,000 K to 4,000 K in nighttime. The mean attitude gradient in the lower plasmasphere is much larger in daytime than in nighttime [5]. Higher T_e at higher altitudes may appear to suggest the downward heat flow, but the net flow of heat through the thermal electrons is upward because of the anisotropy of temperature. In the daytime polar cap at ~800 to 2,400 km, T_e is larger in the upward direction than in the downward and perpendicular directions with a ratio of 1.5 to 2. The net heat flux is upward and ~10^{-2} erg cm^{-2} s^{-1}, which is about 5 times that of atmospheric photoelectrons above 10 eV. This suggests that thermal electrons heated by photoelectrons are carrying more energy upward than the heat conducted downward [53].

Photoelectrons can influence the outflow in two ways. The first is to increase the ambipolar electric field. This electric field is generated along field lines as the electrons having a lighter mass stream faster and further than the ions having heavier mass, and it is strengthened by the presence of photoelectrons with superthermal energies. The second is to raise energies of the thermal electrons and the ions by imparting their energy. Models of the outflow process have been built that incorporate thermal electrons, superthermal electrons, and ions. In one of the advanced models the superthermal electrons are treated kinetically and are coupled to the other species by the ambient electric field and the Coulomb collisions [34 and 35]. However, it uses the Spitzer-Harm conductivity which would not be appropriate in the plasmasphere where the mean free path is rather long. By combining the *Akebono* observation of the plasmasphere and the Millstone Hill observation of the ionosphere it has been suggested that both additional heating and decrease in the thermal conductivity have to be assumed to explain the observations [46].

In order to explain the acceleration of the polar wind which shows distinct day/night difference it seems necessary to realistically address the non-Maxwellian nature of the electrons characterized by the presence of the superthermal tail in the velocity distribution function. The background electrons are not really 'thermal'. It has also been noted that Coulomb collisions are not always adequate to transfer energy from electrons to ions, and sometimes the ions seem to consist of cold and warm components in the outer plasmasphere [13].

CONCLUSION

The convection theory of the plasmapause formation combined with supersonic acceleration of upflowing plasma has been supported by numerous observations over the past four decades, but some basic questions have also emerged.

The most basic of them is the energetics. Photoelectrons play a vital role in heating and lifting the plasma from the ionosphere, so that it is important to clarify the process of energy transfer from photoelectrons to thermal electrons producing non-thermal tail in the distribution function. Kinetic treatment is needed for both photoelectrons and low-energy "thermal" electrons. The energy transfer from electrons to ions also seems to be governed by non-thermal processes.

Another is the origin and maintenance of the plasmaspheric notches. The origin of the plasma depletion extending below 2 R_e is not easy to understand, but its lifetime of several tens of hours is even more puzzling in view of the competing effects of non-uniform convection, instability, and upflow that would act to deform or smooth the structure. Coupling of dynamics between the plasmasphere and the neutral atmosphere is an interesting possibility and need to be explored.

ACKNOWLEDGMENTS

I would like to thank T. Abe, T. Kikuchi, Y. Nishimura, R.A. Wolf and I. Yoshikawa for useful comments.

REFERENCES

1. Abe T., Whalen B.A., Yau A.W. et al. EXOS-D (Akebono) SMS observations of the polar wind // J. Geophys. Res. 1993a. V. 98. P. 11, 191.
2. Abe T., Whalen B.A., Yau A.W. et al. Altitude profile of the polar wind velocity and its relationship to ionospheric conditions // Geophys. Res. Lett. 1993b. V. 20. P. 2825.
3. Abe T., Watanabe S., Whalen B.A., Yau A.W., Sagawa E. Akebono observations of thermal ion outflow and electron temperature in the polar wind region // Phys. Space Plasmas. 1995. V. 14. P. 3.
4. Abe T., Watanabe S., Whalen R.A., Yau A.W., Sagawa E. Observations of polar wind and thermal ion outflow by Akebono SMS // J. Geomag. Geoelectr. 1996. V. 48. P. 319.
5. Abe T., Balan N., Oyama K.-I., Bailey G.J. Plasmasphere electron temperature — Observations and theory // Adv. Space Res. 1997. V. 20. P. 401.
6. Axford W.I., Hines C.O. A unifying theory of high-latitude geophysical phenomena and geomagnetic storms // Can. J. Phys. 1961. V. 39. P. 1433.
7. Banks P.M., Holzer T.E. High-latitude plasma transport: the polar wind // J. Geophys. Res. 1969. V. 75. P. 6317.
8. Block L.P. On the distribution of electric fields in the magnetosphere // J. Geophys. Res. 1966. V. 71. P. 855.
9. Brice N.M. Bulk motion of the magnetosphere // J. Geophys. Res. 1967. V. 72. P. 5193.
10. Burch J.L., Goldstein J., Sandel B.R. Cause of plasmaspheric corotation lag // Geophys. Res. Lett. 2004. V. 31. P. L05802. DOI:10.1029/2003GL0190164.
11. Carpenter D.L. Whistler evidence of a ‘knee’ in the magnetospheric ionization density profile // J. Geophys. Res. 1963. V. 68. P. 1675.
12. Carpenter D.L. Lightening whistlers reveal the plasmapause, an unexpected boundary in space // Discovery of the Magnetosphere / Eds. C.S. Gillmor, J.R. Spreiter. — AGU, 1997. — P. 47.
13. Comfort R.H. Thermal structure of the plasmasphere // Adv. Space Phys. 1996. V. 17(10). P. 175.
14. Drakou E., Yau A.W., Abe T. Ion temperature measurements from the *Akebono* suprathermal mass spectrometer: application to the polar wind // J. Geophys. Res. 1997. V. 102. P. 17,523.
15. Dungey J.W. Interplanetary magnetic field and the auroral zones // Phys. Rev. Lett. 1961. V. 6. P. 47.
16. Fok M.-C., Ebihara Y., Moore T.E., Ober D.M., Keller K.A. Geospace storm processes coupling the ring current, radiation belt, and plasmasphere // Inner Magnetosphere Interactions / Ed. J. Burch, Schulz M., Spence H.. — AGU, 2005. — P. 207.
17. Foster J.C., Vo H.B. Average characteristics and activity dependence of the subauroral polarization stream // J. Geophys. Res. 2002. V. 107(A12). DOI:10.1029/2002JA009409.
18. Gallagher D.L., Adrian M.L., Liemohn M.W. Origin and evolution of deep plasmaspheric notches // J. Geophys. Res. 2005. V. 110(A09201). DOI:10.1029/2004JA010906.

19. Galperin Y.I., Ponomarev V.N., Zosimova A.G. Plasma convection in the polar ionosphere // Ann. Geophys. 1974. V. 30. P. 1.
20. Goldstein J., Shiro R.W., Reiff P.H., Wolf R.A., Sandel B.R., Freeman J.W., Lambour R.L. IMF-driven overshielding electric field and the origin of the plasmaspheric shoulder // Geophys. Res. Lett. 2002. V. 29(16). DOI:10.1029/2001GL014534.
21. Goldstein J., Sandel B.R., Hairston M.R., Reiff P.H. Control of plasmaspheric dynamics by both convection and subauroral polarization streams // Geophys. Res. Lett. 2003. V. 30(24). DOC:10.1029/2003GL018390.
22. Goldstein J., Sandel B.R., Forrester W.T., Thomsen M.F., Hairstonl M.R. Global plasmaspheric evolution 22–23 April 2001 // J. Geophys. Res. 2005/ V. 110(A12218). DOC:10.1029/2005JA011282.
23. Gonzalez W.D., Mozer F.S. A quantitative model of the potential resulting from reconnection with arbitrary interplanetary magnetic field // J. Geophys. Res. 1974. V. 79. P. 4186.
24. Grebowsky J.M. A model study of plasmapause motion // J. Geophys. Res. 1970. V. 75. P. 4329.
25. Gringauz K.I., Bezukikh V.V., Ozerov V.D., Rybehinsii R.E. Study of the interplanetary ionized gas, energetic electrons and solar corpuscular radiation with the aid of three-electrode traps for charged particles on the second Soviet space rocket // Dokl. Akad. Nauk SSSR. 1960. V. 151. P. 1301.
26. Gringauz K.I. Low-energy plasma in the Earth's magnetosphere // Rev. Geophys. 1969. V. 7. P. 339.
27. Hashimoto K.K., Kikuchi T., Ebihara Y. Response of the magnetospheric convection to sudden interplanetary magnetic field changes as deduced from the evolution of partial ring currents // J. Geophys. Res. 2002. V. 107(A11). DOI:10.1029/2001JA009228.
28. Heinemann M., Pontius D.H., Jr. Representations of currents and magnetic fields in isotropic magnetohydrostatic plasma // J. Geophys. Res. 1990. V. 95. P. 251.
29. Huang C.-S., Foster J.C., Kelley M.C. Long-duration penetration of the interplanetary electric field to the low-latitude ionosphere during the main phase of magnetic storms // J. Geophys. Res. 2005. V. 110(A11309). DOC:10.1029/2005JA011202.
30. Kikuchi T., Hashimoto K.K., Kitamura T.-I., Tachihara H., Fejer B. Equatorial counterelectrojets during substorms // J. Geophys. Res. 2003. V. 108(A11). DOI:10.1029/2003JA009915.
31. Kikuchi T., Hashimoto K.K., Nozaki K. Penetration of magnetospheric electric fields to the equator during a geomagnetic storm // J. Geophys. Res. 2008. V. 113. P. A06214. DOI:10.1029/2007JA012628.
32. Kotova G., Bezrukikh V., Verigin M., Similauer J. New aspects in plasmasphere ion temperature variations from *Interball-2* and *Magion-5* measurements // J. Atm. Sol. Terr. Phys. 2008. V. 70. P. 399.
33. Liemohn M.W., Brandt P.C. Small-scale structure in the stormtime ring current // Inner Magnetosphere Interactions / Eds. J. Burch, M. Schulz, Spence H.. — AGU, 2005. — P. 167.

34. Liemohn M.W., Khazanov G.V., Moore T.E., Guiter S.K. Self-consistent superthermal electron effects on plasmasphere filling // J. Geophys. Res. 1997. V. 102. P. 7523.
35. Liemohn M.W., Khazanov G.V., Craven P.D., Kozyra J.U. Non-linear kinetic modeling of early stage plasmaspheric refilling // J. Geophys. Res. 1999. V. 104. P. 10,295.
36. Marubashi K. Escape of the polar ionospheric plasma into the magnetospheric tail // Rept. Ionosph. Space Res. Japan. 1970. V. 24. P. 322.
37. Murakami G., Hirai M., Yoshikawa I. The plasmapause response to the southward turning of the IMF derived from sequential EUV images // J. Geophys. Res. 2007. V. 112(A06217). DOC:10.1029/2006JA012174.
38. Nishida A. Formation of plasmapause, or magnetospheric plasma knee, by the combined action of magnetospheric convection and plasma escape from the tail // J. Geophys. Res. 1966. V. 71. P. 5669.
39. Nishida A. Geomagnetic DP2 fluctuations and associated magnetospheric phenomena // J. Geophys. Res. 1968a. V. 73. P. 1795.
40. Nishida A. Coherence of geomagnetic DP2 fluctuations with interplanetary magnetic variations // J. Geophys. Res. 1968b. V. 73. P. 5549.
41. Nishida A. Average structure and storm-time change of the polar topside ionosphere at sunspot minimum // J. Geophys. Res. 1967. V. 77. P. 6051.
42. Nishida A. Deformation of the duskside plasmapause // Cosmic Electrodynamics. 1971. V. 2. P. 340.
43. Nishida A., Iwasaki N., Nagata T. The origin of fluctuations in the equatorial electrojet // Ann. de Geophys. 1966. V. 22. P. 478.
44. Nishimura Y., Shinbori A., Ono T., Iizima M., Kumamoto A. Evaluation of ring current and radiation belt particles under the influence of storm-time electric field // J. Geophys. Res. 2007. V. 112. P. A06241. DOI.:10.1029/2006JA012177.
45. Nishimura Y., Wygant J., Ono T. et al. SAPS measurement around the magnetic equator by CRESS // Geophys. Res. Lett. 2008. V. 35. P. L10104. DOI:10.1029/2008GL033970.
46. Pavlov A.V., Abe T., Oyama K.-I. Comparison of the measured and modeled electron densities and temperatures in the ionosphere and plasmasphere during 20–30 January, 1993 // Ann. Geophys. 2000. V. 18. P. 1257.
47. Sandel B.R., Goldstein J., Gallagher D.L., Spasojevic M. Extreme ultraviolet imager observations of the structure and dynamics of the plasmasphere // Space Sci. Rev. 2003. V. 109. P. 25.
48. Spasojevic M., Goldstein J., Carpenter D.L., Inan U.S., Sandel B.R., Moldwin M.B., Reinisch B.W. Global response of the plasmasphere to a geomagnetic disturbance // J. Geophys. Res. 2003. V. 108(A9). P. 1340. DOC:10.1029/2003JA009987.
49. Su Y-J., Thomsen M.F., Borovsky J.E., Lawrence D.J. A comprehensive survey of plasmaspheric refilling at geosynchronous orbit // J. Geophys. Res. 2001. V. 106. P. 23,615.
50. Toffoletto F., Sazykin S., Spiro R., Wolf R. Inner magnetospheric modeling with the Rice convection model // Space Sci. Rev. 2003. V. 107. P. 175.

51. Vasyliunas V. Mathematical models of magnetospheric convection and its coupling to the ionosphere // Particles and Fields in the Magnetosphere / Ed. B.M. McCormac. — Reidel, 1970. — P. 60.
52. Wolf R.A., Sazykin S., Xing X., Spiro R.W., Toffoletto F.R., De Zeeuw D.L., Gombosi T.I., Goldstein J. // Inner Magnetosphere Interactions / ed. J. Burch, M. Schulz, and H. Spence. — AGU, 2005. — P. 127.
53. Yau A.W., Whalen B.A., Abe T., Mukai T., Oyama K.I., Chang T. *Akebono* observations of electron temperature anisotropy in the polar wind // J. Geophys. Res. 1995. V. 100. P. 17,451.
54. Yau A.W., Abe T., Peterson W.K. The polar wind: recent observations // J. Atmosph. Solar Terr. Phys. 2007. V. 69. P. 1936.
55. Yoshikawa I., Yamazaki A., Shiomi K. et al. Evolution of the outer plasmasphere during low geomagnetic activity observed by the EUV scanner onboard Planet-B // J. Geophys. Res. 2000. V. 105. P. 27777.Abe, T., B.A. Whalen, A.W. Yau, R.E. Horita, S. Watanabe, and E. Sagawa (1993a), EXOS-D (Akebono) SMS observations of the polar wind, J. Geophys. Res., 98, 11,191.

S.F. SINGER
University of Virginia,
Science and Environment Policy Project (SEPP),
Arlington, USA
singer@sepp.org
J.F. LEMAIRE
Belgian Institute of Space Aeronomy,
Brussels, Belgium
jfl@astr.ucl.ac.be

GEOMAGNETICALLY TRAPPED RADIATION: HALF A CENTURY OF RESEARCH

The study of charged particles trapped in the geomagnetic field was stimulated by an attempt to understand the physical mechanism of magnetic storms. Following the first observations of a radiation belt, the cosmic-ray neutron-albedo theory successfully explained the energy spectrum and spatial distribution of trapped high-energy (>100 MeV) protons of the inner belt. Since these early achievements much of the research effort has been devoted to account for the acceleration of lower-energy electrons, relationship to solar activity, etc., and to the practical problems of protecting equipment and astronauts from the damaging ionizing effects of trapped particles[1].

1. THE EARLY STEPS

The study of the motion of a charged particle in a magnetic dipole field was initiated in attempts to explain the polar aurora in laboratory experiments by Birkeland and analyses by Poincaré. A major advance came from Carl Störmer in 1903, who developed the theory of particle motion in a dipole field. One constant of motion is the particle's kinetic energy. Since the Hamiltonian is independent of the azimuthal angle, the angular momentum of a charge in the cylindrically symmetric dipole is another constant of motion, referred to as Störmer's constant. A third constant of motion cannot be found and Störmer had to resort to numerical integrations. He derived a general result as follows: a particle coming from infinity cannot be trapped; it will either strike the earth or be reflected and return to infinity. He demonstrated the existence of allowed and forbidden zones around the earth dipole. He showed that there exists an inner allowed region within the normally forbidden region, but inaccessible for particles coming from infinity. Störmer and others therefore assumed that the allowed regions would be empty. Some suggestions of particles in inaccessible regions of an (assumed) solar dipole field came from Hannes Alfvén and was followed up by John Wheeler and his students [around 1949].

An important advance also came from experiments in Alfvén's institute by Malmfors and by Brunberg and Dattner, and from similar studies by Bennett at the US Naval Research Laboratory.

1 The historical introduction is a condensation of a review article by S.F. Singer and A.M. Lenchek, forming chapter 3 in Progress in Elementary Particle and Cosmic Ray Physics. (Edited by J.G. Wilson and S.A. Wouthuysen) North Holland Publishing Company, Amsterdam, 1962.

2. FIRST IDEAS ABOUT THE RING CURRENT

The impetus for studying particle motion in the geomagnetic field came from efforts to evaluate the contribution of cosmic ray secondaries ('albedo') to rocket experiments of the primary cosmic ray flux [75]. Griem and Singer [27] investigated the motion of albedo particles that were completely trapped in the Earth's magnetic field before being removed by collision loss.

An independent approach came from efforts to explain the origin and character of the Ring Current, thought to be responsible for the main phase of geomagnetic storms, the large decrease in magnetic field intensity lasting one or two days. Early work was carried out Chapman and Ferraro [9] and criticized by Alfven [1]. This ongoing debate [see here a discussion by S. Akasofu in *Eos* 84, 22 July, 2003] prompted Singer [63] to look for another mechanism, based on the hypothesis that solar particles might be completely trapped in the earth's magnetic field. He argued that even though a single particle could not enter the trapped region, the collective action by a number of particles might perturb the dipole field sufficiently to permit entry.

Using the perturbation theory developed by Alfvén, Singer then described the motion of these trapped particles in the Earth's dipole field and calculated the azimuthal drift velocity, protons drifting towards the west and electrons towards the east. This drift produces a current which he identified with the Ring Current postulated to account for the main phase of magnetic storms. He further suggested that a small number of these particles might be accelerated through collisions with magneto-hydrodynamic waves, perhaps connected with the observed micro-pulsations of the geomagnetic field, and be accelerated to auroral energies.

3. THE DISCOVERY OF THE RADIATION BELTS

In order to verify the hypothesis of the existence of trapped particles, Singer [64] suggested a specific experiment, to be carried out in the *Farside* vehicle, a four-stage rocket launched at the equator from a high-altitude balloon, and designed to reach an altitude of 4,000 miles. Unfortunately, the single successful flight in November, 1957 did not carry the Geiger counter intended for it, which would have detected trapped radiation one year earlier.

The discovery of geomagnetically trapped radiation was made by Geiger counters in artificial Earth satellites. *Sputnik-2* (November 1957) showed a fairly sharp counting-rate increase starting at 400 km; at 700 km the rate was ~40 percent above the 500 km intensity [81].

For *Explorer-1*, the maximum altitude was high enough to give an unmistakable increase in counting rate (by more than 100 times), which actually blocked the Geiger counter. The correct interpretation of this radiation in terms of trapped particles was given by Van Allen in May, 1958 [78]. In initial papers, the trapped radiation was interpreted as of solar origin and held to be responsible for magnetic storms and aurorae [79]. Differing views concerning the source of energy were put forth by Dessler [14], who assumed local acceleration, and by Gold [25], who assumed that fast particles could be 'conveyed' from the Sun.

4. THE NEUTRON-ALBEDO HYPOTHESIS

However, it seemed difficult to explain the existence of trapped solar radiation so very close to the Earth near the equator [65]. Hence Singer fell back on cosmic rays as a possible injection mechanism, and suggested fast cosmic-ray albedo-neutrons as a likely source for the observed radiation, which he assumed to consist of energetic protons. (The popular acronym given to this source of trapped particle in the magnetosphere is *CRAND*, standing for *Cosmic Ray Albedo Neutron Decay*).

Singer showed that this injection mechanism would give intensities that were in reasonable agreement with the observations. Assuming that the atmospheric density determines the lifetime of trapping, he calculated the expected intensity vs. altitude distribution and also the energy spectrum of the trapped high-energy protons [see Appendix], which result from the decay of the high-energy neutrons [66]. Kellogg investigated the possibility that decay electrons from thermal cosmic-ray albedo neutrons form a portion of the radiation belt [34, 35]. Independently, Vernov et al. [82] suggested cosmic-ray neutrons as a source for the geomagnetically trapped radiation and pointed out also that magnetic scattering would provide a limit on the lifetime at high altitudes. Also independently, N. Christofilos may have made a similar suggestion [unpublished; C. McIlwain, private communication].

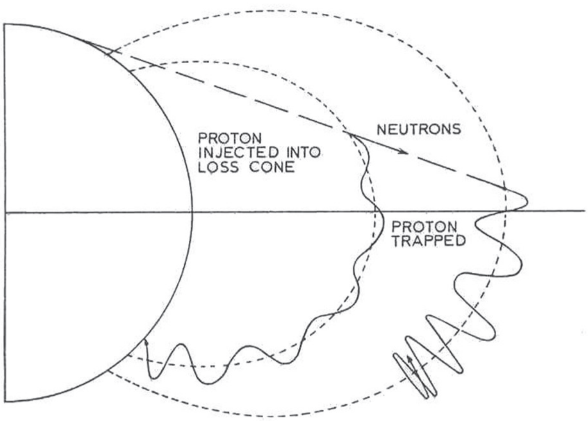


fig. 1. Schematic diagram to illustrate the neutron-albedo theory. The primary cosmic radiation, incident on the earth's upper atmosphere, produces nuclear disintegrations, which emit both energetic protons and neutrons. The energy spectrum of the protons has been measured (e.g. in photographic emulsions carried in balloons). The neutrons are assumed to have a similar energy spectrum. A tiny fraction of neutrons will decay in flight while traversing the Earth's magnetosphere (typical decay rate is 10...13 $cm^{-3}\ sec^{-1}$) and release a decay proton of the same kinetic energy. If the released proton is injected into a loss cone it will be immediately lost and not contribute to the trapped radiation flux. Magnetically trapped protons will eventually disappear by energy loss in the Earth's exosphere (typical lifetime is 10^{13} sec).

With the demonstration that cosmic-ray albedo-neutrons could furnish a reasonable radiation belt, Singer [67, 69] in November 1958 explicitly predicted the existence of two separate radiation belts, an inner belt (at 1...2 Earth radii) of cosmic-ray origin and containing penetrating protons, and an outer one mainly of solar origin containing 'soft' particles. Both the *Pioneer-3* rocket and *Lunik-1* observed two distinct belts with maxima at 1.5 and at about 3.4 Earth radii, with a rather pronounced minimum in between [76, 83].

5. THE SPATIAL AND SPECTRAL DISTRIBUTIONS OF TRAPPED RB PARTICLES

The existence of a minimum ('slot') does not by itself indicate separate origins for the two belts. Dessler [15] made the important point that large anomalies of the Earth's field could affect the trapped particle intensity, and Gold [26] held that both belts could be of solar origin with particles drifting inward, and with the slot caused by an instability of the geomagnetic field near a distance of 2 Earth radii. However, an important conclusion concerning the separate origin of the two belts was drawn by Christofilos [10] from the observed fact that electrons, artificially injected by atomic bomb explosions, produced shells whose position did not change with time. He argued that a radial drift of trapped particles is very unlikely, and that therefore the two observed belts are of separate origin.

On the other hand, the decrease in intensity of the inner belt beyond the first maximum may be due to a breakdown of adiabatic invariance of the magnetic moment of trapped high-energy protons; as a further consequence the maximum energy of the trapped protons should decrease with increasing altitude [17, 68, 88].

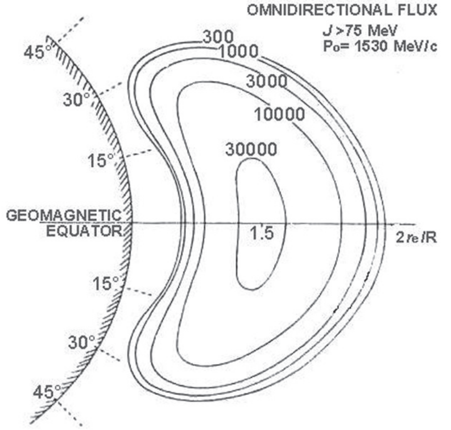


fig. 2. The omni-directional proton flux with energies >75 MeV, as computed from first principles. The figures on the iso-intensity contours are in units of protons-cm^{-2} sec^{-1}. The upper limit to the spectrum is set by the breakdown of adiabatic invariance of the magnetic moment [68]. The atmospheric density model assumes T=1,500 K at 530 km. The shape of these contours seems to depend markedly on the atmospheric model chosen; thus the radiation belt data may be used to give information on exospheric structure. The energy spectrum of the trapped protons can be derived from the energy spectrum of the albedo neutrons (see Appendix).

Intensity contours, approximately kidney-shaped, have been determined for the inner belt by *Explorer IV* [76], and show very clearly the control exercised by the actual geomagnetic field (approximately represented as an eccentric dipole). Intensity contours, more crescent-shaped, have been determined for the outer belt by *Pioneer* and *Lunik* probes [76, 77, 81, 83]. The neutron albedo theory accounts quite well for the spatial distribution of protons in the inner belt [70].

The nature of the trapped particles has been measured with various degrees of refinement. Krassovsky et al. [39, 40], in *Sputnik-3*, identified the bulk of the outer belt particles as low energy (10...50 keV) electrons. Vernov et al. [83] have given their energy spectra as $E^{-5}...E^{-3}$ from measurements in *Lunik-1*. (The pioneering contributions of Acad. Sergey Vernov in space sciences and specially in cosmic-ray physics have been recalled by Panasyuk [58]).

For the inner belt, the most detailed knowledge of the trapped protons comes from photographic emulsion measurements [22]. With more detailed theoretical analyses by Lenchek and Singer [45] one finds good agreement with the spectrum predicted by the neutron albedo theory (see Appendix).

Finally, the inner belt also contains a large flux of low-energy electrons [85]. Only a tiny fraction of them arise from the decay of thermal cosmic-ray albedo neutrons [29]. Quantitative calculations of the energy spectrum and intensity produced by this mechanism [36, 46, 86] do not give good agreement with the observations. Since their magnetic moments are smaller than those of the inner belt protons, these electrons will remain adiabatically trapped to higher radial distances and extend beyond the proton belt. They contribute to the trapped radiation beyond 2 Earth radii. They may even form all of the outer radiation belt, as suggested by Dessler and Karplus [16], who argue this view as against a solar origin for most of the outer-belt electrons.

Cosmic rays from solar flares can also contribute to the trapped radiation through the intermediary of albedo neutrons from the polar cap [3]. This phenomenon has been analyzed by Lenchek and Singer [45], who investigated in particular the peculiar spatial distribution and dynamics of the energy spectrum.

After the early successes of Radiation Belt (RB) and Ring Current (RC) experimental and theoretical studies in the 60s, the interest of the space community shifted toward particles of smaller kinetic energy and magnetospheric plasmas of lower temperatures. It was only in the 90s that the community became again concerned with modeling the RB environment.

6. THE MARCH 24, 1991 EVENT

A revival of interest for the Radiation Belts can be observed in 1988, when McCormack [52] noticed that the NASA models badly failed to predict the energetic particle fluxes which were measured at low altitudes in the late 80s.

The renewal of interest for RB studies has also been prompted by an exceptionally large enhancement of relativistic electron flux, and of energetic proton flux, measured by the *CRRES* satellite deep inside the magnetosphere. These extreme flux enhancements occurred simultaneously following the Sudden Storm Commencement (SSC) of March 24, 1991. By chance, *CRRES* was then in the slot region, when the unexpected injection and acceleration of charged particles suddenly took place.

Electrons were accelerated up to 10 MeV, and protons reached energies larger than 50 MeV. *CRRES* recorded with unprecedented detail the abrupt initial flux enhancement in all energy channels, as well as subsequent drift-echo flux enhancements. The latter correspond to successive revolutions of bunches of energized electrons and protons, drifting in opposite directions around the Earth. The relativistic electrons observed at $L = 2.5$ had been injected at large equatorial distance $(>10R_E)$ over a limited range of local times. They were accelerated while surfing deeper into the geomagnetic field, and drifting around the Earth with a period of 3 minutes [5, 80].

The counting rates of these electrons for different energy thresholds are shown in the top panel of fig. 3a. The dawn-dusk component (E_Y) of the electric field, and the northward component (B_z) of the magnetic field perturbation measured by *CRRES*, are shown in the lower panels of fig. 3a.

Such powerful injections of bunched RB particles had never been noticed before. However, in a retrospective search, it was found that similar events had occurred earlier, with smaller magnitudes. Subsequently, similar injections of relativistic electrons at $L<3$ have been recorded following violent storms of February 21, 1994. September 7, 2002; October 29, 2003, but all of them had lower intensities [50, 74].

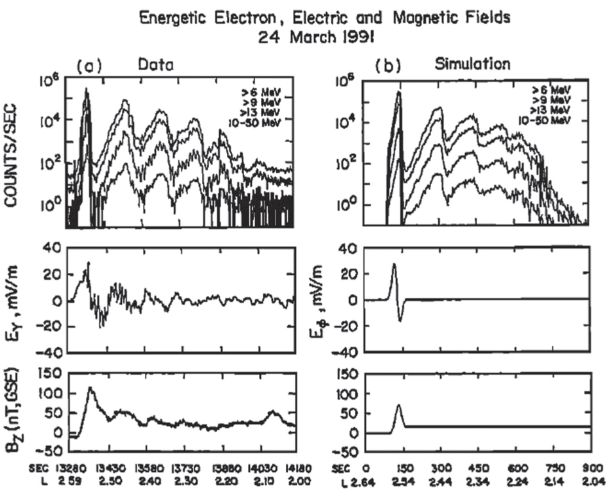


fig. 3. (a) Data from the CRRES satellite at the time of the March 24, 1991, Sudden Storm Commencement. Top panel shows count rates as a function of time from four energetic electron channels [5]. Middle and bottom panels show the measured electric field E_Y in a co-rotating frame of reference, and the B_z magnetic field component with a model magnetic field subtracted, in GSE coordinates over the same time interval [89]. (b) Simulated results in the same format as (a) determined at the same location corresponding to the trajectory of the CRRES satellite, including the *Aerospace* detector responses in the top panels. Time is measured from start of the numerical simulation by Li et al. [47].

The remarkable March, 1991 event definitely lent new momentum to studies of the Radiation Belts and raised the question: what is the mysterious mechanism by which RB particles are accelerated to such high energies?

7. WHAT PHYSICAL MECHANISM ACCELERATES PARTICLES DURING STRONG STORMS?

Detailed analysis of the unique *CRRES* observations, combined with kinetic particle simulations indicated that the acceleration of ambient trapped particles (or untrapped solar flare particles) can be explained by the electric field of a collisionless shock passing through the magnetosphere.

A solar Coronal Mass Ejection (CME) hit the front of the magnetosphere, compressed it, and produced the Sudden Storm Commencement observed on the ground, as well as an enhancement of B_z observed at *CRRES* (see bottom panel of fig. 3a). Solar flare particles were also present at the time of the CME impact.

Starting with an ensemble of 1 to 9 MeV equatorially mirroring electrons, distributed between 3 and 9 R_E, and using an adjusted time-dependent electric field model, Li et al. [47] simulated the motion of over 300,000 test particles in the magnetosphere. The 'adjusted' analytic model of the shock-induced electric pulse is displayed in the mid-panel of fig. 3b, for the position of *CRRES*. The electrons that were in the dayside magnetosphere at the time of the interplanetary shock passage experienced the accelerating electric force during an appreciable portion of their eastward drift path about the Earth. The simulated fluxes of electrons at the positions of *CRRES* are shown in the top panel of fig. 3b for the same energies as in fig. 3a. It can be seen that the flux values obtained with Li et al. [47] kinetic model match very well the experimental results.

Using an almost similar model for the induction electric field pulse, Hudson et al. [31] simulated the rise in the flux of protons drifting westward and inwards onto deeper drift shells. The simulated proton flux increases abruptly by several orders of magnitude on a drift time scale. This was followed by quasi-periodic echo enhancements fitting very well the proton fluxes measured at the positions of CRRES, during the March, 1991 event (not shown).

Equally good results have been obtained by Hudson et al. [32] with MHD test particle simulations, where B-field and E-field pulses are used from MHD models of the magnetosphere induced by the solar wind variations, instead of the 'adjusted' analytic E-field model adopted by Li et al. Other Global-MHD test particle simulations for another geomagnetic storm event have been performed by Kress et al. [41]. They obtain again the formation of a new belt of >10 MeV electrons deep in the magnetosphere. Furthermore, they showed that the pitch angle distribution of the accelerated electrons is first significantly peaked at 90°: i.e. predominantly equatorially mirroring electrons. This initial pancake pitch-angle distribution eventually diffuses as a result of pitch angle scattering due to VLF wave-particle interactions. Kress et al. [41] found a diffusion timescale in agreement with the delay in appearance of peak fluxes of relativistic electrons at the positions of *SAMPEX*, in low Earth orbit.

Using the Li et al. analytic E-field model, Gannon et al. [24] performed a parametric study for the peak location (the L-parameter where the flux is maximum in the new belt) as a function of parameters characterizing the SSC electric-field pulse model. Their parametric study shows that a pulse propagation speed >1,200 km/s is required in the equatorial plane to achieve the flux levels in a new belt, which are comparable to those observed by *CRRES* during the March 1991 storm. Furthermore, they conclude that a pulse peak amplitude >120 mV/m is required to form new electron belts inside *L=3*.

We close this historical review by noting, with Walt [84], that the rather remarkable fit of all these different kinds of simulations leads us to conclude that: 'a most important lesson from the March, 1991 event, is the realization that, even after decades of experimental and theoretical efforts, all important physical processes producing energetic Van Allen belt particles had not yet been described, let alone understood'.

There is still an open question: how the bulk of the ambient electrons can coherently acquire much larger post-storm fluxes and energies, than before recovery phases of geomagnetic storms — and this solely by resonant wave-particle interactions with a wide spectrum of VLF or ULF waves — in spite of the unavoidable presence of random-phase non-resonant wave-particles which tend to increase the entropy of the system, instead of decreasing it.

8. FUTURE PERSPECTIVES

What other surprises and discoveries can we now expect for the coming years? Are we not overlooking other key physical mechanisms: e.g. possible direct entry of Solar Energetic Electrons (SEE) into the geomagnetic field — for instance as the result of lowering of the magnetic barrier separating the inner and outer allowed zones of Störmer's theory, when the IMF turns southward, as shown by Lemaire [43, 44]. Who knows what are the additional physical mechanisms relevant to inject and accelerate RB particles within magnetospheres — that will be examined in the years to come by our descendants?

APPENDIX

Dimensional analysis of the energy spectrum of trapped protons[2].

The equation of continuity describing the particle distribution is

$$\partial n/\partial t = 0 = q(E) - d/dE[n\, dE/dt] - \sigma(E)\, \rho_a\, n\, \beta\, c \qquad (1)$$

where n is the number of protons [cm^{-3} MeV^{-1}] having directions within a unit solid angle, q(E) is the number of protons produced [cm^{-3} MeV^{-1} $ster^{-1}$ sec^{-1}], σ(E) is the catastrophic absorption cross section, ρ_a is the averaged density (cm^{-3}) of atmospheric nuclei with the above cross section, βc is the proton velocity. Note that n refers to a particular line of force, labeled by the geocentric distance r_e, and to a particular equatorial pitch angle; we omit the symbols r_e, and μ (=cos α_e), which are implied. The dependence on r_e, and μ is contained in the source term,

2 From Singer and Lenchek, 1962

$$q(E) = j_n^{eff}(r_e, \mu, E)\, \Lambda_n/\gamma\beta c, \qquad (2)$$

Lambda is the decay rate $MeV^{-1}\ cm^{-3}\ ster^{-1}\ sec^{-1}$. [$\Lambda_n$ is inverse of the neutron lifetime]

We have found that the equilibrium energy spectrum of differential intensity is given by a certain integral over the source function. In general, the result is not expressible as a simple power law nor indeed can a closed analytical expression be given. Nevertheless, considerable insight into the understanding of the shape of the spectrum can be gained by a simple dimensional analysis. The analysis proceeds by approximating all functions of energy in the continuity equation (1) by power laws, reducing the equation to an identity and thereby determining approximately, the logarithmic slope of the spectrum. Since the power law approximations for β(E), dE/dx vs. E, etc., are good only over small intervals of energy, the result is valid only over a small interval. However, since in the exact calculation, the source function, q, is a rapidly varying (monotonically decreasing) function of E, most of the contribution to the intensity at E (the lower limit of the integral) comes from q evaluated near E. It is found that the exact result is, in fact, fairly closely approximated by the power-law technique now to be described.

We set all constants equal to unity and write
$j(E) = E^{-\nu}$, $j_n(E) = E^{-\alpha}$, $\beta = E^b$, $\gamma = (1-\beta^2)^{-1/2} = E^g$, and $\eta = E^{-\nu}$.

All the exponents, p, α, g, b, ν, are functions of energy. For example, for sufficiently low energy b = ½ and g =0. At energies >> Mc^2 we have b~0 and g~1.

At low energies (~<200 MeV) the effect of nuclear interactions is negligible and the steady-state continuity equation reduces to

$$q(E) = d/dE\,[n\, dE/dt] \qquad (3)$$

With the relations

$n = j/\beta = E^{-p-b}$, $q = \eta j_n/\beta\gamma = E^{-\alpha-b-g-\nu}$,

and

$dE/dt = k/\beta = E^{-b}$,

we have, upon substitution:

$E^{-\alpha-b-g-\nu} = E^{-p-2b-1}$

This implies

$p = \alpha + \nu + g - b - 1 = \alpha + \nu - 1.5$,

and hence, in the nonrelativistic regime,

$$j(E) \text{ is proportional to } E^{-(\alpha + \nu - 1.5)} \qquad (4)$$

Since the effective neutron albedo energy spectrum was $\eta(E)\, j_n(E)$ is proportional to $E^{-(\alpha + \nu)}$ we see that the effect of energy loss is to flatten the spectrum by a power of 1.5 ($\alpha + \nu > 1.5$).

Conversely, for low energies, we may deduce the albedo spectrum from the equilibrium spectrum; the albedo spectrum is steeper by $E^{-1.5}$.

This is in agreement with the earlier result [66] where an albedo spectrum $E^{-1.8}$ led to an equilibrium velocity spectrum proportional to $\beta^{0.4}$. Note that $\beta^{0.4}d\beta = E^{-0.3}dE$ non-relativistically.

An equilibrium spectrum between $\sim E^{-1.4}$ and $\sim E^{-1.8}$ has been observed between 20 and 100 MeV [3, 30, 57]. These observations imply an effective neutron-albedo spectrum between $\sim E^{-2.9}$ and $\sim E^{-3.3}$.

The values $\alpha = 1.8$, $\nu = 1.1$ used here satisfy this requirement. This emphasizes the importance of some form of modulation of the 'isotropic' neutron spectrum, as by the energy-dependent anisotropy effect contained in $\eta(E)$.

In the limit of 'very high' energy, pure absorption by nuclear interactions applies. The energy at which slowing down can be neglected relative to nuclear interactions is not well defined and depends on the ratio of oxygen to hydrogen in the atmosphere. The 'cross-over energy' may vary from ~300 MeV for a pure oxygen atmosphere to ~900 MeV for a pure hydrogen atmosphere. It is also slightly dependent on the degree of ionization.

If we assume pure absorption, with a cross section independent of energy, then neglecting energy loss, the continuity equation is simply

$$q = \sigma\rho_b j \tag{5}$$

Therefore $E^{-\alpha-\nu-b-g} = E^{-p}$ or $p = \alpha + \nu + b + g$

For $E >> Mc^2$ we have $b \sim 0$ and $g \sim 1$. For $E \sim Mc^2$ we find $b + g \sim 0.5$.

Since $b + g$ is always positive, we see that the effect of nuclear interactions is to steepen the spectrum.

One outcome of such a dimensional analysis is the conclusion that, relative to the effective albedo spectrum, the equilibrium spectrum of the trapped protons is flatter by ~1.5 in the power law exponent at E ~<200 MeV and is steeper by ~0.5 in the exponent at E ~1 GeV; thus the equilibrium spectrum does not follow a single power over the entire range 10...1000 MeV. The observations show in fact a spectrum whose power steepens continuously at increasing energies [3, 22].

The effective neutron albedo spectrum $E^{-2.9}$ thus leads to an equilibrium spectrum asymptotic to $E^{-1.4}$ at E ~50MeV, and steepens continuously to $E^{-3.4}$ at ~1 GeV. The observed spectra have an exponent of 1.4 to 1.8 in the range 20 to 100 MeV and agree very well with the calculated slope at the higher energies.

ACKNOWLEDGMENTS

We wish to thank C.E. McIlwain for his input regarding the origin of the value adopted for the magnetic moment M_d of his reference dipole magnetic field used to calculate the invariant coordinates (B & L) , as well as for additional information on the history of the CRAND mechanism. We thank M.I. Panasuyk for shedding new light concerning the contributions of S.N. Vernov and his coworkers in the early interpretation on the origin of the corpuscular radiation discovered above the atmosphere.

REFERENCES

1. Alfv´en H. // Tellus. 1955. V. 7. P. 50.
2. Alfv´en H. // Phys. Rev. Lett. 1959. V. 3. P. 459.
3. Armstrong A.H., Harrison F.B., Heckman H.H., Rosen L. J. Geophys. Res. 1961. V. 66. P. 361.
4. Baker D.N. et al. // Geophys. Res. Lett. 1998. V. 25. P. 2975.
5. Blake J. B., Kolasinski W.A., Filius R.W., Gussenhoven E.G. // Geophys. Res. Lett. 1992. V. 19. P. 821.
6. Buhler P., Desorgher L. // J. Atmospheric and Solar-Terrest. Phys. 2002. V. 64. P. 593–599.
7. Chang D. B. Some plasma instabilities of the magnetosphere, Radiation Trapped in the Earth's Magnetic Field. — Proc. NATO Adv. Study Inst., 1966. — P. 491.
8. Chapman S., Ferraro V. C.A. // Terr. Mag. 1931. V. 36. P. 77,171.
9. Chapman S., Ferraro V. C.A. // Terr. Mag. 1932. V. 37. P. 147.
10. Christofilos N. C. // J. Geophys. Res. 1959. V. 64. P. 869.
11. Cornwall J.M. // J. Geophys. Res. 1964. V. 69. P. 1251.
12. Cornwall J.M. // J. Geophys. Res. 1966. V. 71. P. 2185.
13. Davis L. Jr., Chang D.B. // J. Geophys. Res. 1962. V. 67. P. 2169.
14. Dessler A. J. // Phys. Rev. Lett. 1958. V. 1. P. 68.
15. Dessler A. J. // J. Geophys. Res. 1959. V. 64. P. 713.
16. Dessler A. J., Karplus R. // Phys. Rev. Lett. 1960. V. 4. P. 271.
17. Dragt A. J. // J. Geophys. Res. 1961. V. 66. P. 1641.
18. Falthammar C.-G. // J. Geophys. Res. 1966. V. 71. P. 1487.
19. Fermi E. // Phys. Rev. 1949. V. 75. P. 1169.
20. Fok M.-C., Horne R. B., Meredith N. P., Glauert S.A. // J. Geophys. Res. 2008. V. 113. P. A03S08. DOI:10.1029/2007JA012558.
21. Frank L.A. // J. Geophys. Res. 1965. V. 70. P. 3533.
22. Freden S. C., White R. S. Phys. Rev. Lett. 1959. V. 3. P. 9.
23. Fujimoto M., Nishida A. // J. Geophys. Res. 1990. V. 95. P. 4265.
24. Gannon J.L., Li X., Temerin M. // J. Geophys. Res. 2005. V. 110. P. A12206. DOI:10.1029/2004JA010679.
25. Gold T. // Nature. 1959a. V. 183. P. 355.
26. Gold T. // J. Geophys. Res. 1959b. V. 64. P. 1219.
27. Griem H., Singer S.F. // Phys. Rev. 1955. V. 99. P. 608
28. Harris I.R., Jastrow R., Kelsall T. // Refereed to in *Scientific American*. 1959. V. 43 (Aug. 1959).
29. Hess W.N. // J. Geophys. Res. 1960. V. 65. P. 3107.
30. Holly F.E., Allen L.A. Jr., Johnson R.G. // J. Geophys. Res. 1961. V. 65. P. 2500.
31. Hudson M.K., Kotelnikov A.D., Li X., Roth I., Temerin M., Wygant J., Blake J.B., Gussenhoven M.S. // Geophys. Res. Lett. 1995. V. 22. P. 291.
32. Hudson M.K., Elkington S.R., Lyon J.G., Marchenko V.A., Roth I., Temerin M., Blake J.B., Gussenhoven M.S., Wygant J.R. // J. Geophys. Res., 1997. V. 102. P. 14087.

33. Hudson M.K., Kress B.T., Mazur J.E., Perry K.L., Slocum P.L. // J. Atm. Sol.-Terr. Phys. 2004. V. 66. P. 1389.
34. Kellogg P.J. // Nuovo Cimento Suppl. (Ser. 10). 1959. V. 11. P. 48.
35. Kellogg P.J. // Nature. 1959. V. 183. P. 1295.
36. Kellogg P.J. // J. Geophys. Res. 1960. V. 65. P. 2705.
37. Kennel C.F., Petschek H.E. // J. Geophys. Res. 1966. V. 71. P. 1.
38. Kim H.-J., Chan A. J. Geophys. Res. 1997. V. 102. P. 22107.
39. Krassovsky V.I. et al. // Uspekhi Fiz. Nauk. 1958. V. 64. P. 425.
40. Krassovsky V.I. et al. // Proc. 10th Int. Astronautic Cong. (London). — Vienna: Pringer Verlag, 1960.
41. Kress B.T., Hudson M.K., Looper M.D., Albert J., Lyon J.G., Goodrich C.C. // J. Geophys. Res. 2007. V. 112. P. A09215. DOI:10.1029/2006JA012218.
42. Kruglanski M. // Radiation Belts: Models and Standards. Geophys. Mon. 97 / Eds. J. F. Lemaire, D. Heynderickx, D.N. Baker. 1996. P. 195–200.
43. Lemaire J. // Adv. in Space Res. 2003. V. 31(5). P. 1131—1153.
44. Lemaire J.F., Batteux S.G., Slypen I.N. // J. of Atmosph. Solar-Terrestrial Phys. 2005. V. 67. P. 719—727.
45. Lenchek A.M., Singer S.F. // J. Geophys. Res. 1962. V. 67. P. 1273.
46. Lenchek A.M., Singer S.F., Wentworth R.C. // J. Geophys. Res. 1961. V. 66. P. 4027.
47. Li X., Roth I., Temerin M., Wygant J.R., Hudson M.K., Blake J.B. // Geophys. Res. Lett. 1993. V. 20. P. 2423.
48. Li X., Temerin M., Baker D.N., Reeves G.D., Larson D. // Geophys. Res. Lett. 2001. V. 28. P. 1887.
49. Liu W.W., Rostoker B., Baker D.N. // J. Geophys. Res. 1999. V. 104. P. 17391.
50. Looper M.D., Blake J.B., Mewaldt R.A. // Geophys. Res. Lett. 2005. V. 32. P. L03S06. DOI:10.1029/2004GL021502.
51. McAdams K.L., Reeves G.D. // Geophys. Res. Lett. 2001. V. 28. P. 1897.
52. McCormack P.D. // Acta Astronaut. 1988. V. 17. P. 231.
53. Nagai T. // Geophys. Res. Lett. 1998. V. 15. P. 425.
54. Nishida A. // Geophys. Res. Lett. 1976. V. 81. P. 1771.
55. O'Brien T.P. // J. Geophys. Res. 1962. V. 67. P. 3687.
56. O'Brien T.P. // J. Geophys. Res. 1964. V. 69. P. 13.
57. Naugle J.E., Kniffen D.A. // Phys. Rev. Lett. 1961. V. 7. P. 3.
58. Panasyuk M.I. // Acta Astronomica. 1998. V. 43(1–2). P. 51–56.
59. Paulikas G.A., Blake J.B. // Geophys. Res. Lett. 1976. V. 3. P. 227.
60. Paulikas G.A., Blake J.B. // Quantitative Modeling of Magnetosperic Processes. Geophysical Monograph V. 21 / Ed. W. P. Olson. — Washington D.C.: AGU, 1979. — P. 181.
61. Paulikas G.A., Blake J.B., Freden S.C. // J. Geophys. Res. 1966. V. 71. P. 3165.
62. Preszler A.M., Simmett G.M., White R.S. // Phys. Rev. Lett. 1972. V. 28. P. 982.
63. Singer S.F. // Bull. Am. Phys. Soc. Series II. 1956. V. 1. P. 229 (A).
64. Singer S.F. // Missiles and Rockets. 1957. V. 2. P. 120 (Oct.).

65. Singer S.F. // Phys. Rev. Lett. 1958c. V. 1. P. 171.
66. Singer S.F. // Phys. Rev. Lett. 1958d. V. 1. P. 181.
67. Singer S.F. // Adv. in Astronautical Sci. 1959a. V. 4. P. 335.
68. Singer S.F. // Phys. Rev. Lett. 1959b. V. 3. P. 188.
69. Singer S.F. // Symp. on Space Physics & Medicine. Ch. IV. San Antonio, Nov. 1958. — N. Y.: J. Wiley, 1960a.
70. Singer S.F. // Phys. Rev. Lett. 1960b. V. 5. P. 300.
71. Singer S.F., Lenchek A.M. // Progress in Elementary Particle and Cosmic Ray Physics // Eds. J.G. Wilson, S.A. Wouthuysen. — Amsterdam: North Holland Publishing Company, 1962. — Ch. III.
72. Summers D., Ni B., Meredith N.P. // J. Geophys. Res. 2007. V. 112. P. A04207. DOI:10.1029/2006JA011993.
73. Theodoris G.C., Paolini R.R., Frankenthal S. // J. Geophys. Res. 1969. V. 74. P. 1238.
74. Ukhorskiy A.Y., Anderson B.J., Brandt P.C., Tsyganenko N.A. // J. Geophys. Res. 2006. V. 111. P. A11S03. DOI:10.1029/2006JA011690.
75. Van Allen J.A., Singer S.F. // Nature. 1952. V. 170. P. 62.
76. Van Allen J.A., Frank L.A. // Nature. 1959a. V. 183. P. 430.
77. Van Allen J.A., Frank L.A. // Nature. 1959b. V. 184. P. 219.
78. Van Allen J.A., Ludwig G.H., Ray E.C., McIlwain C.E. // Jet Propulsion. 1958. V. 28. P. 588.
79. Van Allen J.A., Mcllwain C.E., Ludwig G.H. // J. Geophys. Res. 1958. V. 64. P. 271.
80. Vampola A.K., Korth A. // Geophys. Res. Lett. 1993. V. 19. P. 625.
81. Vernov S.N., Grigorov N.L., Logachev Yu.I., Chudakov A.E. // Dokl. Akad. Nauk SSSR. 1958. V. 120. P. 1231.
82. Vernov S.N. et al. // Sov. Phys. Dokl. 1959a. V. 4. P. 338(T).
83. Vernov S.N., Chudakov A.E., Vakulov P.V., Logachev Yu.I. // Dokl. Akad. Nauk SSSR. 1959b. V. 125(2).
84. Walt W. // Radiation Belts: Models and Standards. Geophys. Mon. V. 97 / Eds. J. F. Lemaire, D. Heynderickx, D.N. Baker. 1996. P. 1–13.
85. Walt M., Chase L.F., Cladis J.B., Imhof W.L., Knecht D.J. // Space Research // Ed. H.K. Kallman-Bijl. — Amsterdam: N.-Holland Publ. Co., 1960. — P. 910–920.
86. Walt M., MacDonald W.M. // J. Geophys. Res. 1961. V. 66. P. 2047.
87. Welch J.A., Whitaker W.A. // J. Geophys. Res. 1959. V. 64. P. 909.
88. Wentzel D.G. // J. Geophys. Res. 1961. V. 66. P. 359, 363.
89. Wygant J.R., Mozer F., Temerin M., M, Blake J.B., Maynard N., Singer H., Smiddy M. // Geophys. Res. Lett. 1974. V. 21. P. 1739.

J. BLAMONT
Centre National
d`Etudes Spaciales (CNES),
France
jacques.blamont@cnes.fr

THE EXPLORATION OF THE PLANETARY ATMOSPHERES

RESULTS OBTAINED WITH SPACE METHODS SINCE THE LAUNCH OF SPUTNIK-1

The impact of space experimentation over our knowledge of the Solar system and planets is perfectly illustrated by a book which has greatly influenced me when I was a small boy, «*On the other worlds*», by the French astronomer Lucien Rudaux which contains a discussion of our knowledge of the planets at that time, well informed and thorough scientifically. This book presents indeed a recording of total ignorance, and therefore it can be asserted that everything we know today on the telluric planets, to which my purpose shall be limited Mercury, Venus, Mars and by extension Titan, the satellite of Saturn, (leaving aside Mercury, which has been found without atmosphere by *Mariner-10*), everything has been brought by space methods since 1962.

What is surprising is that the exploration of the Solar system has proceeded so easily. Only twenty years after the first successful launch of an A-4 (later called V-2) at Penemuende by German engineers on October 3, 1942, the Great Powers, that means the Soviet Union and the United States, were in the possession of the necessary tools; what was missing was quickly developed, on a continuous and regular path. It is not to be forgotten that, seen at short term, that path was painfully rocky. As soon as October, 1960, Soviet specialists had tried unsuccessfully to launch two probes to Mars, then, in February 1961, two first attempts to reach Venus, also without success. The Soviet Union has launched 18 planetary probes without scientific return from 1960 to 1965. It will have to wait for 1967 to at last reap the fruit of its labour with *Venera-4*. NASA was more lucky since out of two probes launched to Venus in 1962, one of them, *Mariner-2*, reached Venus without, it is to be reckoned, telling us any new information, and *Mariner-4* flew by Mars on July, 1964, measured the density of the atmosphere down to the ground and revealed the presence of craters on the surface. From now on, the planets were not sacred bodies anymore, but objects of science. With ups and downs, planetary exploration has continued without respite, and the two Great Powers of space, joined by the European Space Agency, have launched fly-bys, satellites, aerostats, landers and rovers.

A priori Venus, the Earth and Mars should be similar, since their distance to the Sun are comparable, their orbital dimensions different by a factor inferior to two, and their density very near, however they bear no resemblance whatsoever. Space exploration has shown again and again, that all anthropomorphic projection leads to error.

THE ATMOSPHERE OF VENUS

The axis of rotation of Venus is quasi perpendicular to the ecliptic plane (inclination 20.8), its period 243 Earth days in the inverse sense of the Earth's rotation, the duration of its day 117 terrestrial days.

It's atmosphere is an exotic and complex chemical system, dominated at the top by photochemistry, in all it's depth by radiation transfer and at the bottom by thermochemical cycles including the surface. Its major constituents are CO_2 (96.5%) and N_2 (3.5%). Sulfur, carbon, chlorine, fluorine, phosphorus exist as traces, with an abundance varying from a few to hundreds parts per million. Their high variations in altitude and latitude show that they enter into chemical and dynamical cycles.

At the ground, the pressure reaches 90 bar and the temperature 475°C, because of the greenhouse effect generated by polyatomic molecules, solid and liquid aerosols. The region between 60 and 70 km of altitude is a photochemical plant, where the solar ultraviolet radiation induces dissociation and reactions of molecules, hence the production of aerosols of sulphuric acid which constitute between the altitudes of 70 to 50 km thick clouds preventing the observer placed outside the atmosphere to see the ground.

Below these ubiquitous clouds, we do not understand the chemistry. Cycles including all the minor constituents SO_2, CO, COS, H_2O, P, dominate the equilibrium of a very active atmosphere. Venus has probably received as much volatiles species as the Earth from the protoplanetary nebula, then outgassing, escape, cometary impacts and interaction with ground surface have modified the inventory. It is thought that for the Earth and Mars, outgassing and cometary supply have been the major processes, completed by strong erosion on Mars, after a quasi-complete loss of the primeval atmosphere. The evolution of Venus may have been completely different since it's atmosphere contains much more argon and neon, as do meteorites and the Sun. In order to improve our understanding, it would be helpful to measure the concentration and the isotopic composition of xenon.

Today H_2O is relatively scarce on Venus but its abundance must have been higher in the past, as shows a ration D/H 150 times higher in the low atmosphere of Venus than its value on Earth. Enrichment in deuterium implies the existence at some time of H_2O at a level equivalent to an ocean tens of meters deep covering all the planet.

The processes of escape through interaction with either the solar wind or the solar radiation in the absence of a magnetic field (since Venus has none) are unknown: the escape fluxes of O and O^+, as well as for H are uncertain by two orders of magnitude.

Venus is provided with the cloud system the most developed of all telluric planets. It covers all latitudes between 50 and 70 km altitude with an average density between 1 and 10 mg/m^3. The top is made of 75% H_2SO_4 droplets of micronic diameter. At lower levels chlorine and phosphorous provide possibly substantial components. One of the most irritating enigma touching these clouds is the still unknown nature of an ultraviolet absorber observed as elongated regions of great dimensions, in the shape of a letter Y parallel to the equator. The French astronomers Boyer and Camichel discovered this phenomenon in 1957 and did demonstrate by following the apparent motions that the top of the atmosphere rotates in four terrestrial days in the retrograde sense. The chemical origin and the physical mechanisms dominating the existence and the behaviour of the clouds are today not understood in the absence of data on the interaction with the ground surface. The idea of replenishment of the atmospheric soup by recent volcanism is highly controversial.

Below the clouds, the atmosphere is more or less clear, but visibility is limited by scattering. Below 20 km, where the temperature reaches 300°C the light becomes redder: an unknown molecule absorbs blue and green. At the ground level, reached by 2% of the radiation incident at the top of the atmosphere, strong mirages limit the horizontal sight to about 1 km.

Windows exist in the spectrum of the absorbing species, near the wavelength of 1 or 2.2 microns among others, and the ground can be observed through them, even from Earth. Venus Express has obtained a map of the surface with a 100 km resolution which is now compared with the data obtained with a much higher resolution by the *Magellan* radar.

The general circulation of the atmosphere is the most extraordinary feature of Venus and gives to this planet its singularity. Are present two global dynamical regimes:

At high altitudes, between 200 and 100 km, a circulation from the subsolar region to the antisolar region. There is nothing surprising in this typically convective structure.

But at lower altitudes, below a transition zone placed near 100 km of altitude, all the atmosphere at all altitudes and latitudes rotates rapidly in the rotation sense of the planet, presenting the phenomenon called zonal superrotation. The velocity which reaches 100 m/s at the top of the clouds decrease with altitude (still 60 m/s at 50 km) and approaches zero at the ground level. The superrotation was discovered in 1972 by Viktor Kerzhanovitch by measuring the Doppler effect of the signals transmitted by *Venera-8*. The ultraviolet marks of Boyer and Camichel express therefore the motion of the atmosphere which spins at their altitude sixty times faster than the surface of the planet. This situation differs completely from the terrestrial circulation where the strong zonal motions are limited to middle latitudes to form the jet streams. To this zonal circulation is added a meridional component of 10 to 20 m/s from low latitudes to the pole where the air is recycled towards lower altitudes in extended vortices of 2 to 4 days period. The mechanisms maintaining this circulation are unknown, since we ignore everything touching heat and momentum transfers in the inferior atmosphere: below the clouds, we know nothing. Lacking data, we have no explanation.

Radar images obtained by *Venera-15* and *-16*, and principally by Magellan, revealed that the surface is relatively young and has been modified by volcanism and tectonic motions different from the terrestrial plates, and they present landscapes with a morphology encountered nowhere else in the Solar system. The relatively small number of impact craters and their uniform distribution seem to indicate a radical change of the surface all over the globe seven hundred million years ago.

Thus forty years of space exploration have provided a physical and chemical description sufficiently complete to provide the possibility of asking the real questions:

- What is the explanation of the superrotation, or rather of the general circulation of the atmosphere? Remember that the kinetic energy of the Venus superrotation is four hundred times the kinetic energy of ours, for an insolation only double.
- What is the nature of the chemical cycles relating sulfur, oxygen, carbon, phosphorus, chlorine?
- Can we advance in the search for the formation conditions of the planet?

As for the future exploration my personal conviction is that all those problems have to be approached by methods providing long duration in-situ measurements at all latitudes and altitudes. Such methods exist, adapted to the conditions of the atmospheric research on Venus: the balloons. Two aerostats were placed in the Venus atmosphere by the Soviet Union in 1985 during the VEGA mission, following a proposal I made to the Presidium of the Academy of Sciences in November 1967, and drifted each over 11,000 km during 43 hours. Many consider this mission as one of the great achievements of the Soviet space program. It had opened the way, as we believed at that time, to a new approach to planetary exploration, and it is to be regretted that its potential was not exploited. I do hope that the EVE proposal of multiple Venus aerostats made to the European Space Agency for its *Cosmic Vision* program, with a strong participation of Russia, France and Japan, will be favourably received.

THE ATMOSPHERE OF MARS

The composition of the Martian atmosphere is quasi identical to Venus's, that is to say 96% CO_2 and 3.5% N_2, but its density is much smaller: the average pressure at ground level is 7 mbars (one thousand for the Earth and hundred thousand for Venus). The greenhouse effect is therefore negligible and the temperature low (-100°C to 20°C). The major characteristic is the very weak thermal inertia towards diurnal and seasonal changes, which are fast since the duration of the day is only longer than our own day of 24 hours, by 37 minutes, with an inclination of the polar axis of 24° (23.5° for Earth). Two specific phenomena derive from this geometry:

- The contrast between day and night creates major air mass motions between the sunlit and the dark regions, similar to tides.
- The temperature of the poles decreases during winter below the sublimation temperature of CO_2, and therefore an important fraction of the atmosphere solidifies: in-situ probes have discovered seasonal changes of the atmospheric pressure on the ground, reaching 30% related to the annual variations of the white zones observed around the poles called polar caps. Since their temperature is always inferior to 200°C the poles are covered with glaciers a few kilometers thick, constituted of water ice mixed with

dust. In our times, the North cap is essentially water ice and the South cap H_2O covered by a CO_2 ice layer of thickness varying with the season (tens of meters). Large meridional transfers of CO_2 and H_2O from one polar cap to the other take place with the change of the orbital configuration which modify the insulation and therefore the temperature of the polar regions.

Superimposed to these transfers which we may call cryospheric, atmospheric phenomena related to the topography dominate the weather: seasonal sand storms, born in the great depressions of the south hemisphere, fill in a few weeks and for many months all the atmosphere up to 50 or 60 km altitude at all latitudes and generate notable thermal effects. And also, as in the terrestrial stratosphere, undulating motions of various nature as tides, relief or gravity waves dominate the wind structure and the heat transfers.

Thermal conditions of today do not authorize the existence of liquid water. However, the recent observations of *Mars Global Surveyor* have shown evidence for liquid flows originating from the slopes of canyons, and introduce a new idea: the existence of water adsorbed in a quasi liquid state which is conceivable if the thickness of the stratum is larger than two monomolecular layers. The quasi liquid interfacial H_2O can provide transport and maintenance of solutions.

The question everybody asks is the evolution: did Mars possess at sometime an atmosphere sufficiently dense of CO_2 (at least a pressure of 2 bars) to create a greenhouse effect able to support the phenomenon of liquid water as geomorphology seem to indicate? The abnormal ratio N^{14}/N^{15} shows that a considerable erosion of the atmosphere happened in the past, and the ratios of noble gases point towards a loss of the primeval atmosphere, as was mentioned before. This may have taken place very early, and what happened afterwards? Recent measurements by *Mars Express* of the infrared reflectance of this surface exhibit a great diversity of composition due to the coupled evolution of magnetic, volcanic and alteration episodes, without CO_2 intervention: no indication has been found of any long term CO_2 greenhouse after the first billion years. Early escape would have left the quantity of CO_2 observed today and H_2O would have staid stable in the present reservoirs as polar caps, supposed permafrost in the northern hemisphere and hydrated superficial materials synthesized during the primitive evolution.

THE ATMOSPHERE OF TITAN

Titan, the largest satellite of Saturn, possesses a dense nitrogen atmosphere, four times more dense than that of the Earth (ground pressure 1.5 atm and temperature 94K). The methane CH_4 is present under solid, liquid and gaseous phases, and plays a role similar to H_2O's on Earth. Fluvial forms, lakes, clouds and maybe rain exist: Titan has 5 m of precipitable CH_4 and the Earth 2.5 cm of H_2O.

The Titan atmosphere is a chemical plant where methane is irreversibly dissociated at high altitudes by the solar ultraviolet radiation with production of hydrocarbides (C_2H_6, C_2H_2), nitriles (HCN, C_2N_2, HC_3N) and also large molecules as benzene or naphtalene which condense to form mist and clouds, falling towards the ground and fill the stratosphere. Their absorption of light heat the ambient air and generates a wind system of the middle atmosphere similar to the effect of ozone in the Earth's stratosphere and mesosphere. Finally these organic molecules cover the ground. It has been said that it rains oil accumulating at the rhythm of 500 meters in 4.5 billion years.

The vertical profile of the atmosphere is similar to the profile in Earth with a stratosphere and a troposphere presenting a 20K greenhouse effect to be compared to 30K on Earth.

The quick disappearance of small clouds and the fluviale morphology of the equatorial landscapes seem to indicate the occurrence of rain. At high and middle latitudes in the North appear lakes and even seas of methane and probably ethane which can be recycled in the atmosphere and create erosion of ground features. Dunes regularly striated are present over thousands of kilometers in the equatorial region and demonstrate the existence of wind systems. Clouds appear south of 60°S and also as a wide strip around 40°S.

The themes of the future exploration can be listed as:

- The methane cycle and its analogies with H_2O on Earth,
- The chemical transformations of organic molecules in the atmosphere and on the ground,
- The nature of the general circulation.

We are today confronted to more questions than answers. For instance, does Titan have a superrotation as some specialists imagine? From the point of view of the geological evolution if it agreed that the origin of nitrogen is to be attributed to the photodissociation of primordial ammonia, the origin and permanence of CH_4 are a mystery. And also, we wonder if a liquid reservoir containing H_2O, NH_3 maybe sulphur, does not subsist under the surface.

More than Venus or Mars Titan presents types of coupling similar to what is found on Earth between various parts of the atmosphere and the ground surface, because of the apparent similitude of the phase changes of H_2O and CH_4. It is obvious that the first task is to obtain more in-situ data and, here again, an aerostat is the solution. A beautiful idea is to use a hot air balloon or montgolfiere, using local nitrogen heated by the thermal loss of the RTG needed anyway for any mission to Titan.

CONCLUSION ON PLANETARY ATMOSPHERES AND PERSPECTIVES

Not so long ago it was surmised that the comparative study of the atmospheres of Venus, the Earth and Mars, would provide us with deep insight on the evolution of Earth. The work of the last fifty years has shown that these three planets, so near spatially, are however so dissimilar that their resemblance is limited to the phenomena depending on the basic principles of physics. But the devil is in the details. If these atmospheres were originally comparable, small differences did amplifiy as it happens in a chaotic situation, which is the case for the orbital parameters over periods inferior to the age of the Solar system. On the other hand, they can have been widely different already in their origin.

The interest of this comparative planetology has become different but not smaller, with the discovery of exoplanets of which 240 examples have been identified in only twelve years. The question of habitability that is to say, of the conditions of the birth and permanence of life is now set not as a function of us, but as a function of stellar populations, whose number and variety appear limitless.

A new tendency emerges, the ambition of Space Agencies desiring to participate in the exploration of the Solar system, as those of China, India and Japan. It is a bit surprising to see so many uncoordinated plans for lunar missions. Nevertheless the study of planets has shown fine examples of international cooperation. One can call to mind in Russia the efforts of Joseph Delisle, founder of the Pulkovo observatory upon the invitation of Peter the Great, in order to enroll all the European astronomers in the observation of the 1760 transit of Venus. The objective was to measure with great accuracy the solar parallax, that is the distance from Earth to Sun following an idea of Edmund Halley. Among the governments convinced by Delisle was the one of tsarina Elizabeth Petrovna. On the top of a number of Russian expeditions which were specifically organized for this purpose, a French astronomer, Chappe d'Auteroche received the help of five imperial sledges and of an escort to reach Tobolsk. There he performed some of the best observations of the transit. Notwithstanding the fury of the Seven Years war, 120 usable observations were obtained in tens of stations most of them in remote areas. If the value of the parallax was not better, after the dust settled, than the value obtained at Cayenne by Richer in 1672, the operation led to the discovery of the atmosphere of Venus by a score of observers including Mikhail Lomonosov. Such a tradition has been resurrected touching, surprisingly Venus and Halley, by Roald Zinnurovitch Sagdeev, director of Space Research Institute of USSR Academy of Sciences, when in 1980 he decided to internationalize the Soviet mission VEGA to Venus and comet Halley, with as a result, a harvest of scientific data much superior to what would have been achieved by a mission limited to the scope of a national program.

It would he desirable to return to this tradition by bringing nearer the Space Agencies so they would move further from their practice of coordination limited to exchange of information, towards real cooperation. Cooperation could be imagined in the following areas:

- Joint development of specific products as components, instruments, system or software.
- Distribution of space operations among various national Deep Space Networks.
- Deployment of joint space systems around planets as communication and navigation networks, and even Mars or Lunar rovers managed by a group of partners.
- Joint missions as sample returns from Mars, montgolfiere in the Titan atmosphere or a balloon fleet in the atmosphere of Venus or Mars.

I strongly believe that today 50 years after Sputnik we are ready to make these crucial steps and begin our mutual trip towards the future wonders of the Solar system.

V.V. SHEVCHENKO
Shternberg Astronomical Institute
Lomonosov Moscow State University
vladislav_shevch@mail.ru

MOON RESEARCH FOR HALF A CENTURY

INTRODUCTION

A little more than in a year after launching of the First Artificial Earth's Satellite, the Soviet Union tested a new carrier rocket stage designed to overcome the second escape velocity of 11.2 km/s. That opened the era of flights to the Moon and planets of the Solar system. The launching of the *Luna-1* space station on January 2, 1959, opened a new stage of research. The station reached the second escape velocity and passed at the distance of about 6,000 km from the Moon, and became the first artificial planet.

The space era of the Moon research provided such amount of new information about the Earth's natural satellite that exceeds all known of earth-based telescopic observations scores of times qualitatively and quantitatively.

1. MOON'S MAGNETISM

The study of the Moon's magnetism became practicable only using spacecraft. The first station to reach the Moon's surface — *Luna-2* was launched on September 12, 1959. By readings of the station's instrumentation when approaching the Moon, it was discovered that the Moon's magnetic field having intensity within 50…100 gamm is absent [12]. A three-component flux-gate magnetometer of the station performed measurements up to the altitude of 50 km above the Moon's surface. The subsequent analysis of the measurements showed that the Moon does not have the dipole magnetic moment that does not exceed, at least, 10^{-4} of the Earth's magnetic moment [8]. These results were proved by measurements of the *Luna-10* Moon's first artificial satellite launched in 1966. On altitudes of 350 km and over, disturbance effects and features of distribution of the magnetic field and the interplanetary plasma in the Moon's environment were explored in details. Based on the field topology observed, a conclusion was drawn that it has interplanetary nature taking into consideration the distortion by the Moon.

The *Explorer-35* Moon's artificial satellite began more detailed researches of the lunar magnetic field intensity in 1967. Magnetometers installed aboard this spacecraft discovered neither the shock nor the lunar field in the perilune, and the centralized dipole moment upper limit corresponded with the dipole field intensity on the lunar surface of less than 4 γ [13]. Further magnetometric researches proved the absence of the regular magnetic field but discovered local extended residual fields of different orientation in the intensity range of 30 to 327 nT. Measurements of the residual magnetism intensity variations of the lunar rocks by means of the three-component flux-gate magnetometer, when the *Lunokhod-2* self-propelled vehicle (1973) moved on the lunar surface, were of the particular interest. The magnetic survey was executed along the vehicle 30-kilometers route with the purpose of detailed study of the local anomalies.

The magnetic field in whole in the Le Monnier Crater, where the measurements were executed, was weak and its intensity did not exceed 20...30 nT. However, the measurements discovered characteristic anomalies of 10...15 nT connected with craters having sizes of up to 50 m and over. At that, the magnetic field intensity increased with increase of the crater diameter crossed by the *Lunokhod* [8]. That important result became the first experimental indication of possible impact nature of the magnetic anomalies on the Moon.

Apparently, the experimentalists faced with another nature of the local magnetism intensity of the lunar rocks when analyzing magnetometric results by stationary instruments located in landing places of the *Apollo* spacecraft (SC). A residual magnetic field with intensity of up to 327 nT was discovered in the Descartes Crater area (the *Apollo-16* landing area) that had existed on the Moon when the crust material had cooled below the Curie point, i.e. 3.7 to 4.2 billion years ago [7]. The overall conclusion per the stationary measurements results came to that fields with characteristic extension of 50 to 100 km were magnetized uniformly in the early epoch of the lunar surface forming, while anomalies with linear dimensions of 5...10 km had random direction magnetization originated as the result of impact processes. Anomalies of small dimensions have high intensities often, and their sources are relatively close to the surface [7, 8].

The next stage of study of the Moon's magnetic field was acquisition of data from subsatellites of the *Apollo-15* and the *Apollo-16* circulating in own orbits with perilune altitudes of <100 to 150 km above the lunar surface. Many anomalies discovered by means of plasma sensors were simultaneously identified per the magnetometric survey results that provided detection of the local fields. Identification of the same anomalies, when repeated flights, provided no random errors and increased reliability of separation of useful signal from noises in the cases of low field intensity.

The first indications of connection of very considerable magnetic anomalies with diffuse structures, i.e. the albedo anomalies (swirls), on the Moon appeared during the researches described [34]. Such dual anomaly on the visible hemisphere was the Reiner-γ Formation located in the western part of the Oceanus Procellarum attracting attention of observers and analysts long before that as well. Fig. 1 shows the local magnetic field intensity isolines (in nanoteslas – nT) superposed on the image of the Reiner-γ diffuse structure [28]. The *Lunar Orbiter* (NASA) took the Reiner-γ area picture. The magnetic anomaly structure was determined per the magnetometric survey results of the *Apollo* subsatellites at the altitude of about 20 km. Maximum field intensity at this altitude was 22 nT. During further simulation, Hood and Williams [35] determined that this anomaly reached 1,700 nT when reducing to the lunar surface level.

Magnetometric measurements executed per the *Apollo-15* and *Apollo-16* manned flights program related to a narrow near-equator area and numbered just 10,000 individual measurements.

The *Lunar Prospector* magnetometric measurement results (NASA, 1998—1999) covers the whole lunar surface and are based on 700,000 individual

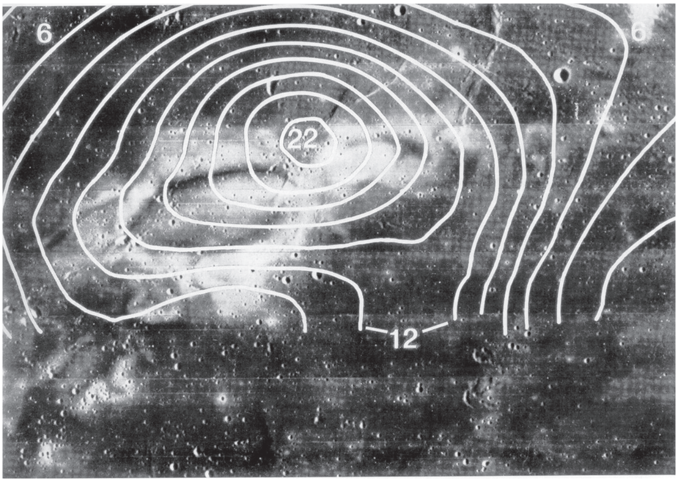


fig. 1

determinations of the local magnetic field intensity. These data are enough to draw up the magnetism map of the whole Moon with resolution of about 3° on the surface. The resolution of the magnetometric survey is only 0.5°, i.e. 15 km [36] for individual areas researched from low altitudes.

The data proves existence of areas of the local magnetic field anomalous intensities. It turned out that four largest magnetic anomalies are on the back side of the Moon in areas being approximately antipodal to Mare Imbrium, Mare Serenitatis, Mare Crisium and Mare Orientale respectively. However, there were no magnetic anomalies discovered in areas being antipodal to ring structures being similar by ages and sizes such as Mare Nectaris or Hertzsprung. Though the experiment authors insist that the origin of the lunar magnetic anomalies is unambiguously connected with the origin of the youngest large ring structures — basins, which became circular seas (mascons) of the visible hemisphere of the Moon in the following stages of development. In the context of such hypothesis, the origin of the largest magnetic anomaly of the Moon coincided, by the location, with the Reiner-γ formation — the enigmatic albedo structure remains inexplicable. This area is not an antipode to any basin on the opposite hemisphere of the Moon.

We could suppose that the Moon's 'spotty' field owes its origin to the combination of different processes taking place in different epochs of the lunar globe evolution. For instance, there were fragments having considerable magnetism intensity among samples of lunar rocks brought to the Earth. Most of them have absolute age of 3.6 to 3.8 billion years. In this connection, many authors do no exclude the dynamo mechanism as the source of the Moon's magnetic field origin in the early stages of the lunar globe forming [37].

Since there is the correlation of many magnetic anomalies with place of intense releases surrounded by large basins of the Moon, it is natural to connect occurrence of such local fields with forming of the impact ring structures about 4 billion years ago.

The origin of the diffuse structures (the albedo anomalies) connected with largest magnetic anomalies of the lunar globe is still one of the most enigmatic features of the Moon's nature. Since the diffuse structures are very young formations, which are about 10 million years old, simultaneous origin of albedo and magnetic anomalies in areas, located antipodally to points of bodies fall, formed large basins of the Moon about 4 billion years ago, remains inexplicable.

Remains of magnetic fields of probably different origins are currently classified as the Moon's paleomagnetism.

However, magnetic anomalies related to the diffuse structures, which are young formations, probably, have nothing to do with the Moon's paleomagnetic fields.

2. NEAR-MOON IONIZED RADIATION

Since the Moon has no magnetic field of dipole nature and practically no atmosphere, the character of the lunar surface exposure is considerably different from corresponding phenomena observed near the Earth's surface. Types of radiation reaching the lunar surface and interactions of each of them with the cover material are more various.

The solar wind ions, due to their low energy, can penetrate only into the very thin stratum of the lunar material — not more than one micrometer. Per some evaluations, the saturation of the solar wind particles near the Moon is such that, for the time of more than 4 billion years, the total number of atoms reached it can be equivalent to the lunar material surface stratum of up to 10 m thick [29]. The solar wind stream density near the Moon is usually accepted as $1{\cdot}10^8$ to $8{\cdot}10^8$. In spite of the fact that a considerable part of the solar wind atoms leaves the lunar surface finally, it is considered that the solar wind is the source of such elements uncommon for chemical compositions of the lunar rocks as H, C, N and some other gases.

Electrons with energy of about 0.5...1.0 MeV, after an intense solar flare, reach the Moon's environment in the time of 10 minutes to 10 hours. Solar protons with energy of 20 to 80 MeV, moving along lines of the interplanetary magnetic field, appear in the circumlunar space in up to 10 hours after [5]. Most of solar cosmic rays do not penetrate into the lunar material deeper than centimeters. In the top stratum, these particles can produce reactions, which leave cascade traces. A stratum of about 100 g/cm^2 is usually a barrier for penetration of the secondary stream particles. Many of the lunar rocks samples brought to the Earth have the solar cosmic ray particle traces, by which the solar wind intensity in the past for the period of about 10 million years and the exposure age of the lunar rocks themselves can be determined.

Heavy nuclei in the galactic cosmic rays do not penetrate into the lunar rocks usually deeper than ~10 cm. Despite the fact that these particles produce reactions in the lunar material and induce cascade phenomena, a stratum of several g/cm^2 is enough for complete attenuation of this process. On the contrary, light nuclei in the galactic cosmic rays, which usually include protons and alpha-particles, can penetrate into the lunar ground deeply and initiate cascades of secondary particles spreading for meters around. The number of the secondary particles, as a rule, exceeds the primary stream by several times. For instance, a stream of the primary particles of the galactic cosmic rays with density of 2 particles/cm²·s can induce a stream with density of about 13 neutrons/cm²·s [27]. One of the processes accompanying bombardment of the lunar cover material with particles of the galactic cosmic rays is 'beat-out' of gamma-particles and neutrons, which form a radiation stream from the Moon, the energetic center of which indicates the chemical composition of the initial material. This phenomenon was the basis of a remote content determination method of such elements as Th, Ti, Fe, Mg, K and others by means of orbital spacecraft [33].

3. MOON'S GRAVITY FIELD

The Moon's external gravity potential — in the form of sum of spherical harmonics is as follows [40]:

$$U = (g\,M_{☾}/r)\,[1+ \sum_{n=2}^{\acute{\omega}} \sum_{m=0}^{n} (R_{☾}/r)^{n}\, P_{nm}(\sin\varphi)\,(C_{nm}\cos m\lambda + S_{nm}\sin m\lambda)]\,,$$

where $M_{☾}$ — Moon's mass; $R_{☾}$ — Moon's average radius; φ, λ — spherical coordinates (latitude and longitude); r — current space point radius-vector; $P_{nm}(\sin\varphi)$ — associated Legendre functions; C_{nm}, S_{nm} — independent expansion coefficients.

The external gravity field spatial structure, as a rule, is described by means of equipotential surfaces. The U value is constant in every point of such surface. In case of distribution uniformity of gravitating masses, the equipotential (daturence) surfaces are spherical, and the potential value depends just on remoteness of the current space point (r-value). Since real mass distribution in the Moon's body is not uniform, local excess or shortage of the material mass result in deformation of embedded equipotential surfaces modeling the system gravity field in the anomaly point environment.

Currently, the main method of study of the Moon's gravity field remains research of gravity disturbances of orbits of its artificial satellites. The first experience of such measurements became results of analysis of trajectory data of the first Moon's artificial satellite (MAS) — *Luna-10* (1966). These pioneer researches allowed determining values of 11 coefficients of expansion of the Moon's gravity field [1]. Fig. 2 presents three-dimensional diagram of plane sections in mutually perpendicular planes of the gravity field plotted per data of the *Luna-10*. It should be noted that general character of distribution of the gravity field intensity values anticipates discoveries of the following years. An anomaly in the Moon's visible

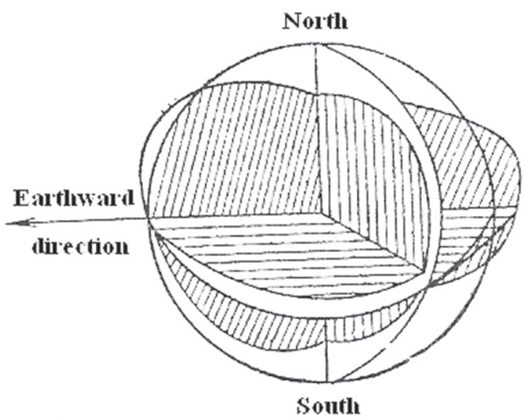


fig. 2

side northern hemisphere corresponds exactly to mascons described in detail later per measurements carried out aboard the MAS of the *Lunar Orbiter* series (1967). Muller and Sjorgen [21], measuring Doppler change of the MAS radio signal frequency, determined components of velocity and acceleration of satellites toward the Earth. The first detailed gravimetric map of the Moon's visible hemisphere central area was plotted basing on these data.

Anomalous intensification of the gravity field in the equatorial area of the Moon's back side, discovered per data of the *Luna-10* (Fig. 2, p. 137), was explained considerably later, when researches were executed by means of MAS of new *Lunar Prospector* generation. One of achievements of that mission was higher accuracy of trajectory measurements by about 5 times, owing to which new more detailed models of the gravimetric field of the Moon's both hemispheres could be constructed [14].

According to researches of the *Lunar Prospector* MAS (1998—1999) with surface resolution of up to 30 km, there were up to 100 harmonics in the expansion the Moon's gravity field. In addition to the mascons, the model discovered gravity anomalies, which have not had the reliable interpretation yet. Fig. 3 gives the gravity map representing the results. The gravity field model for the visible hemisphere has been constructed taking into account 110 expansion terms. The model for the invisible hemisphere has 60 terms. The field intensity is in milligals, isolines on the map are drawn in 100 milligals. We should pay attention to gravity anomalies in the equatorial area of the Moon's back side, which are not connected with the mascons, as it is on the visible side. If the generally accepted mascon structure model of the visible hemisphere is internal rise of mantle masses in circular marine places, then the origin source of the gravity anomalies of the back side remains unknown for the time being.

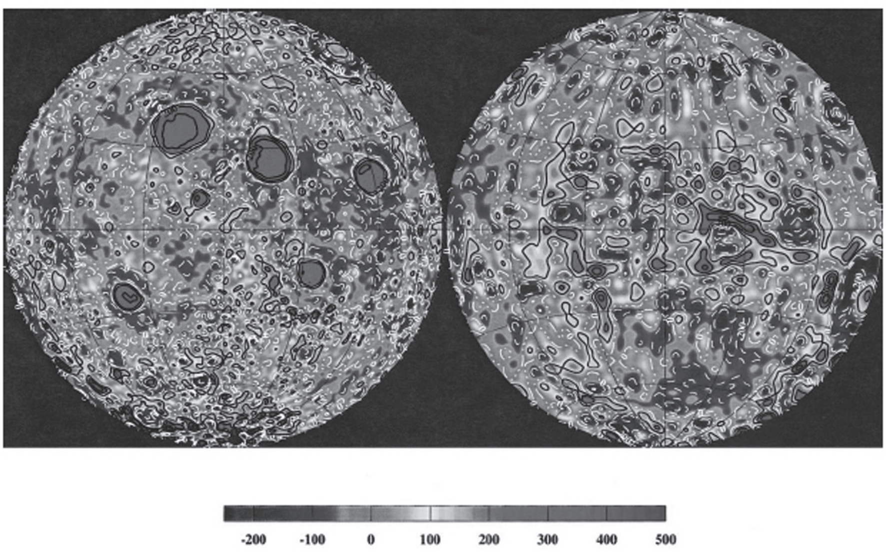


fig. 3. For coloured image, please, refer to [14]

Thus, in addition to general non-centrality, the Moon's gravity field has considerable local anomalies causing the equipotential surfaces deformation in its special model. The largest mascons of the visible hemisphere have local mass excess of about $20 \cdot 10^{-6}$ masses of the Moon.

Fig. 4 shows the generalized model of the Moon's gravity field — lunar geoid. The geoid surface relative heights above the sphere with standard radius of 1,738 km are in meters. 110 expansion terms of the Moon's gravity potential were used to construct this model [14].

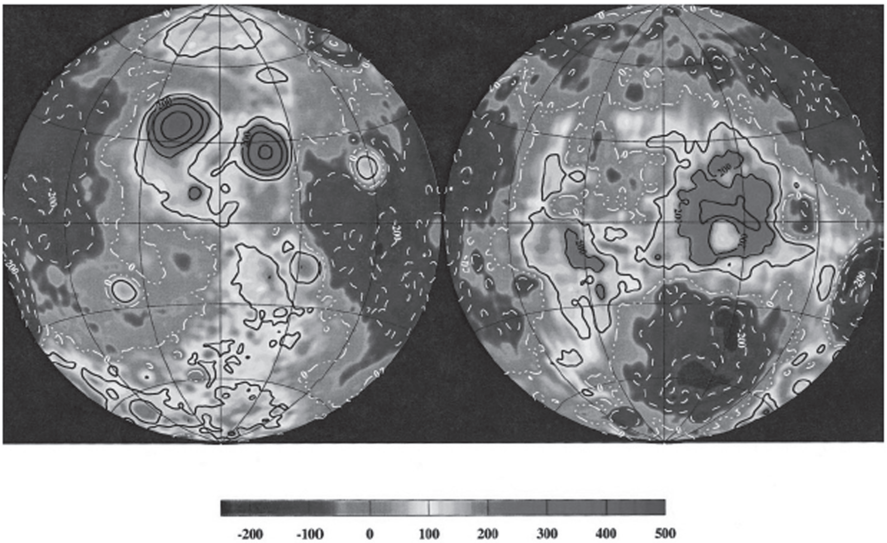


fig. 4. For coloured image, please, refer to [14]

Generalization of the new gravity model allowed pioneering the evaluation of specific dimensions of the lunar metal core, the radius of which is probably within 250 to 430 km and its mass does not exceed 4% of the lunar globe total mass (for more information refer to [42]).

4. METEOROID AND DUST COMPONENTS AT MOON'S SURFACE

With no gas blanket on the Moon, even the smallest meteoroid particles reach the lunar surface causing intense erosion of surface strata. Estimated velocities of fall onto the lunar surface of impact particles are 13 to 18 km/s [11]. According to evaluations of various authors, the total stream of solids falling onto the Moon is about $4 \cdot 10^{-16}$ g/cm$^2 \cdot$s taking into account objects with weight of 10^{-16} g (micrometeorites) to 10^{18} g (large meteorites and asteroids) [40]. Availability of particles of different sizes is usually presented by the $N = aD^b$ dependence, where N — number of particles of D diameter, falling onto a unit of area per a unit of time. The same dependence is used to present distribution of the falling particles by weight:

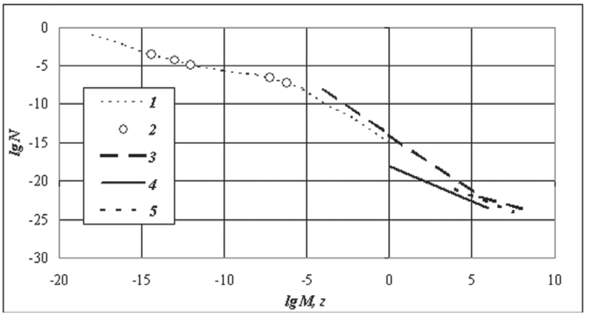


fig. 5

$N = cD^d$.

The *b* and *d* exponents are negative for real distributions. Fig. 5 presents the results of various series of the meteoroid density observations in the near-Moon space in the form of mass distribution of the particles falling onto the lunar surface. The vertical axis gives quantity of the particles *N (the logarithmic scale)* falling onto the area of *1 m² per second. The horizontal axis gives masses of the particles (the logarithmic scale)*

The curve 1 and points 2 are generalizations of various series of earth-based observations and results found aboard *Pioneer-8*, *Pioneer-9*, *Pegas* and GEOS spacecraft reduced to the distance of 1 AU [6; 17]. All of these data characterize the micrometeorite stream with masses of ~10^{-17} to 10^2 g.

The distribution 3 in Fig. 5 is the generalized analysis of the data found per the meteorite erosion degree of the samples of the lunar surface rocks delivered to the Earth [39]. The results of the passive seismic experiment executed on the lunar surface by *Apollo* program allowed to evaluate the meteorite substance stream, which falls onto the Moon really [45]. The dependence found per these data is presented in Fig. 5 by the distribution 4. The recorded stream was 10...1,000 times less than the predicted one per earth-based observations. Later data found by means of the passive seismic experiment and related to a stream with particle mass interval of 10^3 to 10^5 g are shown by the distribution 5 [10].

The values of the meteorite stream falling onto the Moon allow to suppose constant presence, in the near-Moon space, of diffused fine substance — a peculiar 'aerosol component' of the Moon's exosphere. Some observations of excess luminescence of the lunar sky confirm such suppositions. Per data of measurements executed directly on the lunar surface, the particle stream density with mass of more than 10^{-13} g and falling velocity of about 25 km/s is $2{\cdot}10^{-8}$ $cm^{-2}{\cdot}s^{-1}$ [2]. The effect of higher concentration of the microparticles near moments of local sunrises and sunsets at eight lunations was recorded during this experiment. The particle recording velocity increased by almost 100 times for the time of several hours up to 40 hrs before the sunrise and within 30 hrs after the sunrise. It was determined that the particles move predominantly from the Sun. The predicted mechanism of such horizontal transfer of the particles along the lunar surface consists in the interaction of electrostatic charges of the particles with electrostatic fields generated on the lunar surface under action of the solar radiation.

5. MOON'S THERMAL FIELD

The lunar radiation includes a reflected component, i.e. the solar radiation diffused by the surface in visible and near infrared spectral regions and the self-radiation, which appears in the farther infrared region.

The spectral investigations of the Moon became all-wave ones with the beginning of the space era.

The curve 1 in Fig. 6 (the logarithmic scale) shows schematically the distribution in the solar radiation spectrum in the range of from the X-ray radiation to the infrared radiation. Value *E* is in erg/cm²·s per the wavelength interval of 1 μm. Some details in the solar radiation spectrum are smoothed.

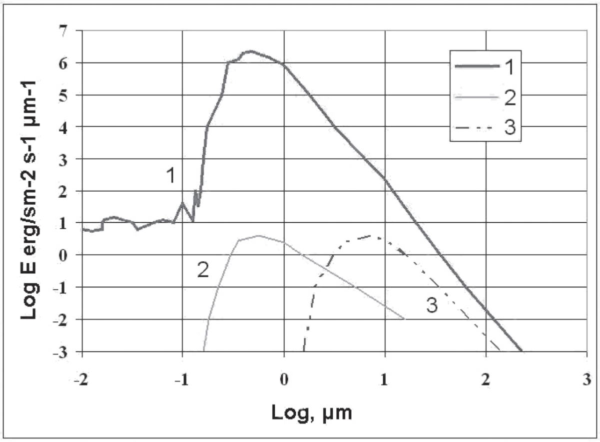


fig. 6

The curve 2 (the solar radiation diffused by the lunar surface) repeats, in the main, the energy distribution in the solar spectrum taking into account the change of the spectral geometric albedo [40]. This last circumstance explains the sharp decrease of the radiation in the ultraviolet spectral region and more slanting curve in the infrared spectral region.

Low reflectivity of the lunar surface stratum results in that about 90% of the solar radiation falling onto the Moon turns into the heat. As the result, the Moon has thermal self-radiation in the infrared spectral region and partially in the radio region. The Moon's self-radiation can be presented by a Planck curve calculated for T = 400 K (in a sunny point of the illuminated hemisphere). The curve 3 in Fig. 6 presents the results of such calculations. The radiation coefficient of the lunar surface is near to 1.

The maximum reflected radiation of the Moon is λ ~0.6 μm while the maximum energy distribution in the solar spectrum is near λ = 0.47 μm.

It follows from this that the sunlight reflected by the lunar surface has the reddish tint. The maximum self-radiation of the Moon is about λ = 7 μm. The Moon's thermal field, on a global scale, can be presented per results of the space infrared survey of the visible disc of the Moon at various phase angles.

The experimental absorptance of different areas of the lunar surface can be evaluated by intensity of the self-radiation of the Moon in the infrared spectral region, since the case in point is the reradiated solar radiation absorbed by the surface stratum of the lunar regolith.

The rugged relief, the surface stratum thermal conductivity and conditions of lighting and survey form the photometric non-uniformity of the Moon's surface in the infrared spectral region. High-accuracy measurements of the thermal radiation of the Moon's surface, carried out by means of present-day germanium high-resolution detectors, have showed considerable deviations of actual temperatures from empirical models existed before. The empirical data discrepancies with actual measurements increase up to 50% when slanting beams and considerable increase towards the limb.

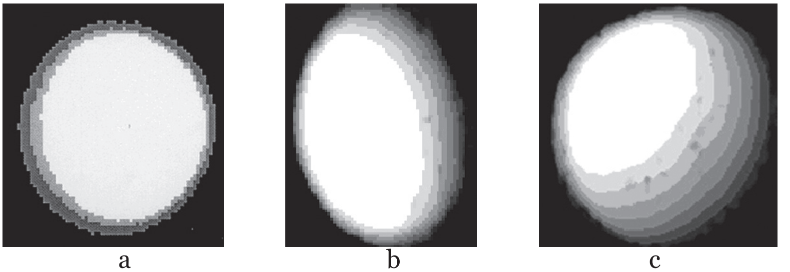
The Moon's thermal radiation in the infrared spectral region was studied by both earth-based and space facilities [42].

Fig. 7 (a, b, c) show isotherms of the visible lunar disc taken when different phases aboard the GOMS Russian stationary satellite in March-July, 1995, and in January, 1996 [26].

The phase presented in Fig. 7a is -26.5^{o}. The isotherms show the temperature distribution of the surface within 261.1 to 397.0 K. Fig. 7b gives the temperature distribution of the surface within 250.1 to 394.9 K at the phase angle of +34.4^{o}. Fig. 7c shows the infrared image of the Moon taken as the previous ones, i.e. by the GOMS satellite at the phase angle of +34.4^{o} and in the surface temperature rage of 250.0 to 394.9 K.

When comparing the infrared images, there is the absence of any details of the marine type and the continental type shown in the isotherm configurations. This circumstance indicates that the absorptance of different landscape types differ insignificantly, since it depends mainly on the regolith nature, which is practically equal by its mechanical properties in the seas and the continents.

Measurements of the thermal radiation of the unlit part of the lunar disc during the phase changing or the lunar eclipses allow evaluating the thermal inertia of the cover matter. The thermal inertia of the lunar ground is two orders less than of the terrestrial rocks. This characteristic allows evaluating the matter fragmentation degree, since such low thermal inertia is peculiar to only highly fine rocks being in high vacuum conditions.



a b c

fig. 7

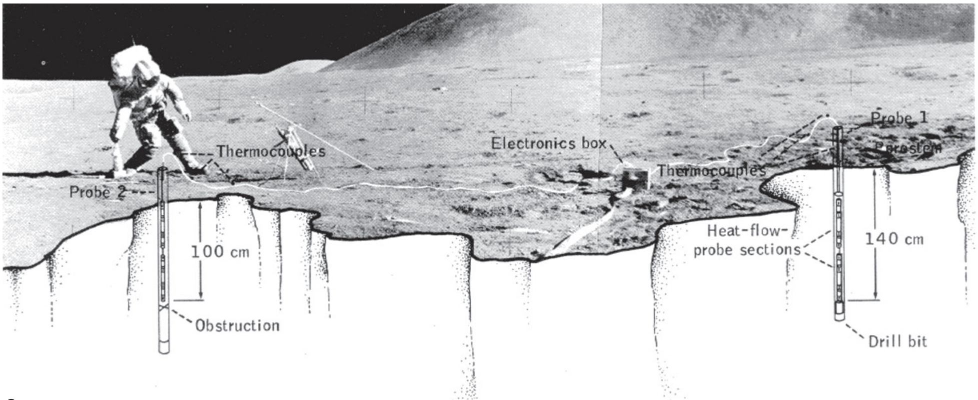


fig. 8

The first branch of the curve 3 in Fig. 6 (p. 141) can be actually extended into the radio wave region. However, low energy level of the Moon's radiation in the radio region does not permit detailed research. At the same time, measurements of radio brightness temperature contain information, which allows determining the thermal conditions of strata located at the depth of several wavelengths under the Moon's surface. The radio measurements of this type, in particular, determined that the lunar regolith temperature is not changed appreciably at the depth of about one meter.

This conclusion was subsequently confirmed by experimental ground probing by *Apollo* crews [33]. The thermal current was measured directly on the Moon by American astronauts in the moon landing places of the *Apollo-15* and *Apollo-17* spacecraft in areas of Vallis Headly (26°26´N, 3°39´E) and Taurus-Littrov (20°10´N, 30°45´E). Fig. 8 presents the experiment photo diagram on deep temperature measurements during the *Apollo* expeditions.

The Moon's surface radiation temperatures measured by spacecraft coincide with data of the earth-based measurements. Temperature differences of the Moon's surface stratum indicate very low thermal conductivity and extremely high porosity of the matter.

As the calculations showed, the brightness temperature of the lunar surface areas at the equator varies from 400 K in the sunny point to 94 K in the anti-sunny point. The maximum temperature of the Moon's illuminated disc surface in latitude 80° is 325 K and the surface temperature in the subpolar point is 185 K.

Fig. 9 shows the temperature probing results in latitude + 26°.

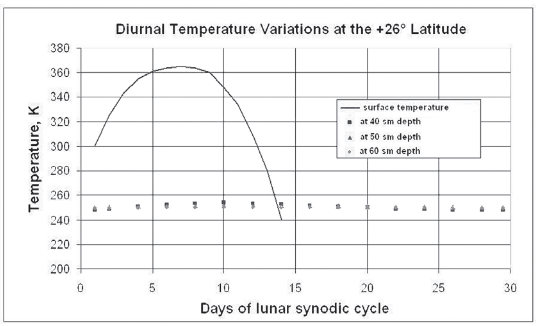


fig. 9

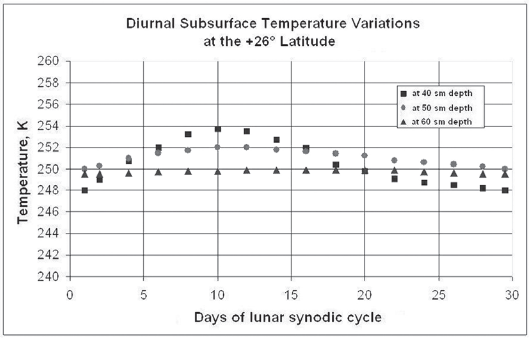


fig. 10

As it follows from Fig. 9, a stratum of the lunar regolith of 50 cm thick makes practically absolute barrier for radiation even in periods of flashes on the Sun. As the Moon's surface direct measurements show, the temperature remains practically constant and does not depend on external daily changes already at the depth of 60 cm. Fig. 10 shows comparison of daily temperature variations of the surface ground stratum shallow under the surface stratum when increasing the diagram vertical scales for better visualization (the *Apollo* data).

6. MOON'S FAR SIDE AND LUNAR GLOBE STRUCTURE

Two years after launching of the First Artificial Satellite of the Earth – October 4, 1959 — the third lunar unmanned station started, the main scientific task of which was solving the age-old mystery of the natural satellite of our planet – the structure of the Moon's back side. Synchronous rotation of the Moon around the Earth and around its axis did not allow to solve this problem by other facilities. During this flight, so-called gravity-assist maneuver was used for the first time, which was applied with success many times afterwards when many interplanetary missions. The *Luna-3* unmanned station passed along a complex trajectory flying around the Moon and returning into the near-earth space. When maximum approach to the Moon at the distance of 60...70 thousand km, the station was directed so that the visible disc of the Moon was in field of view of phototelevision cameras. During the photographing session, a series of pictures was taken, and then they were processed aboard the station and were transmitted to the Earth when the spacecraft returned into the near-earth space.

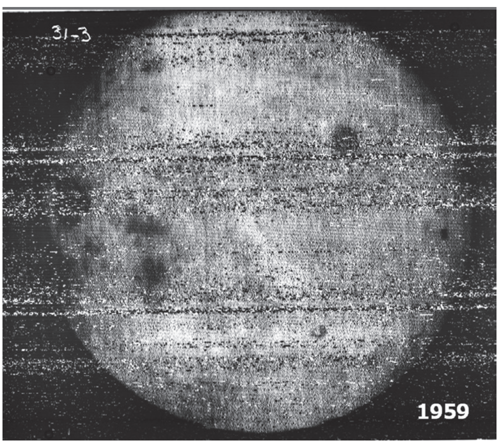


fig. 11

Fig. 11 gives one of the pictures of this series before processing in the Shternberg Astronomical Institute of the Lomonosov Moscow State University. The pictures transmitted to the Earth by the *Luna-3* unmanned station showed, for the first time, such a phenomenon in the Earth-Moon system as morphological asymmetry of the lunar globe. It turned out that mare formations concentrate on the hemisphere facing the Earth. The Moon's side invisible from the Earth is practically a single highland plate with small depressions filled up with lava.

This conclusion was supported by shooting results from the flight trajectory of the Moon's back side part, not been covered when the first photographing, which were taken by the *Zond-3* space station (1965). Further detailed shooting of the whole lunar surface executed during work of the Moon's artificial satellites of the *Lunar Orbiter* series proved these conclusions in full.

Thus, the main landscape type of the lunar globe was the surface of the highland type with elevated relief and higher density of craters per a unit of area than within the mare formations. The mare structures cover on the Moon only about 16% of the surface. The global survey of the lunar globe showed that the highland surface is the Moon's ancient crust undergone by the strongest asteroid-meteoroid bombardment in the initial period of the Moon as a self-body. Occurrence of the mare formations is the further evolution stages of the terrestrial satellite and reflects individual periods of its further history.

General evolution stages of the Moon's surface are often observed by the example of regional areas of combination of different landscape forms. Fig. 12 presents pictures taken during the *Lunar Orbiter* project. The picture of middle resolution

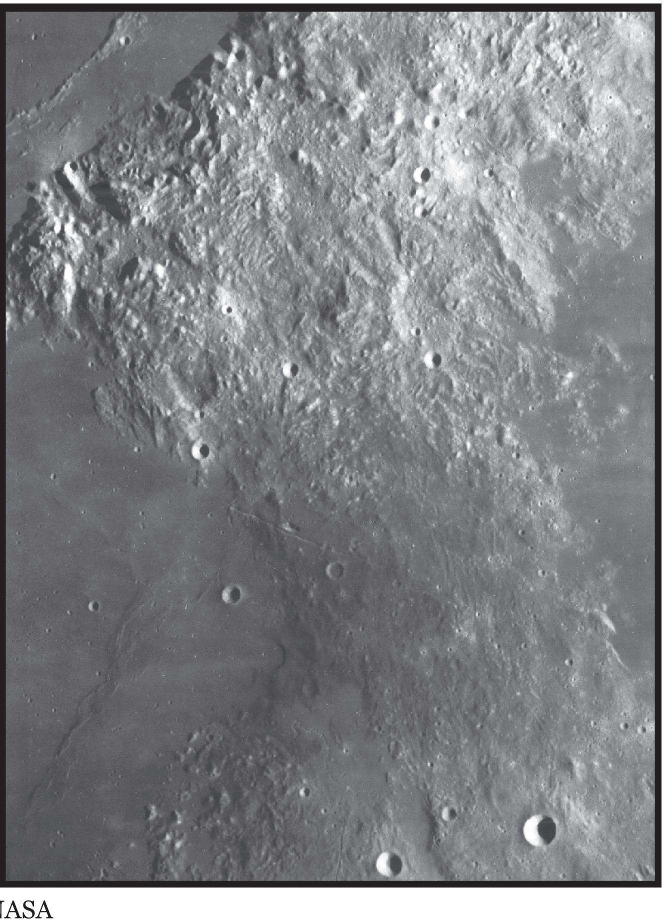


fig. 12 © NASA

covers considerable territory including a highland area (on the right) and a part of the adjacent sea (on the left). Swells on the mare surface indicate processes of gradual infill of the depression existing before with molten lava and then solidified lava. Generalization and analysis of such material allowed producing the map of global geological transformations of the lunar surface showed in Fig. 13. The mare formations are marked on the map by the spectrum of red colors, areas of comparatively young outliers from large impact structures are marked by blue-cyan colors, craters of the highland landscape, the occurrence of which relates to different geological epochs of the Moon, have the corresponding colors.

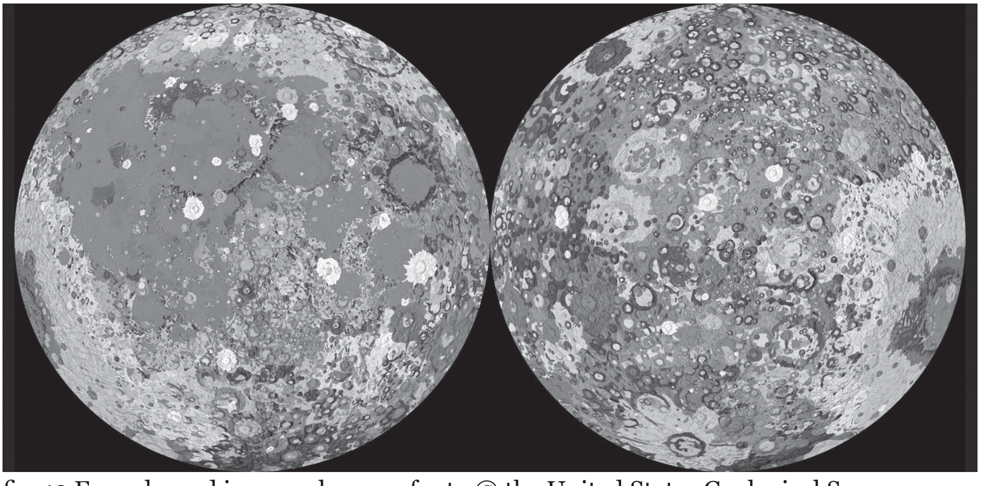


fig. 13 For coloured image, please, refer to © the United States Geological Survey

The oldest structures have the age of about 4.0...4.5 billion years. The seas were generated within several stages and took, in large, quite long period of the lunar history — from ~3.9 to 2.0 billion years. It should be taken into consideration that these evaluations are not final, since the correct absolute age of rocks of different genesis has been determined only in 9 limited areas of the lunar surface. Samples of the lunar matter were delivered to the Earth from those areas per the *Apollo* program and the *Luna* program. Individual craters have the age of a wider range. The overwhelming majority of the continental impact formations are of the ante-maria period. However, a number of large formations of this type were generated comparatively not long ago. The suppositional age of such crater as Copernicus is several hundred million years. Since meteorite bombardment of the lunar surface is going on, formations of modern origin are also to be within both the mare landscapes and the highland landscapes.

Morphological survey of the lunar globe by means of spacecraft resulted in discovery of the largest ring impact structure in the Solar system, which is probably one of the oldest as well. Preliminary age evaluations of this formation correspond to about 4 billion years [25].

The nature and the origin of the unique formation, which is conventionally called 'The South Pole – Aitken Basin' until now, remains one of the most important problems of modern researches of the Moon. Not only its dimensions in the absolute scale draw attention but the fact that its diameter nearly coincides with the Moon's diameter as well. There is no such correlation on other bodies of silicate or ice compositions in the Solar system.

The history of discovery of the largest ring planetary structure return us to results of the first space survey of the Moon. A spacious dark area in the south-eastern part of the limb being visible when shooting was interpreted as the mare-type formation and was named 'Mare Somnium' (see Fig. 11, p. 144). Further shootings aboard the *Lunar Orbiter* Moon's artificial satellite did not reveal this structure. However, the later mutual analysis of the pictures from the *Luna-3* and the *Zond-3* unmanned stations showed the correctness of the first interpretation of the picture taken in 1959. Fig. 14 gives the mosaic of transformed pictures by the *Luna-3* (on the left) and the *Zond-3* (on the right) executed by Stooke [16]. The arrows indicate internal slopes of the formation eastern swell, the images of which in such foreshortening have not taken so far.

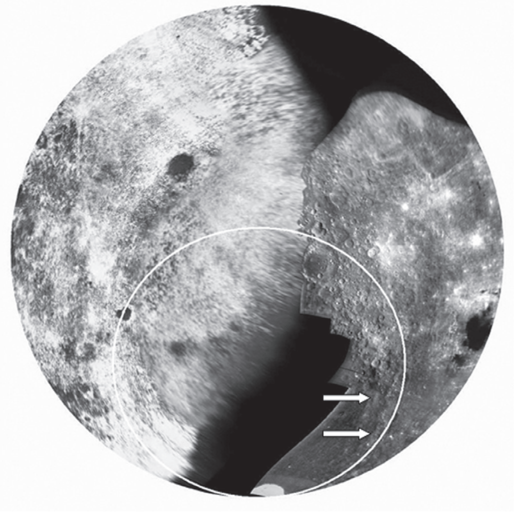


fig. 14

It is interesting to compare two images of the back side presented in Fig. 15. On the left of the image, there is the picture taken in 1959 by the *Luna-3* unmanned station and processed by the Shternberg Astronomical Institute of the Lomonosov Moscow State University. The arrow indicates the dark area on the eastern limb

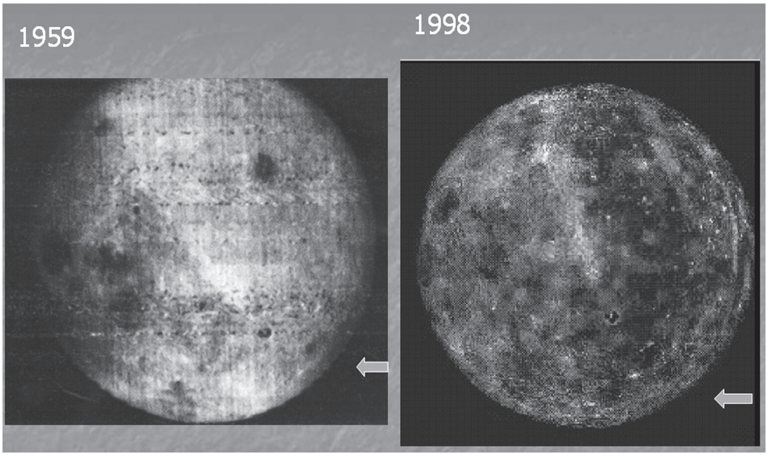


fig. 15

named 'Mare Somnium'. On the right of the picture, there is the image in the same foreshortening showing iron spreading in the surface stratum according to data of the *Lunar Prospector* spacecraft. An arrow indicates the area corresponding to Mare Somnium. It easy to note that the same formation is indicated on the left image and on the right image. According to modern notions, the noted area concerns the internal part of the South Pole — Aitken Basin. The surface rocks here have high iron content and low albedo respectively [32]. Thus, it can be claimed that the first information about the largest impact structure in the Solar system was in the first pictures of the Moon's back side.

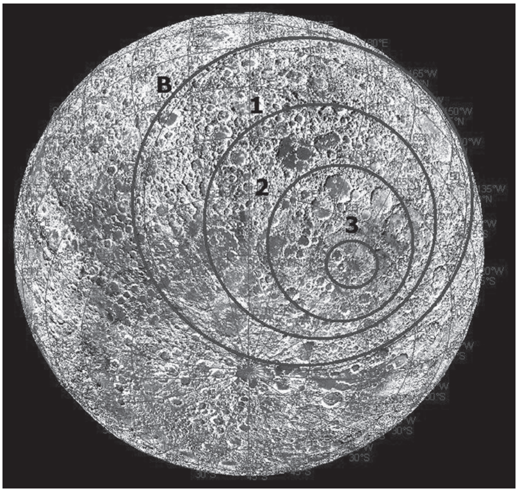


fig. 16

On the grounds of the generalized data on attitude measurements within the hemisphere containing the ring structure of the South Pole — Aitken Basin, the Sternberg Astronomical Institute of the Moscow State University generated the hypsographic map and the altitude profiles of the basin based on the spherical surface of pertinence [43]. The derived dependencies of spreading of the main chemical elements (Fe and Th) on the structure altitude levels discovered total content correlation of these indicators of the lunar rocks with the altitude levels of their prevailing spread. The basin external ring (B) and the structure of rings of its central depression (1, 2 and 3) were discovered on the combined hypsometric and geochemical basis (Fig. 16). The basin total size of about 3,500 km was authentically determined for the first time. It turned out that the unique feature of the basin structure is difference from central-ring symmetry of its internal rings. Such feature can indicate that a hypothetic impact agent moved along a trajectory (or an orbit), which was oriented almost normally to the ecliptic plane.

Without going into modeling details of the impact process itself, two reliable facts can be indicated. The hypothetic impact agent moved along a trajectory (or an orbit), which was oriented almost normally to the ecliptic plane. Among currently known large objects, long-period comets or objects of the Kuiper Belt, what is probably the same have such orbits. Since asteroids or planetesimals in the past had orbits close to the ecliptic plane (or the pre-planet disc plane) [9], it significantly decreases the probability of impact generation of the South Pole — Aitken Basin as a result of fall of such a body.

Recent Byrne's research works [3, 4] contain the direct indication that the impact agent, which formed the South Pole — Aitken Basin, is of another population of bodies than those impact agents falling onto the Moon generated all other basins.

Works of O'Keef and others [23, 24] showed that decrease of the 'depth-diameter' correlation in impact structures takes place under other equal conditions as the result of decrease of the impact agent matter density. Shevchenko's work [41] stated some problems of application of this theory to conditions of generation of impact structures on the Moon.

Based on these two statements, a hypothetic conclusion can be made that unique features of the nature of the South Pole — Aitken Basin can be originated from an unusual process of its generation as the result of fall of a comet-type body. At the same time, a very prevalent notion about the nature of the Edgeworth-Kuiper Belt bodies is their classification as comet nuclei of considerable sizes in inactive states. Known evaluations of quantity of transneptunic objects state the considerable quantity of them. However, positions of these objects at the periphery of the Solar system place limits on their total mass, which results inevitably in the conclusion that the considered objects have extremely low average density.

The substantiation that objects of the Edgeworth-Kuiper Belt or huge comet bodies from the Oort Cloud predominated among main types of impact agents during the supposed period of forming of the South Pole — Aitken Basin can be found in works of Morbidelli [19, 20], Murray and others [22], Fernandez and others [30] and Schmitt [44].

7. LUNAR REGOLITH

The substantial part of knowledge about the Moon's nature was obtained by way of research of the lunar globe surface stratum – the regolith. The regolith is a loose cover of shattered rocks consisting of fragments of various sizes including a fine dust-like fraction. Since the regolith was mainly formed during impact processes, the substantial part of the surface matter fragments are breccias – caked rock fragments, which are partially covered with glass generated as a result of high temperatures in the impact epicenter.

The first historic soft landing of the *Luna-9* unmanned station onto the Moon opened the feasibility of direct research of the lunar regolith in natural conditions and laboratories. Panoramas of the surrounding landscape with resolution of individual details of centimeter sizes allowed making quite full notion about the Moon's cover stratum structure on the first stages of the imagery analysis (Fig. 17). A scanning camera of the *Luna-9* unmanned station, for the period of the lunar day, transmitted several cycloramas, which allowed to reconstruct the surrounding

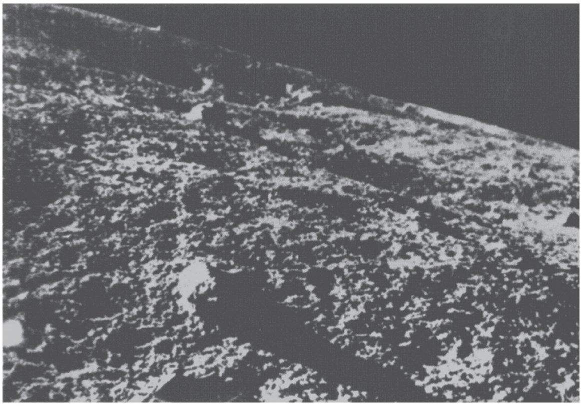


fig. 17

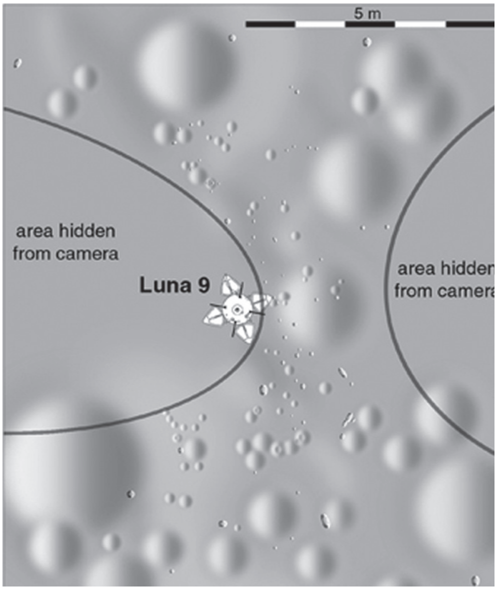


fig. 18

area relief. Fig. 18 presents the diagram of the first shooting on the Moon generated by Stooke [16].

Later on, multiple-kilometer routs of the *Lunokhod-1* and the *Lunokhod-2* accompanied by shooting of the surrounding landscapes showed that the nature of the regolith surface stratum in the place of landing of the *Luna-9* unmanned station is typical for both marine and continental types. There are differences, in the main, in density of large fragments (stones) per the unit of area (Fig. 19). These features depend on closeness of the shooting place to the impact structure surrounded by outlier field. It is noted, in the research area of the *Lunokhod-2*, that the richest stony outliers appear when increase of crater diameters up to 20…30 m (see the review in the book [40]). The regolith average thickness of 4 to 6 m corresponds to it. Similar investigations executed for the *Lunokhod-2* working area in the Le Monnier Crater allowed to evaluate the regolith average thickness of 2 to 3 m at extreme values of 1 to 10 m. The deepest stratum of the regolith was in the area of landing of the *Surveyor-6* spacecraft (Sinus Medii). Its thickness in this area is to 25 m. Calculated data based on comparison of measured depth of an eroded crater about 500 m in diameter with its probable depth in the moment of generation evaluate the regolith stratum of to 50 m in the crater centre part.

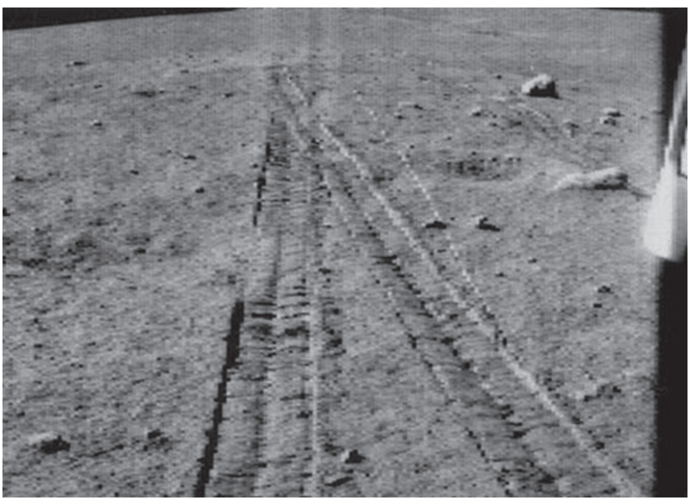


fig. 19

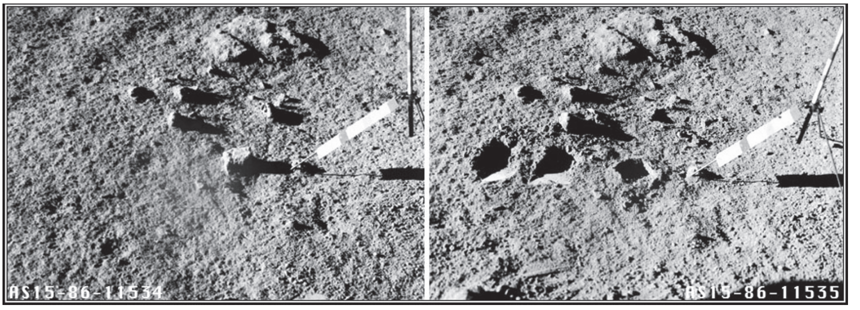


fig. 20 © NASA

There is also a notion of megaregolith implying a stratum of rock divided by extensive fracture system. Such structure of upper part of the lithosphere can have depth of hundreds meters and more.

The regolith fine structure was mainly researched on ground samples delivered to the Earth. Fig. 20 shows an surface site in the place of *Apollo-15* landing before (on the left) and after (on the right) sampling of rock and ground by the astronauts. The regolith fine-grained fraction (less than 1 mm) consists of particles with average sizes of 60…70 µm. There are several particle types. Some of them are fragments of primary rocks and monomial grains. Relative content of these particles depends on the regolith maturity, i.e. the degree of its processing by external factors of the surrounding space environment. For example, there were about 20% of fragments of primary rocks among fragments of more than 0.5 mm in a mature marine regolith sample delivered to the Earth by the *Luna-16* unmanned station. A sample delivered by the *Luna-20* unmanned station had younger exposure age (the maturity) and the primary rock content in its fine-grained fraction increased up to 50% [31].

Another type of particles, the already mentioned breccias, is a product of processing of the lunar fine matter, and their content in the regolith increases with increase of exposure age. For the fraction of particles larger than 0.5 mm in samples delivered by the *Luna-16* and the *Luna-20*, the part of breccias is about 40% and 20% of the particles respectively.

A special type of particles is agglutinates — fragments of minerals, rocks and glasses, and pure glass fragments of spherical, cylindrical and dumbbell-like shapes caked into heterogeneous mass. The total content of the glass particles is low. Their masses do not exceed 0.01% of total mass of the ground samples. However, variations of contents of these components allow predetermining the considered sample evolution stages.

In the general case, the fine-grained part of the regolith is in loose state. The surface stratum porosity is 32% to 58%. The regolith density increases with the depth increase. If the regolith maximally packed sample (1.8…2.0 g/cm^3) is taken as 100% of its density, then on the surface in the natural bedding, the lunar ground has density of 48%, the fine fraction density is already 82% at the depth of about 10 cm, and then it increases up to 99% at the depth of about 60 cm.

Density of stone fragments varies from 2.4 to 3.4 g/cm^3 and its upper value, probably, approximates to the density of underlying rocks [38]. It should be reminded that the lunar globe average density is 3.343 g/cm^3 .

The lunar matter samples were researched by direct methods on the lunar surface in nine separate regions, which differ depending on the area of sampling. Six landing places of the *Apollo* spacecraft are characterized by wide coverage and variety of samples. Three landing places of the *Luna* spacecraft give information about confined areas of sampling of the ground fine fraction and individual small fragments of rocks. In common, this information characterizes peculiarities of content of the lunar rocks in marine and continental regions of different types.

All areas researched by the direct methods are located within the Moon's visible hemisphere and can serve as reference ones when spreading the information onto more spacious areas by means of remote methods.

The main information about elemental chemical composition of the surface stratum on the Moon's visible and back sides has been obtained by remote methods of the planetary astrophysics. These methods, when over-fly or orbital shooting, are based on the analysis of distribution of reflecting surface spectral albedo. Since the main parameter determining the spectral albedo of rocks is their chemical composition, these data are used, when solving the inverse problem, in case of remote determination of elemental and mineralogical composition of the lunar rocks.

When using the orbital shooting of the lunar surface, in a number of cases, methods of recording and analysis of alpha-particles and gamma-particles were used, by the spectral distribution of which the cover matter elemental composition is also determined.

The analysis of the average elemental composition of the lunar ground surface stratum shows that main characteristic features of chemical composition of the lunar rocks are contents of iron, titanium, aluminium and magnesium. Individual indicators of anomalous prevalence of these or those minerals in the lunar rocks are such elements as calcium.

The main luminophor is iron and its compounds, mainly oxides. Correlation dependences of iron and other elements serve often as secondary indications to determine chemical and mineralogical compositions of the rocks having different albedo.

Mafic (dark) rocks of the Moon consist mainly of minerals differing by high contents of iron oxides. Furthermore, by shooting materials of the *Clementine* spacecraft (NASA), direct data about presence of the correlation (inverse) between the albedo of the rock surface stratum and contents of iron oxides in them.

The analysis of the geochemical survey of 20% of the lunar surface implemented by the *Apollo* Program showed that high contents of Al_2O_3 in continental rocks serve as an additional characteristic of differences of basalts (marine rocks) and anorthosites (continental rocks). It is natural that comparison of contents of iron and aluminium in the Moon's surface matter results in discovery of the inverse correlation. Thus, the dependence global analysis (without taking into consideration fine petrologic features) results in the conclusion that the aluminium content in the lunar rocks does not have independent significance as the additional index of the rock types but just repeats global differences between marine and continental surface materials.

The contents of titanium in the rocks have another significance. Determination of elemental contents of titanium in the surface rocks makes for their identification by different types of basalts.

In addition to rocks of anorthosites and basalt types, rocks of the norite series stand out in the global scale on the Moon, which differ by high contents of potassium and rare-earth elements (KREEP). The characteristic feature of rocks of this type is high contents of thorium, which ranks among elements determined remotely and, as a rule, serves as an indicator of the norite rocks.

The anorthosite (highland) rocks are the lightest in the lunar albedo terms. The norite rocks have intermediate transient albedo values. The darkest (mafic) rocks are basalt (mare) lavas with the lowest albedo values.

Generalized results of the direct laboratory research of the lunar rock chemical compositions are given in the table. Designations of basic oxides and such individual elements as potassium, sodium and others are given in the first column. The contents of corresponding oxides (in %) and individual elements (in ppm) are given in the next columns; the contents of oxides for typical samples of the marine ground enriched with titanium delivered from the region of Mare Tranquillitatis are given in the second column; the contents of oxides in the ground of Oceanus Procellarum are given in the third column; the contents of oxides in the ground of the anorthosite-type continental sample delivered

Compound	**Sample 10084 Apollo-11 Mare Tranquillitatis**	**Sample 12070 Apollo-12 Oceanus Procellarum**	**Sample 60601 Apollo-16 Decart Crater**
SiO_2, %	42.1	46.0	45.4
TiO_2	7.2	2.7	0.55
Al_2O_3	13.0	12.7	26.3
Cr_2O_3	0.27	0.33	0.11
FeO	15.4	16.3	5.6
MnO	0.20	0.21	0.07
MgO	8.0	9.7	6.5
CaO	11.3	10.6	16.9
Na (ppm)	3150	3090	3510
K	1090	1900	890
Co	27	42	31.2
Ni	280	200	400
La	15	33	13
U	0.35	1.7	0.57

from the region of the Descartes Crater are given in the fourth column. Such data served as the reference ones when calibrating information about the lunar regolith surface stratum chemical composition obtained by the remote methods.

The data generalized in the table characterize the lunar globe background surface in respect of abundance of chemical elements for basic morphological categories (types of marine and continental matters). These data, on the current stage of study of the Moon's nature, can be used as reference ones when discovering anomalous regions differing by contents of chemical elements in surface strata of the regolith.

The multispectral survey from the board of the *Clementine* spacecraft showed that this method is effective to determine mineralogical characteristics of the surface rocks with high detail resolution. Fig. 21 gives the diagram of passbands of filters used when surveying from the board of the *Clementine* spacecraft.

Fig. 21 shows spectra of grounds from the landing places of the *Apollo-12* (AP-12) and the *Apollo-16* spacecraft (AP-16) to compare with details of spectra of specific samples. Chemical composition of these samples is indicated in the table.

It should be noted that spectral characteristics of different minerals stand out more clearly for samples exerted by the space erosion. In this regard, the data for non-shattered fragments distinguish by greater definiteness. The more a sample is shattered (e.g. a ground sample having high maturity degree) the less its mineralogical and petrological characteristics are determined by means of the remote methods.

The sample 67455, the spectrum of which is shown in the figure, is such an example of a continental rock fragment (a stone), which is seen from strongly pronounced spectral lines being typical for corresponding minerals.

Fig. 22 presents the map of the whole surface of the Moon (excluding near-pole areas) showing distribution of iron contents (in percents). The map was made per data of the multispectral survey from the board of the *Clementine* spacecraft.

Fig. 23 shows the analogous map of titanium distribution.

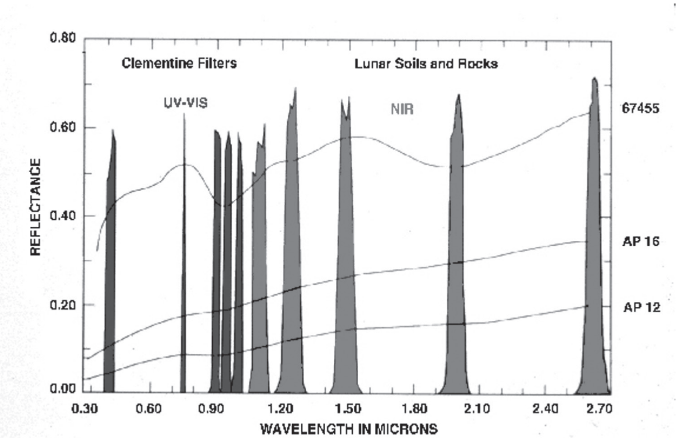


fig. 21

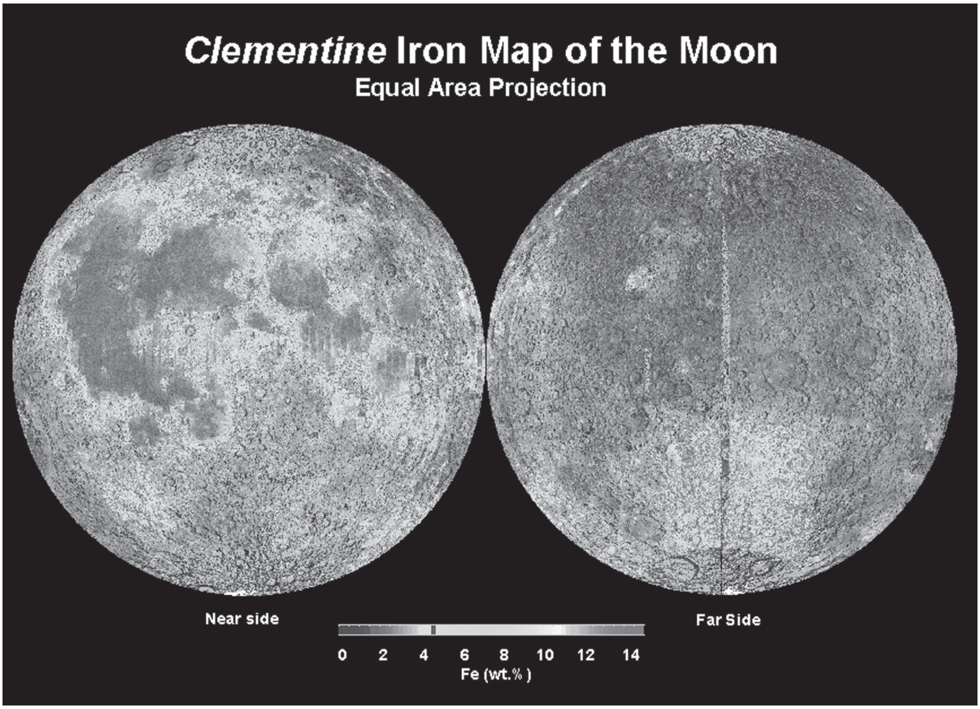


fig. 22. For coloured image, please, refer to © NASA

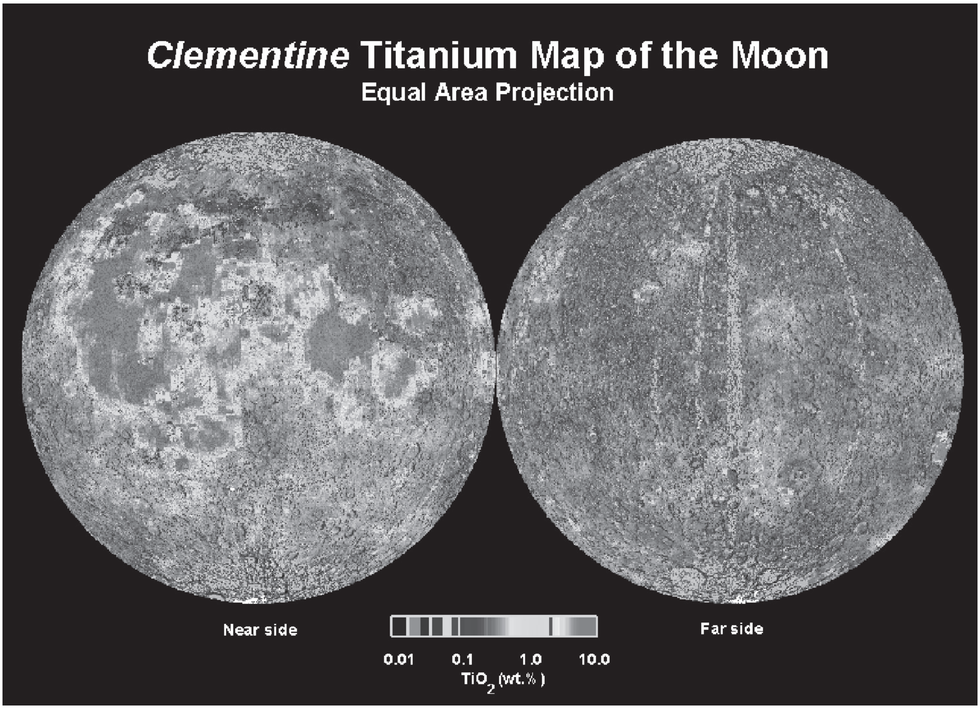


fig. 23. For coloured image, please, refer to © NASA

CONCLUSION. UNSETTLED FUNDAMENTAL PROBLEMS OF LUNAR RESEARCH.

After almost half a century of lunar research by space facilities, a number of the greatest problems relating to the cosmogony remain unsettled so far. In spite of the long-term program of seismic investigations in the context of the *Apollo* Project, there is no quite reliable information about sizes and physical characteristics of the lunar nucleus. Perhaps, seismometers-penetrators, installation of which is planned on the *Luna-Glob* Russian spacecraft, and making of the international network of such instruments on the Moon's surface will allow approaching to reliable model of the lunar globe internal structure.

The hypothesis of existence of the lunar polar ices in 'cold traps', which are areas of permanent shadow near the Moon's poles, waits for its final solution. Perhaps, the launch of the LRO spacecraft (*Lunar Reconnaissance Orbiter*, NASA), aboard of which the LEND Russian instrument has been installed, will allow approaching the understanding this feature of the lunar nature as well [18].

In spite of multitudinous investigations, there is no absolute clarity in origin and evolution of the South Pole — Aitken Basin. However, we could hope that new data acquired by lunar satellites launched by the ESA (the SMART-1 spacecraft, 2003—2006), Japan (the *Kaguya* space vehicle, launched in 2007) and China (the *Chang'e I* spacecraft, launched in 2007) already have solutions of some current problems [15]. India started lunar investigation (*Chandrayaan-1* spacecraft was launched in 2008).

The problem of the Moon's origin remains ambiguous as before, which, seemingly, will require both revision of existing data interpretation and acquirement of new information about the Earth-Moon system nature.

REFERENCES

1. Akim E.L. Determination of lunar gravitational field by the motion of the *Luna-10* artificial satellite of the Moon // Kosmicheskie Issledovaniya. 1966. V. 4. № 6. P. 823—826. (in Russian).
2. Berg O.E., Richardson F.F., Rhee J.W. et al. Preliminary results of a cosmic dust experiment on the Moon // Geophys. Res. Lett. 1974. V. 1. № 7. P. 289—290.
3. Byrne C.J. Size distribution of the lunar basins // Lunar and Planet. Sci. XXXVI. 2005. № 1260.
4. Byrne C.J. The near side megabasin of the Moon // Lunar and Planet. Sci. XXXVII. 2006. №n 1930.
5. Van Hollebeke M.A.I., Ma Sung L.S., McDonald F.B. The variation of solar proton spectra and size distribution with heliolongitude // Solar Phys. 1975. V. 41. P. 189—223.
6. Grun E., Zook H.A., Fechtig H. et al. Collisional balance of the meteoritic complex // Icarus. 1985. V. 62. P. 244—272.
7. Dayl P., Parking C.W., Daily W.D. Lunar electrical conductivity, permeability, and temperature from Apollo magnetometer experiments // The Soviet-American

Conference on Cosmochemistry of the Moon and Planet. Eds J.H. Pomeroy and N.J. Hubbard. NASA SP-370, Washington, D.C., Part 2, 447-492, 1977.

8. Dolginov Sh.Sh. Planetary magnetism // Itogi nauki i tekhniki. Ser. Issledovaniya kosmicheskogo prostranstva. V. 18. Moscow, VINITI, 1982. 130 p.
9. Dorofeeva V.A., Makalkin A.B. Early Solar System evolution. Moscow, Editorial URSS, 2004. 264 p. (in Russian)
10. Duennenbier T., Dorman J., Lammlein D. et al. Meteoroid flux from long period lunar seismic data // Lunar Science VI. Houston. LPI. 1975. P. 217—219.
11. Zook H.A. The state of meteoritic material on the Moon // Proc. 6th Lunar Sci. Conf. 1975. P. 1653—1672.
12. Keldysh M.V., Marov M.Ya. Space researches. Moscow, Nauka, 1981. 192 p. (in Russian)
13. Colburn D.S. Lunar Magnetic Field Measurements, Electrical Conductivity Calculations and Thermal Profile Inferences // The Moon. Eds S.K Runcorn and H.C. Ury. IAU. D.Reidel Publishing Company/Dordrecht – Holland, 1972. P. 355 – 371.
14. Konopliv. A. S., *et al*. Recent gravity models as a result of the *Lunar Prospector* mission // *Icarus*. Mar. 2001. V. 150. № 1. P. 1—18.
15. Spudis P.D., Plescia J., Bussey B., Josset J.-L., Beauvivre S. and the AMIE team. The geology of the South pole of the Moon and age of Shackleton crater // Lunar and Planet Sci / XXXIX. 2008. № 1626.
16. Stooke P.J. International Atlas of Lunar Exploration. Cambridge Univ. Press. 2007. P. 400.
17. McDonnell J.A.M. The role of accretionary particles on lunar exposure and aging processes – lunar dust slow lunar clocks. COSPAR 20th Plenary Meeting. 1977. P. 6.
18. Mitrofanov I.G., Sanin A.B., Kozyrev A.S., Litvak M.L., Mokrousov M.I., Tretyakov V.I. Lunar Exploration Neutron Detector for NASA Lunar Reconnaissance Orbiter // Lunar and Planet. Sci. XXXVI. 2005. № 1879.
19. Morbidelli A. Chaotic diffusion and the origin of comets from the 2/3 resonance in the Kuiper Belt // Icarus. 1997. V. 127. Issue 1. P. 1—12.
20. Morbidelli A. New insights on the Kuiper Belt // Science. 1988. V. 280. № 5372. P. 2071—2073.
21. Muller P., Sjogren W.L. Mascons: Lunar mass concetrations // Science. 1968. V. 161. № 3848. P. 680—684.
22. Murray N., Holman M. The origin of chaos in the Outer Solar System // Science. 1999. V. 283. № 5409. P. 1877—1881.
23. O'Keef J.D., Ahrens T.J. Cometary impact calculations: flat floors, multirings and central peaks // Lunar and Planet. Sci. XI. Part 2. 1980. P. 830—832.
24. O'Keef J.D., Ahrens T.J. Cometary and meteorite swarm impact on planetary surfaces // J. Geophys. Res. 1982. V. 87. № B8. P. 6668—6680.
25. Petro N.E., Pieters C.M. Reconstructing the stratigraphy of the ancient South Pole – Aitken basin interior // The Moon beyond 2002: next steps in lunar science and exploration. № 1. 2002. P. 47.
26. Pugacheva S.G., Shevchenko V.V., Yakovlev S.G. et al. Development of the calibration method of IR images of the Moon transmitted be AES GOMS // Astronomichesky Vestnik. 1999. V. 33. № 1. P. 29—35. (in Russian)

27. Reedy R. C., Arnold J. R., Lal D. Cosmic-ray record in Solar system matter // Annu. Rev. Nucl. Part. Sci. 1983. V. 33. P. 505—537.
28. The status and future of lunar geoscience // NASA SP-484. Washington. 1984. P. 29.
29. Walker R.M. Interaction of energetic nuclear particles in space with the lunar surface // Annu. Rev. Earth Planet. Sci. 1975. V. 3. P.99—128.
30. Fernandez J.A., Gallardo T. From the Oort cloud to Halley-type comets // Evolution and Source Regions of Asteroids and Comets. Proc. IAU Coll. 173 / Eds. Svoren J. et al. 1999. P. 327—338.
31. Florensky K.P., Basilevsky A.T., Ivanov A.V. The role of exogenous factors for the formation of the lunar surface // The Soviet-American Conference on Cosmochemistry of the Moon and Planet. Eds J.H. Pomeroy and N.J. Hubbard. NASA SP-370, Washington, D.C., Part 2, 1977. P. 571—584.
32. Hiesinger H., Head J.W. III. Lunar South pole – Aitken impact basin: topography and mineralogy // Lunar and Planet. Sci. XXXV. 2004. № 1164.
33. Heiken G.H., Vaniman D.T., French B.M., eds. Lunar Sourcebok, Cambridge Univ. Press. 1995. 736 p.
34. Hood L.L., Coleman P.J., Wilhelms D.E. The Moon: sources of the crustal magnetic anomalies // Science. 1979. V. 204. P. 53—57.
35. Hood L.L., Williams C.R. The lunar swirls: distribution and possible origins // Proc. 19th Lunar and Planet. Science Conf. 1989. P. 99—113.
36. Hood L.L., Yingst A., Mitchell D.L. et al. High-resolution mapping of lunar crustal magnetic fields: correlations with albedo markings of the Reiner gamma class // Workshop on new views of the Moon II. LPI Contribution. 1999. № 980. P. 28—29.
37. Hood L.L., Zuber M.T. Recent refinements in geophysical constraints on lunar original evolution // Origin of the Earth and Moon / Eds. Canup R.M., Righter K. / Univ. Arizona Press. 2000. P. 397—409.
38. Cherkasov I.I., Shvarev V.V. The Lunar soil. Moscow, Nauka, 1975. 142 p. (in Russian)
39. Chapman C.R. Surface properties of asteroids. Massach. Inst. Of Technology, 1972.
40. Shevchenko V.V. The Modern selenography. Moscow, Nauka, 1980. 288 p. (in Russian)
41. Shevchenko V.V. On the cometary origin of the lunar ice // Solar System Research, V. 33. № 5. 1999 P. 400—409.
42. Shevchenko V.V., Pugacheva S.G., Novikov V.V., Kvaratskheliya O.I. Optical and thermal parameters of the lunar surface // Trudy GAISh. V. LXIX. Moscow: «Yanus-K», 2001. 152 p. (in Russian)
43. Shevchenko V.V., Chikmachev V.I., Pugacheva S.G. Structure of the South Pole – Aitken lunar basin // Solar System Research, V. 41. № 6. 2007 P. 447—462.
44. Schmitt H.H. Source and implications of large lunar basin-forming objects // Lunar and Planet. Sci. XXXI. 2001. № 1821.
45. Latham G.V., Dorman H.J., Horvath P. et al. Passive seismic experiments: A summary of current status // Proc. Lunar Planet. Sci. Conf. 9th. 1978. P. 3609-3613.

M.I. PANASYUK
Skobeltsyn Institute
for Nuclear Physics
Lomonosov Moscow State University
Moscow, Russia
panasyuk@sinp.msu.ru

COSMIC RAYS ARE WANDERERS OF THE UNIVERSE

Cosmic rays (CR) were discovered almost 100 years ago. Since then the scientific world has learned a lot of their nature: the particles nascent in the Universe, both in our Galaxy and outside. The basic mechanisms of their acceleration, transfer in interstellar environment and interaction of the primary cosmic rays with the atmosphere surrounding the Earth. Before 1957, i.e. the beginning of the Space Era, researchers' capabilities were limited only by experiments performed on ground, under ground and in near-Earth atmosphere up to flight altitudes of aerostats, airplanes and rockets, i.e. where only secondary radiation is in existence, that is the result of interaction of cosmic rays with the Earth's atmosphere. Launching of spacecraft allowed the scientists to commence on explorations of the Universe primordial matter itself outside the atmosphere, i.e. the primary cosmic rays.

INTRODUCTION

R. Millican is an English scientist who invented the 'cosmic rays' term for high-energy particles filling our Universe, and he was in the wrong. These particles have nothing in common with a light ray that is photons spreading along the straight line from their source. On the contrary, the cosmic particles, traveling in the Universe and, in particular, within our Galaxy, deviate in magnetic fields existing therein. As the result, their movement trajectories take complex forms determined by their impulse (p) and the magnetic field itself (B) (astrophysicists deem that, within Galaxy B, about several microgausses, μGs). In spite of such small magnitude of the magnetic field, it is enough for protons, which are the most common particles of the cosmic rays with energy of up to 10^{17} eV, cannot escape from our Galaxy. It must be emphasized that an important for the cosmic rays conclusion follows from this, i.e. the source and the accelerator of particles of such energies are to be located within our Galaxy, and the source of particles of higher energies must be outside it.

Today we know that the energy range of the cosmic ray particles is immensely wide, i.e. from hundreds megaelectron-volts (1 MeV or 10^6 eV) to, at least, 10^{20}...10^{21} eV. The cosmic ray energy spectrum (see Fig. 1a) is almost straight line in the log-log scale and its spectral index is $\gamma \approx 2.7$. To imagine 'energy scopes' of the cosmic rays, it is enough to compare them with maximum energies of the most powerful of our planet *Tevatron* collider working in the USA and LHC being constructed in CERN in Switzerland. If the maximum energy of colliding in the mass center system protons in the first collider slightly exceeds 10^{15} eV, then in the second one, which will be put into operation in 2008[1], it will be 10^{17} eV. The cosmic rays maximum energy measured by onground instruments is several degrees higher.

1 Put into operation in 2008, but due to the accident delayed until 2009.

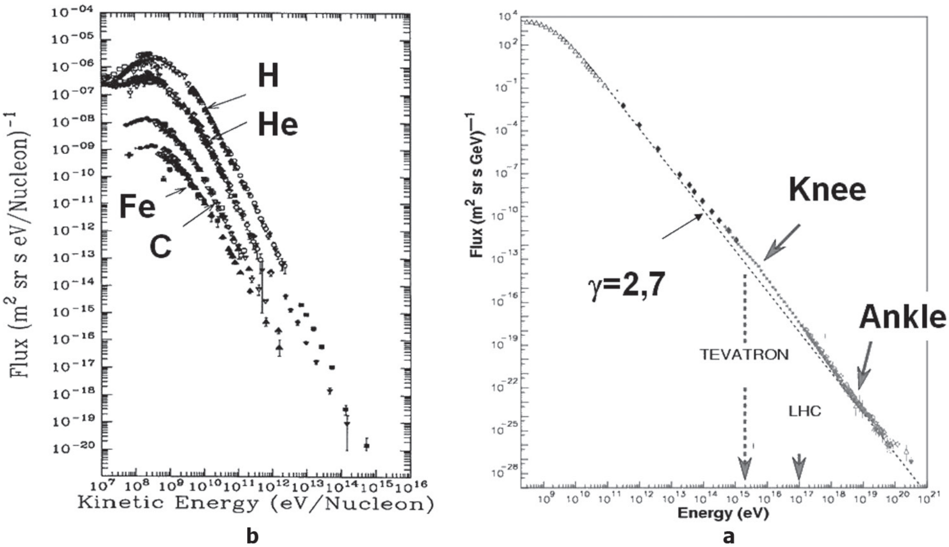


fig. 1. Energy spectra of cosmic rays: (a) spectrum of all particles in the energy range studied and (b) spectrum of individual nuclear components in the low-energy part of the spectrum

Now we know that the cosmic rays composition (see Fig. 1b) is complex. Among the particles of the cosmic radiation, there are nuclei: from protons, which are the most common particles in the Universe, to heavy and super-heavy elements. To a first approximation, the cosmic rays chemical composition reflects average abundance of the elements in our Galaxy. However, it must be stipulated that today we know the cosmic rays chemical composition up to ~10^{14} eV only. When higher energies, it is unknown for us[2].

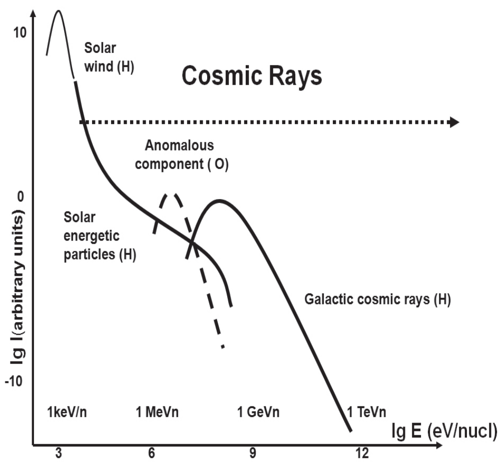


fig. 2. Qualitative representation of energy spectra of cosmic rays (dotted line indicates their energy range) of different origin: solar cosmic rays (protons), anomalous cosmic rays (oxygen) and galactic cosmic rays (protons)

There are several sources of the cosmic rays (see Fig. 2). Depending on the origin, CR subdivide into ‘galactic cosmic rays’ (GCR), the sources of which are within our Galaxy, ‘solar cosmic rays’ (SCR), which are the solar wind accelerated plasma, ‘anomalous cosmic rays’, the sources of which are probably neutrals of local

2 The ‘cosmic rays’, usually, includes a number of elementary particles as well, among which electrons and protons are the most widespread. As for photons, i.e. X-rays and gamma-rays, they are the prerogative of ‘X-ray’ and ‘gamma’ astronomy.

interstellar media outside the Galaxy, and finally, cosmic rays of extragalactic origin (so-called 'tail' of GCR, which is not shown in Fig. 2). The last ones are often called 'ultra high energy cosmic rays' (UHECR) (>10^{18}...10^{19} eV). These particles get into the Earth's atmosphere and cause cascades of secondary particles as we can see later.

However, all these facts are at the level of our up-to-date knowledge, and before 1912, when the cosmic rays were discovered, nothing but excess radioactivity of terrestrial rocks causing air ionization on low altitudes above the Earth was known.

1. BRIEF 'PRE-SATELLITE' HISTORY OF COSMIC RAYS PHYSICS

V. Hess [1], who ascended by a balloon in 1912 (Fig. 3), discovered considerable increase of the air ionization rate at altitudes of more than 1 km. He was the first who supposed that the air ionization at such altitudes could relate to the radiation of cosmic nature. He was honored with the Nobel Prize many years later (in 1926) for this very supposition.

P. Auger [2] (Fig. 4), in 1938, proved by experiments that the air ionization are created by a cascade of secondary particles born as the result of interaction of primary ones, which intrude into the atmosphere from the outer space, with the air atoms.

D.V. Skobeltsyn [3] and G.T. Zatsepin [4] (Fig. 5), in the 30s—40s, developed the mechanism of interaction of the primary radiation with the atmosphere, i.e. the cascade theory, which became the basis of the physics of the cosmic rays in the atmosphere, and in other mediums afterwards. Experimental discovery of the

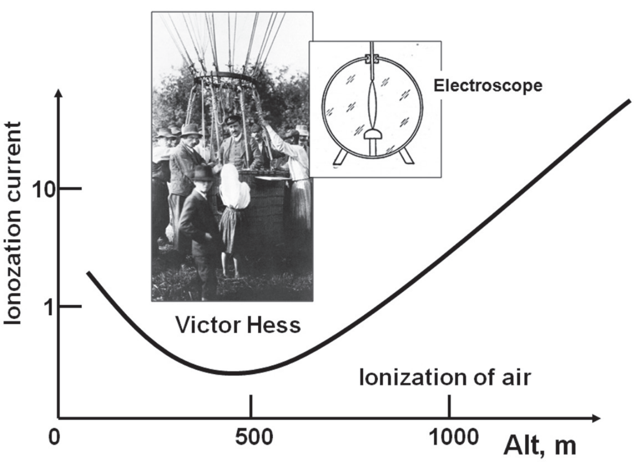


fig. 3. V. Hess has launched an electroscope on the balloon and discovered an excess ionization of air above the Earth. He linked this phenomenon with the influence of particles of cosmic origin over the Earth's atmosphere

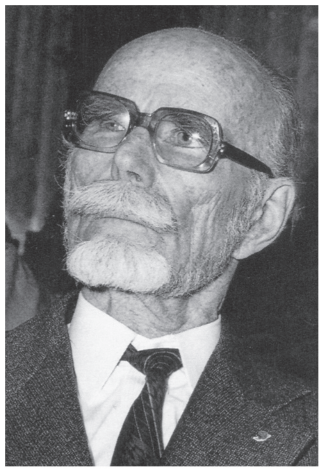


fig. 4. P. Auger has experimentally proved that showers of secondary particles from cosmic rays exist in the atmosphere, which were called extended air showers (EAS)

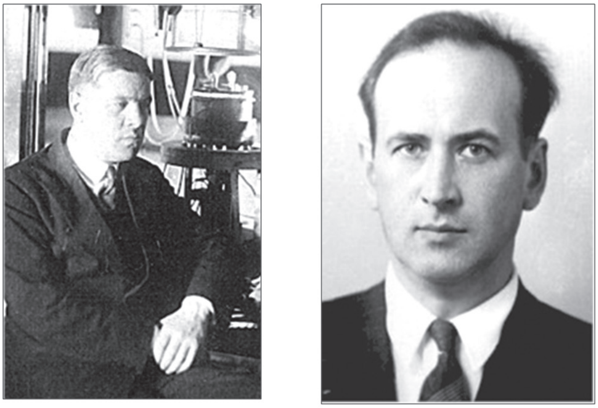


fig. 5. D. Skobeltsyn and G. Zatsepin have developed a theory of extended air showers

cascade process is also connected with D. Skobeltsyn. He 'saw' cosmic-ray showers for the first time when experiments with the Wilson chamber and discovered tracks of two-three particles born simultaneously therein. Per se, they were the first observations of multiple processes in the high-energy nuclear physics [5].

Due to the cascade theory, it became feasible to measure the primary particle energy experimentally per the number of the cascade secondary particles, i.e. so-called 'extended air shower' (EAS).

However, right up to the 40s, the primary cosmic radiation nature and its composition were not known. Understanding that the primary cosmic rays

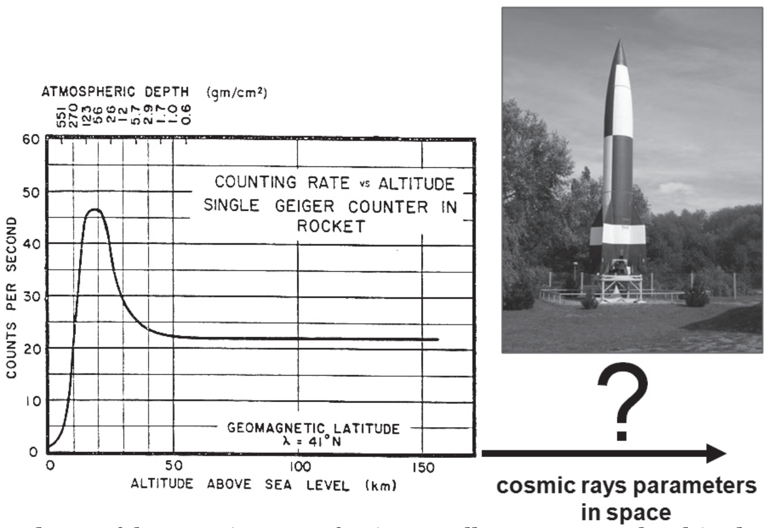


fig. 6. Dependence of the counting rate of Geiger-Muller counter on the altitude, obtained in the rocket flight up to ~150 km altitude. After the maximum observed at the 15…20 km altitude due to the maximum of EAS particle number, plateau is observed. Prior to satellites launched beyond the atmosphere boundaries characteristics of cosmic rays were never known

consist of positive-charged particles mainly, i.e. the protons, became feasible owing to works implemented series of task-oriented experiments on study of the East-West effect related to the particles intrusion into the Earth's magnetic field [6] and also experiments using balloons [7].

After the Second World War, the cosmic rays researches began using rockets at altitudes of more than 100 km. Such experiments were implemented both in the USSR (S.N. Vernov, A.E. Chudakov and others) and in the USA (J. Van Allen, F. Singer and others). One of the examples of the particles changes at high altitudes is shown in Fig. 6 [8]. The altitude dependence of the cosmic rays intensity is clear herein, following the maximum at altitudes of 15...20 km, where there is the greatest number of the EAS secondary particles, and then the intensity decrease occurs when coming out to the plateau.

The rocket-based experiments allowed scientists to approach the atmosphere boundary but not to come out into the outer space where the cosmic rays are to be observed, which have had no time yet to interact with the terrestrial atmosphere. The cosmic rays study outside the atmosphere became feasible due to the Space Era, i.e. the launching of the Earth's First Artificial Satellite in 1957.

2. FIRST SATELLITE EXPERIMENTS

S.N. Vernov and his colleagues from the Moscow University [9] were the first who succeeded to arrange equipment for the cosmic rays study onboard the Second Soviet Satellite launched on November 2, 1957 (Fig. 7).

fig. 7. S.N. Vernov and his colleagues from the Moscow University — N.L. Grigorov, A.E. Chudakov and Yu.I Logachev — succeeded to perform the first physical experiment in space onboard the Second Soviet artificial Earth satellite — cosmic rays measurements using Geiger-Muller counter

Using KS-5 instrument based on the simple Geiger-Muller counter, the first in the world physical experiment in the outer space was performed. The scientists expected to research the spatial distribution of the cosmic rays outside the atmosphere experimentally by means of it.

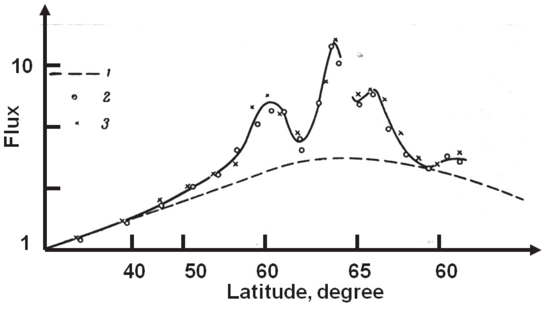
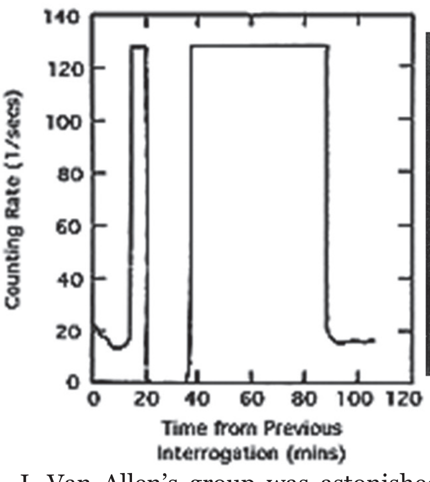


fig. 8. Researches of S.N. Vernov's group performed onboard the Second Soviet satellite resulted in a discovery of bursts of the particle currents, with their amplitudes exceeding the particles' current expected on the base of calculated cosmic rays latitudinal dependence.

The cosmic rays latitudinal dependence (the dashed line in Fig. 8), expected from the calculations, and caused by the nature of their intrusion into the Earth's magnetic field, did not prove true in the first satellite circuit already. There were the counting rate bursts at latitudes of ~60…65^{o}. The first hypothesis to explain the observed deviation became possible SCR from the solar flare, which had occurred just several hours before observing the bursts. It was incorrectly, and it had a dramatic effect onto the first experiment studying CR in the outer space. The scientists faced with an absolutely new natural phenomenon, i.e. energetic charged particles captured into the Earth's magnetic field, i.e. radiation belts (RB). The space exploration pioneers observed so-called 'eruption' of the particles from the RB, however, they could not explain the phenomenon due to lack of information (see the review [10]).



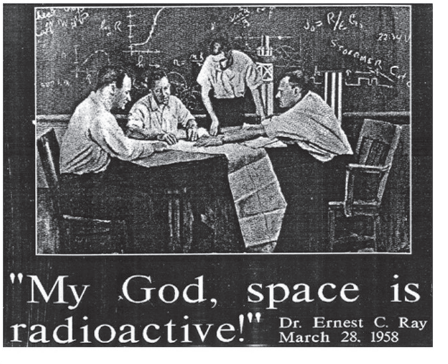


fig. 9. J. Van Allen's group was astonished with the unusually big counting rate of Geiger-Muller counter onboard the artificial Earth satellite Explorer-1. 'My God, space is radioactive!' exclaimed E. Ray, one of J. Van Allen's colleagues

As soon as two months later, American scientists, headed by J. Van Allen, launched the *Explorer 1* satellite. The scientists wondered so much taking readings (Fig. 9) of a saturated Geiger-Muller counter (the same as S.N. Vernov had used), and E. Ray, one of J. Van Allen's colleagues, exclaimed 'My God, space is radioactive!' That was a response to the unexpected high counting rate of the instrument in individual intervals of the satellite orbit [11]. However, the interpretation following this observation was incorrect, i.e. J. Van Allen referred the observed phenomenon to auroral particles causing auroras, i.e. low-energetic radiation intruding from high latitudes to the equator. Perhaps, this idea was dominating for J. Van Allen due to the fact that, during his previous career, he had studied the auroral radiation exactly by means of the rocket experiments.

It was even more so striking that another American scientist F. Singer, a year before the *Explorer-1* satellite launching, had published an article [12], in which he had been directly proving possible existence of the charged particles 'ring currents' in the Earth's magnetic field, i.e. the RB. F. Singer had no luck; he lost a contest to J. Van Allen to place his instrument onboard *Explorer-1* satellite. Otherwise, launched his instrument aboard *Explorer-1*, he could have declared about the captured energetic radiation around the Earth. However, J. Van Allen could not but know about those F. Singer's works!

It is currently known that the RB is natural phenomenon inherent in all planets with magnetic fields and having just mediated relation to the cosmic rays themselves. There are several RB courses. This is, for instance, the solar wind plasma injecting into the magnetosphere and accelerated in it. Along with the solar plasma, the terrestrial ionospheric plasma is also a powerful source of energetic particles, i.e. the RB. The SCR can be injected into the magnetosphere and are partially captured in the magnetic field. But it occurs randomly, i.e. during magnetic storms when there is depression of the Earth's magnetic field allowing relatively low-energetic solar particles to intrude inside the captured radiation belt.

As for high-energetic CR, i.e. anomalous and galactic ones, they are also the RB sources but indirect ones, since their Larmor radius (ρ_L) in the Earth's magnetic field is too long to provide stable capture formulated by H. Alfven [13] $\rho_L/\rho_m<<1$ (ρ_m – magnetic force line curvature radius). Such as the ACR ions (oxygen, neon and others) having minimum charge state of 1+, penetrate into low altitudes freely and can be undergone by charge-exchange process in the atmosphere upper layers. Newly-generated multiple-charged ACR ions, flying out the atmosphere, replenish the RB. A similar mechanism for the GCR is observed.

The GCR nuclei, having the impulse enough to overcome the geomagnetic barrier, can cause nuclear reactions with the atmosphere atoms with generation of secondary particles replenishing the radiation belts. As for the GCR protons, they, interacting with the atmosphere, generate a flux of albedo neutrons, which, flying out the atmosphere, generate energetic protons and electrons, which are then captured by the magnetic field.

The initial stage of the outer space exploration i.e. discovery of the RB, refers to this mechanism of the RB particles generation by means of the CR generating

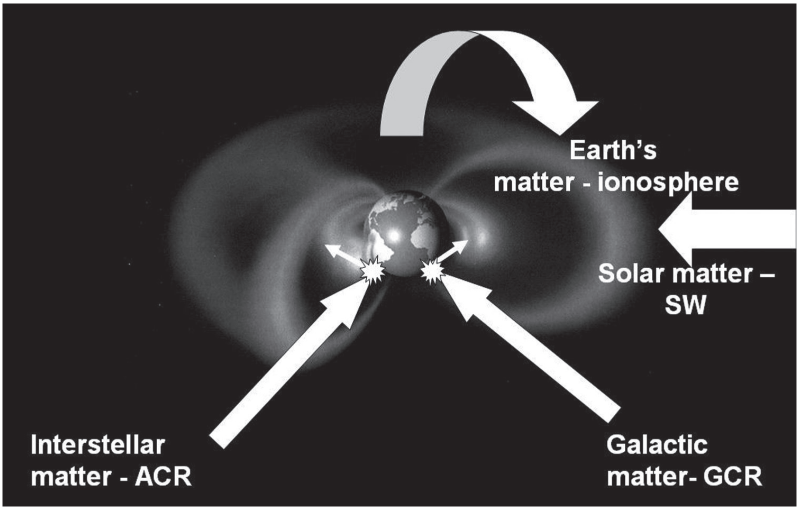


fig. 10. Main sources of the Earth's radiation belts' particles: ionospheric plasma, solar plasma and cosmic rays (galactic and anomalous component)

the albedo neutrons. S.N. Vernov and A.I. Lebedinsky were the first to formulate it [14]. In the USA, F. Singer published a work describing similar mechanism [15]. It happened in July—August, 1958, as soon as several months after the first experiments in outer space. The Cosmic Ray Albedo Neutron Decay (CRAND) model became the final milestone of the first stage of the near-Earth space researches concerned with the RB discovery.

The initial stage of the space researches can be differently evaluated, intending priority problems, however it is currently obvious that two groups of scientists in the USSR and in the USA working independently on each other during a very short period of time, discovered new natural phenomenon, i.e. the RB, and developed the first physical model of their forming.

3. FIRST MEASUREMENTS OF HIGH ENERGY COSMIC RAYS IN OUTER SPACE

Cherenkov integral-type detectors to study chemical compositions of the cosmic rays were installed already onboard the 3rd soviet artificial Earth satellite launched on May 15, 1958 (Fig. 11). Scientists of the P.N. Lebedev Physics Institute headed by V.L. Ginzburg and L.V. Kurnosova initiated those experiments. Then, they proceeded with the researches using similar methods onboard other spacecraft and presented results of one of the first measurements of chemical compositions of the GCR and SCR nuclear components in the space [16, 17]. Prior to proceeding to subsequent experiments on studying the CR in space, I'd like to draw the reader's attention to the following facts.

Year 1957 became actually the striking and significant milestone in the CR physics. One outstanding discovery was already mentioned above. That is the Earth's radiation belts. However, in the same 1957, there was one more discovery

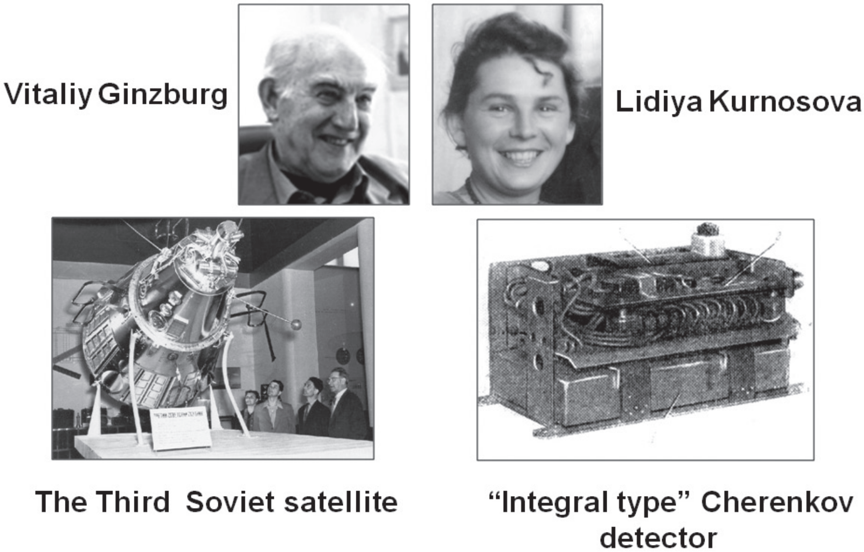


fig. 11. First measurements of high energy cosmic rays in space were performed by the group of Soviet scientists from P.N. Lebedev Physics Institute headed by V.L. Ginzburg and L.V. Kurnosova onboard the Third Soviet satellite using Cherenkov detector

in the CR physics, i.e. a ‘knee’ in the CR energy spectrum at energy of ~$3 \cdot 10^{15}$ eV (see Fig. 1, p. 160). It is imperceptible in the common log-log scale but it becomes apparent in the vertically ‘extended’ scale (Fig. 12). G.B. Christiansen and G.V. Kulikov from the Moscow State University were the first who observed that ‘knee’ [18]. Currently, it is often referred to as ‘Christiansen astrophysical knee’. Attempts to interpret the ‘knee’ at $3 \cdot 10^{15}$ eV played exclusively important role in developing the cosmic ray physics. First of all, it concerned the CR acceleration mechanisms in the Galaxy. It turned out that the ‘knee’ position is close to

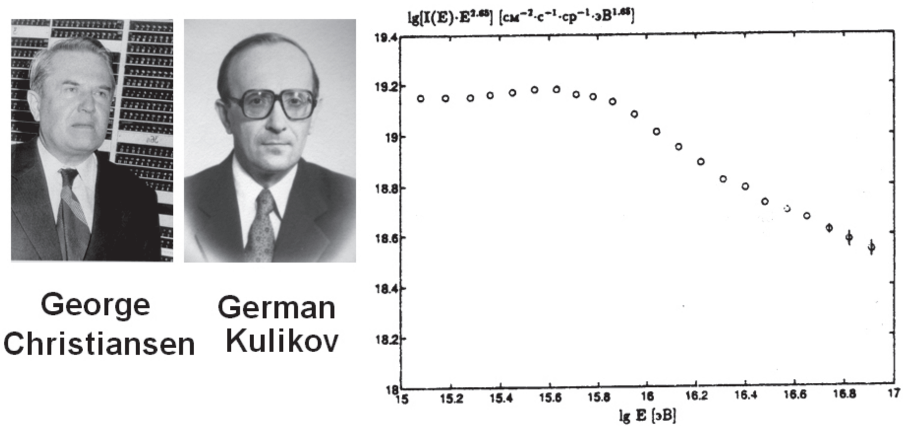


fig. 12. G.B. Christiansen and G.V. Kulikov from Moscow University discovered a sharp curve in the cosmic rays energy range at the energies of ~$3 \cdot 10^{15}$ eV — so-called ‘knee’, that played a major role in the development of cosmic rays astrophysics. Represented on the picture is the spectra from their first paper, devoted to the discovery of the ‘knee’

the upper energy limits of the particle accelerated in remnants of supernovas. Indeed, it can be shown that, when the Fermi-type stochastic acceleration on the shock front of the supernovas (Fig. 13), the accelerated particle maximum energy is E_{max}~BLZ (here L – particle Larmor radius, Z – charge). It follows that when the interstellar magnetic field is B~3 μGs, $E_{max} \approx 10^{14}$ eV. Up-to-date models based on G.F. Krymsky's ideas give $E_{max} \sim 10^{16} ... 10^{17}$ eV but not more. Therefore, the 'knee' area is the key energy area for the cosmic ray astrophysics. To get experimental demonstrations of the standard model [19, 20] in this energy area means to prove predominance of the CR acceleration process in supernovas' explosions comparing with other models. Energetic 'grounds' are more than profound for it. In supernovas explosions up to 10^{52} erg, i.e. 10%, can be released; this energy is enough to accelerate all particles existing in the Galaxy.

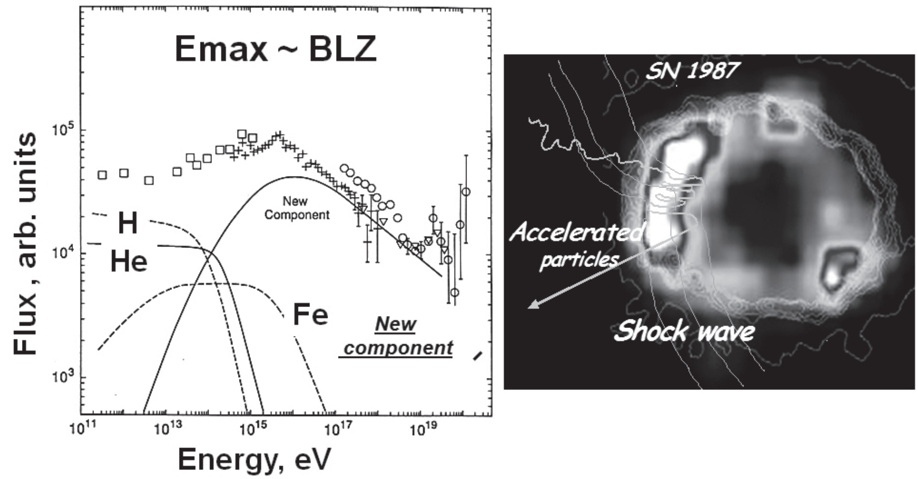


fig. 13. (a) The main point of the 'standard' model of CR acceleration — stochastic acceleration of particles at the shock fronts, generated in the supernovas' explosions. (b) The 'knee' in the energy spectra of single CR

The problem of the experiment in that energy area is that the 'knee' very area is, in a sense, the 'watershed' between earth-based experiments on the one hand, and balloon or space experiments on the other hand. The EAS methods do not allow studying the CR at energies of $<10^{14}$ eV since such energy particles cannot intrude into the atmosphere. Therefore, it is necessary to apply 'direct' methods of the CR measurement installing the equipment aboard balloons and satellites.

However, the difficulty is the method of such particles energy measurement. Scientists N.L. Grigorov, V.S. Murzin and I.D. Rappoport of the Moscow University proposed a fundamentally new method of energy measurement of high-energy particles, i.e. the ionization calorimeter method (Fig. 14) in the same 1957. Currently, this method is widely applied in the physics of high energies both in Earth-based acceleration experiments per the particle physics and the physics of cosmic rays. Ionization calorimeters for CR measurement were produced for the first time for mountain researches and then were launched into the space. The history of the first space experiments with calorimeters is significant.

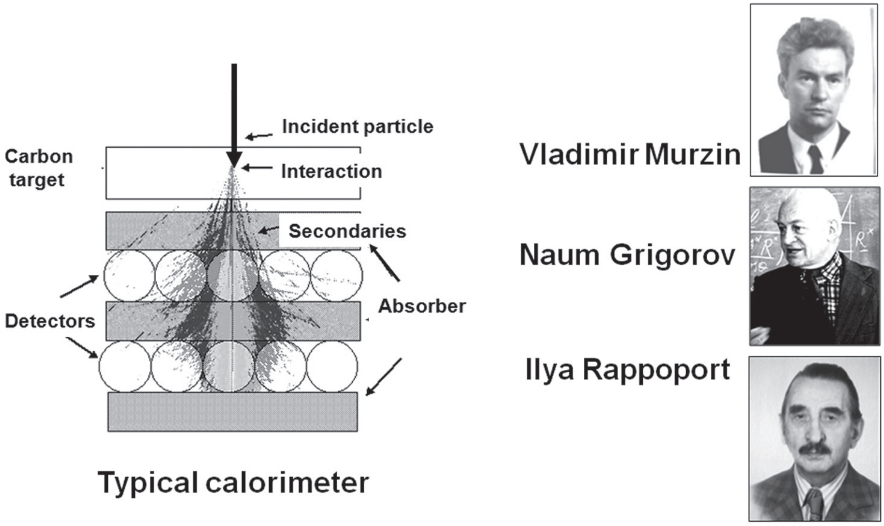


fig. 14. In 1957 N.L. Grigorov, V.S. Murzin and I.D. Rappoport from the Moscow University proposed an idea of and created the first ionization calorimeter — instrument for measurements of cosmic rays (high energy particles) energy

The balloon experiments with their short time flights have upper limits of measured energy because of insufficient statistics. In this regard, satellites with longer existence time have advantages over the balloons. However, large dimensions and, as the consequence, heavy weight of the instruments required to provide good statistics of recorded events, are of fundamental importance for these experiments.

Soviet scientists S.N. Vernov and N.L. Grigorov got the opportunity to implement large-scale experiments with calorimeters in the 60s. They succeeded to arrange scientific equipment, i.e. ionization calorimeters, in place of dimension-weight models of nuclear war-heads when test launching of military ballistic missiles (Fig. 15). There were four launches of rockets with *Proton 1—4* satellites having

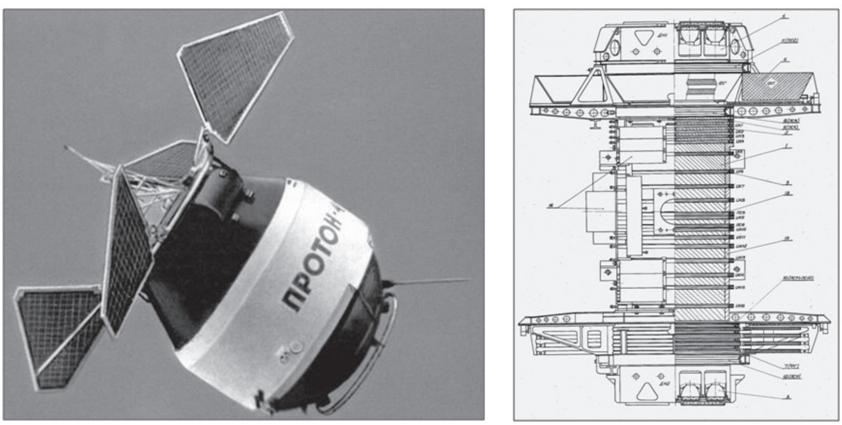


fig. 15. *Proton-4* — the heaviest ionization calorimeter, launched to space in 1968 by the initiative of S.N. Vernov and N.L. Grigorov. Per se, this experiment was the beginning of the 'space stage' of ultra-high cosmic rays researches

ionization calorimeters onboard from 1965 up to 1968. The weight of the heaviest SEZ-14 calorimeter (*Proton-4*) was 12,5 tons! It has been no success to launch such heavy instrument for the CR study so far.

It must be mentioned that a 'blank spot', i.e. the area of energies before the 'knee' not filled with experimental data, remained in the CR energy spectrum right up to 60s. The *Protons*' experiments, for the first time, overlapped the spectrum in the widest range of 10^{11} to 10^{15} eV (Fig. 16). The 'direct' experiment onboard the satellites closed with results of the earth-based EAS units [22]. It was a great success of the national science. The *Protons* measured the spectrum of all particles. However, information about the CR chemical composition in the range of energies before the 'knee' remained limited.

Determination of nucleus chemical compositions in the 'knee' area must play the key role in identification of the specific acceleration mechanism of the particles and their transport in the interstellar medium. The CR composition changes in this area of energies are expected both from models of shock acceleration and from models of spreading, for which E_{max} is proportional to Z. In other words, in this interesting area of energies, the CR composition 'weighting' is possible, i.e. increase of relative content of heavy nuclei over protons. On the other part, when energies after the 'knee', particles are expected, which are probably accelerated in other astrophysical objects (a 'new component'), which are more powerful than the supernovas are.

During years following the *Protons* epoch, three experiments on study of the CR chemical composition with energies before the 'knee' were executed. They were

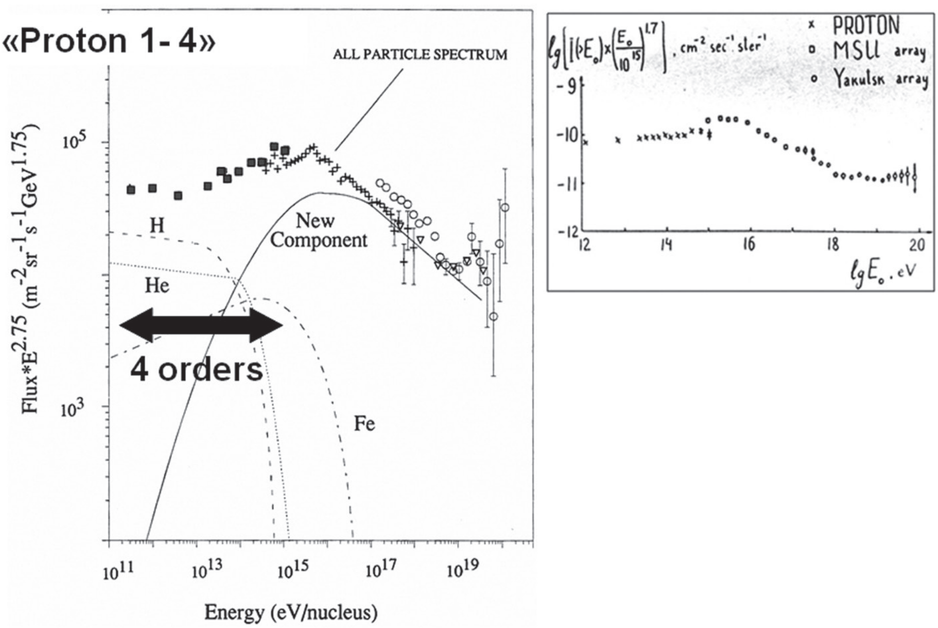


fig. 16. The main result of the Proton's 'space odyssey' that haven't been repeated up to now — measurements of spectra of all CR particles in the wide energy range from 10^{11} to 10^{15} eV

the American CRN onboard the *Spacelab* orbital station in the 80s [23] and soviet *Sokol-1* and *Sokol-2* onboard satellites in the beginning of the 90s [24, 25]. Along with the balloon CR measurements (there have been 9 of them up to date), these experiments have considerably advanced our knowledge about energetic and mass characteristics of the CR (Fig. 17). In spite of divergence of some experiments and limited statistics (especially in the area of energies >10^{13} eV), these data are evidence of a tendency of weighting of the CR chemical composition before the 'knee'. However, the 'knee' area has not been reached yet and information about the chemical composition in this area is severely limited. This conclusion is also correct for results of the EAS CR measurements, the data are also discrepant here.

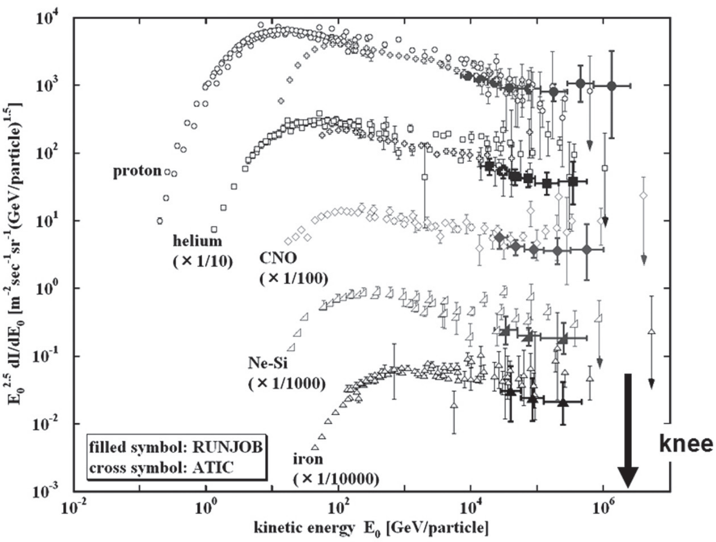


fig. 17. 'Modern' chemical composition of CR with the energies below the 'knee'

Therefore, the future of the CR researches in the 'knee' area is, of course, large-scale space experiments with large geometrical factors to provide statistics and with good resolution per weights to identify wide spectrum of nuclei to iron and further. Because of mass limitations for useful loads of spacecraft, these requirements contradict real capabilities. Therefore, at the present-day stage, the International Space Station (ISS) is the most perspective platform to execute such experiments. Unfortunately, it should be stated that there have been no large astrophysical experiments onboard the ISS so far. The project of large ACCESS calorimeter for the ISS discussed till recently, which could provide 'progression' towards the 'knee', was declined by the NASA. It is quite possible that the CALET Project [26] of the Japan Aerospace Exploration Agency (JAXA) will fill up existing gap, which can execute measurements of both chemical compositions of CR in the concerned area of energies and of the electrons.

It is important here to study the electron component in TeV area of energies (1 TeV = 10^{12} eV) in addition to the nuclear one. The point is that electrons are liable to effective energy losses during the process of transport in the interstellar space filled with magnetic fields. They are the ionization, the bremsstrahlung and the synchrotron radiation, and the Compton scattering. Therefore it is expected that cutting must occur when energies of ~TeV, i.e. the 'knee' is in the electron energetic spectrum. It cannot be asserted yet that the electron spectrum is cut at energies of 1 TeV since the statistics is small here (Fig. 18). It can be noted that the first event related to recording a number of electrons with energy of ~1 TeV was executed in the *Interkosmos-6* experiment [27].

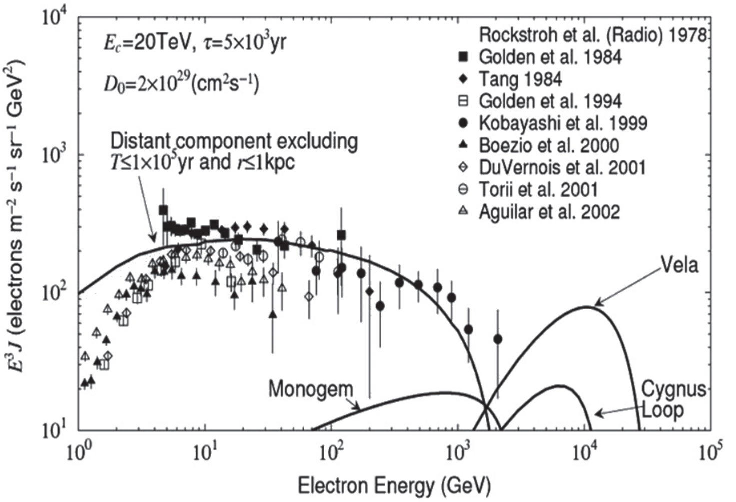


fig. 18. Electrons of cosmic rays. Their spectra in the TeV energy range is not known by now. A 'synchrotron cutoff' is expected to occur in this range in case there are no 'near' sources, located at a relatively small distances

Of course, the probability exists that there is no electron 'cutting' in this energy range. This means absence of sources generating electrons of such energies at distances more than a few kiloparsecs. Indeed, evaluations show that the frequency of occurrence of supernovas stars at distances of less than that distances is very low, i.e. not more than 2...5 per 300 000 years. Therefore, the problem of existence of the 'electron knee' in the TeV area of energies is interesting for the physics of cosmic rays.

In addition to the CALET experiment onboard the ISS, the AMS-02 experiment is expected [28]. The interest to it is connected not so much with the problem of the CR nuclear component chemical composition (the AMS-02 will not be able to reach the 'knee' for several years of measurements) as, to a greater degree, because of the main target of this experiment, i.e. studies of existence of antimatter and dark matter in the Universe.

If the AMS-02 experiment will be able to identify antihelium nuclei with good statistics then it will be the convincing proof of the existence of the antimatter in the Universe. Undoubtedly, this is the 'problem of century'. Antihelium and other antinuclei can be generated as the result of thermonuclear reactions in 'antistars'. Moreover, the antihelium could remain after the Big Bang.

Attempts to discover primary antiprotons[3] rather then products of interaction of the CR with the interstellar medium have not been successful yet. All antiprotons (Fig. 19) recorded up to date can be explained by this physical process, i.e. they are secondary ones.

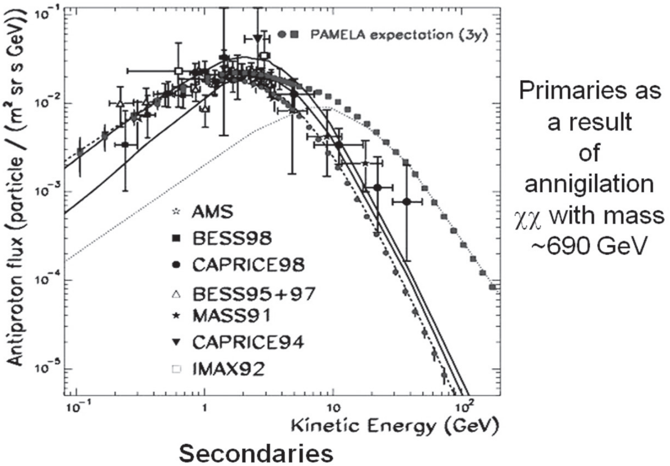


fig. 19. The spectra of antiprotons known. In the low-energy range (up to maximum) it is explained with the help of 'secondary particles' produced due to CR interactions with the interstellar medium. In the high energy range the form of the spectra still remains undefined. Dotted curve — calculated spectra of antiprotons for PAMELA experiment in the assumption of neutralino annihilation.

Researches of the antiprotons in a range of energies more than several GeV give, nevertheless, another important information, i.e. identification of the dark matter. If the hypothesis of weakly interacting massive particles (WIMP) of the dark matter is true, then, as the result of their annihilation, the antiprotons and positrons can be generated. Therefore, changes of the antiprotons and the positrons in the range of energies of tens GeV and over are the key to solve the dark matter nature.

In this regard, the beginning of the RIM-PAMELA Russian-Italian Space Experiment in 2006 [30] passed ahead of the AMS-02 experiment can play a significant role in solving this important problem of the astrophysics. If an 'excess' of the antiprotons, comparing with the secondary component, will be discovered (see Fig. 19), this will become the convincing proof in favor of the WIMP (Weakly Interacting Massive Particles), i.e. the dark matter hypothesis.

3 E. Bogomolov executed the first experiments to discover the antiprotons aboard balloons in 1969—1971[29].

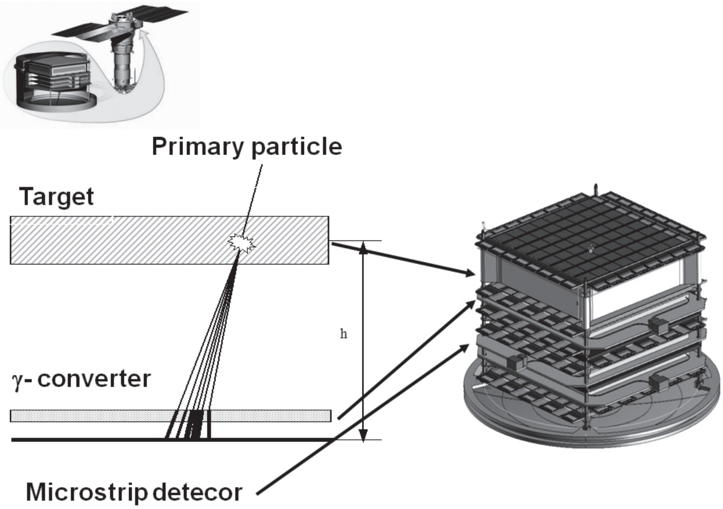


fig. 20. Russian experiment NUCLEON — one more step towards the research of chemical composition of CR with the energies below the 'knee'

The above experiments do not limit future projects of experiments to study the CR of high energy. In the context of the Federal Space Program of Russia, the NUCLEON experiment [31] (Fig. 20) is prepared being based on the energy measurement method different from the conventional ionization calorimeter one that allows to decrease the instrument weight considerably. The chemical composition of the CR basic nuclei right up to the 'knee' will be measured at 5-years exposure of the statistics set.

4. COSMIC RAYS OF ULTRAHIGH ENERGIES

In the magnetic field with intensity of ~3 μGs, the Larmor radiuses for particles with energy of >10^{20} eV are about 1,000 pc. It is more than thickness of the Galaxy's disk. Therefore, it can be asserted that the CR particles of ultrahigh energies are born outside our Galaxy. Undoubtedly, the most important objective of the UHECR physics is to ascertain the existence of objects in the nature, which can accelerate the particles up to such gigantic energies. On the other part, there are limits for the accelerated particles maximum energy Emax determined by the acceleration area volume, the shock velocity and the magnetic field value. A.M. Hillas [32] showed that when known parameters of neutron stars and active galactic nuclei, which are the most powerful astrophysical objects, an evident candidate to be the particles' accelerator up to energies of 10^{20}...10^{21} eV is not 'seen'. Nevertheless, searches in this direction go on, and just here, experimental data are to play the key role to determine the acceleration sources and mechanisms.

There is another exclusively important problem in this area of CR energies, i.e. 'GZK cutoff' of the CR energetic spectrum in this energy range. K. Greizen [33], G. Zatsepin and V. Kuzmin [34] showed in 1966 that interaction of the CR with the relic background with temperature of 2.7 K in the interstellar medium results in generation of new particles (N), and π-mesons, electron-positron pairs and others due to photonuclear processes:

$$\rho+\gamma_{2,7K}\rightarrow N+\pi.$$

Such a process, i.e. disintegration, is to be for heavy nuclei as well, their interaction with the relic background results in the nucleus fission.

As the result, the proton (nucleus) energy decreases rapidly down to a value, below which the generation of new particles becomes impossible. Evaluations performed in the [34] leads to the maximum proton energy to be E_{max}~$5\cdot10^{19}$ eV and d_{max}>50 Mpc, the sources of these particles cannot be beyond this limit. Similar evaluations for the nuclei make E_{max}~10^{20} eV and d_{max}>100 Mpc.

Probably, recently (in 2007) published experimental results of *HiRes* [35] and *Auger* [36] Earth-based facilities 'close' the 'GZK cutoff' problem. The *HiRes* and *Auger* facilities have given close values of energy spectra γ inclination at energies of >$5\cdot10^{15}$ eV, i.e. -5.3 и -4.1 respectively (Fig. 21). These values are the evidence of existence of the 'GZK cutoff' and the truth of the theory of K. Greizen, G. Zatsepin and V. Kusmin in 40 years after the first publications.

However, the CR nature in this area of energies remains unknown. Firstly, due to the fact that the composition of these particles is unknown, and secondly, owing to the fact that the astrophysical objects themselves, which are responsible for generation of the particles in the 'GZK cutoff' area, have not been identified yet.

The space experiments can play the decisive role in identifying the UHECR sources and their composition. Their main advantage over the Earth-based units is considerably lager effective areas of recording of the EAS particles and, as the consequence, providing necessary statistics of events in the energy range where the quantity of the particles is extremely small, i.e. there is only 1 particle per 1 km² per year!

The idea of space experiments on observation of the UHECR goes back to works of A.E. Chudakov [37] who published, in 1959, the experimental method of the

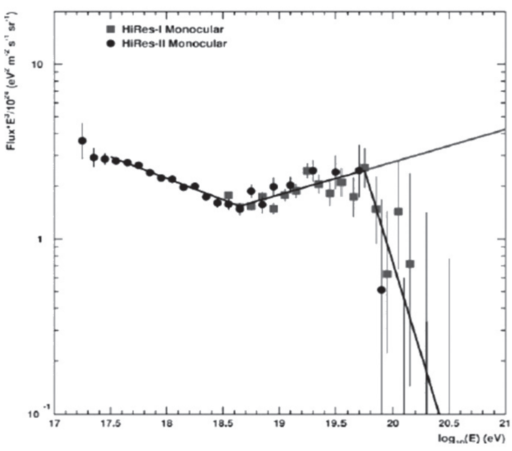


fig. 21. Recent results of ultra-high energy CR spectra research using earth-based facilities HiRes: 'GZK-cutoff' predicted by K. Greizen, G. Zatsepin and V. Kuzmin exists!

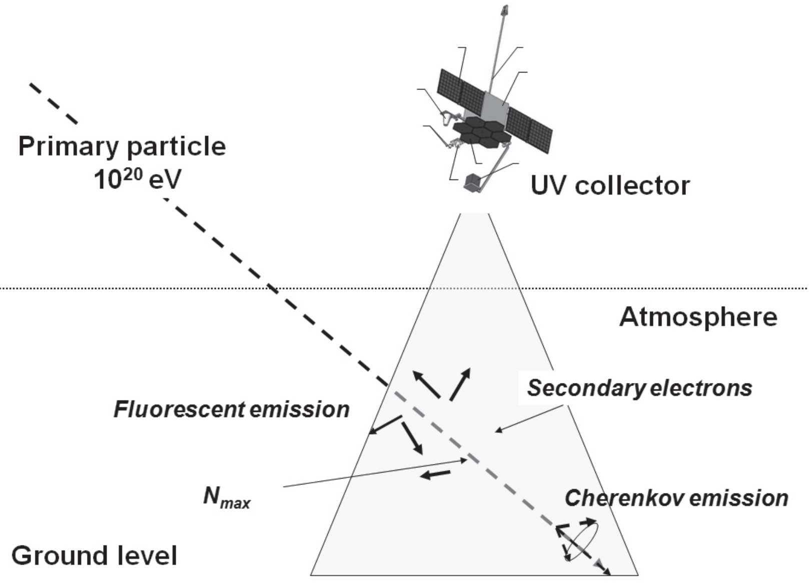


fig. 22. Russian project TUS will become the first space experiment on measurements of UHECR

EAS in the atmosphere by means of measurement of atmospheric scintillations, i.e. ultraviolet flashes in the wavelength range of 300...450 nm. Then, in 1980, J. Linsley proposed the method [38] of measurement of the ultraviolet radiation (UV) of the EAS from the space. His AIRWATCH project became the first among the following propositions; however, it has remained unrealized.

Currently, Russia performs works on the TUS project [39], i.e. a unit based on the paraboloidal mirror-concentrator of the UV-radiation with area of 2 m^2, which will be installed onboard a small spacecraft with the orbit altitude of about 600 km (Fig. 22). The atmosphere view area will be 4,000 km that will exceed dimensions of the currently largest Earth-based *Auger* facility. A mosaic of photoreceivers located in the mirror-concentrator focus will execute recording of effective UV-signals. The TUS unit threshold energy will be $5 \cdot 10^{19}$ eV. The expected commencement of the TUS experiment is in 2011.

Currently, JAXA examines a larger project that is JEM EUSO (*Extreme Ultraviolet Space Observatory*) [42] for the Japanese Module (JEM) aboard the International Space Station. Its capabilities to observe the UHECR will become much more impressive, they will allow to progress in the area of energies of $>10^{20}$ eV.

There is one more project on the UHECR observations, which will be capable to exceed the near-earth space experiments (Fig. 23) that is LORD (Lunar Orbiter Radio Detector). The idea is that the Cherenkov radioemission is recorded, which is generated when the UHECR interacts with the lunar regolith [41]. This project is developed in the context of the Federal Space Program of Russia for the lunar orbital spacecraft. Probably, its implementation will allow to increase statistics in the UHECR area of energies.

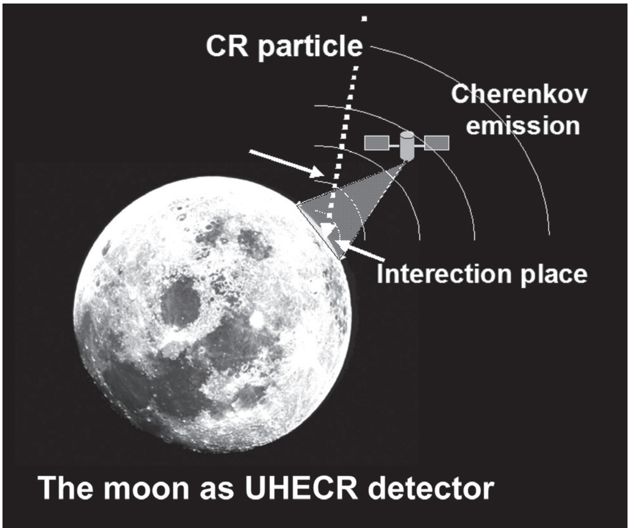


fig. 23. LORD — experiment on UHECR study inboard lunar orbiter basing on Cherenkov emission, originating as a result of CR interaction with lunar regolith

The near-earth space experiments have another important objective in the field of astrophysics that is study of neutrinos of ultrahigh energies. It is expected that, by means of the space experiments, it will be a success to record showers in the night atmosphere coming in at a small angle to the Earth's surface. They can be generated by the neutrinos, which, in contrast to the charged particles, have passed a considerable substance stratum. When $E>10^{20}$ eV, the stream of 'cosmological' τ-neutrinos from such objects as active galactic nuclei or born as a result of the photon-photon collision can be great and make the particles composition outside the GZK cutoff.

5. SOLAR COSMIC RAYS

One of the first observations of the CR in the outer space was executed aboard the Soviet Third Artificial Satellite in May 1958 using scintillation and Cherenkov detectors designed in the Moscow State University (the Skobeltsyn Institute of Nuclear Physics of the Moscow State University) [10] and the Lebedev Institute of Physics of the Academy of Sciences [16]. It happened in 16 years after the first experimental proofs of acceleration of particles on the Sun up to a several GeV had been obtained by S. Forbush in 1942. Since then, the problem of the solar particles acceleration, i.e. the mechanism of the acceleration and its localization, remains the central one for the physics of the SCR.

It is obvious today that the particles of the solar origin can accelerate up to gigantic energies comparable with the GCR energies (up to ~10 GeV) as the result of various in its nature acceleration mechanisms both in the Sun's chromosphere and during their spreading in the interplanetary medium.

If the possibility of the particles acceleration on the Sun is considered, then it, evidently and first of all, is to be connected with energy release during development of the solar flare burst process (up to 10^{42} erg that is a degree less than in the supernova explosion).

The fundamental physical concept of the particles acceleration in the solar flares was founded in S.I. Syrovatsky's works [42] and is connected with generation of electric fields during the process of reconnection of magnetic force lines. Induction electric fields generated when sharp changes of the magnetic field are essentially responsible for the particles acceleration. In essence, this idea is currently the basis of standard model of the particles acceleration in the solar flares existing in various modifications (in combination with the turbulent acceleration and the betatron acceleration in closed magnetic loops in the active area of the Sun's chromosphere for instance, see Fig. 24). The magnetic reconnection acts as the injector in this model.

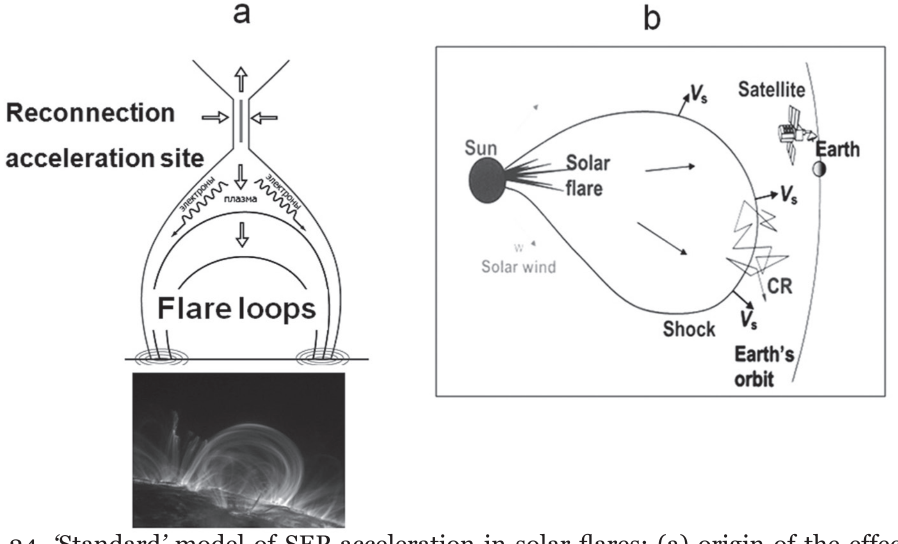


fig. 24. 'Standard' model of SEP acceleration in solar flares: (a) origin of the effect of magnetic reconnection in loop structures connected with the active regions on the Sun and (b) model of stochastic SEP acceleration in shock, connected with CME, propagating in the interplanetary medium

An alterative approach to consider the SCR acceleration process is their acceleration on frontal shock waves of coronal mass ejections (CME) spreading from the Sun in the interplanetary space after flare processes [43]. This scenario, probably, has the same physical approach to the particles acceleration as when acceleration of the GCR in supernova remnants, i.e. the Fermi-type stochastic acceleration on the shock waves.

However, whatsoever the acceleration model would be, it is to explain the experimental facts. It must be noted that main experimental results on the SCR were obtained during the satellite experiments, i.e. after 1957. It was infeasible to get such information about the SCR during the Earth-based experiments since the low-energetic component of the SCR cannot reach the Earth because of its atmosphere and magnetic field.

It should be noted that the experimental data are the evidence of possibility of the electrons acceleration up to energies of ~100 MeV within ~0.5 s and the protons acceleration up to several GeV per 1...10 s. These very values are to be the key ones to validate one or another model of the SCR acceleration. There is no an impression that modern models can explain such 'stiff' experimental data realized in the nature. As for the reconnection models, then their application faces the problem of simultaneous acceleration of both protons and electrons up to the above energies. The electrons energies of about several MeV are seemingly marginal for the reconnection models. On the other part, both the reconnection models and the acceleration models on the CME shock front can accelerate the protons up to hundreds MeV and even several GeV (see [43, 44] for instance, Fig. 24b, p. 177).

As for the SCR electronic component, then convincing proofs have been recently developed that there is generation of the particles with energies of tens MeV during the pion decay process:

$$\pi^{\pm}\rightarrow\mu^{\pm}\rightarrow e^{\pm}+\nu+\tilde{\nu},$$

which are generated as the result of the nuclear reactions in the Sun's atmosphere [45].

In conclusion of this section it can be noticed that further researches of the SCR must be directed to the detailed study of mechanisms of their acceleration in various mediums. This acceleration, probably, happens both in the Sun's chromosphere in the area of maximum energy liberation and during the process of their spreading in the interplanetary space on the CME shocks. The research problem is the necessity to separate the particles transport effects in the near heliosphere, which are capable to distort the energy distribution typical for the source, from sought effects of the acceleration mechanisms themselves. On this account, it is utterly important to observe the SCR onboard spacecraft located at different distances from the Sun.

6. ANOMALOUS COMPONENT OF COSMIC RAYS

The anomalous cosmic rays (ACR) were discovered in 1972 [46] as the result of measurements of the GCR nuclear component with energies of 10...20 MeV/nucl. Their main difference from the 'conventional' GCR is the spectrum form, i.e. the maximum at energies of 10...20 MeV/nucl. for oxygen nucleuses when minimum of the solar activity. Besides, the ACR contain certain elements only (^{16}O, ^{14}N, ^{20}Ne and others first of all), practically no ^{12}C and, that is the most interesting, the charge state of these ions is close to maximum, i.e. 1+, and it makes this component solitary among the cosmic rays of high energies. This very physical characteristic of the ACR served as the basis to search experimental proofs of the model of their origin proposed by L. Fisk et al. [47]. Experimental indications, that the ACR ions are not completely ionized, were published in works of R. Nymmik et al. [48], Biswas et al. [49]. The charge state of 1+ for ^{16}O and ^{20}Ne ions of the ACR was exhaustively proved as the result of joint Russian-American experiments aboard *Cosmos* and IMP-8 series satellites. It served as the corroboration for the hypothesis [47] of the ACR nature. According to this standard model, the interstellar neutrals penetrate into the heliosphere where they are ionized under the action of the solar UV-radiation. Then, these ions,

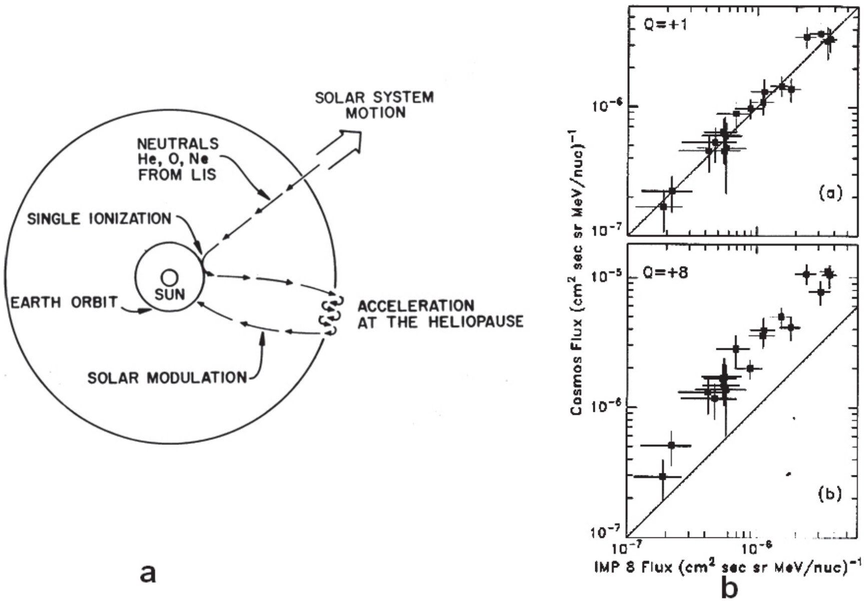


fig. 25. Model of the CR's anomalous component origin, by L. Fisk: interstellar neutrals penetrate into the heliosphere, are ionized by the solar UV-radiation and subsequently accelerated at the shock front — the boundary of heliosphere

entrained by the solar wind, reach the bow shock of the heliosphere where they accelerated under the action of the same Fermi-type stochastic mechanism (see Fig. 25).This very fact was the main one or required to final approval of the model proposed by L. Fisk et al. [47].

The *Voyager-2* interplanetary probe crossed the termination shock wave at the distance of 90...100 AU [50] in 2005. However, the particle spectrometers installed onboard did not record the expected spectrum 'straightening', i.e. its transformation into the power spectrum typical for the Fermi-type mechanism. Further researches including those aboard the *Voyager-1* crossing the heliosphere boundary in other place will probably help to solve the problem, i.e. whether the particles acceleration on the heliosphere termination shock wave exists.

CONCLUSION

The beginning of the Space Era, i.e. launch of the Earth's First Artificial Satellite in 1957, is the beginning of a new stage of the cosmic ray physics as well and research of them by means of spacecraft. This stage of the cosmic ray physics development has brought much new and unexpected in our understanding of the origin of high energy particles in the Universe, the processes of their transport, acceleration and destruction, and their sources. Suffice it to say that, without the space experiments, we could not conceive that the energy spectrum of high energy particles or the most important physical characteristic of them appears as it is presented in Fig. 2. This means that remaining on the Earth we could not approach to the understanding of the cosmic rays nature.

ACKNOWLEDGMENT

The author thanks the symposium organizers for the opportunity to deliver the invited report and to publish it in this book.

REFERENCES

1. Hess V. // Phys. ZS. 1911. V. 12. P. 998.
2. Auger P., Ehrenfest P. // J. de Phys., 1935. V. 6. P. 255.
3. Skobeltsyn D. Kosmicheckie luchi. Moscow. 1936. (in Russian)
4. Zatsepin G. // DAN SSSR. 1949 V. 67. № 6. P. 993. (in Russian)
5. Skobeltsyn D. // ZS, f.Phys., 1927. V. 46. P. 354.
6. Johnson J.H. // Phys. Rev., 1933. V. 44. P. 856.
7. Shein M., Jesse W.P., Wollan E.O. // Phys. Rev., 1941. V. 59. P. 615.
8. Van Allen J.A., Singer S.F. // Phys. Rev., 1950. V. 78. P. 819.
9. Vernov S.N., Grigorov N.L., Logachev Yu.I., Chudakov A.E. // Doklady Akademii Nauk SSSR. 1958. V. 120. № 6. P. 1231. (in Russian)
10. Logachev Yu.I. Issledovaniya kosmosa v NIIYaF MGU. Moscow, 2007. (in Russian)
11. Van Allen. J.A. Transcript of 1958 lecture. // IGY Satellite Rep., № 13. Nat. Acad. Sci, Washington, D.C. 1961.
12. Singer F. // Bull. Amer. Phys. Soc. Series II, 229 (A), 1956.
13. Alfven H. Cosmic Electrodynamics. Oxford Univ. Press, 1950.
14. Vernov S.N., Grigorov N.L., Ivanenko I.P., Lebedinskii A.I., Murzin V.S., Chudakov A.E. // DAN. 1959. V. 124. № 5. P. 1022. (in Russian)
15. Singer F. // Phys. Rev. Letters, 1958. V. 171. P. 1.
16. Kurnosova L.V., Razorenov L.A., Fradkin M.I. // In “Iskusstvenye Sputnki Zemli”. 1958. № 2. P. 70. (in Russian)
17. Ginzburg V.L., Kurnosova L.V., Rasorenov L.A., Fradkin M.I. // UFN. 1964. V. 82. № 4. P. 585. (in Russian)
18. Kulikov G.V., Khristiansen G.B. // J. Exper. Theor. Phys. 1958. V. 35. P. 635.
19. Krymski G.F. // DAN SSSR. 1977. V. 234. P. 1306. (in Russian)
20. Berezhko E.G., Elshin V.K., Ksenofontov L.T. // ZhETF. 1996. V. 109. P. 3. (in Russian)
21. Grigorov N.L., Murzin V.S., Rappoport I.D. // ZhETF. 1958. V. 34. № 2. P. 506. (in Russian)
22. Grigorov N.L., Nesterov V.E., Rappoport I.D., Savenko I.A. Space Research XII. Academic Verlag, Berlin. 1972.
23. Swordy S. et al. // Nucl. Instrum. Methods, 1982. V. 193. P. 591.
24. Ivanenko I.P. et al. // 23d ICRC, 1993. V. 2. P. 17.
25. Grigorov N.L. // Yad. Physica. 1990. V. 51. P. 157. (in Russian)
26. Torii S. Advances in Polar Upper Atmospheric Research, 2006. V. 20. P. 51.
27. Nymmik R.A. // Cosmic Research (Kosmicheskie Issledovania), 1981. V. 19. № 4. P. 233.

28. AMS collaboration // Nuclear Inst. and Methods A, 2005. V. 5. P. 251.
29. Bogomolov B.A. et al. // Proc. 20th Intern. Cosmic Ray Conf., Moscow, 1987. V. 2. P. 72.
30. Piccardi S. for the Pamela Collaboration // Proc. 26th Intern. Cosmic Ray Conf., 1999.
31. Podorozhny D.M. et al. // Proc. 26th ICRC, 2001. P. 2188.
32. Hillas A.M. // Ann. Rev. Astron. Atrophys., 1984. V. 22. P. 425.
33. Greizen K. // Phys. Rev. Letters, 1966. V. 16. P. 748.
34. Zatsepin G.T., Kuzmin V.A. // ZhETF Letters. 1966. V. 4. P. 78. (in Russian)
35. HiRes collaboration // Phys. Rev. Lett. 2004. V. 92. P. 151101.
36. Cronin J. et al. // Proc. 30th ICRC, 2007.
37. Chudakov A.E. et al. // Proc. 6th ICRC, 1959. V. 2. P. 46.
Belyaev V.A., Chudakov A.E. // Izvestiya Akademii Nauk SSSR. 1966. V. 30. № 10. P. 1700. (in Russian)
38. Linsley J. // Proc. 26th ICRC, 1999. V. 2. P. 423. Preprint, Leeds, 1979.
39. Abrashkin V. et al. // Adv. Space Res., 2006. V. 37. P. 1876.
40. Ebisuzaki T. et al. // Proc. 30th ICRC, 2007.
41. Kotelnikov K.A., et al. // Bull Russ. Acad. Sci. Phys., 2002. V. 66. P. 1810. (in Russian)
42. Syrovatsky S.I. // ZhETF, 1971. V. 33. P. 933. (in Russian)
Syrovatsky S.I. // Solar Phys. 1982. V. 76. P. 3. (in Russian)
43. Berezhko G., Taneev C.N. // Pisma v Astronomicheskii Zhurnal. 2003. V. 29. № 8. P. 601. (in Russian)
44. Somov B.V. // Phys. Proc. in Solar Flares. Dordrecht: Kluwer Academic, 1992. V. 121. P. 77.
45. Kuznetsov S.N. et al. // Int. J. of Modern Physics A., 2005.
46. Magro-Campero A., Simpson J.A. // Phys. Rev. Lett., 1970. V. 25. P. 1631.
47. Fisk L.A. et al. // Astrophys. J. Lett., 1974. V. 190. P. L35.
48. Nymmik R.A. Препринт НИИЯФ МГУ N87-06. 1987.
49. Biswas S. et al. // Space Sci. Rev. 1993. V. 62. P. 3.
50. Stone E. et al. // COSPAR proceedings (in press). Bejging, 2006.

V.E. FORTOV
Joint Institute of High Temperatures,
Russian Academy of Sciences,
Moscow, Russia
fortov@ihed.ras.ru

DUST PLASMA CRYSTALS AND LIQUIDS ON THE EARTH AND IN SPACE

INTRODUCTION

Plasma with fine particles placed in it or dust plasma is characterized by such particles can be charged by streams of electrons and plasma ions and by means of photoemission, thermoemission and secondary emission of electrons [1, 2]. The emission of electrons from surfaces of the particles can result in their positive electric charge, at that the particles emitting electrons can increase concentration of the electrons in gaseous phase and its electroconductivity. If the particles capture the electrons, then their charge is negative and the adverse effect, i.e. decrease of the concentration of the electrons, can arise.

By virtue of high great charge of dust particles (about 10^2...10^5 electron charges), potential energy of electrostatic interaction between them, which is proportional to the product of the interacting particles' charges, can exceed their average thermal energy considerably, that means the strongly coupled (non-ideal) plasma. Theoretical calculations of equilibrium properties of such plasma show that, on certain conditions, strong interparticle interaction results in liquid-solid phase transfer and spatially ordered structures of dust particles similar to those of liquids or solids [4, 5]. Such structures are called 'Coulomb or plasma crystal'.

At present, the dust plasma is actively researched in laboratory environment. The dust particles can be introduced into the plasma not only purposely but also be generated spontaneously as a consequence of various processes. The dust plasma is widespread in the space. It is present in planet rings, comet tails, interplanetary clouds and interstellar clouds. The dust plasma has been discovered near the Earth's artificial satellites and space vehicles, and in thermonuclear units with magnetic retention. Widespread of dust plasma systems and also a number of their unique and extraordinary properties (the capability for self-organization and formation of ordered structures, the system openness, the simplicity of recovery, observation and control, and the possibility to research at the kinetic level) make the dust plasma an extremely attractive and interesting matter to explore.

Among the first crystal-like structures experimental observations are of iron and aluminium charged particles of micron sizes retained by certain configurations of alternating and static electric fields. Coulomb crystals for atom ions were also discovered in traps of other types, e.g. in Penning traps. Subsequently, experimental works appeared [6, 7], which described observations for the dust particles Coulomb crystallization with strong interaction in weakly ionized plasma of high-frequency discharge at low pressure. Ordered particle structures were explored in the thermal plasma at atmosphere pressure in the positive glow gap of the direct current and in the nucleus-induced dusty plasma [8—10]. It should be noted that the increase of interest in the dust plasma in the beginning of the 90s related, first of all, to discovery of plasma-crystal structures. It gave incentive to rapid development of investigations in this area, which is currently going on (it is sufficient to indicate exponential increase of the number of publications on this research area).

The Joint Institute for High Temperatures (JIHT) of the Russian Academy of Sciences (RAN) has been studying the dust plasma and developing various methods of its diagnostics during the last one and half decades. For this period, JIHT has implemented a series of experimental works on studying the plasma-crystal structures in the glow-discharge plasma of direct current and the high-frequency discharge at low pressure, and in the thermal plasma (at temperatures of 2,000 to 3,000 °K) and the cryogenic plasma (at temperatures of 4 to 7 °K). The nuclear-induced dust plasma and the dusty plasma generated by proton beams are intensively explored jointly with the State Scientific Center of the Russian Federation – Institute for Physics and Power Engineering named after A.I. Leypunsky (the city of Obninsk). JIHT also runs experiments on the dust plasma when actions of the electron bunch and the ultraviolet radiation. JIHT implemented pioneer experiments on study of plasma-dust structures in microgravitation conditions aboard the *Mir* space station in 1998 jointly with S.P. Korolev Rocket and Space Corporation *Energia*, and *Plasma Crystal* Joint Russian-German Experiment is actively running at the International Space Station (ISS) since the beginning of 2001, which has been prepared by JIHT of RAN with the assistance of the Institute of Extraterrestrial Physics of the M. Plank Society (Germany) and S.P. Korolev Space Rocket Corporation *Energia*.

1. STRUCTURAL PROPERTIES OF DUST PLASMA LIQUID

Having a number of unique properties, the dust plasma is a good experimental model both for studying the greatly non-ideal plasma and for deeper understanding the self-organization phenomena in the nature. So experimental investigations of the dust plasma can play an important role in verifying existent and developing new phenomenological models for liquids since, due to strong interparticle interaction, the liquid theory does not have the small parameter, which can be used for analytical description of the state and the thermodynamic characteristics of the liquid as it is possible in case of gases.

As is known, the liquid physical properties such as pressure, energy density and compressibility are fully determined by a pair correlation function g_2 ($|r_1$-$r_2|$) characterizing probability of a particle being at a distance $r = |r_1$-$r_2|$ from the considered particle and, thereby, the particle spatial arrangement, i.e. chaotic or ordered. Information about the three-particle correlation function, which determine probability of simultaneous detection of three particles near points r_1, r_2 and r_3, is important for calculating such physical characteristics of the environment as entropy, thermal expansion coefficients, etc.

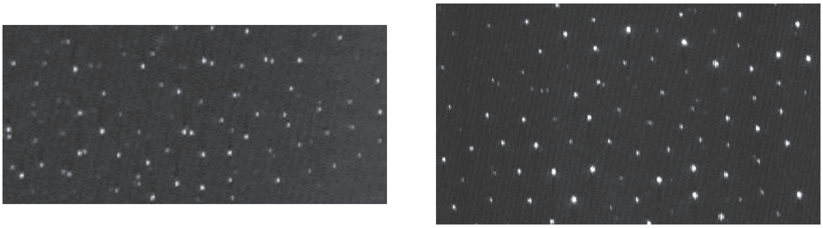


fig. 1. Videoimage of dust cloud particles in the near-electrode layer of the radio-frequency discharge for different experiments: (a) – P=5 Pa, W =9 W; (b) – P=7 Pa, W =10 W

For real liquids and gases, direct determination of the three-particle correlation function is impossible without information about coordinates of individual particles. For analyzing the three-particle correlation in liquids, indirect diagnostics methods are used, e.g. measurement of the structure factor for several pressures of the environment at its constant temperature. Dust particles in the dust plasma can be recorded with a video camera, therefore the dust plasma is a good experimental model to study physical properties of such non-ideal as liquids.

One of important correlations in the liquid theory is the so-called superposition approximation or Kirkwood correlation, which interconnects triple and binary correlation functions: $g_3(r_1,r_2,r_3) \approx g_3^{сп}(r_1,r_2,r_3) = g(r_1-r_2)\, g(r_2-r_3)\, g(r_3- r_1)$. This correlation is often used when calculating integral equations in the kinetics of interacting particles and also when recovering potentials of the interparticle interaction by various methods.

The dusty plasma allowed to perform experimental check of the superposition approximation for the first time. In experiments performed in JIHT, the dusty

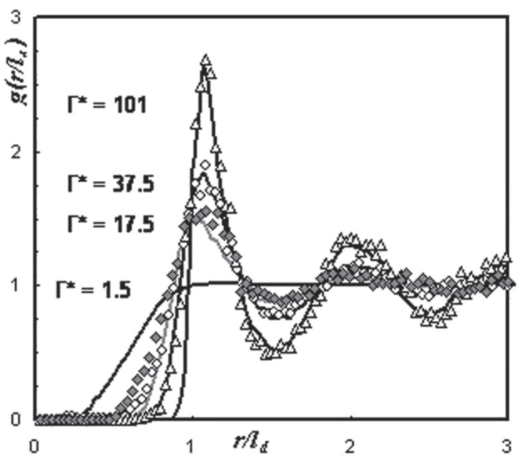


fig. 2. Binary correlation functions g(r/ld) measured in the experiments: (◇) – P=5 Pa, W =9 W; (o) – P=3 Pa, W =2 W; (Δ) – P=7 Pa, W =10 W, — and obtained as a result of numerical simulations for various values of coupling parameter Γ* (firm lines), indicated on the picture

plasma was explored in the argon high-frequency gas discharge at the pressure of 2 to 10 Pa. Monodisperse particles of micron sizes were used as the dust component. For the diagnostics, the dust cloud was illuminated with a flat laser beam and recorded with a video camera. Fig. 1 shows video images of the dust particles in the illuminated dust layer. The dust structures being seen were systems of the liquid type with the average interparticle distance l_d in the range of 260 to 350 micron [11].

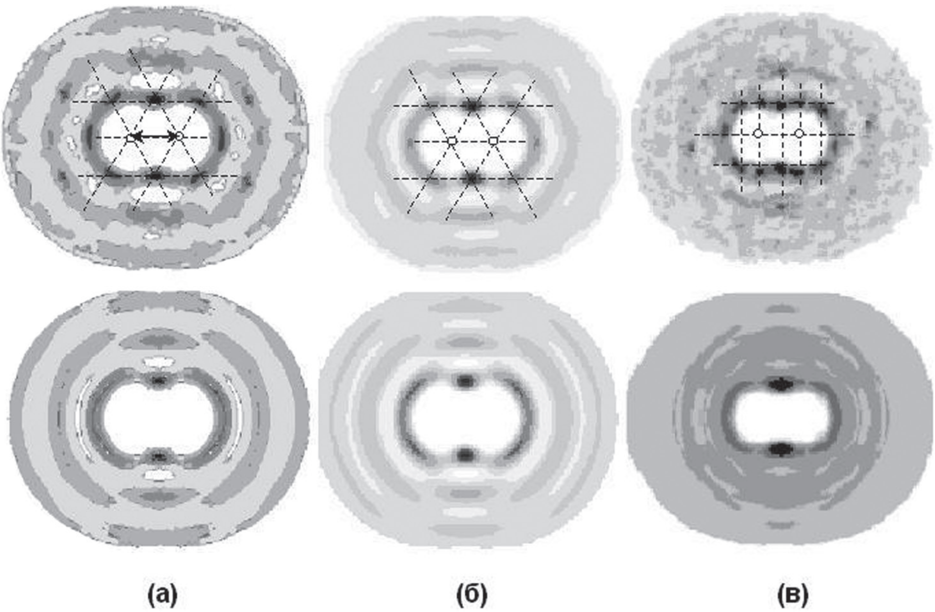


fig. 3. Illustration of the measured three-particle correlation functions of g_3 (upper row) and superposition approximation $g_3^{сп}$ (lower row) for different experiments: (a) P=7 Pa, W=10 W, δ=0.61; (b) – P=3 Pa, W=2 W, δ=0.28; (c) – P=5 Pa, W=9 W, δ=0.3

Pair $g(r/l_d)$ and three-particle $g_3(r_{12},r_{23},r_{31})$ correlation functions were got as the result of the video-record processing. The pair correlation functions $g(r/l_d)$ for various gas pressure P and discharge power W values are given in Fig. 2. Sections of three-particle correlation functions $g_3(r_{12},r_{23},r_{31})$ at the fixed value r_{12} equals to the most probable interparticle distance r^d_{max} for the same discharge parameters are given in Fig. 3. The results of the three-particle function $g_3^{сп}(r_{12},r_{23},r_{31})$ calculation within the superposition approximation are shown in the same figure. To present data of functions ($g_3(r_{12},r_{23},r_{31})$ and $g_3^{сп}(r_{12},r_{23},r_{31})$) in pictorial 'two-dimensional' form convenient for comparison, they were normalized by the maximum value of $g_3(r_{12},r_{23},r_{31})$: the black color corresponds to the one and the white color corresponds to the zero ($g_3 = g_3^{сп} = 0$). In the experiment conditions, the mean-square deviation δ of the function $g_3^{сп}(r_{12},r_{23},r_{31})$ from the calculation results of $g_3(r_{12},r_{23},r_{31})$ and thus the deviation from the superposition approximation was in the range of 30 to 60%.

2. NONLINEAR WAVES IN DUST PLASMA STRUCTURES

New low-frequency oscillations (the dust sound) and instability resulting in self-oscillation of running nonlinear waves were experimentally discovered practically at the same time with the dust plasma crystallization. It highly interested in theoretical study of wave processes in the dusty plasma. The study of the nonlinear waves in the dusty plasma in laboratory conditions continues to be of extremely great interest today. So there were Mach cones in the dusty plasma, which are generated in the environment by an object moving with the supersonic speed; propagation of nonlinear compression impulses over a two-dimensional plasma crystal and the shock wave in microgravitation conditions [12] were studied.

JIHT pioneered to get high-amplitude waves in the direct current glow discharge dust plasma during the experiments [13]. The experiment unit is a discharge tube filled with neon at low pressure, in which the direct current glow discharge with standing strata is generated, i.e. nonuniform luminosity immovable zones regularly alternating with dark gaps having the typical scale about a few centimeters [9, 10]. The electron concentration, their energy distribution and the electric field is highly nonuniform along the stratum length. The electric field is relatively large in the stratum head (the luminous part) and small outside this area. The discharge tube walls also have a high floating potential. Thus, each stratum head has an electric trap, which is capable to capture fine particles in the area of the discharge positive column when the discharge tube is in the vertical position, and strong radial field prevents their precipitation on the discharge tube walls. The dust particles are entered into the gas discharge plasma, are charged and retained by the stratum electric field against the gravity force. Levitating dust particles are illuminated with a laser sheet, and the scattered light is recorded by video camera with shooting rate of 1,000 shot per second.

The particles of several types were introduced into the discharge positive column that are 50 to 60 µm in diameter borosilicate glass hollow microspheres and 3 to 5 µm sized Al_2O_3 polydisperse particles and 1.87 µm in diameter melamine formaldehyde monodisperse particles. The particles were seen as a cloud in the stratum center. There are usually several particle clouds (three-dimensional structures) in adjacent strata. The cloud diameters were 5 to 10 mm (for the glass microspheres) and increased up to 20 mm for Al_2O_3 particles. The particles located in 10 to 20 (for the glass microspheres) and more (for Al_2O_3 particles) flat layers. The particles formed chains ('one-dimensional' structures). The range of distances between the layers was 250 to 400 µm, the range of distances between the particles in the horizontal plane was 350 to 600 µm that corresponds to the particle concentration of ~10^3 to 10^4 cm^{-3} [10].

Fig. 4a shows an image of the horizontal section of the dusty plasma crystal composed of melamine formaldehyde monodisperse particles 1.87 µm in diameter in neon-hydrogen mixture discharge at 0.8 Pa pressure and 1.1 mA current. Fig. 4b shows the particle distribution function g(r) in the distance r from a particle, which confirms long range ordering in the particle arrangement that means forming of the crystal structure. Minimum interparticle distance in the structure was about 315 µm.

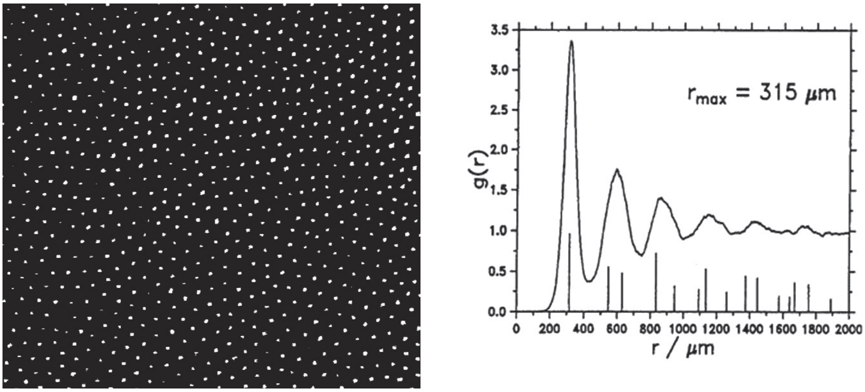


fig. 4. Image of a horizontal section of a ordered dust structure in the glow discharge positive column (a) and pair correlation function g(r) for dust structure (b)

The particle cloud shape could be changed by varying the discharge parameters (pressure and current). So decrease of the discharge current and the discharge pressure results in the fusion of two nearest elliptic clouds into a cylindrical structure, the vertical dimension of which is several tens of centimeters. There can be transfer from the crystalline state into liquid and then into gas, i.e. the crystal 'melting' occurs.

In case of small particles, the increase of their number when certain discharge parameters results in forming of structures where there are various areas: areas of strong ordering (plasma crystals), areas with convective and oscillatory movement of the particles (dusty plasma liquid). At that, as a rule, in the lower part of the structure, there is the oscillatory movement of the particles in vertical direction (the particle density wave) with the frequency of 25 to 30 Hz and the wavelength of about 1 mm at average distance of 200 μm between the particles. Self-exciting oscillations of such kind can correspond to instability of the dust sound oscillations.

The magnetic field is used to excite waves, which is generated by a current impulse in a flat coil, which is coiled around a tube through a ballast resistor (Fig. 5). When enabling the magnetic field impulse, the stratum goes upward in a time less than the one-shot duration, i.e. 1 ms. The levitating dust particles

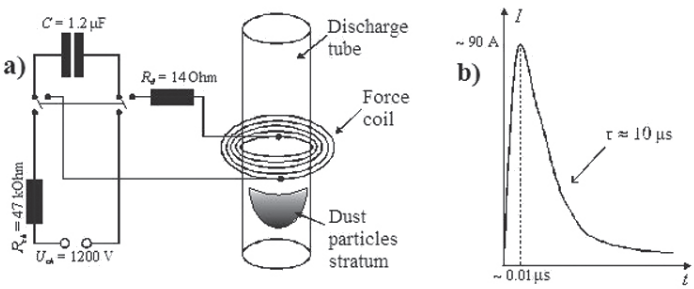


fig. 5. Experiment layout (a) and coil current pulse shape (b), τ – characteristic time of decay

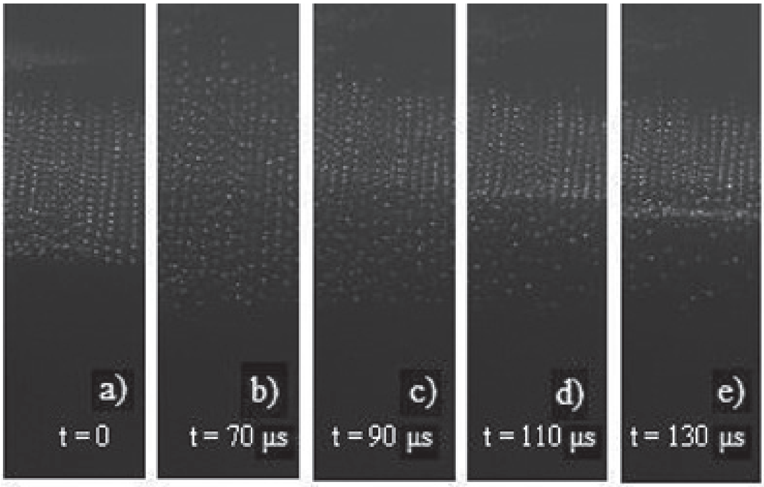


fig. 6. Series of images depicting the shock front development in the dusty plasmas. Frame width 5.8 mm

can not move together with the stratum. Consequently, the dust particles lose the balance and start to fall. The stratum returns and catches up the falling dust particles as the current decreases. Since the strong electric field area reaches the lower particles later than the upper ones, the dust plasma structure stretches out. The upper particles take initial configuration quickly, while there is underpressure in the lower part for quite a long time, and a sharp boundary is formed between areas of different density of the dust particles (a nonlinear wave front) (Fig. 6). The brightness sequential diagrams show the forefront steepening (Fig. 7).

The particles velocity in the underpressure area is directed upward and is on average 1.3 cm/s, the front itself moves at velocity of 2.5 cm/s in the laboratory reference system, and hence, the velocity of relative movement of the front and the environment is 3.8 cm/s. Evaluation of the dust acoustic velocity for the underpressure area results in values of 0.5 to 1.5 cm/s, thus the disturbance is supersonic.

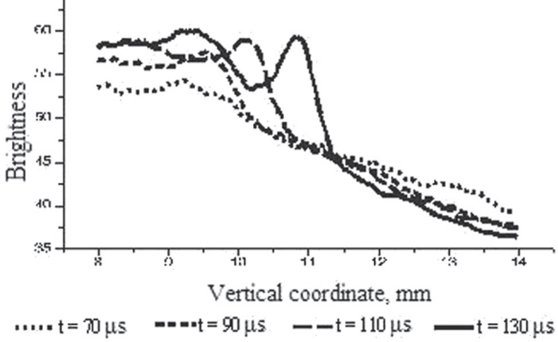


fig. 7. Brightness spatial profiles for videoframes on the Fig. 5 (p. 187)

So it was successful to fix the density disruption in the dust component, the supersonic velocity of the disturbance propagation and the steepening of the wave front. These characteristics are evidence that there was forming of the shock wave in the dust component.

3. DUST PLASMA IN MICROGRAVITATION CONDITIONS

In many cases, the gravity force limits capabilities of laboratory experiments essentially in earth-based conditions, therefore experimental investigations of the dusty plasma in the microgravitation conditions are significantly focused upon recently. Such experiments allow to study a wide scope of phenomena (photoemissive charging of atmospheric aerosol, ambipolar diffusion, dynamics of massive dust particles having sizes of more than 100 μm, etc.), which can not be observed in laboratories on the Earth. One of considerable advantages of the microgravitation experiments is the capability of exploration in a wide range of the dusty plasma parameters, which is not limited by conditions providing the particle levitation in the gravity field. Such experiments have been recently performed on several plasma types. JIHT implemented pioneer experiments on study of plasma-dust structures in microgravitation conditions aboard the *Mir* Space Station in 1998 jointly with S.P. Korolev Rocket and Space Corporation *Energia*, and *Plasma Crystal* Joint Russian-German Experiment is actively running at the International Space Station since the beginning of 2001, which has been prepared by JIHT of RAN with the assistance of the Institute of Extraterrestrial Physics of the M. Plank Society (Germany) and S.P. Korolev Rocket and Space Corporation *Energia*.

3.1. DIFFUSION OF DUST PARTICLES IN PLASMA INDUCTED BY UV RADIATION (EXPERIMENTS ABOARD 'MIR' SPACE STATION).

One of mechanisms of the dust charging in the space is photoemission. The dust particles can get positive electric charges of about 102 to 105 electron charge and form crystalline, liquidal and gaseous dust structures. The phase states of these structures are tightly bound up with the diffusion, which is one of main sources of power losses in the dusty plasma. Both suspended dust particles of a substance in buffer gases (Brownian motion) and gas molecules themselves or particles of the plasma component (self-diffusion) can diffuse. For dusty plasma clouds consisting of charged dust particles, ions and electrons, joint diffusion of unlike charged particles (ambipolar diffusion) can take place. The works [14, 15] presents the first exploration results of such phenomenon for the dust particles plasma induced by the solar radiation and photoelectrons emitted by these particles.

The experiments were run in microgravitation conditions aboard the *Mir* Space Station with bronze particles (covered with a caesium monolayer) having average size $2a \approx 150$ μm in neon at pressure $P \approx 5.3$ kPa. The dust particles were in a glass cylinder, one end of which was the uviol window designed for exposing the dust cloud by the solar radiation. A video camera recorded the image. Fig. 8 shows the evolution of the pair correlation function for the illuminated dust cloud.

The diffusion coefficient behavior during initial observations was similar to the diffusion coefficient behavior of non-interacting particles, the values of measured thermal diffusion coefficients of the dust particles were about $D_o^x \approx 1{,}4 \cdot 10^{-5}$ cm²/s and $D_o^y \approx 6.2 \cdot 10^{-6}$ cm²/s. In a cloud transparent for photoelectrons, determination of the charges Z_d of the dust particle can be based on the analysis of change of their concentration $n_d(t)$ by time [14], from where such charge can be get as: $Z_d = (4.3 \pm 0.2) \cdot 10^4$ electron charges.

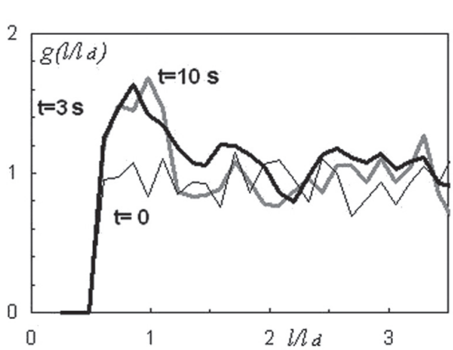


fig. 8. Pair correlation functions $g(l/l_d)$ for illuminated dust cloud at various observational moments t

In the course of these experiments, it was discovered that the concentration of the dust particles exposed by the solar radiation is 3 to 5 times higher than their concentration recorded in the dark. This means the charges-in-plasma polarization, which hinders withdrawal of the charged dust particles onto the cylinder walls and their precipitation on there. The polarization effects are possible only when slight disturbance of the neutrality: $\delta n = |n_e - Z_d n_d| \ll n_e \approx Z_d n_d$. For a cylinder with the radius R, it takes place when $\delta n/n \approx (\lambda/R)^2 << 1$.

Fig. 9 presents observed time dependence λ/R for measured parameters (Z_d, n_d) of the particles at electron temperature T_e = 2 eV. Ibidem dash-dot lines show limits λ/R = 0.5 and λ/R = 0.33 corresponding extreme cases of transparent and non-transparent (for photoelectrons) dust clouds. The ambipolar diffusion coefficient $D_a \approx 1.3 \cdot 10^{-2}$ cm²/s was derived from experimental data approximation for the dust particles concentration evolution $n_d(t)/n_o$ with the function characterizing the velocity of the particles diffusion withdrawal to the cylinder walls. Its value

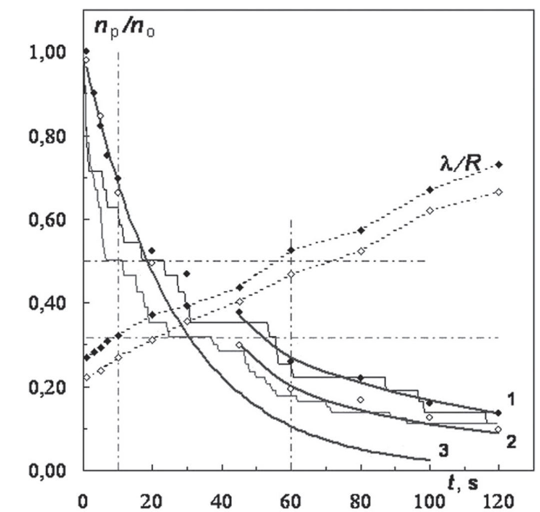


fig. 9. Dependences of relative concentration n_d/n_d^o on time t: (•) — experiment for n_d^o = 195 sm⁻³ and (o) — for n_d^o = 300 sm⁻³; thin line — molecular dynamics method; curves (1)-(2) — approximation; curve (3) — function $n_d/n_d^o = \exp(-\nu_d t)$, and also modification of the λ/R ratio for n_d^o: (▪) —195 sm⁻³; (▫) — 300 sm⁻³

is in accord with theoretical evaluations $D_a \approx (1.2...2.4)\cdot10^{-2}$ cm²/s performed for the particles measured parameters: $Z_d \approx 4.3\cdot10^4$ electron charges at T_e = 1...2 eV.

Thus the measurement results analysis showed that, during initial phases (t < 10 s), there was the particles ambipolar diffusion, i.e. the two signs densities were high enough to generate a considerable spatial charge as the result of their division. The unlike charges polarization effects were reflected in the dust particles drift velocity decrease relatively to their movement in the system transparent for photoelectrons and resulted in dust oscillations after dynamic forcing onto the system when open solar radiation. It should be emphasized that direct experimental observations of phenomena concerned with the charge polarization in the two-component dusty plasma system are impracticable in common laboratory conditions when there is the Earth gravity.

3.2. EXPERIMENTS ON DIRECT CURRENT GAS DISCHARGE (EXPERIMENTS ABOARD 'MIR' SPACE STATION).

Experiments on investigation of the dusty plasma generated in the direct current gas discharge [16] were implemented aboard the *Mir* Space Station. The main difference from similar unit in earth-based experiments was presence of a dual mesh electrode between the anode and the cathode. During the experiments, the electrode was under floating potential and prevented from withdrawal of negative-charged particles onto the anode.

When experiments, a discharge was generated in neon at pressure p = 133 Pa. Polydisperse bronze spheres with average diameter $2a \approx 130$ µm were used as microparticles. The electrons temperature and the plasma density were as $T_e \sim 3...7$ eV, $n \approx n_e \approx 2.0\cdot10^9$ cm^{-3}.

The experiment was run per the following scheme. Initially, the particles were on walls of a gas discharge chamber. Therefore, the system undergone the dynamic forcing after enabling a discharge with the specified discharge current I. Being in the plasma, the particles were being charged (as usual for the gas discharge plasma due to absorption of electrons and ions) and were moving toward the anode. In the mesh electrode environment, a portion of the particles were absorbed and formed a steady-state three-dimensional structure (a cloud), the typical image of which is presented in Fig. 10. This image was recorded using a video camera. The particles returned on to the chamber walls when disabling the charge. The experiment was repeated with new values of the discharge current.

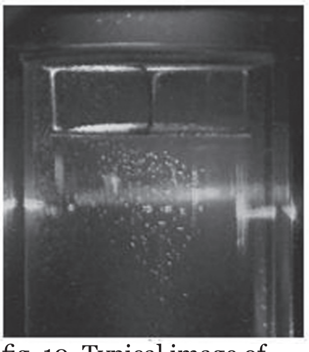


fig. 10. Typical image of a stationary 3D structure (cloud) in the vicinity near the net-shaped electrode

The analysis of video images of the steady-state dust structure generated near the mesh electrode allowed to measure static (the pair correlation function) and dynamic (the diffusion coefficient) characteristics of the dust particles system. Comparison of them with results of numerical simulation of dissipative Debye dust systems was used for the dusty plasma diagnostics. Besides, the analysis of drift movement of the particles to the mesh electrode during initial phase of the experiment allowed evaluating their charge. Basic results are given below.

Measured pair correlation functions were evidence of generation of the ordered dusty plasma structure of the liquid type. It is in accord with results of evaluations of the coupling parameter (decreasing from ~75 to ~25 when the discharge current increase), which were done on basis of changes of the microparticles diffusion coefficient. The particles chaotic movement kinetic energy was evaluated as $T_d \sim 10^5$ eV. Finally, the particles charge was evaluated as $Z_d \sim -2 \cdot 10^6$ electron charges that corresponds to the negative surface potential ~40 V and considerably exceeds the value corresponding to approximation of limited orbital motion. It should be noted that the physics of charging and interparticle interaction between large particles ($\alpha \geq \lambda_{Di}, l_i$) is studied quite badly, mainly since experiments with such large particles are impracticable in Earth-based conditions. It makes the above experiments in zero-gravity conditions to be unique.

3.3. VORTEX GENERATION IN RADIO-FREQUENCY DISCHARGE DUST PLASMA LIQUID (EXPERIMENTS ABOARD INTERNATIONAL SPACE STATION).

Experiments on investigation of the dusty plasma in the radio-frequency gas discharge were implemented at Russian Segment of the International Space Station within the *Plasma Crystal* Program implemented by JIHT jointly with the Institute of Extraterrestrial Physics of the M. Plank Society (Germany). These experiments allowed to investigate the non-ideal dusty plasma not being limited to conditions of the dust particles levitation in the gravity field.

The basic element of the experiment unit was vacuum plasma chamber; the working gas (argon) pressure was 36 to 98 Pa. The experiments were performed for monodisperse particles 3/4 μm in diameter and having mass $m_d \approx 3.1 \cdot 10^{-11}$ g. Fig. 11a represents the dust structures image. In all cases, an area formed in the structure center, which was not filled with the dust particles ('void'), and symmetric vortex motions formed in the cloud boundary sections. There was

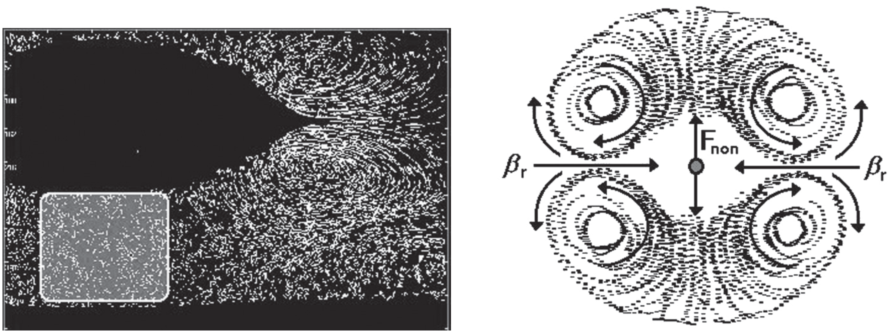


fig. 11. (a) Illustration of dust structures, observed in the camera FOV (2,8x2,1) sm^2 for W = 0.25 W and P = 98 Pa, isolated is the area of microscopic transport characteristics measurements; (b) results of numerical simulation with the parameters: $F_{non} = F_I \approx 0.4\, m_d g$, $\beta_r /(eZ_d) = -0.03\ sm^{-1}$, $v_{fr} = 200\ s^{-1}$ и $Z_d = 6{,}000$

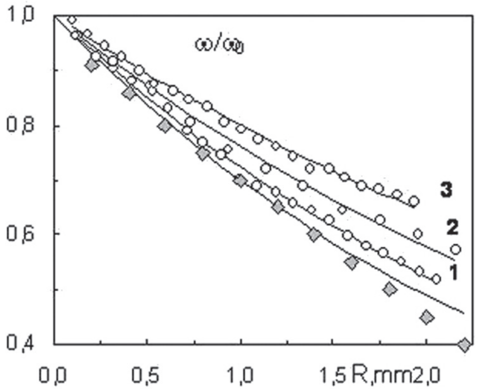


fig. 12. Experimental dependence w/w_o (circles) on R, and it's approximation with the function exp(-R/Ro) (firm line) for W = 25 mW и P: 1 – 36 Pa; 2 - 49 Pa; 3 – 98 Pa, and w/w_o ratio (diamonds), obtained as a result of numerical simulations with the parameters

no regular dust movement along the vacuum chamber axis in areas under and above the 'void', and the dust structures were the structures of the liquid type with average interparticle distance of several hundred microns. The pair correlation functions g(l), the thermal diffusion coefficient D and the dust particles temperature T_d characterizing the kinetic energy of their chaotic thermal movement [17, 18] were measured for these areas.

Experimental investigations of the dust rotation are of considerable interest in respect to develop suitable theoretical models to describe vortices in greatly non-ideal dissipative systems. One of the models allowing to explain the vortices in the dusty plasma originated from existence of the charge gradient $\vec{\beta} = e\Delta Z_d$ in the nonelectrostatic forces field (the gravity force, the thermophoretic force or the ion entrainment force) forcing on the dust cloud particles and being orthogonal to this gradient $\vec{\beta}$. Since the dust particles charge is self-consistent with parameters of the surrounding plasma, it also changes when they change; at that, the charge gradients of the dust particles can reach 50% per centimeter solely due to slight disturbance of the plasma electroneutrality, e.g. in the discharges when the ambipolar diffusion.

Theoretical analysis and numerical simulation show that the dust vortices are possible when there is even a minor charge gradient of the dust particles: $|\beta/Z_d| \sim 1...2$ % cm^{-1}. The ion entrainment force produced by ordered motion of the ions at velocity u relative to a dust particle can be considered as the nonelectrostatic force actuating the mechanism of the dust vortex forming. In some cases ($a_d/\lambda << 1$ and $\nu_{Ti}/u \approx 4.8$, where ν_{Ti} is the ion thermal velocity), the ion entrainment force can be evaluated as: $F_I \approx (0.42...1.25) \cdot 10^{-8}$ dyne [or in units $m_p g$, $F_I \approx (0.15...0.4)\ m_p g$]. The regular velocity value $u \approx (0.12...0.25) v_{Ti}$ taken for evaluations corresponds to the ion ordered motion in electric fields E ~ 1...2 V/cm. Such fields can be the consequence of polarization of the plasma charges in discharges when the ambipolar diffusion of the plasma particles to the working chamber walls. The illustration of simulation of the dust particles rotation in the orthogonal vector field of the force (which forces from the center along the axis of the cylindrical system) and the charge gradient β_r for conditions similar to the experiments when $F_{non} = F_I \approx 0.4$ mdg, $\beta_r/(eZ_d) = -0.03$ cm $^{-1}$, $\nu_{fr} = 200$ s^{-1} and $Z_d = 6{,}000$ is shown in Fig. 11b. At that, average kinetic energy of the particles ordered motion (rotation) was ~0.02 eV.

Fig. 12 shows measurement results of radial distribution of rotation angular velocities ω(R) of the dust particles when different discharge parameters.

Assuming $\omega_o = \omega(0)$ and the ion entrainment force $F_I \approx 0.3\ m_d g$, it is possible to evaluate the dust charge gradients required for rotation of the dust particles with recordable frequencies. In this case, it is enough to change the charge from 5 to 20% per centimeter. The rotation angular velocity ω can be approximated using the exponential function: $\omega(R)=\omega_o \exp(-R/R_o)$, where $\omega_o=\omega(0)$, R — dust rotation radius, and R_o — dust vortex typical scale. The exponential form of the angular velocities distribution ($\omega(R)\sim\omega_o \exp(-R/R_o)$ for $R<R_o/2$) experimentally discovered was also derived in numerical simulation of systems with parameters similar to experiment conditions.

4. TECHNOLOGICAL APPLICATIONS

4.1. NUCLEAR-INDUCED PLASMA-BASED PHOTOVOLTAIC ELEMENT

Compact self-contained current sources of ~1…10 kW with lifetime of a number of years are required for electrical supply of space vehicles, automatic meteorological stations and other similar consumers. Currently, solar energy photoelectric converters, thermoelectric sources with heat-generating elements made of Sr^{90}, Pu^{238} or Po^{210}, and thermoemission converters (TEC), where the heat source is a U^{235} nuclear reactor, are used as such sources. All these sources have a number of disadvantages, very low performance factor in particular. Besides, a nuclear reactor is very complicated in fabrication.

The work [19] offered conversion of the nuclear energy into electrical one by means of the photovoltaic effect in wide-band-gap semiconductors based on the diamond and the boron nitride. Generation of such current sources became practicable as the result of diamond film synthesis investigations resulting in semiconductor structures, and investigations in the scope of the dusty low-temperature plasma physics.

The operation principle of sources converting energy of radioactive isotopes into electricity by means of the photovoltaic effect is as follows. Specially selected gas mixture is excited and radiates in the ultraviolet (UV) range under the action of the ionizing radiation. This UV radiation induces the electromotive force in a wide-band-gap semiconductor by means of the photovoltaic effect. For this purpose, it is the best to use semiconductors based on diamond structures, since they have high radiation resistance and high conversion efficiency factor (up to 70%). β-active isotopes having comparatively long half-decay periods (10 to 30 years), e.g. Sr^{90}, or similar solid isotopes, e.g. α-active Pu^{238}, can be used as the radioactive isotopes.

When using solid isotopes in a photovoltaic converter, it is necessary to get as large isotope surface area as practicable. The best is homogeneous mix of gas and isotope dust, at that the surface/volume ratio is the highest practicable. The gas mix excitation is carried out by means of β-radiation or α-radiation of the radioactive dust. The evaluations show that it is possible to get the specific power of ~1 W/l when the particle size is 1 to 20 μm and the dust-in-gas concentration is 10^5 to 10^9 cm^{-3}. The gas pressure for effective energy conversion of β-radiation or α-radiation into UV radiation is to be about 1…10 atm.

The main physical problem arising when producing a battery of such type is generation of homogeneous gas-dust medium at pressures of a number of atmospheres. Such feasibility follows from results of investigations of the dusty plasma and crystallization processes of such medium. As the result of self-consistent processes in such plasma, the ordered steady state settles. Such state is required to transmit the radiation from the excited gas to the photoconverters.

4.2. REMOVAL OF DUST PARTICLES FROM INDUSTRIAL UNITS

One of interesting line of investigation of the dusty plasma, from the point of view of practical application, is concerned with studying conditions, under which the plasma dust formations move, forced by the external electric field and keeping the internal structure, and precipitates onto certain surfaces. Results of these investigations can serve as the method development basis for removal of dust particles from the static low-temperature plasma. Application of this technique allows to separate the particles by sizes, to remove them from wall areas of the plasma, from processed surfaces when plasma etching, and to control processes of the dust plasma structure generation.

Micron dust particles are present in magnetic confinement thermonuclear devices [20]. Their generation is caused by interaction of the plasma with surfaces of walls. In the scope of International Experimental Thermonuclear Reactor (ITER) design, it has been obvious that, when high densities of plasma particles stream (in a steady state), generation of dust particles is a grand safety problem. ITER will have carbon-based wall components (graphite, carbon composites) like majority of existing large devices. Intrusion of tritium into the graphite dust can result in dust particles, in which there are two tritium atoms for one carbon atom, and total tritium weight in the dust component can reach a number of kilograms. The dust particles can cause tritium emissions into environment in case of a major emergency and contaminate considerable areas.

JIHT, jointly with the State Scientific Center of the Russian Federation — Institute for Physics and Power Engineering named after A.I. Leypunsky, implemented experimental works, in which feasibility to remove the dust particles from the dust structure when it is forced by the external electrostatic field. The experiments were run with the nuclear-induced plasma, which have the track structure. It means that the quasineutrality of such plasma takes place only for volumes containing a great number of nuclear particle tracks, i.e. near a radioactive source. In tracks themselves, quasineutrality is disturbed very fast because of the big difference between diffusion coefficients of electrons and ions. A track decays into moving bunches of electrons and ions in the electric field. Interaction of these bunches with a dust particle forms its charge.

Experiments on removal of the dust particles from nuclear-induced plasma were run using a laboratory bench, the schematic diagram of which is represented in Fig. 13 [21]. The experiments were run in neon at pressure of 25 to 100 kPa. As the radioactive source, a thin circular layer of californium-252 was used, which was 7 mm in diameter and had activity of $4 \cdot 10^6$ fissions per second, and it was located in the centre of the lower grounded electrode. Various upper

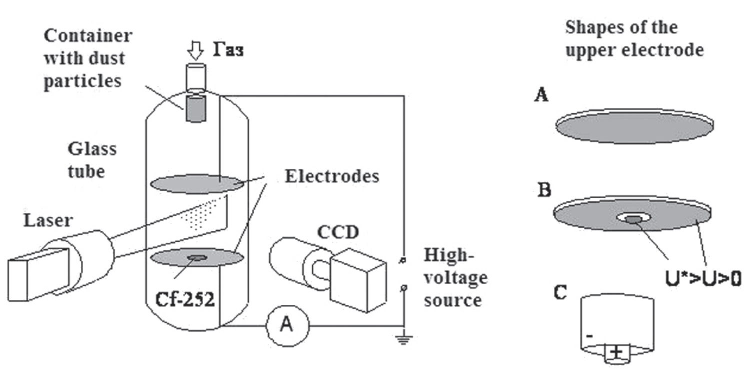


fig. 13. The experimental set-up for removal of dust particles from the volume occupied by nuclear-induced plasma

(high-voltage) electrode designs were used to generate electric fields of various configurations. The distance between the electrodes was ≈ 40 mm.

Polydisperse zinc particles of micron size were used in the experiments. When using a solid electrode (Fig. 13, A), the dust particles produced a cloud having sharp boundaries for some minutes after applying voltage of 160 V to the upper electrode and injecting gas-dust mixture. The cloud had a shape of the truncated cone with the base being on the upper electrode plane and the vertex near the radioactive source. Radiuses of the dust particles in the cloud were within 1.0 to 1.4 μm (weights of the particle were $3 \cdot 10^{-11}$ to $8 \cdot 10^{-11}$ g). The electric charge values range was from 400 to 1000 units of electron charge depending on the radius of the particles. If, after the cloud is produced of the dust particles, the upper electrode potential is increased, then the dust particles rush toward it at a speed, which is the higher, the greater the potential is. When the upper electrode is up to 200 V, not the whole structure rushes upward. One or several streams with changeable in time shapes and positions of the bases on the electrode are generated in its upper section (Fig. 14).

Use of an additional electrode, made in the form of the hemisphere having the diameter of 3 mm (Fig. 13, B) and being isolated from the main electrode, let to remove the dust particles from the cloud (Fig. 14b). When the upper electrode

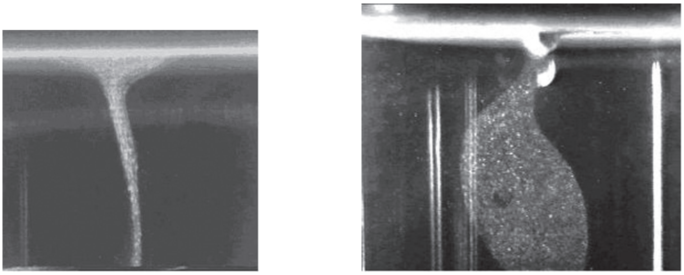


fig. 14. (a) Motion of dust particles to the high-voltage electrode in the form of dust stream; (b) dust particles' drift to the auxiliary hemispherical electrode

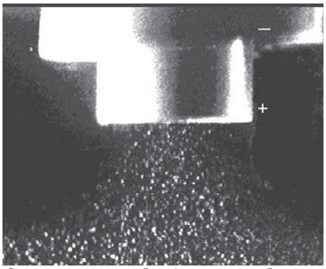


fig. 15. Particles' removal with the use of coaxial probe, diameter of the inner tube — 3 mm, outer — 5 mm

potential U ≈ 200 V and the additional electrode potential U ≈ 300 V, one dust stream is generated, which is directed toward the additional electrode. The dust particles gravitating to the electrode remain on its surface.

When the electric field is generated by two thin-wall tubes (Fig. 13, C) having different potentials, it succeeds to get the electric field configuration, where the negative-charged dust particles gather from the dust cloud into the internal tube with the positive potential of 150 V (Fig. 15), at that the particles remain on the internal surface of the tube.

4.3. SPATIAL SEPARATION OF DUST PARTICLES IN PLASMA

Measurements of spatial size distribution of the particles, levitating in the low-pressure radio-frequency inductive discharge, were carried out in an experiment, the schematic diagram of which is given in Fig. 16. Monodisperse transparent microspheres 0.98 ± 0.03 и 1.87 ± 0.05 µm were used in the investigation. The radio-frequency inductive discharge was induced in a vertically directed cylindrical glass tube filled with neon. The power applied to the discharge was tenth fractions of watt. The measurements carried out with the plasma-generating neon gas under two pressures of 80 and 150 Pa. Microspheres of different sizes were entered in the discharge from a special container suspended in the upper section of the tube. The particles hovered in an electrostatic trap in the lower portion of the induction discharge. Such trap is generated due to a combination of the ambipolar diffusion fields and the charged surface of the discharge tube.

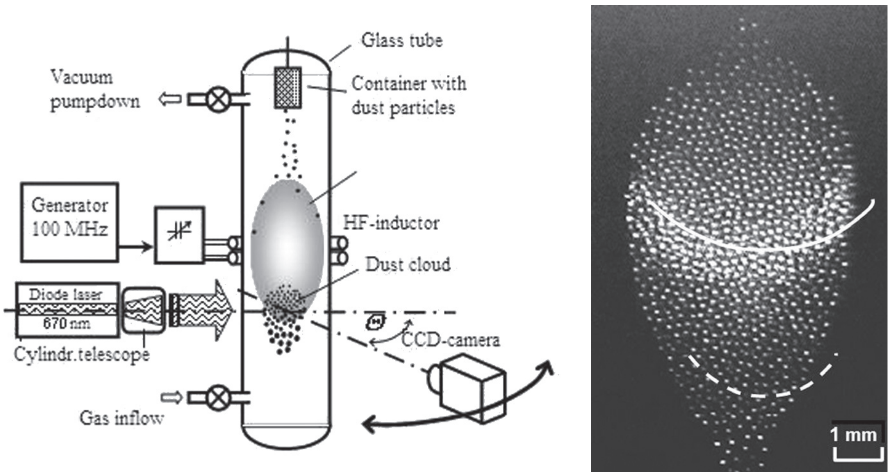


fig. 16. (a) Scheme of the experiment for measuring of polymeric microspheres spatial distribution in the inductive discharge; (b) view at the angle θ = 74° of dust cloud levitating in the inductive discharge in neon with the pressure of 150 Pa and consisting of transparent microspheres with the diameter ~1,8 µm. Solid/dotted line indicates the area with maximum/minimum intensity of scattered laser radiation

For transparent particles of micron sizes, practically the only method of measurements of their sizes in situ is measurements of the scattering phase functions of the monochromatic polarized light with further interpretation of the findings by the scattering theory Mie. The measurement results of the microspheres spatial size distribution in the dust cloud of particles with rated diameters of 0.98 ± 0.03 and 1.87 ± 0.05 µm are given in Fig. 17. The accuracy of the measurements of the microspheres diameters was ~1%. As evident from the figure data, particles contained in the cloud have considerable size spread, and there are dust particles in the lower portion of the cloud, the scattering indicatrix of which does not have pronounced intensity maxima or minima. To all appearances, these are agglomerates consisting of two or more microspheres.

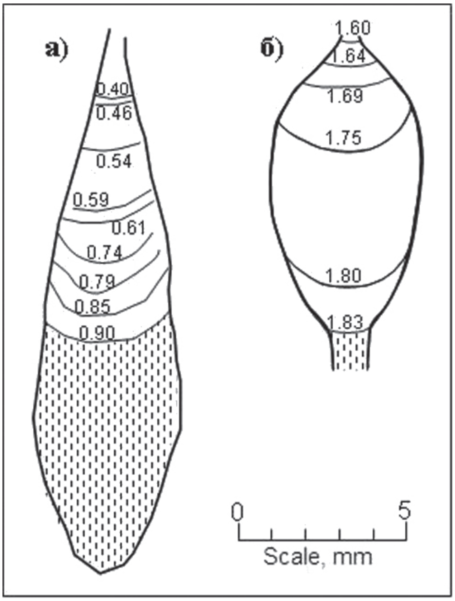


fig. 17. Measured spatial distributions of microspheres into sizes in dust clouds formed by the injection of monodisperse particles with diameters 0.98±0.03 µm (a) и 1.87 ±0.05 µm (b) into plasma. Digits are measured microspheres' diameters along respective lines on the picture. Hatched areas — area of dust cloud formed by particles' agglomeration.

The gradient of the particle size dispersion, to all appearances, is determined by the gradient of the electric field in the induction discharge. In standing strata of the direct current glow discharge, the electrons temperature nonuniformity within the stratum can also make an additional impact upon the gradient of the particle size dispersion. Thus dusty plasma structures can be used for separation of fine particles and recovery of individual fractions in the gas discharge conditions.

SUMMARY

In spite of almost century-old history, the study of the dusty plasma properties has taken on particular scope just in the last ten years after experimental discovery of the dusty plasma crystals. Due to the unique properties, the dusty plasma is successfully used to solve both fundamental and application problems. Visualization simplicity allows to measure dust components on the kinetic level, at that, a detailed analysis of thermodynamic and kinetic properties of dislocations and other defects of the dust crystal lattice, having much common with general crystal lattices of solids, is feasible. The study of easily induced linear and nonlinear low-frequency oscillations and their unsteadiness is of great interest. The study of phase transfers in systems of symmetric and asymmetric dust particles gives useful information about critical phenomena and processes of self-organization, in particular about feasibility of natural generation of dusty plasma ordered structures in the Universe. The first space experiments have been carried out in microgravitation conditions aboard the *Mir* Station and the International Space Station, where a number of important and sometimes unexpected results have been obtained.

Among application problems, one of the main ones is the problem of dust particles removal when producing computer microcircuits by plasma technology methods, and to solve it, fundamental understanding of physical processes in the gas discharge plasma is required. Furthermore, unique practicability of retention and control of physical-chemical properties of the dust particles makes the plasma excellent medium to generate dusts with specified properties and their modifications. Currently, the dusty plasma is one of the most intensively developed fields of the physics, on average more than one articles in this field are published during the last years. Therefore, there is no doubt that many other interesting and important results will be obtained in the future.

REFERENCES

1. Sodha M., Guha S. Physics of Colloidal Plasmas // Adv. Plasma Phys. 1971. V.4.
2. Zhukhovitski D., Khrapak A., Yakubov I. Ionization Equilibrium in the Plasma with the Condensed Disperse Phase / Ed.Smirnov B.M. — Chemistry of plasma. —Iss. 11. Moscow, Energoatomizdat, 1984.
3. Fortov V.E., Yakubov I.T. Non-Ideal Plasma. Moscow: Energoatomizdat, 1994. P. 282.
4. Ichimaru S. Strongly Coupled Plasmas: High-Density Classical Plasmas and Degenerate Electron Liquids // Rev. Mod. Phys. 1982. V.54. P. 1017.
5. Ikezi H. Coulomb Solid of Small Particles in Plasmas // Phys. Fluids. 1986. V.29. P. 1764.
6. Chu J., I L. Direct observation of Coulomb Crystals and Liquids in Strongly Coupled rf Dusty Plasmas // Phys. Rev. Lett. 1994. V. 72. P. 4009.
7. Thomas H., Morfill G., Demmel V. et al. Plasma Crystal: Coulomb Crystallization in a Dusty Plasma // Phys. Rev. Lett. 1994. V. 73. P. 652.

8. Nefedov A.P., Petrov O.F., Fortov V.E. Crystallike Structures in the Plasma with Strongly Coupled Macroparticles // UFN. 1997. V. 167. № 11. P. 1215—1226. (in Russian)
9. Fortov V.E., Molotkov V.I., Nefedov A.P., Petrov O.F. Liquid- and Crystallike Structures in Strongly Coupled Dusty Plasmas // Physics of Plasmas. 1999. V. 6. №. 5. P. 1759—1768.
10. Nefedov A.P., Fortov V.E. Liquid- and Crystallike Structures in Non-ideal Dusty Plasma // Vestnik RAN. 2000. V. 70. № 10. P. 934. (in Russian)
11. Vaulina O.S., Petrov O.F., Fortov V.E. et al. Three-particle correlations in nonideal dusty plasma // Phys. Rev. Lett. 2004. V. 93. №. 3. P. 035004.
12. Samsonov D., Morfill G., Thomas H. et.al. // Phys. Rev. E. 2003. V. 67. P. 036404.
13. Fortov V.E., O.F. Petrov O.F., Molotkov V.I., Poustylnik M.Y., Torchinsky V.M., Khrapak A.G., Chernyshev A.V., Nonlinear Dust Waves Excited by the Gas-Dynamic Impact in a DC Glow Discharge Plasma // Phys. Rev. E 69. 2004. P. 016402.
14. Fortov V.E., Nefedov A.P., Vaulina O.S. et al. Dusty Plasma Induced by Solar Radiation under Microgravity: Experiment onboard "Mir" Space Station // ZhETF. 1998. V. 114. Iss. 6 (12). P. 2004—2021. (in Russian)
15. Vaulina O.S., Nefedov A.P., Fortov V.E., Petrov O.F. Diffusion in microgravity of macroparticles in dusty plasma induced solar radiation // Phys. Rev. Lett. 2002. V. 88. P. 035001.
16. Nefedov A.P., Vaulina O.S., Petrov O.F. et al. Dynamics of Macroparticles in DC Discharge under Microgravity // ZhETF. 2002. V. 122. № 4. P. 778—788. (in Russian)
17. Fortov V.E., Vaulina O.S., Petrov O.F. et al. Transport of Microparticles in Weakly Ionized Gas-Discharge Plasmas under Microgravity Conditions // Phys. Rev. Lett. 2003. V. 90, № 24. P. 245005.
18. Fortov V.E., Vaulina O.S., Petrov O.F. et al. Dynamics of Macroparticles in Dusty Plasma under Microgravity (First Experiments onboard ISS) // ZhETF. 2003. V. 126. № 4. P. 798—815. (in Russian)
19. Baranov V.Yu. et al. // In book: «Isotopes: properties, production, utilization» / Ed. Baranov V.Yu. Moscow, Atomizdat, 2000. 626 p.
20. Tsytovich V.N., Vinter J. Dust in the Devices of Controlled Thermonuclear Fusion // UFN. 1998. V. 168. P. 899. (in Russian).
21. Vladimirov V.I., Deputatova L.D., Molotkov V.I., Nefedov A.P., Rykov V.A., Filinov V.S., Fortov V.E., Khudyakov A.V. Ordered Dusty Structures in Nuclear-Track Neon and Argon Plasmas // Plasma Physics Reports. 2001. V. 27, № 1. P. 36—43.

A.I. GRIGORIEV,
A.N. POTAPOV
Institute of Biomedical Problems,
Russian Academy
of Sciences,
State Scientific Center
of the Russian Federation
Moscow, Russia
grigoriev@imbp.ru

ADVANCES AND PERSPECTIVES OF SPACE BIOLOGY AND MEDICINE

FORMATION OF SPACE BIOLOGY AND MEDICINE

History of space biology counts over 50 years. In Russia, this scientific discipline evolved in the 1950s with the appearance of opportunities to carry out experiments with various biological objects aboard altitude rockets. These investigations were primarily aimed to evaluate functioning of these bio-objects in the extreme conditions of the upper atmosphere which are drastically different from those on Earth and similar to the conditions of space. As a result, the principle possibility for terrestrial organisms to survive in these conditions was established.

The next fundamental phase was implementation of biological investigations directly in space. In November of 1957, a month after launch of the first Earth's artificial satellite, the second artificial satellite was launched with a living being — dog Laika, and equipment to record its physiological parameters and provide life support (Fig. 1).

fig. 1. Dog Laika — first living being to accomplish space flight on November 3, 1957 on 2nd artificial satellite of the Earth

These were followed by launches of returnable vehicles-satellites that carried dogs and other bio-objects and had the goal to study the effects of spaceflight factors on living organisms. Academician N.M. Sisakian with coauthors appraised results of these unprecedented investigations: 'A series of outstanding experiments flown on the second, third, fourth and fifth vehicles-satellites gave answers to many questions. Obtained were the key starting data for the most important conclusion that from the standpoint of biology and medicine flights in a round orbit below the Earth's radiation belts will be safe for human health and life' [1]. This audacious scientific foresight was confirmed.

April 12, 1961 the historical flight of Yu.A. Gagarin paved the way to space for mankind. This remarkable day gave birth to space medicine.

Every next flight posed new nontrivial scientific and applied problems. Results required profound scientific analysis and ensuing development of methods and technologies of providing medical safety and maintaining cosmonauts' health and ability to work.

Retrospective overview of space biology and medicine we can point to the following basic strategies:

- through scientific analysis of data concerning cosmonauts' health in and after space flight;
- gradual extension of piloted missions;
- implementation of a comprehensive research program in space and ground-based simulation studies in order to elucidate mechanisms of the spaceflight effects on human organism and consistent patterns of adaptation;
- R&D, and testing countermeasures during space flights and in simulation experiments;
- substantiation of space biology and medicine methodology using the data obtained in space and also in a broad variety of simulation and laboratory investigations.

RESEARCHES ON BIOLOGICAL SATELLITES

Studies with animals have not lost their significance despite the beginning of piloted space flights. We can use the example of the unique experiment conducted in the 22-day mission of biosat *Kosmos-110* (1966) with dogs Veterok and Ugolek [2]. The main goal for this flight (aside from scientific objectives) was set by S.P. Korolev and consisted in functional verification of all equipment of vehicle *Voskhod-3* before its launch. The experiment on *Kosmos-110* provided a wealth of valuable information taken into consideration during preparation for an 18-day piloted mission.

In 1973 on the initiative and leadership of the Institute for Biomedical Problems (IBMP) a broad cooperation of Russian and international investigators was started that has the purpose to consolidate efforts within the framework of the BION program, a series of launches of biosats. These biosats functioned virtually as a venue for biological researches in space. Eleven biosats were launched in the period between 1973 and 1997 for missions from 5 to 22.5 days in duration. Bio-objects were different species of animals and plants as various levels of evolutionary development including primates, rats, birds, amphibians, fishes, worms, insects, and unicellular organisms, isolated cell cultures, fungi, lower and higher plants.

The range of investigations undertaken within the BION program was surprisingly large but was predominantly focused on studying the effects of microgravity on biological and physiological processes. Inquiry into the fundamental processes including cell metabolism, cell division, transmission of genetic information, frequency of spontaneous mutations, embryogenesis and ontogenesis at large showed absence of irreversible damaging effect of microgravity on the basic life processes [3].

Physiological investigations on biosats were concerned with micro-g impacts on the cardiovascular system, muscles, skeleton, vestibular and sensory systems, locomotion control and immunity. These investigations gave support to many phenomena observed in cosmonauts in flight and after flight completion, and shed light on specific mechanisms underlying the noted changes. No pathological, irreversible and delayed shifts were identified in separate organs and systems and yet adaptation to microgravity was shown to degrade the functional activity of organism attested by its deconditioning [3, 4].

Experiments on three biosats were dedicated to artificial force of gravity (AFG). It was K.E. Tsiolkovsky who was first to propose AFG as a countermeasure against microgravity. On *Kosmos-936* AFG was generated by rotating centrifuge with rats at 1 g. It was found that AFG moderates atrophy and prevents metabolic disorders in skeletal muscles and decrease of calcium, phosphorus and mineral density in trabecular bones [3, 4].

Biosat *Kosmos-690* had the goal to study combined effects of radiation and microgravity on organisms. It was shown that in rats irradiated by source Ce^{137} at the doses endangering with radiation disease, microgravity did not significantly modify the course of radiation pathology and post-radiation recovery. The significance of these investigations grows in view of future exploration missions.

Biological space researches in Russia and other countries are performed on dedicated satellites, transport vehicles and orbital space stations. They enriched substantially life sciences by broadening the spectrum of experimental conditions and phenomena to be studied, and enabled acquisition of scientific data valuable to fundamental sciences. The most ponderable outcome was demonstration of the universal role of gravity in the functioning of living systems on different levels of organization, i.e. on the levels of organism, organs, tissues, cells, and molecules.

Space researches ensured virtually reformation of extraterrestrial life science — exobiology addressing the problems of life origin in the Universe and the Solar system (abiogenous synthesis of organic substances, germs travel across space). Long before that implications of these issues had been foreseen by Russian scientist V.I. Vernadsky who wrote: 'New broad horizon is opening for biological finding. Proof, if obtained, of life being space rather than planetary phenomenon will have extraordinary consequences for biological and humanitarian concepts' [5].

Biological researches in space flights provided a wealth of data about patterns of ontogenesis process, genetic modifications in a sequence of generation in various species of animals, lower and higher plants. The results lay the basis for further investigations toward the design of advanced biological life support systems (LSS) for remote space missions and substantiation of future settlements in space [6].

Space investigations stimulated studies of consistent patterns of biosphere formation and sustenance. Artificial biosphere models used in development of biological LSSs permit, within a short space of time, reproduce biospherical events which normally take dozens and hundreds of years [7].

Radiobiologists were given an opportunity to investigate a wide spectrum of cosmic rays.

The major outcome of the biological researches in space was foundation of new rapidly developing disciplines — gravitational biology and gravitational physiology.

PHYSIOLOGICAL INVESTIGATIONS IN PILOTED MISSIONS

The indispensable condition and source for advancement of space biology and medicine were medical observations, biomedical and physiological investigations and tests of means and methods of countermeasures during space flights. The success of space biology and medicine and applied efforts rendered possible long-duration space missions. In 1994—1995 cosmonaut-physician V.V. Poliakov made a record 438-day flight which cleared the way to other cosmonauts for extended performance in space. Today no one is surprised at the now usual space missions lasting six months and longer. The number of cosmonauts who have spent several years flying on space missions is growing (S.V. Avdeev — 750 days in 3 missions, A.Ya. Soloviev — 653 days in 5 missions, V.M. Afanasiev — 547 days in 3 missions, A.V.Viktorenko — 489 days in 4 missions).

This became possible owing to a reliable system of medical care for cosmonauts' health based on the data of the systematic perennial investigations in space physiology and medicine. In the course of its development main risk factors in space flight were defined and the most vulnerable body systems were revealed. The system helps study the key mechanisms that regulate the body functions and adaptation in microgravity.

Medical care in the pre-launch period includes medical selection and training, regular clinical and physiological investigations, and health certification.

During space flight, medical care consists of monitoring and predicting health status of cosmonauts, implementation of the countermeasure program, rendering medical assistance, if necessary, scheduling and control of the work-rest cycle, providing psychological support, radiation safety and environmental monitoring.

On completion of space flight, the programs of rehabilitation and maintenance of professional longevity are put in practice.

Multiple investigations showed that microgravity is the major factor in near-Earth's flights [8]. Microgravity is the cause of considerable changes in the so-called gravity-dependent systems which in the conditions of 1 g works against the force of gravity (cardiovascular and muscular systems, skeletal apparatus) or are constantly guided by the gravity vector (vestibular system).

Microgravity may provoke negative changes in the myocardium such as rhythm disturbances (particularly during extravehicular activities and physical loads), increase in heart rate and reduction in the cardiac beat strength, distorted repolarization evidenced by ECG records, decreases in cardiac output and minute volume. Hemodynamics in microgravity is characterized by venous stasis in visceral organs, reduction in the volume of circulating liquids, dysregulation of the vascular tone and insufficient microcirculation.

The primary triggers of the shifts in the cardiovascular systems and water-salt metabolism are reactions to the microgravity-induced elimination of hydrostatic blood pressure resulting in redistribution of body liquids. This leads to an increase in the intrathoracic blood volume interpreted by the atrium receptors as a surplus of circulating blood. The inevitable consequence of this misinterpretation is activation of the mechanisms of intensive renal excretion of water and electrolytes, and establishment of negative water-salt balance.

In microgravity low loading on the muscular system entails disorders in all components of the neuromuscular apparatus. The most dramatic changes occur to the anti-gravitational muscles bearing the force-resistive function during locomotion and posture maintenance (Fig. 2).

In space flight we observe numerous changes in muscle parameters such as strength properties, power output, contraction-relaxation ratio, histochemical profile of muscular fibers and metabolic supply of muscle effort [9]. These shifts result in atrophy and deconditioning of muscles, and degradation of physical performance.

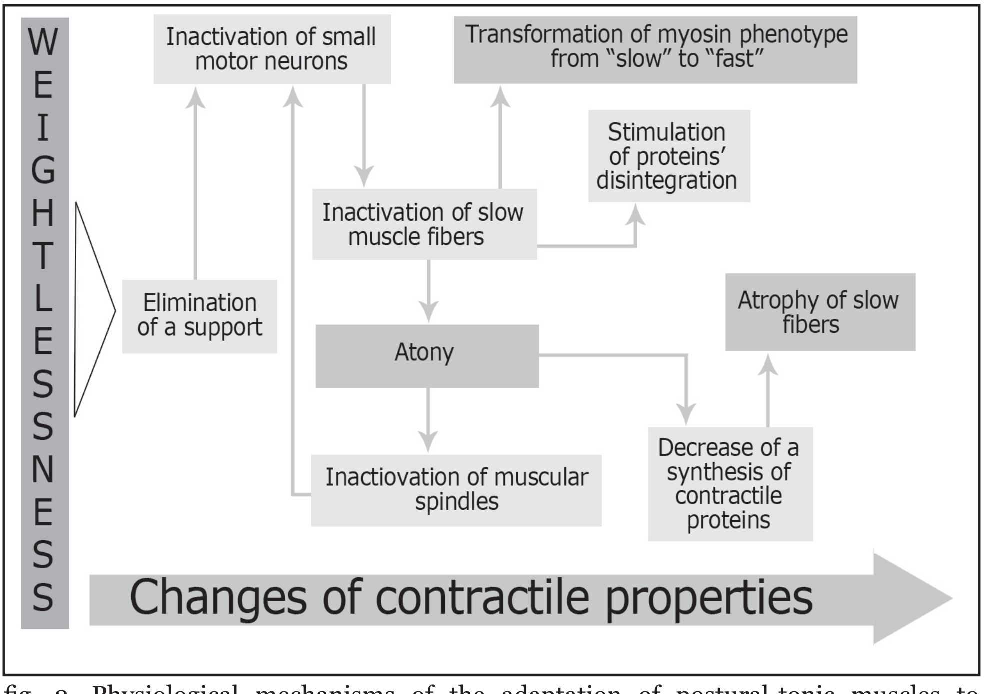


fig. 2. Physiological mechanisms of the adaptation of postural-tonic muscles to weightlessness

Microgravity affects the sensory systems and sensory input to the motor control systems causing reduction in afferentation from the otoliths, muscular receptors including support afferentation from the foot receptors (Fater-Pacchinian bodies), functional deprivation of the otolith apparatus, and vestibular-oculomotor coordination [10]. The aggregate of these changes impacts sensory synthesis, movement coordination, orientation, control of voluntary movements and rises, by way of compensation, role of the ocular analyzer that takes upon itself evaluation of the body spatial position normally performed by the vestibular apparatus.

Significant problems in long-duration space flights may arise from changes in bone tissue (Fig. 3). Segments of the skeleton, especially those that bear most of the weight loading on Earth, respond to load removal by inhibition of osteoporosis (synthesis of new bone tissue) with simultaneous enhancement of resorption

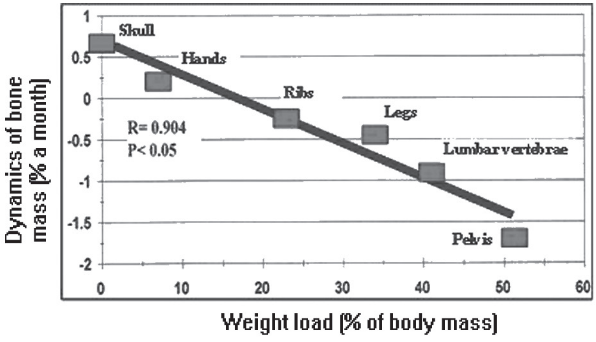


fig. 3. Decrease of mineral density of bone tissue of cosmonauts after space flights

(disintegration) and decrease of bone mineral density due to calcium and phosphorus loss [11]. These developments, if left out of control, increase the risk of possible bone fractures and formation of renal stones. It should be emphasized that recovery of bone mineral density and structure after flight takes a long time. For example, in one cosmonaut bone recovery required about two years.

An integral index of microgravity effect on organism is shifts in protein, carbohydrate, lipid, and energy metabolism much dependent on functionality of the hypophysis.

Data obtained from the cosmonauts and results of the experiments with animals onboard the biosats revealed development of immediate and delayed adaptive reactions and ensuing new level of homeostasis. These were accompanied by numerous metabolic shifts, restructuring of some tissues, establishment of another level of energy processes, and intensification of catabolism [12]. However, these changes were found to activate mechanisms of self-regulation counteracting further negative development and making them somewhat less dramatic.

To prevent and mitigate the adverse effects of microgravity, a special system of countermeasures is used during space flight. The central place in the system is given to physical exercises favorable to keeping health, fitness and performance of cosmonauts at a required level. The purpose of these exercises is prevention of cardiovascular and locomotion systems disorders, and metabolic shifts [13, 14]. On the ISS, the staple training facility is a treadmill, as training sessions on this apparatus are recognized to be the best countermeasure in long-duration space flight. Additional loading on muscles and bones can be produced by wearing a specially developed suit *Penguin*. There are also a bicycle ergometer and a set of expanders to maintain the muscle strength and speed properties. In addition, strength as well as static and dynamic endurance of muscles are maintained by electromyostimulation.

Adequacy and effectiveness of countermeasures will be better evaluated and improved after delivery to the ISS a computer-based system that will be used to adjust amount and intensity of exercise to the actual functional status of organism deduced from the physiological data recorded in previous training sessions. Ground testing of a prototype of this system has demonstrated its potentiality [15].

At present time, preparation for future missions to the Moon and Mars bring into focus an alternative or complimentary countermeasure, i.e. artificial gravity (AG) generated by a short-arm centrifuge (SAC). Prof. A.R. Kotovskaya with coauthors summarized the results of investigations of centrifugation effects in test-subjects exposed to simulated microgravity [16]. It was shown that AG increases the gravitational tolerance reducing the adverse shifts in functioning of the cardiovascular system, external respiration, water-salt metabolism, blood coagulation, and others.

Practical implementation of the idea to use AG in exploration missions will hinge on solving a large number of problems including the development of AG prescriptions preventing the micro-g impacts on various body organs, systems and functions. In experiments with simulated microgravity negative consequences of SAC rotations will be determined and so will be the methods of their mitigation; also, possible synergy of AG and the conventional countermeasures will be evaluated.

CELL GRAVITATIONAL PHYSIOLOGY

It is getting more evident that changes occurring to organism and separate systems in extreme conditions break out when disorders have happened in cells and molecules. These disorders are now the subject of investigations in gravitational physiology of cell and space biotechnology.

Specifically, it became known that microgravity affects the migration activity of cells, impacts resistance to cytotoxic agents, distorts molecules expression, adhesion on cell surface, potentiates effects of cytokins, and modifies the cytoskeleton [17].

The following investigations in this area could be of significant implications:

1. Identification of genes involved in perception and realization of gravitational stimulus;
2. Studies of the micro-g effects on inter-cell interactions and conduction of intra-cell signals;
3. Evaluation of gravitational effects on differentiation of stem and embryonic cells;
4. Construction of three-dimensional tissue structures in microgravity;
5. Exploration of the gravitational stimulus transformation in cells.

PSYCHOLOGICAL SUPPORT TO COSMONAUTS

Main requirements to space psychology is guidance in selection and training of cosmonauts, crew composition, building crew competencies, psychosocial support during and rehabilitation post flight.

The spaceflight factors and absence of habitual life conditions may negatively influence the psychic state of cosmonaut. The most pressing factors are artificial environment, physiological and psychological stresses, isolation and confinement, boredom, separation from customary socium and family, and a small circle of contacts. The extreme environment of space flights challenges proficiency and mentality of cosmonauts, and interactions within culture-mixed and gender-mixed crews.

Long-duration missions provoke asthenia, sleep disorders and other psychophysiological deviations, emotional disorders that in the long run may impact crew efficiency and quality of job, and falling off in relations within crew [18].

Prevention and correction of these problems are objectives of the psychological support system in the pre-launch period, during mission and after return to Earth.

The pre-launch period is the time of psychological selection, crew training and assignment.

Basic criteria of selecting-in are:

- psychic health;
- professionally suitable personality;
- pattern of behavior and ability for team work;
- tolerance of extreme life and work conditions.

In future, the selection procedure will be complemented with genetic criteria correlating, for instance, with resistance to extreme factors.

During the pre-launch training cosmonauts are familiarized with twists of interpersonal relations and learn to build up relations on the principles of mutual respect and mutual aid.

In flight, the psychological support includes monitoring and diagnostics of psychological and emotional status using PC-based methods. Specifically, a psychodiagnostic computerized simulator, currently on the phase of field (space) testing, will be a valuable asset in evaluating and predicting the reliability of manual vehicle control, and testing and refreshing of critical operational skills [19].

Important components of psychological support during flight are methods of behavior monitoring and correction provided to crew from the ground and/or ready at hand [20].

After landing, cosmonauts return to the normal way of living and a previous system of social contacts. Objectives of this period are psychosocial rehabilitation and maintenance of professional longevity of cosmonauts.

LIFE SUPPORT SYSTEM

The tremendous success was achieved in designing and operation of the systems of life support for orbital space stations [21]. Specifically:

- standards, methods and means of environmental monitoring;
- oxygen generation and CO_2 removal systems;
- technologies of water regeneration from biological wastes and biotechnical systems;
- means to provide eco-safety of humans including chemical and microbial decontamination agents;
- technologies of protecting the interior, systems and equipment in pressurized modules against biodamage.

Out of 234 microbial species identified in the space station *Mir* environment, 108 bacteria and 126 fungi were potential pathogens and could also damage chips of board equipment [22].

Tests with microorganisms are aimed to primarily mitigate the risk of infectious diseases in orbit. Besides, microorganisms serve as optimal objects for solving the exobiology and planetary quarantine issues. The *Biorisk* experiment was designed to expose spores of *Bacillus subtilis* and *Aspergillus versicolor* in open space [23]. After a year and a half of exposure outside the space station the spores were found to remain viable (Fig. 4). Similar experiments are planned for launch on spacecraft *Phobos Sample Return*.

Much effort is spent on development of LSS biological components for possible integration into hybrid LSS combining the physical-chemical and biological

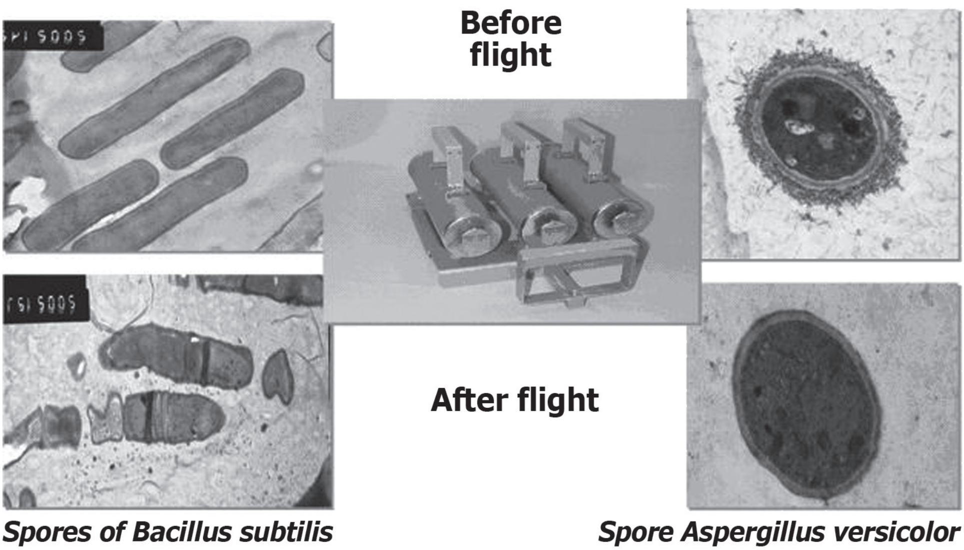


fig. 4. Influence of extreme conditions of a space on microorganisms. *Biorisk* experiment

technologies and characterized by a higher degree of closure. Possibility of avian ontogenesis (Japanese quail) in microgravity was proven experimentally. It was also demonstrated that the spaceflight conditions do not set up a barrier to growth and development of higher plants in a sequence of generations; ontogenesis as well as principle functions, morphology and biometrics of plants remain unaffected [6] (Fig. 5).

fig. 5. Experiments in space with components of a biological life-support system

The genetic experiments with the third generation of ‘space’ peas on the ISS failed to detect a significant increase in the number of chromosomal rearrangements; studies of DNA-polymorphism did not reveal genetic variance on the molecular level [24].

RADIATION SAFETY AND RADIOBIOLOGICAL INVESTIGATIONS

IBMP and the Skobeltsyn Institute of Nuclear Physical of Moscow State University (SINP MSU) have been cooperating in multiyear continuous monitoring of the radiation environment on the orbital station [25]. In-flight and ground-based investigations in radiation physics and biology including the experiments with heavy charged particles on the cyclotron at the Joint Institute of Nuclear Research (JINR) in Dubna yielded the following results:

- description of the radiation conditions along trajectories and inside piloted and robotic spacecraft;
- elucidation of patterns of space radiation dose accumulation in biological tissues (experiments with *Phantom* device on evaluation of dose loads on critical human organs) shielding material;
- estimation of the coefficients of relative biological effectiveness of space radiation;
- development (jointly by IBMP and SINP MSU) of methods for optimal radiation protection including active electromagnetic, electrostatic and dielectric protection;
- determination of acute and delayed somatic consequences of exposure to heavy charged particles [25—28].

Further investigations in radiobiology will look into the damaging effects of space radiation on the central nervous, cardiovascular, blood forming and immunity systems, gastrointestinal tract and other systems of organism.

In view of the high penetrability of heavy charged particles, emphasis will be placed on studying their effects on the central nervous system, visual function, retina and lens, and gene and structural mutations. Findings of these investigations will substantiate the search of methods to protect from galactic space radiation which is important in the context of planned exploration missions. Evaluation of the combined effects of radiation and microgravity will be continued.

Present-day space biology and medicine are concerned with technologies of medical safety in an exploration mission to Mars.

To this end, research and development needs to be pursued in the areas of:

- system of exploration crew selection and training;
- concept of an autonomous medical center with a repository of knowledge and a telemedicine interface;
- advanced system of countermeasures against the gravitational impacts;
- feedback training facilities and systems for objective evaluation of operator's proficiency;
- radiation safety system reliable on all mission phases;
- self-contained regenerative life support systems incorporating physical-chemical and biotechnologies;
- optimization of environment;
- wastes transformation technologies and systems.

This vast range of objectives can be addressed in laboratory research and technology experiments, long-duration orbital missions, and comprehensive simulation projects.

In 2009, IBMP plans to begin an extended international simulation project with the aim to verify the biomedical technologies which are considered as possible candidates for an exploration mission to Mars (*Mars-500*) [29].

The project will specifically have the following objectives:

- verification of improved countermeasures against the adverse effects of microgravity and some other factors of space flight;
- testing of original methods and means of crew health medical monitoring;
- demonstration of new approaches to medical care during space flight, including the telemedicine capabilities.

A prominent place will be given to psycho-physiological experiments on:

- development of methods of testing mental performance and psycho-physiological status;
- optimization of psychological guidance of crew assignment and training;
- appraisal and control of relations within a multinational crew.

In the area of habitability, updated procedures of monitoring and control of environment and life support systems will be tested, and hybrid LSSs technologies evaluated.

CONCLUSION

Complication of cosmonauts' activities in orbit and ideas to explore remote space poses new challenges. Multidisciplinary researches and search for ingenious technologies will, in many respects, be follow-up of the recent achievements in molecular medicine, genetics, genomics, biotechnology and proteomics and form the platform for future progress of space biology and medicine.

REFERENCES

1. Sisakian N.M., Parin V.V., Chernigovsky V.N., Yazdovsky V.I. Some problems of space research and exploration // Probl. Kosmich. Biol. M.: Izdat. AN SSSR, 1962. P. 5—16. (in Russian)
2. Parin V.V., Pravetsky V.N., Gurovsky N.N. et al. Some results of the biomedical experiment onboard biosatellite "Kosmos-110" // Kosmich. Biol. i Med. 1968. V. 2. № 2. P. 7—13. (in Russian)
3. Souza K.A., Ilyin E.A. Main results of biological investigations in space // Space biology and medicine (Joint US-Russian publication). V. III, Bk. 1. Pt. 1. Ch. II. M.: Nauka, 1997. P. 58—108. (in Russian)
4. Ilyin E.A. Investigations aboard biosatellites "Kosmos" // Aviakosm. i Ekolog. Med. 1984. V. 18. № 1. P. 57—66. (in Russian)
5. Vladimir Vernadsky. Discoveries and lives. 1993. M.: Sovremennik. P. 393. (in Russian)
6. Grigoriev A.I., Sychev V.N. Systems of life support of cosmonauts based on biosphere mechanisms // Vestnik RAN. 2004. V. 74. № 8. P. 675—681. (in Russian)
7. Meleshko G.I., Shepelev E.E., Averner M.M., Volk T. Biological systems of life support // Space biology and medicine (Joint US-Russian publication). V. II. Ch. 16. M.: Nauka, 1997. P. 499—558. (in Russian)
8. Grigoriev A.I., Egorov A.D. Long-term space flights // Space biology and medicine (Joint US-Russian publication). V. III. Bk. 2. Pt. VII. Ch. 23. 1997. M.: Nauka. P. 368—447. (in Russian)
9. Shenkman B.S., Grigoriev A.I. Muscles in support-free space // Vestnik Ross. Akad. Nauk. 2008. V. 78. № 4. P. 337—345. (in Russian)
10. Kornilova L.N., Kozlovskaya I.B. Neurosensory mechanisms of space adaptation syndrome // Fiziologiya Cheloveka. 2003. V. 29. № 5. P. 17—28. (in Russian)
11. Oganov V.S., Shneider V.S. Bone system // Space biology and medicine (Joint US-Russian publication). V. III. Bk. 1. Ch. 11. M.: Nauka, 1997. P. 421—259. (in Russian)
12. Grigoriev A.I., Popova I.A., Kaplansky A.S. Minimalization of the regulatory systems functions and metabolism in microgravity // Aviakosm. i Ekolog. Med. 1993. V. 27. № 5—6. P. 4—12. (in Russian)
13. Kozlovskaya I.B. Countermeasures for long-term space flights, lessons learned from the Russian space program // J. Gravit. Physiol. 2002 Jul. V. 9. P. 313—317.
14. Kozlovskaya I.B., Grigoriev A.I. Russian system of countermeasures on board of the International Space Station (ISS): the first results // Acta Astronautica. 2004. V. 55. P. 233—237.
15. Son'kin V.D., Egorov A.D., Zaitseva V.V. et al. Expert system of control of physical training of crew on long-duration spaceflight // Aviakosm. i Ekolog. Med. 2003. V. 38. № 5. P. 41—46. (in Russian)

16. Kotovskaya A.R., Shipov A.A., Vil'-Viliams I.F. Biomedical aspects of creation of artificial force of gravity. M.: Slovo, 1996. 204 p. (in Russian)
17. Buravkova L., Romanov Yu., Rykova M. et al. Cell-to cell interactions in changed gravity: Ground-based and flight experiments // Acta astronautica. 2005. V. 57. P. 67—74.
18. Myasnikov V.I., Stepanova S.I., Salnitsky V.P. et al. The problem of psychic asthenization in long-duration space flight. 2000. M.: Slovo. 224 p. (in Russian)
19. Salnitsky V.P., Myasnikov V.I., Bobrov A.S. et al. Integral evaluation and prediction of cosmonaut's professional reliability in flight // Aviakosm. i Ekolog. Med. 1999. V. 33. № 5. P. 16—22. (in Russian)
20. Kozerenko O.P., Sled A.D., Mirzadzhanov Yu.A. Crew psychological support // Orbital station "Mir". V. 1. M.: IMBP, 2001. P. 365—378. (in Russian)
21. Sinyak Yu.E. Problems of life support of a Martian crew // Piloted mission to Mars (Ed. Acad. A.S. Koroteev) Moscow-Korolev, 2006. P 274—76. (in Russian)
22. Viktorov A.N., Novikova N.D., Deshevaya E.A. Results of microbiological investigations // Orbital station "Mir". V. 1. 2001. M.: IBMP. P. 121—151. (in Russian)
23. Novikova N.D., Deshevaya E.A., Polykarpov N.A. et al. Extermorphylic microorganisms issue of interplanetary transfer on external spacecraft surface // The 36th COSPAR Scientific Assembly, Beijing, China. 16-23 July 2006. ISSN 1815—2619.
24. Gostimsky M.A., Levinskikh M.A., Sychev V.N. et al. Studies of genetic effects in the generation of pea crop cultivated over the whole cycle of ontogenesis in space greenhouse of ISS RS // Genetika. 2007. V. 43. № 8. P. 1050—1057. (in Russian)
25. Petrov V.M., Akatov Yu.A., Arkhangelsky V.V. et al. Scientific radiation physics and biology experiment // Orbital station "Mir". V. 1. 2001. M: IBMP. P. 412—459. (in Russian)
26. Petrov V.M., Mitrikas V.G., Teltsov M.I. et al. Radiation dosimetry in space flights // Model of space. V. 1. Physical conditions in space. Ed. Prog. M.I. Panasyuk. Ch. 3.10. 2007. M.: Izd. Universitet. Knizhnoye delo. P. 642—667. (in Russian)
27. Ostrovsky M.A. Effects of charged particles on the eye structures // Radiation safety for manned mission to Mars. Abstracts for COSPAR Colloquium. 2003. Dubna. P. 59—60.
28. Fedorenko B.S. Radiobiological effects of corpuscular radiations: Radiation safety in space flights. M.: Nauka, 2007. 189 p. (in Russian)
29. Grigoriev A.I., Baranov V.M., Kovachevich I.V., Baranov M.V. Biomedical aspects of simulated missions to Mars. 3rd European Congress Achievements in space medicine into health practice and industry, Berlin, Sept. 28-30, 2005, M.: Copris & M. Ltd. P. 283—285.

Y.B. ZELINSKY
Space Research Center
Polish Academy of Sciences
Warsaw, Poland,
jbz@cbk.waw.pl.
S.K. TATEVIAN
Institute of Astronomy
Russian Academy of Sciences
Moscow, Russia
statev@inasan.ru

CONTRIBUTION OF SPACE TECHNIQUES TO SCIENTIFIC PROGRESS IN GEODESY AND GEODYNAMICS

Plenty of new information about the Earth and the near-Earth environment has been obtained as a result of processing of data of systematic earth-based observations allowing to monitor changes of satellites' orbits. Every artificial celestial body, rocket or satellite irrespective of the fact whether it is equipped with research instrumentation or not, when circling the Earth undergoes drug due to friction in the terrestrial atmosphere, and all irregularities of the Earth gravity field affect its movement, and the Sun, the Moon and other planets have the considerable effect on some orbits. A new branch of science that is satellite (or space) geodesy, born in the beginning of 60s of the last century, allows solving a number of important scientific problems of geophysics, astronomy and geodynamics, which could not be solved by other means.

Application of artificial satellites and spacecraft signifies a qualitatively new stage of development of astronomy, geophysics, geology, meteorology, biology and many other sciences. These new flying research laboratories allow setting and solving problems, which would be impossible before launching of satellites.

Plenty of new information about the Earth and the near-Earth environment has been obtained also by indirect way (as distinct from direct measurements from a satellite's board) as a result of processing of data of systematic earth-based observations allowing to monitor changes of satellites' orbits. Since an artificial satellite circles the Earth and is in its gravitational and magnetic fields, then all results of detailed explorations of its orbital motion have the very direct relation to understand physical and dynamical properties of the near-Earth environment and the planet itself. The first experiments using observations from the Earth's artificial satellites for geodesic determinations were performed as early as in the beginning of the 60s of the last century and gave an impulse to develop a new science named satellite (or space) geodesy.

The important advantage of this new space branch in the field of sciences of the Earth is its global nature in particular. Before 1957, geodesists had available only data of earth-based measurements on comparatively small areas of individual countries or continents (the land constitutes only 30% of the Earth's surface and the water covers the rest 70%). Every country was developing its own reference geodetic network agreed with the nearest neighbors, and therefore, when transferring coordinates for long distances, errors could reach 500 m. The space geodesy allows connecting islands with continents, combining

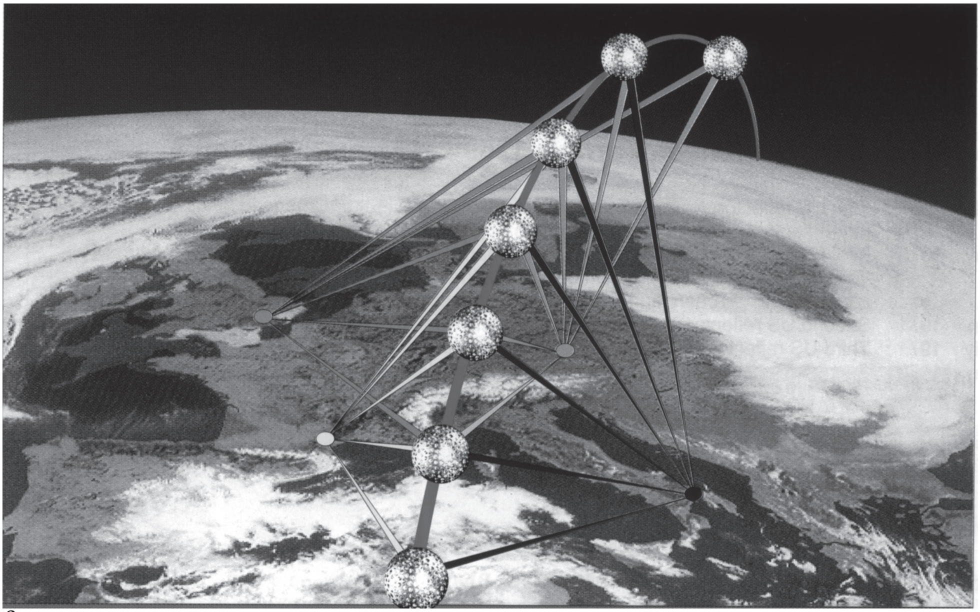


fig. 1

geodetic networks of the continents through oceans and seas, and measuring the Earth's planetary parameters. Theoretical fundamentals and engineering equipment of the space geodesy have been developing very dynamically for the last 50 years. With every new achievement, scopes of application of the results have been widening, both in applied research and when solving fundamental problems of astronomy, geophysics and geology.

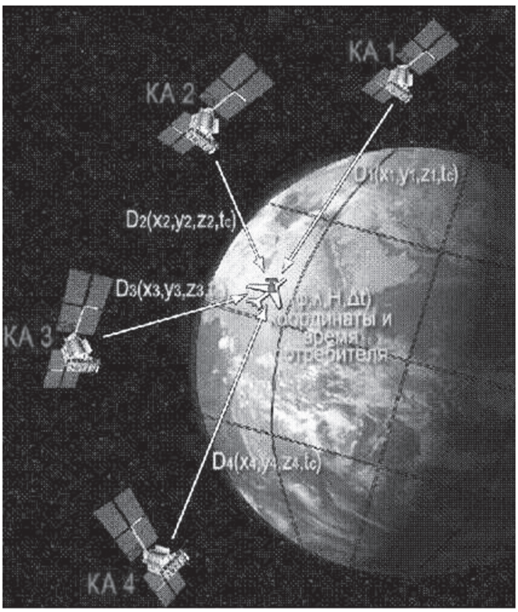


fig. 2

The most considerable increase of measurement accuracy in the space geodesy has been observed for the last 20 years. Laser location of satellites (Fig. 1) equipped with special reflectors has reached centimeter accuracies with fully automated systems of guidance and satellite tracking. Besides, fast development of radio satellite navigation systems such as GPS and GLONASS (Fig. 2) has fully changed the approach to problems of determination dynamic and physical properties of the Earth and allowed increasing accuracies of determination of ground sites geocentric coordinates and their changes in time. An important problem, being solved currently by means of the space geodesy only, is finding and monitoring of short-period variations of the Earth's rotation speed and its orientation in the space, since these parameters determine the Universal Time system and are required to connect the inertial (celestial) frame with geocentric (terrestrial) frame (see table). The International Earth Rotation and Reference Systems Service, regularly, determines and publishes information about the Earth's orientation parameters to within 0.1...0.2 ms of the arc (1 cm) in the pole position and about 0.3 ms in time with resolution of 1 day and less, that is equal to 1.4 cm at the equator. At that, the relative coordinates of the earth-based points and the base lengths at distances of hundreds or thousands kilometers are determined with errors of 1...5 mm horizontally and less than 1.0 cm vertically.

As resolution and accuracy of space and geodesic measurements have been increasing, it has become feasible to record lesser power fluctuations of estimated parameters with higher frequency. Stable oscillations of the Earth's pole movement with periods of weeks up to 1 year have been discovered and correlation dependences of these oscillations with changes of the surface atmosphere pressure have been found. The relation of the Chandler (436-day period) and the annual disturbance of the pole movement with global displacements of atmosphere masses and geologic cycles has been corroborated experimentally. Nutation coefficients (characterizing movement of the Earth's rotation axis in the inertial frame) have been determined to high precision that is necessary to study the Earth's internal structure and free nutation of its liquid external core. Interpretation of long-period fluctuations of the Earth's rotation velocity (day duration) must be probably connected with generation of perfect models of the planet internal structure. These researches go on as observed material is accumulated.

The Earth's gravitation field, before 1957, was known accurate within the third harmonic coefficient of the gravity potential model, i.e. accurate within the Earth's flattening (1/298). The first detailed models of the Earth's gravitation field were constructed per results of photographic and laser observations of satellites using additional satellite radio altimeter and earth-based gravimetric data. The best of these models contain about 150 gravity potential decomposition coefficients.

Years -- Units	σ(X) -- 0.001"	σ(Y) -- 0.001"	σ(UT1) -- 0.0001s	σ(ψδ) -- 0.001"	σ(εδ) -- 0.001"
1962-1967	30	30	20	-	-
1968-1971	25	25	17	-	-
1972-1979	11	11	10	-	-
1980-1983	2	2	3	2	1
1984-1989	.40	.40	.20	.5	.2
1990-2000	.20	.20	.20	.3	.1
2001-2005	.15	.15	.1	.3	.1

Variations of Determination Accuracy of Earth's Rotation Parameters (5-Day Intervals) 1962—2005 Period
(X,Y — the rotation axis pole coordinate, UT1 — the Universal Time,
ψ and ε – precession and nutation of the rotation axis)

This means that the external form of our terrestrial globe has become known hundreds times better than before 1957 due to the satellite geodesy.

Accuracy and reliability of the satellite gravity field models depended on the volume of analyzed data, the number of observed satellites with different orbital

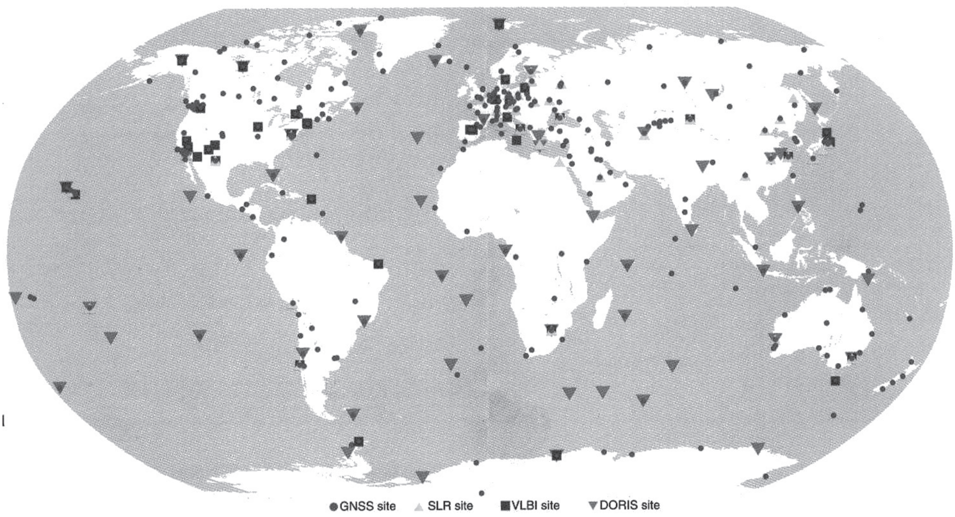


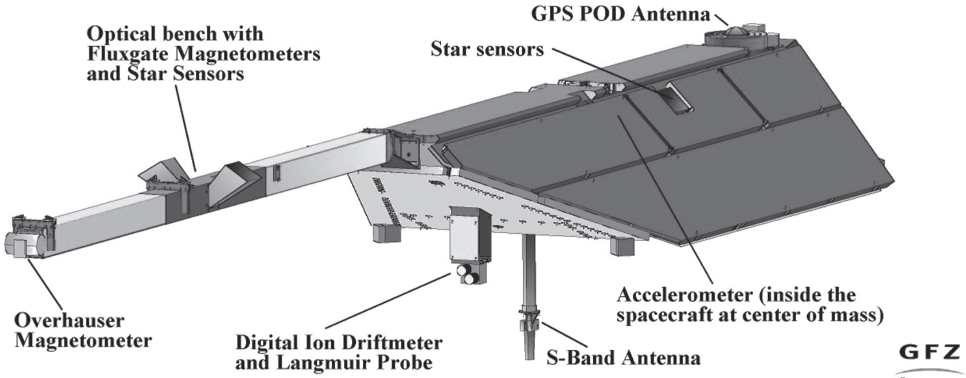
fig. 3

parameters and the observation period duration. However, further detailing of the global models per optical observation data solely is difficult due to a number of reasons. Firstly, when processing data of low-orbit satellites, complexities of considering the atmosphere drag has effect upon changes of the orbit parameters occur. Secondly, nonuniform distribution of observation stations impacts upon the satellite orbit determination accuracy. For instance, density of stations in Europe and North America is considerably higher comparing to Africa and some other regions of the Southern Hemisphere. Furthermore, it should be taken into consideration that geopotential zero-degree coefficient (for the spherical Earth) is three order greater comparing to the following degrees coefficients characterizing local and regional scale gravity anomalies. These coefficients decrease rapidly as the satellite orbit altitudes increase, and therefore, to detail the gravity field model, it is necessary to analyze changes of low-orbit satellites with altitudes of 250...400 km that is impracticable using trajectory measurements by optical facilities.

With the advent of satellite navigation systems such as GPS and GLONASS, it became feasible to run continuous determination of orbit parameters of any satellites, aboard which corresponding receivers are installed. At that, the observation accuracy does not depend on the satellite altitude and the ground stations network configuration. To exclude the atmosphere drag effect upon the orbit parameters changes, high-accuracy accelerometers and 'drag free' systems (excluding the friction-in-atmosphere effect) are installed onboard special satellites.

Application of new satellite observation techniques enables to realize the differential method of geopotential harmonics determination, namely, observations per the 'satellite-satellite' scheme and the gradientometric method. Theoretical justification of these approaches to determine the detailed structure of the Earth's gravity field was done for the first time about 30 years ago, but at that time, there were no engineering equipment accurate enough to realize them. A space project named DIDEX, providing near-earth orbit ascent of two identical satellites being distant from each other at the distance of 100…200 km, was developed in the middle of the 80s by specialists of Poland, Germany and the Soviet Union in the context of scientific collaboration per INTERCOSMOS Program. Unfortunately, political and economical difficulties did not allow realizing this project.

First implementation of the 'satellite-satellite' method succeeded in CHAMP[1] Project in 2000. [1]. The CHAMP satellite has a GPS-receiver and an accelerometer, a satellite orientation and orbit parameters maintenance system, laser reflectors, an orientation sensor, etc. (Fig. 4). The information from this satellite changed the approach to solve the problem of the gravity field determination cardinally. A number of satellites were made per various sets of measurement data. As the result, the Earth's reference surface (geoid) was determined with accuracy of 10 cm and spatial resolution of 350 km (55 geopotential harmonics). Higher 1-centimiter geoid accuracy has resolution of 1,000 km.



CHAMP characteristics	
Eccentricity	0.00368 deg
Inclination	87.3 deg
Local time precession	5.44 min/day
Revolution period	93.5 min
Mass	522 kg
Cross sectional area	0.72 m²

fig. 4

1 http://www-app2.gfz-potsdam.de/pb1/op/champ/

Another successful launch of satellites to determine time changes of the gravity field was in 2002. GRACE[2] Project [2] realized the 'satellite-satellite' method using the newest techniques. One of the main scientific problems of this project is recording of signals of climatic changes induced by large-scale (hundreds kilometers within time intervals of more than 1 month) displacements of the Earth masses. Fig. 5 shows the GRACE Project schematic diagram. Two identical satellites in one orbit are equipped with a distance-measuring system including a microwave K-band distance meter to measure variations of distances between the satellites to high accuracy of 1 μkm/s. These variations depend directly upon the gravity field anomalies forcing the satellites. Moreover, GPS measures coordinates and movement velocities of these satellites. The data flow into receiving stations in Germany and the USA where they are pre-processed and transmitted into research centers for further analysis.

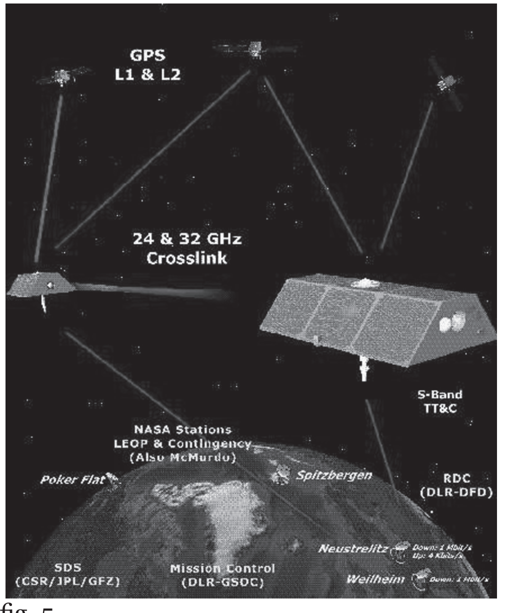


fig. 5

New detailed models of the gravity field allow discovering fine peculiarities of the Earth's tectonic structure, which have not appeared before in global satellite models. These peculiarities are the consequence of various geophysical processes in active tectonic zones of subduction, collision and plate expansion such as, for instance, the Himalayan-Tibetan Region and the Middle-Atlantic Ridge. In so far existing geopotential models, the subduction zones have appeared only as spacious areas with high gravity anomalies due to mass volume increase when shift of the terrestrial crust strata under volcanic arches. More detailed peculiarities of the tectonic structures were discovered only when local gravimetric measurements by means of land-based facilities or from sea vessels.

2 http://www-app2.gfz-potsdam.de/pb1/op/grace/

Currently, the European Space Agency develops the third GOCE[3] Space Project [3] aimed at study of the Earth's gravity field fine structure. Its basis is the principle of gradientometry, i.e. measurement of gravity acceleration gradients. Such technique will be realized for the first time. It is expected, that as a result of this experiment, it will be feasible to construct a model with the resolution of 1°, the geoid determination accuracy of up to 1 cm and the error of $\pm 1 \cdot 10^{-5}$ m/s^2 (1 mGal) in the gravity force anomalies.

The above experiments open new ways to study the Earth's gravity field and its internal structure. It can be the basis for creation of continuous services providing scientists and specialists of many branches of fundamental and applied science with information about the time changes of the Earth's geopotential and dynamic parameters.

To study mechanisms of destructive geodynamic phenomena including determination of places of possible severe earthquakes, volcano eruptions and some other natural hazards , it is important to have means to evolve areas where maximum changes of the displacement velocities and the terrestrial crust vertical movements. Since these displacements appear at the level of centimeters or even millimeters then the measurement accuracy is to be proper one. The experience of the last years has showed that currently the satellite geodesy technique including global navigation systems (GNSS) and laser distance measurements is the most effective (as for accuracy and cost-effectiveness) for research activities in this field [4, 5].

With the help of these measurements, quantitative data have been derived for the first time, which confirm the concept of global and regional tectonic movements and displacements, which are direct indicators of contemporary dynamic processes in the terrestrial crust and the mantle, and which could not have quantitatively evaluated within short time intervals hitherto (Fig. 6).

Permanent control of secular movement (velocities) of GPS-stations of the International GNSS Geodynamic Service located in Russia has allowed improving the reference coordinate frame for North Eurasia since Russian network stations provide representative covering of the largest stable areas (the Siberian and the East European) of the Eurasian plate. Analysis of average values of these velocities shows that the general movement direction of the European part of the continent is north-eastern one [6]. However, as moving along locations of the stations eastward, the northern movement component decreases an, approximately in the longitude of Novosibirsk, the direction changes to the southern one. Movement of the outermost points of the continent (Magadan and Petropavlovsk-Kamchatsky) has strongly pronounced south-western direction, i.e. there is rotation of the Eurasian Continent round the point located within the Himalayas-Tibet massif.

It has been also showed per GPS measurement data that the Eurasian tectonic plate is not a monolithic tectonic block [7]. Along its southern border, there is a

3 http://www.esa.int/esaLP/LPgoce.html GOCE satellite was launched on March 17, 2009 from Plesetsk cosmodrome.

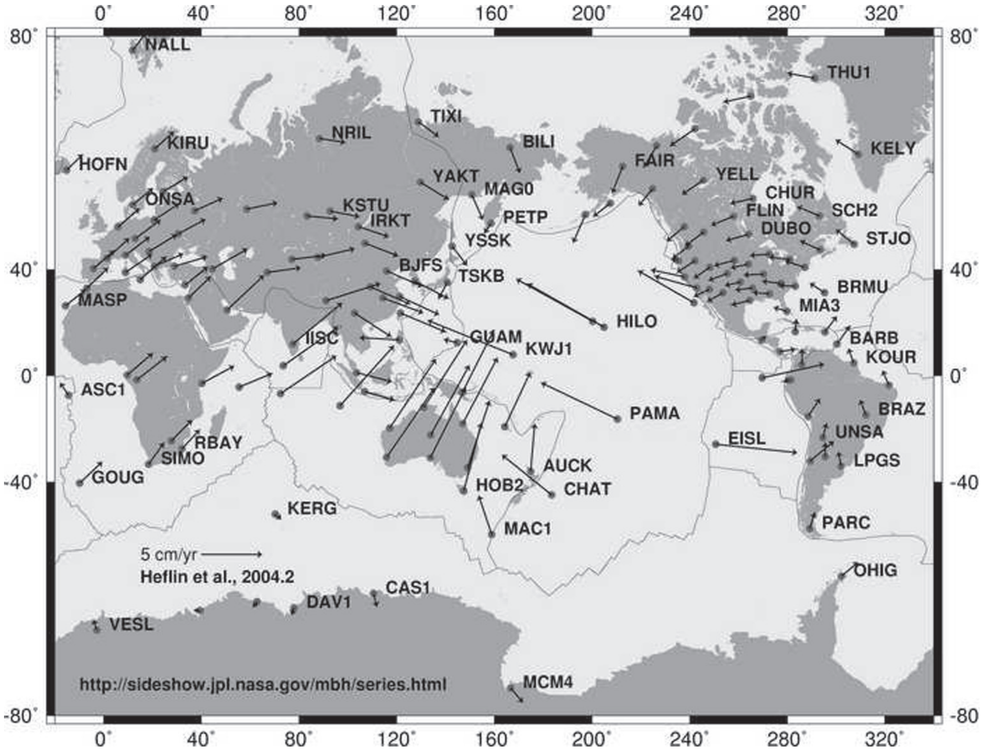


fig. 6

zone consisting of a great number of microplates surrounding the South-Eurasian stable plate. Interaction of these small plates and blocks influences distribution of seismic stresses in internal parts of the continent that is confirmed by the highest seismic activity of the triangle bordered by thrusts of the Himalayas and faults of the Pamirs, the Tien Shan, the Baikal and the North-Eastern China.

Given in this review results refer to just a small part of geodynamical and geophysical researches, for which the satellite geodesy gives unique measurement data. This includes study of the World Ocean level changes and the ground water volume fluctuations that is directly related to global change of the planet climate. The satellite data are also used for monitoring and forecasting of the volcanic activity and control of the Earth's crust movement in seismically hazardous regions. We could say that it has become feasible for the first time to implement the generalized approach to study the Earth dynamics and physics as an integrated planetary system.

REFERENCES

1. Reigber Ch., Luhr H, Grunwaldt L., Foerste Ch, Koenig R., Massmann H, Flack C. CHAMP Mission 5 years in orbit – Observation of the Earth System from Space /Ed. J. Flury et al., Springer Verlag, 2006. P. 3 —16.
2. Schmidt R., Flechtner F., Meyer U., Reigber Ch., Barthelmes F., Foerste Ch., Stubenvoll R., Koenig R., Neymayer K.-H., Zhu S. Static and time-variable gravity from GRACE Mission data — Observation of the Earth System from Space /Ed. J. Flury et al., Springer Verlag, 2006. P. 115—130.
3. Wermuth M., Rummel R., Foeldvary L. Mission simulation and semi-analytical gravity field analysis for GOCE SGG and SST — Observation of the Earth System from Space /Ed. J. Flury et al., Springer Verlag, 2006. P. 193—208.
4. Tatevyan S.K. The main requirements to the earth-based sector of the Time-Coordinate Frame in respect of fundamental geodynamical tasks // Trudy Instituta prikladnoi astronomii RAN. 2005. Iss. 13, P. 81—92. (in Russian)
5. Kuzin S.P., Tatevyan S.K. The use of GPS-measurements for studies of Earth's crust motion // Kosmonavtika i raketostroenie. 4 (37). — Korolev: TsNIIMASh, 2004. P. 37—46. (in Russian)
6. Outkin V.I. About movement inheritance of the Eurasian continent // Proceedings of the "APSG-Irkutsk, 2002" seminar. Moscow, GEOS, 2002. P. 15—20.
7. Gatinsky Yu.G., Rundquist D.V. Geodynamics of Eurasia: Plate Tectonics and Blocks Tectonics // Geotectonics. 38(1). 2004. P. 3—20.

NOVEL SPACE TECHNOLOGY

G.M. POLISHCHUK
S.A. Lavochkin Resarch /
Production Association,
Khimki, Moscow region, Russia

PERSPECTIVE OF RUSSIAN SPACE ACTIVITIES FOR SCIENTIFIC RESEARCH

In October, 2007 all Earth inhabitants celebrated the 50th anniversary of Space Era beginning. During this quite short period the mankind has made a huge progress in space exploration and effective space missile technologies development. Modern life and fundamental science are impossible without usage of spacecraft (SC). The most effective ones for this purpose are automated SC. For leading space countries such as Russia, USA, France and others, as well as for European Space Agency carrying out fundamental and applied scientific research beyond the Earth and its atmosphere is one of the main priorities of their scientific activity.

In Federal Space Program of Russia for 2006—2015 these activities hold a considerable place. For more than 40 years Lavochkin Association has been a prime enterprise of space rocket industry in Russia on creation of key automated SC for space studies. The scope of studies includes: planets; the Moon; small bodies of Solar system; astrophysics; solar physics and physics of cosmic plasma, as well as solar-terrestrial relationships.

In February, 1966 Lavochkin Association launched *Luna-9* SC which performed first soft landing on the Moon, and later this year in April station *Luna-10* was launched from the Earth (first artificial satellite of the Moon). Up to *Luna-14* SC launches were performed using *Molniya* launch vehicle. Then, having switched to more powerful *Proton* launch vehicle, Lavochkin Association developed more sophisticated SC of next generation, designed for performing more complicated tasks: lunar soil samples return; in-situ studies of various surface regions using mobile remote-controlled Moon rovers, as well as studies of the Moon and near-Earth space from the Moon satellite orbit. Task of lunar soil samples return was solved by *Luna-16, -20, -24* SC, mobile autonomous laboratories *Lunokhod-1* and *Lunokhod-2* were delivered to the place of *Luna-17, -21* SC operating. While creating the above mentioned SC at our enterprise, principle of objects construction on basis of unified space modules (platforms) was used for the first time. Orbital-landing unit was developed — multipurpose landing platform. It delivered various payloads to the Moon or near-Moon space: Moon rovers, ascent rockets with soil sampling devices and return vehicles; artificial satellites of the Moon for remote sensing of the Moon and etc. Corrective-braking propulsion system was also unified.

Such approach provided the best value of criteria 'cost-efficiency' of the project, proved by real lunar missions. Therefore it was widely used while designing the subsequent automated SC for Mars, Venus, solar-terrestrial relationship exploration.

Mars-2, *-3*, *-4*, *-5*, *-6*, *-7* SC were successfully launched and SC *Mars-3* for the first time performed soft landing on the planet under study. These SC were created on the basis of the same space platform with necessary modifications for specific objectives.

The most ambitious project in this field was *Mars-96* SC which was designed for complex and comprehensive research of the planet and near-Mars space. This SC was supposed to land on Mars small long-life scientific stations, penetrating probes (penetrators) and carry out research from Mars satellite orbit. However, due to failure of booster it was not injected into interplanetary flight trajectory.

A significant modern scientific problem is a search of relict matter, while the most probable possessors of this are Mars satellites Phobos and Deimos. For this purpose SC of new generation *Phobos-1*, *-2* were developed. Scientific program of the mission, besides comprehensive study of Phobos, provided exploration of Mars, the Sun and space. That was quite successful mission during which rather efficient studies of the planet and its natural satellite were done, although landing failed.

The most successful research of Venus was conducted, and the most significant contribution to that was made by landers of *Venera-9*, *-10*, *-11*, *-12*, *-13*, *-14* SC. They transmitted to Earth panoramas of the planet surface, results of chemical analysis of atmosphere and soil, measurements of temperature and pressure during reentry, wind speed on the surface. While implementation of above mentioned missions transistance in the atmosphere was determined, gamma- and X-ray radiations of the Sun and the Galaxy, characteristics of solar wind, cosmic rays, interplanetary plasma were studied. Using special device soil samples were taken, put inside of descend vehicle and their chemical analysis was made.

Venera-15, *-16* SC performed radar mapping of the planet covered by dense layer of clouds and map of the surface was obtained.

Later on *Vega-1*, *-2* SC a number of unique space experiments have been carried out. Venusian atmosphere circulation was studied, its meteorological parameters and cloud layer using balloon probes, also for the first time comet Halley was examined from short distance. The objectives of this study were: determination of physical characteristics of the comet nucleus; analysis of structure and dynamics of comet near-nucleus area; determination of composition of dust particles, their mass distribution at various distances from nucleus and others. *Venera-9*, *-10*, *-11*, *-12*, *-13*, *-14* and *Vega-1*, *-2* were designed on the basis of one space platform of descent vehicle with modification for target scientific tasks.

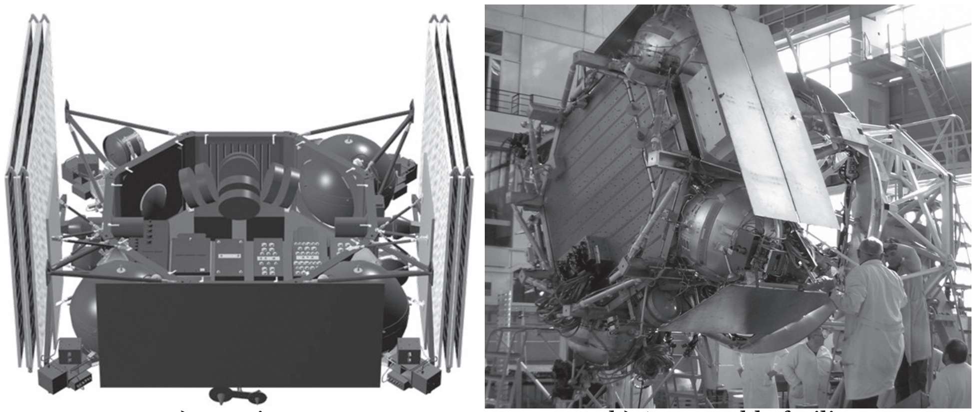
Solar and solar-terrestrial relationships research takes up a strong position in scientific plans of leading space countries of the world. In Russia satellites of *Prognoz* series, developed by Lavochkin Association on the base of the one and the same orbital platform, take special role in research of this kind. This scientific program was successfully realized by 10 satellites. The main scientific objectives were the study of solar-terrestrial relationship, solar activity, Earth ionosphere and magnetosphere, cosmic plasma and Earth plasmasphere, as well as solar wind and others. Further *Interball-1*, *-2* SC provided conducting of unique studies by four spaced-apart satellites, allowing simultaneous observations in different regions of Earth magnetosphere.

Astrophysical and astronomical observations also play significant role in fundamental space exploration. Lavochkin Association have developed and injected into space first domestic full-fledged astrophysical observatory, called *Astron*. Ultraviolet and X-ray telescopes were installed on it. Ultraviolet space telescope was called *Spika*.

In several years we developed and injected into orbit a new space observatory *Granat* with complex of telescopes providing detailed studies of celestial sphere in widest range of spectrum of electromagnetic radiation — from 2 keV up to 100 MeV. Among most important results of these observatories are discovery of flaring X-ray pulsar, monitoring of rotation periods changes of considerable number of X-ray pulsars, detection of tens of X-ray bursts from neutron stars and many others. Above mentioned astrophysical observatories were designed on the basis of orbital platform of *Venera* SC series.

A significant advantage of all mentioned space projects was that flight control of these SC was performed by specialists of Lavochkin Association. Creation of several basic space platforms for several series of SC for various purposes allowed minimizing period of creation, their cost and improving reliability of operation. The same principles are used at Lavochkin Associations while developing all state-of-the-art and perspective space complexes.

Presently all space developed countries pay considerable attention to space fundamental research in their national programs. In the Federal Space Program of Russia till 2015 which is under implementation now a significant position is assigned to creation of space means for conducting of fundamental scientific studies. In coming years *Spektr* projects are to be realized, that include development of space observatories, operating in radio, ultraviolet, X-ray and gamma radiation bands. For these SC and perspective Earth satellites a new orbital platform *Navigator* was built, combining the last scientific and technical developments of service systems. General view of this platform is shown in Fig. 1.



a) Drawing b) At assembly facility

fig. 1. General view of orbital platform *Navigator*

The *Navigator* platform includes housekeeping systems necessary for SC control: on-board control complex, radio complex, power supply system, propulsion system. Structurally it is a octahedral prism inside of which all housekeeping instrumentation are placed on thermo stabilized panel. Outside on its facets propulsion system units, solar battery panels, antennas are placed.

Main characteristics of the platform:

Attitude control:

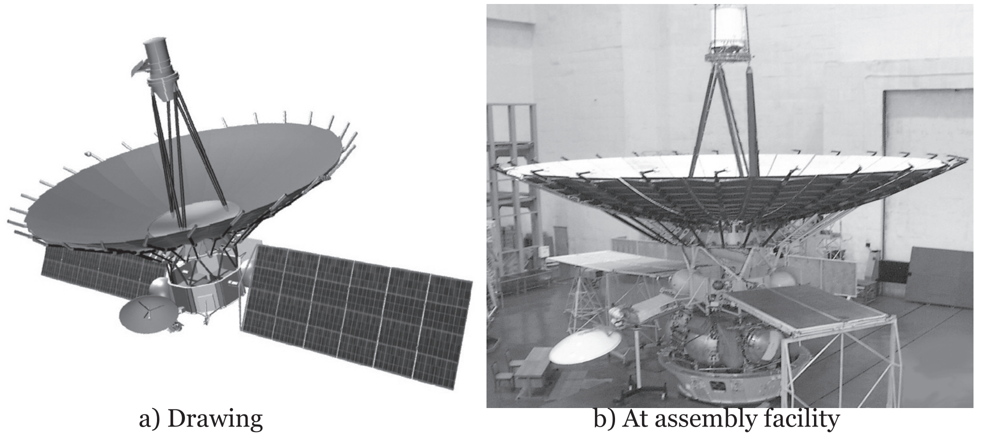
- type of possible orientation   triaxial
- accuracy of axes pointing   1…2 arcmin
- velocity of reorientation   0.25 degree/s
- amplitude of stabilization   ±2.5 arcs
- velocity of stabilization   0.36 arcs/s
- executive devices   complex of flywheel engines

Correction maneuver:

- allowable duration of one burn of engine   1,800 s
- accuracy of generation orbit correction impulse   0.5 degree

Mass of platform (dry)   650…850 kg
Maximum filling (hydrazine)   700 kg
Maximum mass of payload   2,500 kg
Maximum power supply of payload   1,500 W

Currently ground testing of space complex *Spektr-R* is under completion. The main objective of the project is creation of ground-space system consisting of a 10-meter parabolic antenna deployable in space and ground network of the biggest radio telescopes of the world. Space radio telescope is a high-precision structure consisting of central mirror and 27 blades. Such ground-space system (radio interferometer) will be equivalent in terms of resolution capacity to hyper radio telescope with diameter of antenna equal to distance between space and ground objects and will amount to of the order of one-hundred-thousandth of angular second. The task of the experiment will be detection and study of compact radio sources in Universe: quasars, active galactic nuclei, etc. General view of *Spektr-R* space complex is shown in Fig. 2.



a) Drawing   b) At assembly facility

fig. 2. General view of *Spektr-R* space complex

fig. 3. General view of space complex *Spektr-UF*

One of the projects to be completed in the near future is space complex *Spektr-UF / World Space Observatory*. It's main payload is a telescope designed for observing the Universe in ultraviolet band. This is a large-size space telescope with size of primary mirror of 170 sm. Russian telescope will be capable 'to see'space bodies with radiation by 20-times weaker than those observed by the most powerful now Hubble Space Telescope.

The main objectives of this project are: studies of physical and chemical properties of planetary atmospheres and comets; research of physics of hot stars atmospheres and chromospheric activity of cool stars; examination of characteristics of dust particles of interstellar and near-Earth matter; studies of active galactic nuclei nature; determination of ratios of light elements and their isotopes content essentially important for selection of cosmological model; studies of intergalactic gas clouds and gravitational lens, etc.

Spektr-UF space complex consists of multipurpose orbital platform *Navigator* and space ultraviolet telescope T-170 as a payload. Launch is scheduled for 2010. General view of space complex *Spektr-UF* is shown in Fig. 3.

After successful operating of *Granat* space observatory astrophysicists realized the necessity of transition to X-ray oblique incidence optics allowing increasing of telescopes sensitivity by orders of magnitude compared to instruments with coded aperture. Works on a project of that kind are being executed at Lavochkin Association. It is called *Spektr-RG* (*Spektr-X-ray-Gamma*). This is international astrophysical observatory designed for exploration of the Universe in gamma and X-ray bands.

This SC will for the first time carry out full survey of celestial sphere with record sensitivity and quite high angular and energy resolution in high energy range. *Spektr-RG* observatory will allow detecting and studying of hidden population of hundreds of thousands of supermassive black holes with strong absorption in soft X-ray band. According to preliminary estimate this would make good progress in development of cosmology.

Besides above mentioned, the tasks of the project are: research of radiation variability of supermassive black holes; constant observations of sources with weak X-ray luminosity; complex studies of gamma-ray bursts and their X-ray afterglows; observation of supernova outbursts with research of their evolutions; studies of black holes and neutron stars in our Galaxy; localization of hard X-ray emission from extended objects, etc.

Once celestial survey is carried out it is planned to research selected sources during quite long time, in particular, to measure the temperature of clusters of galaxies, spectroscopy and timing analysis of galactic and extra-galactic X-ray sources.

Spektr-RG SC, like previous SC, is being developed on the basis of *Navigator* orbital platform with scientific instrumentation complex installed on it. Launch is scheduled for 2011.

General view of space complex *Spektr-RG* is shown in Fig. 4.

In Federal Space Program considerable attention is also given to development of automated SC for research of the Sun and solar-terrestrial relationship. Necessity of research of our star from short distance (order of 30...40 solar radii) goes without saying. Currently such activities are being performed at NASA and ESA. In Russia project *Interhelioprobe* has been approved for realization. This SC is planned to be sent to the Sun using gravitational maneuver near Venus. Scientific objectives of the project are: study of problems of solar corona heating and research of solar wind parameters from short distances from the Sun; studies of origin of solar flares and coronary discharges which are sources of space weather and etc.

Continuation of researches conducted by SC *Interball-1*, *-2* is project *Resonance* which now is in the phase of development. It envisages the creation of space system consisting of four small satellites operating on magneto-synchronous mutually

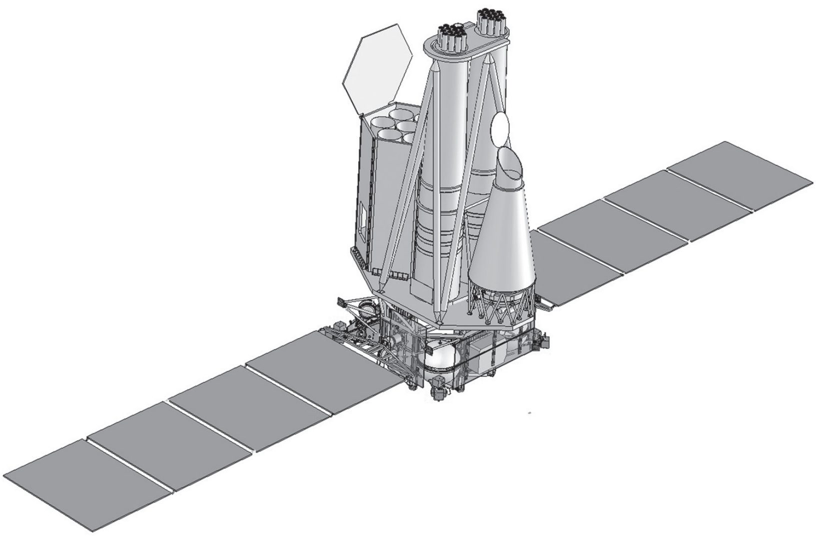


fig. 4. General view of space complex *Spektr-RG*

complementary orbits. Initial parameters of these orbits are selected providing simultaneous positioning of SC in active flux tube adjoint with ground station.

The main objectives of this mission are: long-term observation of natural phenomena in selected flux tube of magnetic field; examination of dynamics and modes of magnetic cyctotron maser; formation and decay of ring current; filling of plasmasphere after magnetic perturbations; role of small-size phenomena in global plasma dynamics and others.

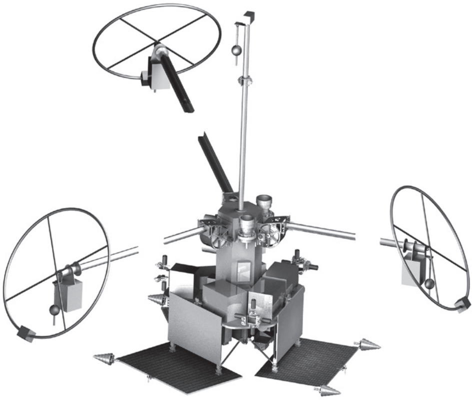
General view of *Resonance* SC is shown in Fig. 5.

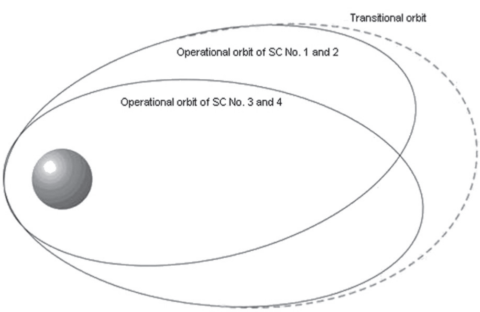
Throughout all periods of space era exploration of the Moon and Solar system planets held key position in programs of scientific space research of leading countries. At present this field of research is also given consideration.

For all times, Mars has been an object of earthlings' great attention. In recent years the world scientific community displays growing interest towards this planet (NASA's missions *Mars Odyssey*, launched in 2001; *Phoenix*, launched in 2007; ESA's *Mars Express*, launched in 2003, etc.). This is a result of new data obtained on it, new phase of space technologies development and extended tasking of fundamental space exploration. Exobiological research on the planet surface is still of interest, although results of Mars remote sensing performed are not too expiring.

Presently space complex of high priority *Phobos Sample Return* is being developed by our enterprise with wide scientific and industrial cooperation. This project is implemented under Federal Space Program of Russia and at the same time it is the first phase of perspective program of Mars exploration using automated space means for the period till 2020 proposed by Lavochkin Association. In our opinion, for coming 2—3 decades the above mentioned tasks can be solved most effectively only by automated space means.

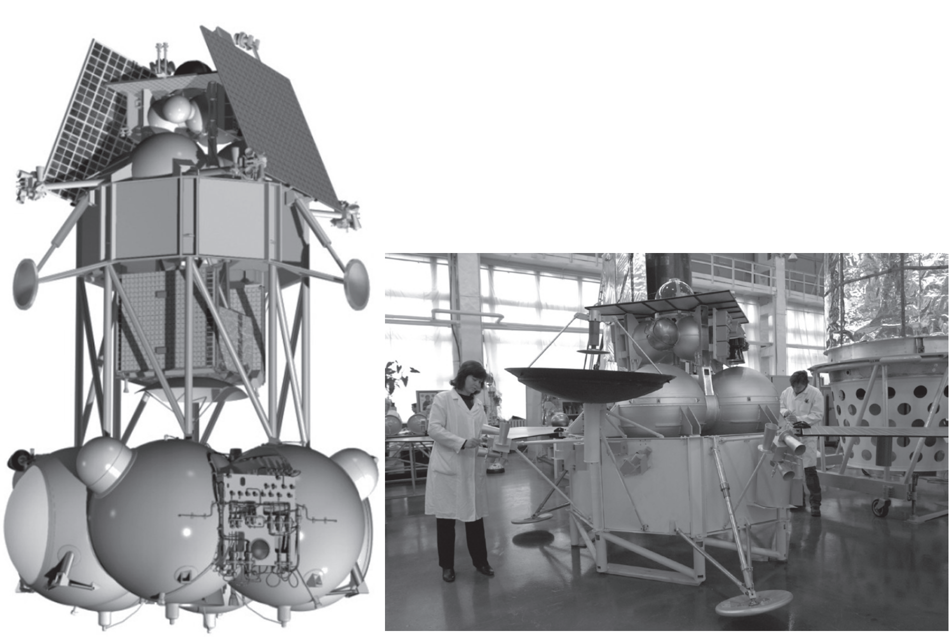
Exploration of the Martian satellite Phobos, classified as potential carrier of relict matter from which Solar system planets were formed, is one of the most critical tasks of modern planetology. The main objective of this mission is Phobos soil





a) General view of *Resonance* SC b) *Resonance* mission profile

fig. 5. General view and mission profile of *Resonance* SC



a) Drawing b) Design test article
fig. 6. *Phobos Sample Return* space complex general view

samples return to the Earth. Moreover, during interplanetary flight and after landing on Phobos, about 20 scientific experiments are planned to be carried out. That includes investigation of soil properties (spectrometry, determination of its optical, mechanical characteristics and others), Mars remote sensing. Scientific studies will begin from the moment of station approaching to the planet. Mars studies will be conducted from satellite orbit for selection of perspective regions for following-up in-situ investigations.

Once Phobos soil samples have been taken they are put into return capsule and pressurized. After take-off of return vehicle the lander with scientific instruments will be left on the surface of Phobos for further long-term research. Also accommodation of piggy-back payload on-board of SC is provided for. This is Chinese automated station designed for Mars exploration from its satellite orbit.

General view of space complex *Phobos Sample Return* is shown in Fig. 6. The SC flight profile is shown in Fig. 7.

Phobos Sample Return space complex consists of the following: cruise and return vehicles, descent module, landing and soil sampling devices, propulsion systems, scientific and service equipment. The launch is scheduled for October, 2009. The mass of space complex is ~8,100 kg.

At next phases of exploration program of Mars and near-Mars space in-situ methods prevail. The second phase of the mentioned program is delivery of small research stations to the specified segmental surfaces to perform large-scale scientific research. The project name is *Mars-MetNet*. The launch is

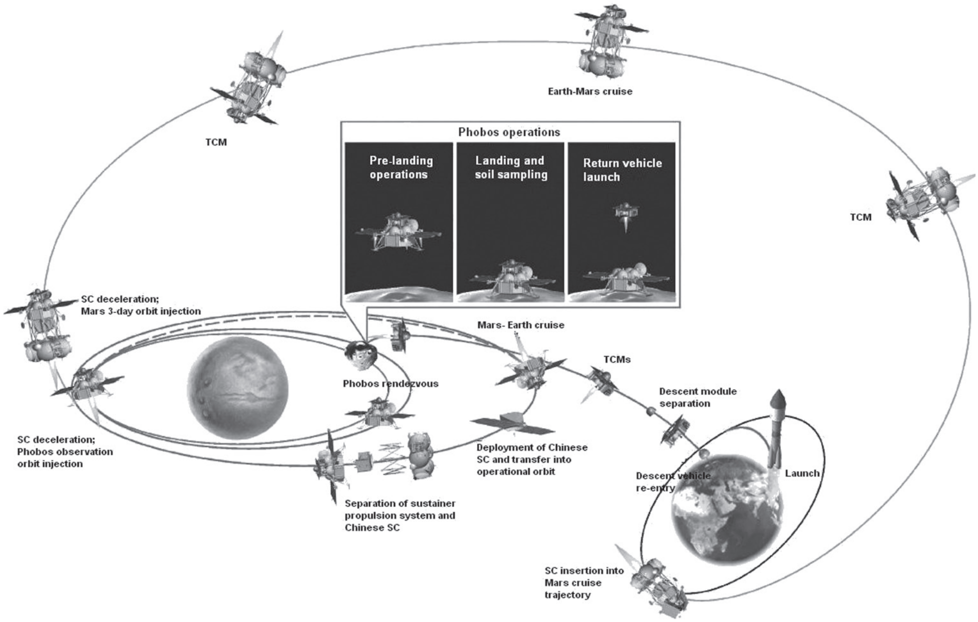


fig. 7. *Phobos Sample Return* mission profile

preliminary scheduled for 2016. The *Mars-MetNet* space complex general view is shown in Fig. 8.

The third phase of the above mentioned program is *Mars Sample Return* mission (a logical continuation of *Phobos Sample Return* mission). Like missions of the first and the second program phases this is the next generation mission for planetary exploration. The mentioned mission will provide the delivery of soil samples from the Mars surface, search of water and biological compounds as well as thoroughly planet exploration and reconnaissance for further manned missions.

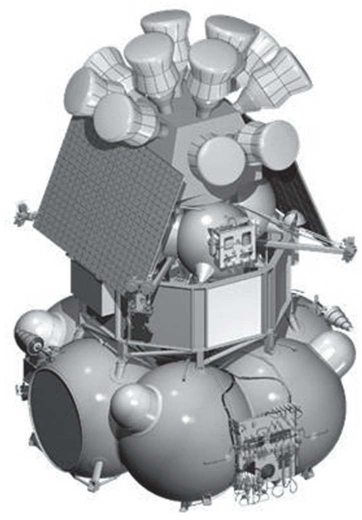


fig. 8. *Mars-MetNet* space complex general view

By present time Lavochkin Association has finalized pre-design activities for creation of scientific and technical grounds for the aforesaid Mars missions' implementation. Mars rover demonstrators have been developed and manufactured, preliminary studies on delivery system development and testing for payload delivery to Mars surface have been done, system analysis of proposed space complex design in order to identify the most efficient variant has been performed.

The conceptual design of SC for Mars missions is combination of baseline and unified modules, such as baseline orbital platform, sustainer propulsion system, landing platform, etc. For this purpose the most efficient are platforms and modules developed in the course of *Phobos Sample Return* mission. *Mars Sample Return* SC general view is presented in Fig. 9.

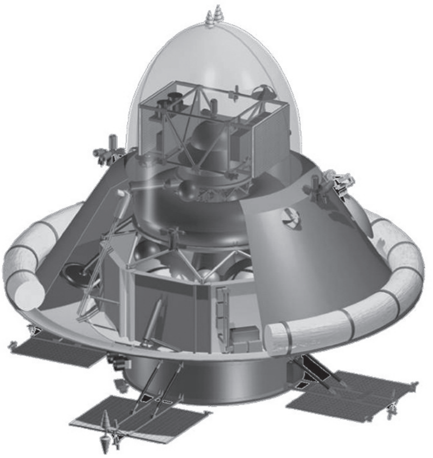


fig. 9. *Mars Soil* SC general view

After a long time gap (over 20 years) Venus exploration was renewed. Today this work is performed by *Venus Express* (ESA, launched in 2005) — the European station with Russian participation: SPICAV/SOIR spectrometer and PFS spectrometer were made with Russian participation.

New quality is a characteristic feature of the renewed Venus studies. Foreseen by Federal Space Program is the development of *Venera-D* SC. At present a preliminary development of a new Venusian mission is being elaborated by experts of our enterprise. That would be a space complex with large-scale capabilities of orbital, atmospheric and surface studies equipped by high-end scientific instruments. Goals of the study are as follows: atmospheric composition and thermal structure; nature, composition and optical parameters of clouds; atmospheric dynamics and zonal circulations dynamics; atmospheric lightning, etc. Studies mentioned will be performed from the orbiter.

Also in the frame of the mission aerostat probes and descent vehicle will be delivered to Venus. Scientific equipment of aerostat probes will study meteorological parameters and chemical composition of the atmosphere as well as clouds chemical composition, and temperature of planet's surface (at night side).

Descent vehicle will provide study of surface composition at landing place (with possible drilling), surface images acquisition during descent, seismicity studies, acoustic signals determination (lightings and thunderstorms), etc.

It is assumed that orbiter will be operated not less than 3 years, and descent vehicle — not less than 1 month. As compared to those landers of previous Venus missions were operated on the planet surface about 1 hour.

By preliminary estimation the space complex mass while rendezvous to Venus will be ~1,900 kg and descent module mass ~1,100 kg.

At present all leading space countries include Moon exploration in their national perspective programs. It is motivated by expediency of applied research devoted to lunar recourses development and construction of both extraterrestrial bases for manned flights to Mars and polygons for scientific purposes, in particular, for astrophysical studies. This line of activities is included into the Federal Space Program of Russia. The Moon exploration program, like those described above, is assumed to be performed on 'step-by-step' basis.

The first national project of the abovementioned program, called *Luna-Glob*, is currently under development at Lavochkin Association. The mission envisages study of planet's inner structure (in particular, to prove the existence of lunar nucleus and under its presence – to determine its dimensions), search of water in the shaded craters at lunar pole, delivery of high-speed penetrators to be introduced into subsurface layers and study of impact of passing corpuscular fluxes and electromagnetic radiation on the Moon (LORD experiment).

Probes, to be penetrated into lunar surface, are separated from the orbiter, which is located in low elliptical orbit (H_π=25 km, H_α=200 km). Two of them are dropped and penetrate at equatorial area of the lunar side faced to the Earth and the other two are dropped at the same area but at back of the Moon symmetrically to the first two. After that orbiter executes a maneuver and transfers into operational polar circular orbit (H=500 km) and carries out the LORD experiment up to the end of its active lifetime.

SC wet mass is ~7,100 kg, SC mass during Moon cruise is ~2,050 kg, active lifetime is 3 years. General view of *Luna-Glob* space complex is shown in Fig. 10.

Similarly with the above-named SC, the *Luna-Glob* SC is developed using modular approach based on usage of the most up-to-date but already proven design and technological approaches of *Phobos Sample Return* mission. In our opinion the project implementation is feasible in 2010—2011, provided that Japanese penetrators are used.

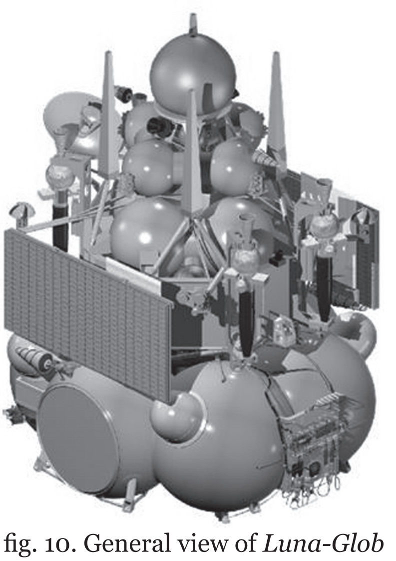


fig. 10. General view of *Luna-Glob* space complex

The extension of this part of the program is the study of the Moon polar areas by means of mobile laboratory — Moon rover. This is a joint Russian – Indian project. In frame of the project we are developing a lander and a Moon rover. This delivery to Moon is planned to be performed by means of Indian space rocket GSLV and orbita-cruise module.

Mass of lander with Moon rover is ~1,200 kg, Moon rover mass is ~60 kg.

Control of Moon rover operation on the Moon surface is performed by means of remote control system from the Flight Control Center. SC operation control can be realized both directly from the Earth (in direct radio visibility environment) and via orbiter as well. Implementation of the mission is scheduled for 2012.

The next phase is *Moon Sample Return* project. Unlike the previous Moon soil sampling missions, current soil sampling shall be performed at strictly defined areas. Those priority areas will be defined by the *Luna-Glob* space complexes during surface reconnaissance and radio beacon installation. Today the most interesting areas are polar areas and the back side of the Moon. In order to solve this problem the following development is foreseen: unified landing platform, Moon rover with large radius of operation, takeoff rocket, return vehicle, soil sampling complex with loading and storage facilities and navigation device to provide precise landing on the radio beacon.

The pre-design study has shown that it is expedient to realize the stated project by two missions. During the first one at given research area the heavy Moon rover with large operational radius will be delivered; the Moon rover will be equipped by special complex for soil samples selection and initial analysis, their loading into return vehicle of the takeoff rocket. During the second mission the takeoff rocket will be delivered to lunar surface.

In order to provide Moon rover operation at the back side of the Moon it is necessary to include lunar satellite-retransmitter in the first mission.

Besides the lunar soil samples return there is another goal for the second mission: verification of technical means for subsequent missions intended to create the lunar research polygon. That is final phase of the above-said unmanned Moon exploration program.

Development and operation of the lunar polygon is of undoubted interest for both fundamental and applied sciences. It is expedient to verify there the technology of minerals mining and processing prior the human colonization of the Moon, to embody the dream of scientists about creation of next generation radioastronomical observatory. The enquiry is that space radiation less than 10 MHz is completely screened by the Earth ionosphere. Moon without its own ionosphere is very convenient ground for astrophysical research in that range, search for exoplanets and many other tasks. And if the radio telescope is installed at the back of the Moon, then the Earth impact and the Sun radiation will be excluded to prevent the observation interference.

Recent progress of space technology made feasible flights to Solar system periphery, towards giant planets and their natural satellites, some of which have atmosphere that is particularly interesting to study. Solution of mentioned tasks as well as flights to small bodies of Kuiper belt, to asteroids (the *Asteroid* project), delivery of cometary nucleus matter samples to the Earth (*Comet Sample Return* project) undoubtedly relay to long-run prospects beyond 2015. Though, currently our western partners have started practical development of similar missions (NASA's NEAR 1996—2001. ESA's *Rosetta* mission, launched in 2004, etc.).

Small bodies of the Solar System (Martian satellites, asteroids and comets) attract special interest of scientists as well. Since all planets have undergone significant changes during their lifetime due to external and internal factors, their matter differs a lot from the firstborn matter. One can say with high degree of probability that the mentioned small bodies consist of initial firstborn matter of protoplanetary cloud, from which all Solar system bodies were formed.

By present time Lavochkin Association has fulfilled pre-design activities and systems research to lay scientific and technical groundwork for future deep space missions. That work generally based upon design and technological approaches gained during the development of *Phobos Sample Return* space complex. It is efficient to take *Phobos Sample Return* module as a basis for development of space platform providing delivery of scientific payload to Solar system small bodies. But, beyond the doubt, it will be upgraded due to considerably longer active lifetime.

Analysis of previous missions, including foreign (*Pioneer*, *Voyager*, *Cassini* etc.) has shown that the most essential design requirement is mass minimization due to considerable power inputs during the flight to Solar system boundary. The trajectory shall be selected in conformity to flight duration no more than 10—12 years, since time factor is critical both for the equipment as well as for organizers teams and scientific support of such kind of mission.

Throughout the history of national cosmonautics Lavochkin Association succeeded in performing tasks of that kind. In 1986 the *Vega-1*, *-2* SC were launched and performed successful studies of Halley comet transmitting unique scientific data to the Earth; in 1988 the *Phobos-1*, *-2* SC were launched. Late in 1980‘s Russian-French SC was developed for Vesta asteroid studies and its scientific and design heritage underlay the *Mars-Aster* project (1995 — preliminary design).

When developing the SC for comet and asteroid studies one shall both proceed from verified scientific and technical approaches and foresee perspective target systems, service systems, equipment and perform survey work. It is necessary for equipment to operate not less than 15 years in open space environment but currently it is rather problematic.

At present we perform pre-design development of space complexes for exploration of Europa (the Jupiter satellite), comets and asteroids. Design approaches for the development of above-said SC are based on design and technological heritage as well as some modules of *Phobos Sample Return* project which is assumed as a baseline.

Europa, the Jupiter satellite is a large celestial body with dimensions compared with Moon size, with atmosphere. It is covered by water ice under which there is a possibility to find considerable volume of water and may be some primitiv life forms. Therefore Europa is of obvious interest for exobiological research. Also it is very important to perform scientific experiments from the Europa's artificial satellite orbit, landing on ice area, studies of ice samples by means of rover and study of surface from landing platform.

The *Jupiter-Europa* space complex consists of the following: orbiter and lander with rover. Currently the opportunity of the project implementation jointly with European Space Agency is under consideration. The general view of proposed lander with rover is presented in Fig. 11.

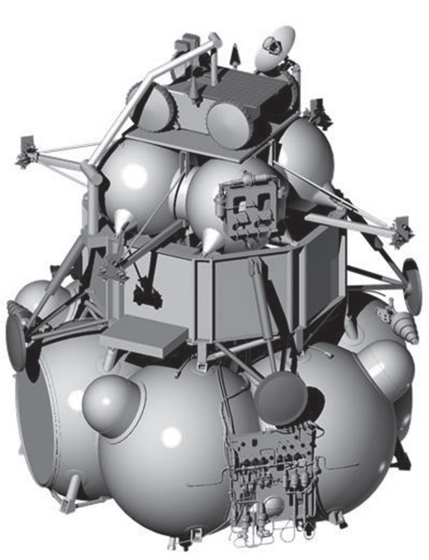


fig. 11. General view of lander of the *Jupiter-Europa* SC

As was mentioned above, comets and asteroids are of a significant interest for fundamental science research. Design and survey works are conducted today by Lavochkin Association for the purpose of development of automated SC for the activity implementation.

The *Asteroid* automated SC is designed to conduct wide spectrum of studies of Solar System small bodies. It shall approach the target asteroid and scientific equipment will photograph the asteroid from different angles, transmit images to the Earth, estimate the physical state and chemical composition and perform other investigations.

The basis of pre-design development is cruise module and insertion propulsion system designed for *Phobos Sample Retur*n space complex, as well as European *Cassini* SC heritage. Also one shall take into consideration special requirements, distinguishing it from interplanetary SC, that are: remoteness of asteroid belt from the Sun, mission duration, and usage of gravitational maneuvers near the planets during the flight. It is efficient to use radioisotopic thermoelectric generator as a power supply.

Based upon this SC design it is logical to develop SC for comets studies. It is proposed to fly to the comet, take soil samples and deliver them to the Earth. By means of the stated SC a complex of remote and contact investigations of the comet will be conducted.

When developing the SC one shall take into consideration a significant amount of high-speed micrometeorites in the concerned SC environment. In order to increase its survival capacity under the stated environment it is necessary to foresee meteorite and dust protection in the design. Special screens will protect the onboard systems and equipment from damage during high-speed

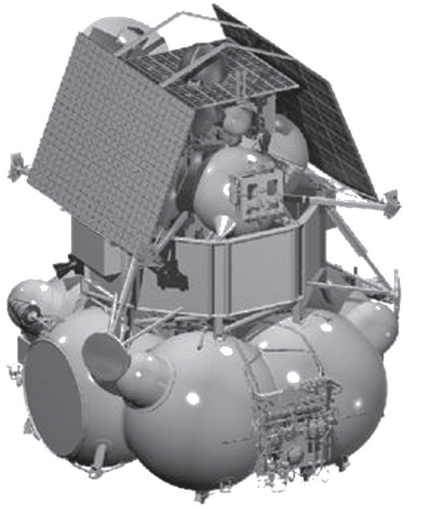


fig. 12. The *Comet Sample Return* Soil SC general view

coma particles attack. The general view of the *Kometa Sample Return* SC is shown in Fig. 12.

At present and for near-term outlook quite promising space activities are fundamental researches by means of automated low-sized (small) SC. A capability of such SC development obviously demonstrates the enhanced manufacturing capabilities of modern cosmonautics. The most attractive factor is their price affordability that makes feasible the organization and conducting of some space experiments by relatively low-budget organizations, for example universities for scientific and educational purposes.

Besides the price factor there is a variety of reasons in favour of development and implementation of small SC as follows: reduction of time and relative simplification of SC development for different unrelated tasks; operational replenishment of orbital groups; small staff for SC development, manufacturing and launch preparation. And the last but not the least: the SC can be inserted into space via small-lift launchers or via heavier-lift launchers both several small SC at once and as a piggyback payload.

The performed analysis has shown that it is efficient to form multi-satellite orbital systems by small SC, particularly for scientific purposes covering simultaneous multipoint measurement of some parameter in near-Earth space. Particularly they are as follows: Sun exploration, studies of solar-terrestrial relationship, physical phenomena in the Earth atmosphere, ionosphere and magnetosphere. For the purpose of fundamental and applied scientific research the remote Earth observations and ocean surface monitoring can be performed by means of the satellites.

For the purpose of programs practical realization by means of small SC Lavochkin Association has developed the *Karat* unified orbital platform which wet mass including design margin is 96 kg. It is capable to carry payload of up to 60 kg onboard. Orientation accuracy is 10 arcmin. Stabilization velocity is $4 \cdot 10^{-3}$ degree/s. In-orbit lifetime is not less than three years.

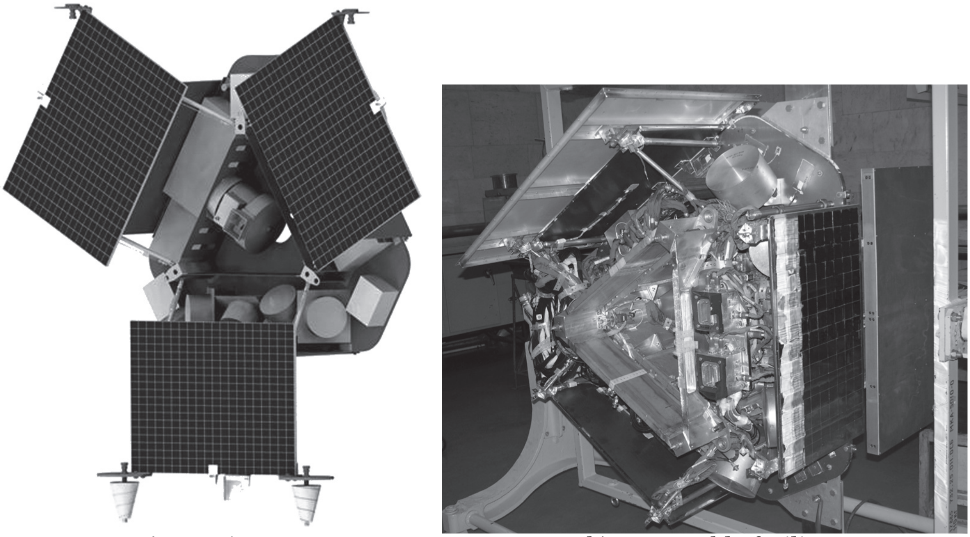
The platform design and development performance based upon long-term experience of Lavochkin Association in development and application of scientific SC including small-sized SC, particularly the long-lived autonomous station in frame of *Phobos-1, -2* program — one of the first open construction SC of that type. General view of *Karat* micro-platform is presented in Fig. 13.

Federal Space Program for 2006—2015 includes development of small SC family capable of conducting studies which recently could be carried out only by full-size satellites. At present a long-term program of fundamental research activities based on small SC is developed. Leading institutes of Russian Academy of Sciences (RAS) world-known for their scientific schools are taking part in arrangement of the space experiments.

As a result of complex analysis of the suggested projects with respect to the scientific importance, availability and other criteria there was made a decision by RAS and Federal Space Agency on target tasks for the first small SC.

The *Zond-PP* SC is intended for studies of Earth surface parameters by means of L-range satellite-born radiometer, soil humidity and water areas salinity mapping and ocean-land-atmosphere system energy exchange. The received results will be used for the purpose of forecast of climate and environment changes.

The *Monika* SC and the *Relek* SC. The first one is designed to study physical mechanisms of cosmic rays generation, forming in active processes on the Sun and in heliosphere. The second one is intended to research physical mechanisms of impact of solar and magnetospheric energetic particles on the Earth atmosphere.



a) Drawing b)In assembly facility

fig. 13. General view of *Karat* micro-platform

Monitoring conducted in the frame of the mentioned studies is necessary both for fundamental and applied science since cosmic rays are one of the principal factors of space weather and radiation hazard for example for SC crews, electronic systems and solar photo-converters.

Besides the above-said experiments practically ready for implementation, there are about twenty space experiments based on the *Karat* platform currently being under development in institutes of RAS. The first launch of the platform carrying one of the above-stated payloads is scheduled for the end of 2008.

System analysis of Lavochkin Association development of perspective space complexes and proposals of institutes of Russian Academy of Sciences on wide spectrum of fundamental research activities based on the stated developmental work has shown that their implementation will certainly provide Russia with proper leading position in world space science.

REFERENCES

1. Eneev T.M. Actual tasks of deep space exploration // *Space exploration*. 2005. № 6. V. 43.
2. Moiseev N.F. Resume and prospect of cosmonautics development in Russia // *Polet*. 2005. № 12.
3. Perminov A.N. Space activities of the Russian Federation and perspectives of development // *Polet*. 2005. № 7.
4. Perminov A.N., Davydov V.A. Current status and perspectives of space activities in the Russian Federation // *Polet*. 2006. № 12.
5. Pichkhadze K.M., Moisheev A.A. et al. Development and application of scientific and design heritage of G.N. Babakin in Lavochkin Association automated spacecraft // *Actual problems of aviation and aerospace systems: processes, models, experiments*. Kazan-Daitona Beach. 2005. № 1(19). V. 10.
6. Pichkhadze K.M., Moisheev A.A., Martynov M.B., Kudryashov V.A. *Karat* unified space platform for scientific micro-satellites development // *Polet*. 2007. № 8.
7. Polishchuk G.M. Prospect fundamental space research activities // *Cosmonautics and rocket production*. 2006. № 1.
8. Polishchuk G.M. Lavochkin Association: on the threshold of new space starts // *Polet*. 2007. № 8.
9. Polishchuk G.M., Pichkhadze K.M., Efanov V.V., Moisheev A.A. Spacecraft for fundamental research // *Polet*. 2006. № 8.
10. Polishchuk G.M., Pichkhadze K.M., Efanov V.V., Podobedov Y.G. Long-term projects of automated space complexes for exploration of giant planets and their satellites // *Polet*. 2005. № 7.
11. Space flight of Lavochkin Association / Ed. Polishchuk G.M. M.: Block-Inform-Express. 2007.
12. Zelenyi L.M. IKI RAN — forty years, new projects // *Polet*. 2006. № 1.
13. Zelenyi L.M. 50th anniversary of space era: some results and perspectives of investigations // *Polet*. 2008. № 1.

L.M. ZELENYI,
A.V. ZAKHAROV
Space Research Institute
Russian
Academy of Sciences,
Moscow, Russia
zakharov@iki.rssi.ru
G.M. POLISCHUK
S.A. Lavochkin Resarch /
Production Association,
Khimki, Moscow region, Russia
E.L. AKIM
Apply Mathematic Institute
Russian
Academy of Sciences,
Moscow, Russia

PHOBOS SAMPLE RETURN MISSION

INTRODUCTION

The first mission to Phobos included two spacecraft *Phobos-1* and *Phobos-2* launched in the USSR in 1988. The history of this mission was dramatic [15]. After a month of flight along the Earth-Mars trajectory communication with *Phobos-1* was lost because of an unfortunate mistake in the mission's control. The second spacecraft identical to the first one continued the flight. In seven months after the launch the *Phobos-2* spacecraft was inserted into an elliptic orbit around Mars and, after several modifications its orbit became the circular one, close to the Phobos's orbit. Scientific investigations of Mars, Phobos and the Martian environment were carried out during operation at the elliptical and circular orbits around Mars. The original plan was to deliver two small landers to the Phobos's surface for its future in-situ exploration right after the phase correction of the spacecraft and Phobos orbits. However, a few days before that culminating stage of the mission, the communication with *Phobos-2* was lost failure of the spacecraft's control system. The mission terminated just prior to the beginning of its basic stage. Nevertheless, studies of Mars, Phobos and the Martian environment that had been lasting for 57 days at the stage of the spacecraft's orbiting around Mars, provided unique scientific results on thermal characteristics of the Phobos's regolith [10], plasma environment of Mars and its interaction with the solar wind (see review by Zakharov [17] and Nagy et al. [13]). For example, the oxygen ions' flow leaving of Mars' atmosphere detected by the ASPERA ion spectrometer, made it possible to estimate the speed Martian atmospheric erosion due to its interaction with the solar wind [11]. These measurements are extremely important for the understanding of the history of Martian water and the evolution of the Martian atmosphere. Resent measurements of the Martian atmospheric erosion made by the ion mass analyzer of ASPERA-3 onboard the *Mars-Express* (ESA) mission [3] show that erosion rate is much smaller in comparison with the measurements made onboard the *Phobos-2* spacecraft, which confirms the dependence of this parameter from the phase of solar activity [4]. Before the *Phobos-2* mission our knowledge about Martian environment had been poorer than that of the more distant planets — Mercury, Jupiter, Saturn. The scientific data gathered during the *Phobos-2* mission are still unique. Analitical reviews of the Martian plasma and magnetic data of the *Phobos-2* and the *NASA Mars Global Surveyor* missions were presented in the review paper 'Mars' Magnetism and its Interaction with the Solar Wind' [12]. A new mission to Phobos named *Phobos Sample Return* (or *Phobos-Soil*) is being preparated in Russia now. The main purposes of this mission are — delivery of Phobos regolith samples to the Earth, studies of Phobos 'in-situ' and by remote sensing methods, investigations of plasma, fields and dust of the Martian environment.

1. MAIN CHARACTERISTICS OF THE MARTIAN SATELLITES

Two Martian satellites Phobos and Deimos can be approximated to ellipsoids with semiaxes of 13.3×11.1×9.3 km — for Phobos and 7.5×6.2×5.2 km — for Deimos [5]. The largest axis of both satellites's ellipsoids is directed towards Mars, and the satellites synchronously rotate around Mars. Their orbits are practically circular with the radii of 9,378 km (2.76 $_{RM}$) and 23,459 km (6.9 $_{RM}$) respectively. Orbital planes of both the satellites are close to the equatorial plane of Mars (~24° inclination to the ecliptic plane). The period of rotation around Mars amounts to 7 h 39 min — for Phobos and 30 h 21 min — for Deimos. Ground measurements of Phobos's orbital parameters show, that it varies slowly — Phobos comes closer to Mars along a very flat spiral. The reasons for such orbit change, the so-called antennial acceleration, are tidal losses of the orbital energy [7; 14]. Phobos's orbit is located inside the so-called Roche radius, and the moon would be destroyed by the tidal forces if it were a fluid. Another interesting feature of Phobos is its proper motion. Having apparently the largest libration amplitude, Phobos is a unique object among the other satellites of the Solar system planets known to be synchronously rotating. The principal cause for such strong effect is the fact, that the period of free libration for this satellite (~10 hours) is close to the period of its orbital rotation (~7.7 hour.). Accurate measurements of Phobos's libration characteristics can help to define its moments of inertia, which is important for determination of Phobos's mass distribution and the planet's internal structure. Precise measurements of Phobos's orbit parameters will allow one to substantially improve some parameters of the celestial mechanics and one will improve our knowledge on the masses of some asteroids behind the Martian orbit [6].

Though the surfaces of both satellites are covered with craters, however, they strongly differ. Phobos has a set of Deep almost direct parallel fosses of 100…200 m in width and 10…20 m in depth. The length of some of these fosses is up to 30 km and almost all of them originate near Stickney — the biggest crater on Phobos [16]. There are no similar surface features on Deimos. Craters on Deimos are much smaller in diameter. Fig. 1 presents the pictures of Phobos and Deimos. The mysterious question, concerning the morphological features of these two satellites is why their surfaces differ so strongly and, what processes had produced so prominent features on the Phobos's surface?

Small bodies of the Solar system, including the satellites of Mars, represent the initial, primary material of the protoplanet clouds from which all planets of the Solar system had been later formed. The influence of various external factors, such as the solar wind, cosmic rays, and meteorites have undoubtely

fig. 1. a) Phobos. This image, taken by the High Resolution Stereo Camera (HRSC) on board ESA's *Mars Express* spacecraft (ESA/DLR/FU Berlin).

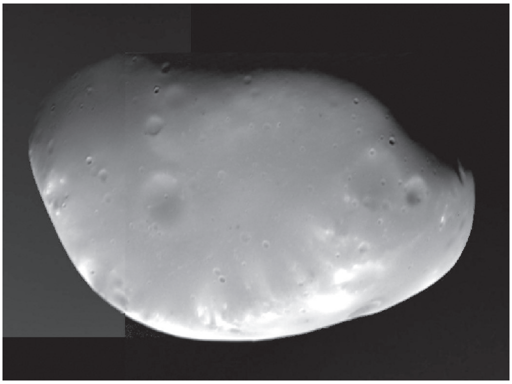


fig. 1 b) Deimos. Computer mosaic constructed from *Viking* orbiter images (NASA)

modified the small bodies' regolith. Thus, investigation of small bodies can give key information on the early stages of the formation of both the planets and small bodies in the Solar system, and their later evolution as well. Taking into account that Phobos and Deimos represent the part of the Martian system, their investigation provides a unique opportunity to determine their relation to Mars and uncover the mystery of their origin.

2. MAIN SCIENTIFIC OBJECTIVES OF THE MISSION

The main purpose of the mission is delivery of the Phobos's regolith samples to the Earth and detailed studies of Phobos's and Martian environment 'in-situ'and by remote sensing methods. The mission will provide an opportunity to solve or to reach better understanding of the following problems of the Phobos — Mars system science:

- physical and chemical characteristics of the Phobos's regolith (this can provide data on properties and evolution of the relict matter in the Solar system);
- understanding of the origin of the Martian satellites and their relation to Mars (the progress in this problem can become a key to understanding of an origin of satellite systems at other planets);
- precise measurements of the orbital and proper rotation parameters of Phobos (which is important for the study of the Phobos internal structure and its orbit evolution);
- characteristics of the physical processes of Martian environment —such as interaction of the solar wind with the ionosphere and remnants of the relict magnetic fields of Mars;
- electric and magnetic fields;
- detecting of a hypothetical dust tori near Phobos's orbit;
- variations of Martian atmosphere.

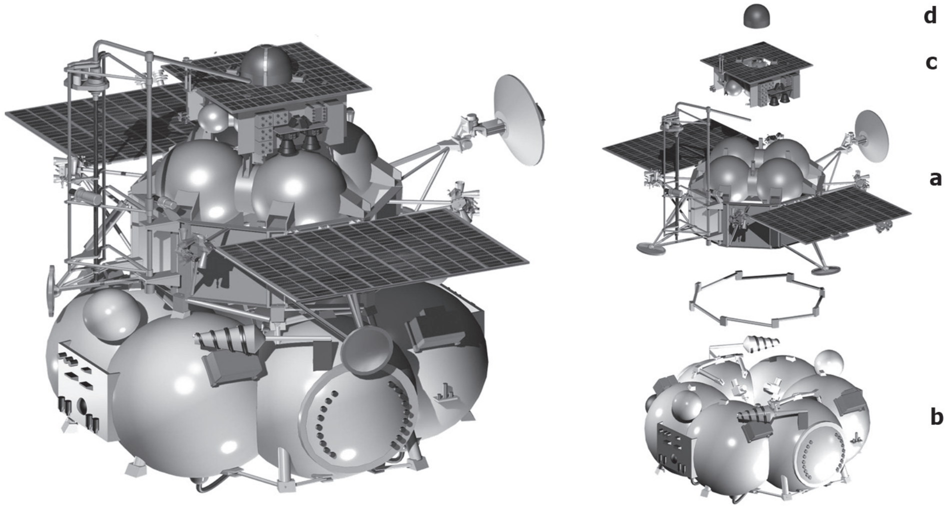


fig. 2. Elements of the spacecraft

3. SPACECRAFT AND MISSION SCENARIO

The spacecraft (SC) (fig. 2) includes four main elements: a transfer-orbital module (TOM) (a), a main propulsion system (MPS) (b), a return module (RM) (c) and a return capsule (RC) (d). Picture of these elements and the assembled spacecraft are presented in fig. 2.

The transfer-orbital module includes two propulsion systems: a braking system (BS) and a low thrust propulsion (LTP) system. BS will be used for the corrections during the flight along the interplanetary trajectory, braking to insert the spacecraft into Martian orbit and other maneuvers near Mars. LTP will be used for small corrections of the interplanetary trajectory, small corrections at the Martians orbits, and perfoming active maneuvers for spacecraft landing on the Phobos surface.

The main contractor for the designing and manufacturing of the spacecraft is the Lavochkin Association, the leading company on robotic science space missions in Russia.

The mission scenario includes several stages, described below.

3.1. LAUNCH AND TRANSFER INTO THE INTERPLANETARY TRAJECTORY

The spacecraft is to be launched from Baikonur and to be inserted into a 200 km circular based orbit with the inclination of 51.8 degrees. The launch can be perfotmed in October 2009.

After a four-hour flight (2.8 revolutions) at this initial orbit the main propulsion system will be switched on to transfer the spacecraft into the elliptical orbit

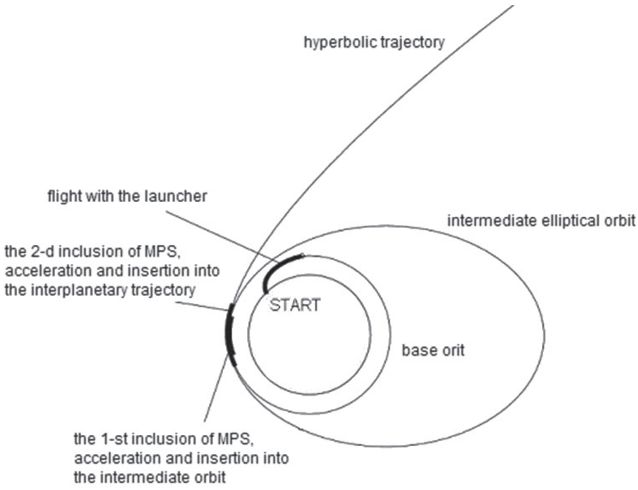


fig. 3a. Launch and transfer into the interplanetary trajectory

(230 km × 11,100 km), with the period 3.65 of hours. After 26 hours (7 revolutions) of orbiting MPS will be separated from the spacecraft. Additional impulse made by the braking engine will transfer the spacecraft into the interplanetary Earth-Mars trajectory. The scheme of the insertion of the spacecraft into the interplanetary trajectory is presented in fig. 3a.

3.2. THE EARTH-MARS INTERPLANETARY FLIGHT

The spacecraft will be directed to the fly-by hyperbolic interplanetary trajectory with the minimum distance to Mars of about 700...1,000 km. During this stage of the mission, which is to take about 11 months, three corrections are to be produced. The SC velocity before arrival to the Mars's vicinity will be about 2.5 km/s respective to Mars and a nominal altitude of the pericenter will be about 800 km. 2009 is a very advantageous year for the launch of this mission because the inclination of the arrival trajectory will be very close to the Martian equatorial plane (and the Phobos's orbital plane). The arrival to the Mars's vicinity should occur in August-September, 2010. The scheme of the interplanetary Earth-Mars trajectory is presented in fig. 3b.

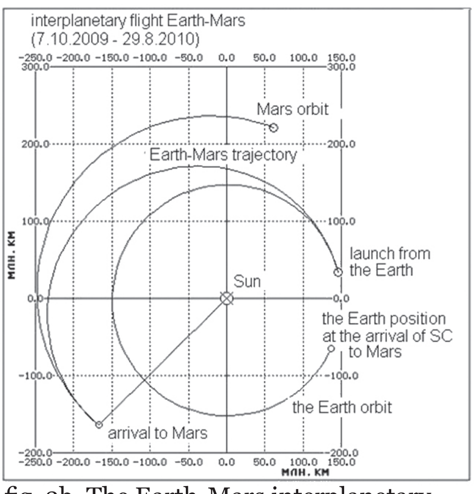


fig. 3b. The Earth-Mars interplanetary flight

3.3. OPERATIONS AT THE MARTIAN ORBITS

The spacecraft will be inserted into the first elliptical orbit with nominal altitude of the pericenter of ~800 km, of ~79,000 km and of the apocenter, having a 3 days orbital period, Fig. 3c.

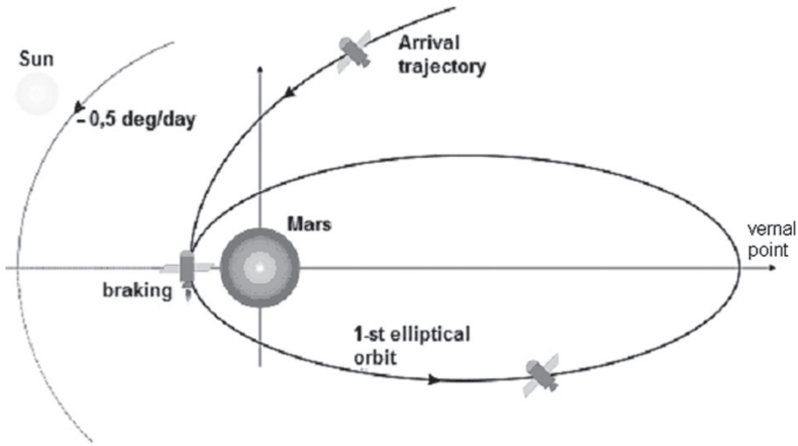


fig. 3c. Scheme of arrival and insertion into the first elliptical orbit around Mars

After 3...4 revolutions (10...12 days) an active maneuver made at the orbit apocenter will transfer the spacecraft into an intermediate elliptical orbit with the pericenter of ~ 9,910 km (about 500 km higher than the Phobos orbit). Then after 12...14 days the braking impulse made at the pericenter will transfer the spacecraft into a circular orbit with the R ≈ 9,910 km, T ≈ 8.3 h (about 535 km higher than Phobos's orbit). The spacecraft will be operating at this 'observation orbit' for several months, and then after a small correction it will be transferred into a quasi synchronous orbit (QSO), very close to the Phobos (distance between the Phobos's surface and the spacecraft will be of 50...150 km) and having the same orbital period. Fig. 3d presents the scheme of the spacecraft orbits around Mars. The spacecraft being at the QSO will make one revolution around Phobos during one period of Phobos's rotation around Mars (~7.66 hours). QSO in the Phobos coordinate system is presented in Fig. 3e.

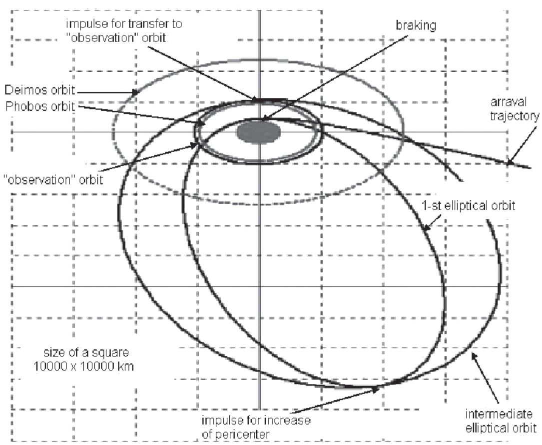


fig. 3d. Orbits around Mars

The quasi synchronous orbit will be used for the detailed measurements of Phobos's

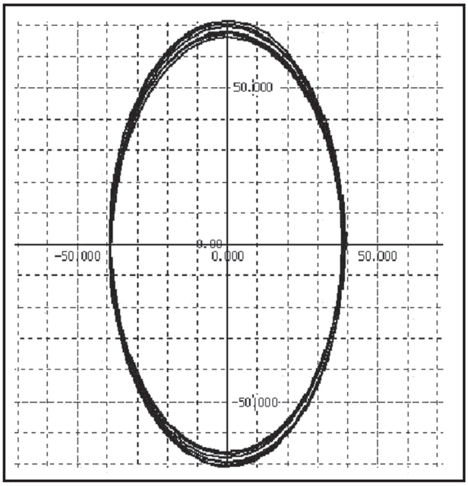


fig. 3e. The quasy-synchronous orbit of the TOM in the Phobos mass center coordinate axis

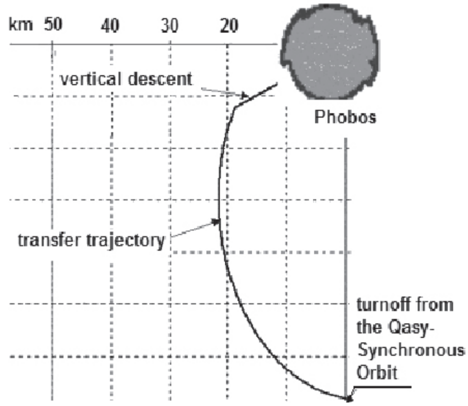


fig. 3f. Active maneuver for TOM landing on the Phobos surface

orbital parameters, and searching for a landing site. Taking into account various conditions of the spacecraft luminosity at the Phobos's surface and communication with the Earth, the landing site of the spacecraft has been chosen at the side of Phobos opposite to Mars (5° N—5° S, 230°—250° W). The most favorable period for landing (from the ballistic point of view) is March—April, 2011. Approaching to the Phobos's surface and landing will be implemented by an autonomous active maneuver using the data of a TV-system, the laser altimeter, and the Doppler instrument. Before landing the axial velocity would be <1 m/s, the side-long velocity < 2m/s. This stage is to take about 2 hours. The scheme of the approach is presented in fig. 3f.

3.4. OPERATION AT THE PHOBOS SURFACE

After the landing the sampling device is to take regolith samples and put them into the capsule of the return module. This is going to be one of the critical stages of the mission. The sampling device will use a robotic arm to collect regolith from the surface. This device has been designed according to the engineering model of the Phobos regolith developed by the Vernadsky Institute of Geochemistry and Analytical Chemistry (GEOKHI) of the Russian Academy of Sciences.

The transfer-orbital module is to stay at Phobos's surface after the launch of the return module and to implement a scientific program of the mission using a package of scientific instruments during one terrestrial year.

3.5. LAUNCH FROM PHOBOS, FORMATION OF THE RM ORBIT, AND THE MARS-EARTH INTERPLANETARY FLIGHT

The return module will be launched from the mother spacecraft (the transfer-orbital module) and after several modifications (the three impulse scheme) of

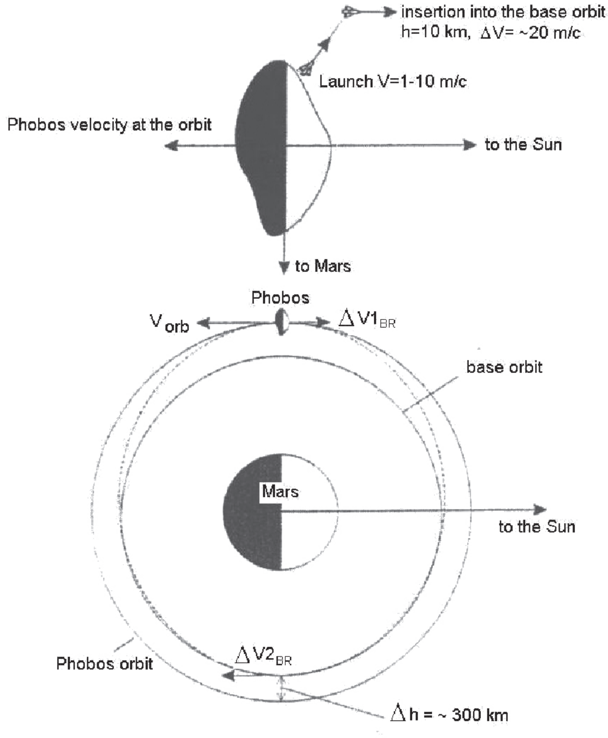


fig. 3g. Launch of the Return Module from Phobos and the base orbit

its orbit around Mars, will be directed to the Earth along the interplanetary trajectory in August, 2012. The schemes of the launch from Phobos and the formation of the base orbit around Mars are presented in fig. 3g.

3.6. RE-ENTER OF RC TO THE EARTH ATMOSPHERE AND HARD LANDING

After about 11 months of the interplanetary flight the return module reach the Earth vicinity, RC will be separated from RM and re-enter the Earth atmosphere with the velocity of 11,8 km/s, Fig. 3h. The landing of RC is to take place in the Kazakhstan republic (hard landing).

The Applied Mathematics Institute of the Russian Academy of Sciences has developed the mission ballistic scenario [1; 2].

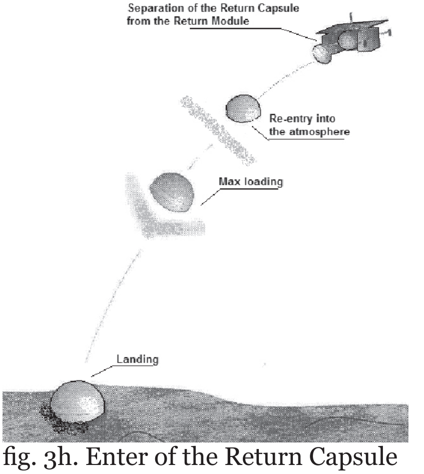


fig. 3h. Enter of the Return Capsule into the Earth atmosphere and landing at the terrestrial surface

4. PROGRAM AND THE SCIENTIFIC EXPERIMENTS

There are several approaches to study the Martian satellites in the *Phobos Sample Return* mission.

First of all, it is sample return. The scientific rationale for obtaining samples of planetary bodies have been discussed in detail by many research groups as well as have been described minutely in many publications (e.g, 8; 9; 18). Phobos's regolith samples delivered to the Earth will be studied according to a science program developed by the Vernadsky Institute (GEOKHI).

Remote sensing of Phobos from the spacecraft can provide global coverage data important for understanding of both the surface and ithe nternal structure of this Martian moon: global geologic mapping; size, shape, mass, bulk density; colour, albedo, photometric scattering and thermal properties; mapping of global elemental and mineral composition, magnetic properties, internal structure, landing site selection and certification.

Operation of the spacecraft at the elliptical orbits and at the circular Phobos-like orbits will provide an opportunity to study the Phobos dust tori (like a planetary ring), Martian plasma environment, interaction of Phobos with the solar wind, and also to investigate the peculiarities of its orbital motion, free and forced librations.

'In-situ' measurements at Phobos's surface will provide an opportunity to verify the remote sensing measurements and to accomplish the detailed study of the regolith near the landing site, specifically: elemental and mineral composition, volatile components, physical and mechanical properties of the regolith, its state of magnetization.

The payload of PhSRM includes scientific instruments, which will provide various data for studies of Phobos as a celestial body, in-situ measurements of chemical and mineralogical composition of the Phobos's regolith, its thermal and mechanical characteristics, and Phobos's internal structure. Other instruments included in the payload are to investigate characteristics of the Martian environment: plasma, fields, waves, dust. The payload includes also two instruments for celestial mechanics experiments. The table below presents the payload of the mission.

Lets have a look at some of the instruments included in the payload.

One of the most complicated instruments of the payload is the Gas-Chromatograph Complex. This complex consists of three instruments: Thermal Differential Analyzer (TDA), with the system of regolith preparation Chromatograph (ChG), and Mass-Spectrometer (MSP). This complex is designed to measure the quantity of individual gas components in a complex gas mixture, which is evolved from the soil sample by pyrolysis, due to their separation by the time of retention in chromatographic columns. The instrument identifies the chemical composition of gas components by their calibrated time of retention and by spectroscopy of specific absorption lines for H_2O, CO_2, and CH_4 gases.

Instruments for the regolith and the Phobos inner structure investigation	
Gas-Chromatograph Complex, GChC Thermo-Difference Analyzer, TDA Chromatograph, ChG, Mass-Spectrometer, MSp	Chemical composition of volatile components in the Phobos's soil
Mossbauer Spectrometer, MS	Mineralogical composition of iron compounds of the Phobos regolith
Gamma-spectrometer, FOGS	Elemental regolith composition
Neutron Spectrometer, HEND	Regolith neutron radiation
Laser time-of-flight Spectrometer, LASMA	Elemental regolith composition
Secondary ions mass analyze MANAGA	Elemental regolith composition
Long-wave Planetary Radar, LWPR	Phobos inner structure, regolith electric characteristics
Seismometer, SEISMO	Phobos inner structure, gravimetry
Optic instruments	
TV system, TSNN	Landing support, Phobos's surface mapping, landing selection
Panoramic and stereo TV-cameras, PANORAMIC and STEREO	Phobos panoramic and stereo images
Fourier Spectrometer, AOST	Monitoring of minor components of the Martian atmosphere
Instruments for Mars environment studies and celestial mechanic experiments	
METEOR and DIAMOND	Micrometeorites monitoring Dust tori at the Phobos's orbit
Plasma set, FPMS	Ion and Electron spectrometers, Magnetometer
Stars and Solar tracker, LIBRATION	Phobos orbital and proper motion
Ultra-stable oscillator, USO	Celestial mechanics experiment

Phobos Sample Return Payload

Neutron and Gamma-ray Spectrometers, included in the payload (HEND-2 and FOGS) will search for hydrated materials or/and water ice in the subsurface of Phobos and the chemical elements concentration at the Phobos's surface: rock-formed elements (from H to Fe) and natural radioactive (K, Th, U) ones.

Laser Time-of-flight Mass Spectrometer (LASMA) is an active experiment. Laser irradiation focused on the surface of regolith evaporates and ionizes a sample. The emerged ion cloud comes in the field of the electrostatic reflector and is directed to the detector where different ions are registered as mass peaks. This instrument is designed for perfoming the quantitative analysis of elemental and isotopic composition of the Phobos's regolith.

The Long-wave Planetary Radar (LPR) mounted at the spacecraft will probe the Phobos's surface and subsurface in order to detect ground layers on the depths of 1…100 m. This instrument will allow to investigate electrical characteristics of the regolith and the Phobos's inner structure together with the data received by the seismometer (SEISMO).

The plasma-wave Martian environment will be studied by Plasma-Magnetic System (PhPMS). This system includes several detectors: ion and electron spectrometers, and magnetometers. The main goals of these investigations include the solar wind interaction with Phobos and Mars, the peculiarity and the effects of this process: erosion rate of the Martian atmosphere; investigation of kinetic processes in the Martian bow shock; the role of planetary ions, captured by the solar wind, the bow shock formation; analysis of physical processes in the vicinity of magnetopause/upper boundary of interplanetary magnetic field pile up region; investigation of plasma boundaries dynamics at Mars; 3D-distribution functions for different plasma components in the Martian magnetosphere tail; identification of physical processes of ion acceleration; chemical composition of secondary ions sputtered from Phobos's soil by the solar wind. The plasma-waves instruments will measure spectral characteristics of the solar wind and the Martian surrounding plasma during the revolution of the spacecraft at different orbits. We hope these data will provide valueable information in addition to plasma-field data received by the Phobos-2 mission and the Mars Global Surveyor [12].

The formation of the scientific program of the mission is governed by scientific priorities and technical limitations from the spacecraft's service systems and measurements conditions. It is accepted that the scientific experiments should have the following order of priorities:

1. Geological and geochemical certification of a place of sampling.
2. Study of physical and chemical properties of the regolith and the internal structure of Phobos.
3. Study of the Martian environment (dust, plasma, field, waves).
4. Study of orbital and proper parameters of Phobos motion, celestial-mechanical experiments.
5. Monitoring of characteristics of the Martian atmosphere.

The Space Research Institute (IKI), the Vernadsky Institute (GEOKHI) and Institute of Radioengineering and Electronic (IRE) of the Russian Academy of Sciences are the main organizations developing the scientific program of the mission and the payload. Scientific teams from several European countries (France, Germany, Sweden, Ukraine) and also from China participate in the development of PhSRM instruments.

REFERENCES

1. Akim E.L., Zaslavsky G.S., Morskoy I.M., Stepanjanz V.A., Tuchin A.G. Ballistic, navigation and flight control of the *Phobos Sample Return* spacecraft // Izvestia RAN. 2002, № 5. (in Russian).
2. Akim E.L., Popov G.A., Tuchin A.G. Mechanics and Motion Control of a Space Vehicle in the Project of Relict Substance Delivery on Earth (The Project *Phobos Sample Return*) // 16th IFAC Symposium on automatic control in Aerospace, Preprints. V. 1. Saint Petersburg, Russia. 2004.
3. Barabash S., et al. Mars Express: The scientific payload / A. Wilson, Ed. // ESA special publication SP-1240. 2004. P. 121—139.
4. Barabash S., Fedorov A., Lundin R., Sauvaud J.-A. Martian Atmospheric Erosion Rates // Science. № 315. 2007. P. 501—503.
5. Batson R.M., Edwards K., Duxbury T.C. Geodesy and cartography of the Martian satellites // Mars / Ed. H.H. Kieffer et al. The University of Arizona press. 1992. P. 1249—1256.
6. Blamont J., Balmino G. Gravity field, shape and internal structure with applications in celestial mechanics // PHOBOS, Scientific and methodological aspects of the Phobos study. Space Research Institute issue. 1986. P. 355—356.
7. Burns J.A., Contradictory clues as to the origin of the Martian moons // Mars / Ed. H.H. Kieffer et al. The University of Arizona press. 1992. P. 1283—1301.
8. Drake M., Boynton W., Blanchard D. The case for planetary sample return missions: Origin of the Solar system // EOS. V. 68. 1987.
9. Gooding J., Carr M., McKay Ch. The case for planetary return missions: History of Mars // EOS. V. 70. №.31. 1989. P. 745—754.
10. Ksanfomality L.V. and Moroz V.I. Spectral Reflectivity of Phobos` regolith within the range 315-600 nm // Icarus. № 117. 1995. P. 383—401.
11. Lundin R., Zakharov A.V., Pellinen R. et al. First measurements of the ionospheric plasma escape from Mars // Nature. № 341. 1989. P. 609—612.
12. Mars` Magnetism and Interaction with the Solar Wind / Ed. D. Winterhalter, M. Acuna, A. Zakharov. Space Science Reviews. V. 111. 2004.
13. Nagy A.F., Winterhalter D., Sauer K., et al. The Plasma Environment of Mars // Space Science Reviews. V. 111. 2004. P. 33—114.
14. Peale S.J. Origin and evolution of the natural satellites // Annu. Rev. Astron. Astrophys. № 37. 1999. P. 533—602.
15. Sagdeev R.Z., Zakharov A.V. Brief history of the Phobos mission // Nature. V. 341. 1992. № 6243. P. 581—585.
16. Thomas P., Veverka J., Bell J., Lunine J., Cruikshank D. Satellites of Mars: geologic history // Mars / Ed. H.H. Kieffer et al. The University of Arizona press. 1992. P. 1256—1282.
17. Zakharov A.V. The plasma environment of Mars: *Phobos* mission results // Geophysical Monograph 66, Venus and Mars: Atmospheres, Ionospheres, and Solar Wind Interactions / Ed. J.G. Luhmann, M. Tatrallyay, R.O. Pepin. 1992. P. 327—344.
18. Zakharov A.V. *Phobos Sample Return* mission (Intention for future) // Planetary systems: The long view / ed. by L.M. Celnikier, J. Tran Tranh Van. 1998. P. 489—498.

M.I. MALENKOV,
V.V. STEPANOV
Russian Mobile Vehicle
Engineering Institute
(J.-St.Co. VNIITransmash)
Saint Petersburg, Russia
malenkovcst@mail.lanck.net

RUSSIAN CREATING TECHNOLOGIES OF PLANETARY ROVER LOCOMOTION SYSTEMS

INTRODUCTION

Soviet and American missions to the Moon in the early 70s of last century confirmed clearly expediency of applying the Moon mobile vehicles to explore the lunar surface using automatons and with the participation of people immediately. During the last years a great body of new scientific information about Mars has been received by American scientists thanks to application of automated Mars rovers. Significance of planetary vehicles will grow in the following stages of study and practical developing of new and new territories of the Moon and Mars.

Undoubtedly, still over a long period of time designers of new lunar and martian vehicles will study experience of creating their first prototypes — the automated lunar rover *Lunokhod-1* [1, 2, 3, 4] delivered to the Moon in 1970 with the station *Luna-17* (Fig. 1) and the manned Moon rover *Lunar Roving Vehicle* (LRV) delivered to the Moon in 1971 as a part of the *Apollo 15* mission. Taking into account it, the most important aspects of creation and Earth-based debugging of locomotion systems for automated planetary rovers.

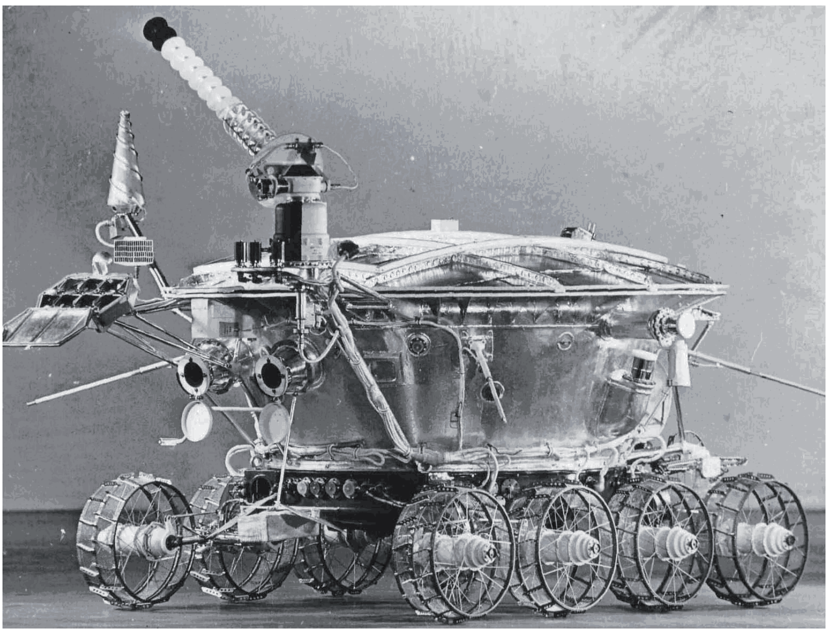


fig. 1

Principal propositions and conclusions about a conception and a design of locomotion systems, are, as rule, just for both the Moon rovers and Mars rovers that allow us to tell about planetary rover technologies.

In the wide sense, the notion ‘a creating technology’ of any new product, article, system, etc. exceeds the limits of merely industrial methods of transformation, initial materials machining. To better understand and bring new technologies to a practical level, other aspects of decision of the problem can be considered, namely, historic, philosophic, scientific, engineering, organizational and technical ones. The present article considers engineering and scientific aspects having the following peculiarities:

- prototyping, physical modeling, running tests of full-scale mock-ups and models on natural testing grounds and in soil canals;
- combination of analytic and heuristic methods of synthesis of locomotion systems taking into account properties of terrain and characteristics of a planetary rover;
- technical realization of new principles of motion on the basis of the combined wheel-walking mover (propulsive device);
- study of properties of the surface soil cover of the Moon and Mars (physical and mechanical properties of soil and peculiarities of the relief) in possible areas of planetary rovers operation;
- investigation of interaction of different movers with weak-coherent soils similar to the lunar regolith.

As to the historic aspect, the greatest importance for forming ideology of planetary rovers projection had the fact that tank-constructors were at the beginning of creating self-propelled automated chassis for *Lunokhod-1* and *Lunokhod-2*. They were enlisted to services by S.P. Korolev. Then G.N. Babakin continued the co-operation. This circumstance was always accentuated by A.L. Kemurdjian who was the founder of the Russian school of projection of locomotion systems for planetary rovers. The philosophy of priority of high cross-country ability compared to other its characteristics of a planetary rover was materialized just in co-operation of the collectives headed by above-mentioned leaders.

Analysis given in the present paper is mainly based on the original investigations carried out at J.-St. Co. VNIITransmash within the framework of the Soviet and international programs [5—10]. This period covers more than 40 years but along with the review of realized projects attention is given to modern understanding of scientific and technical problems of ensuring high effectiveness of planetary rovers and new developments of their locomotion systems [11—16].

1. PURPOSE, COMPOSITION AND CRITERIA FOR ESTIMATION OF LOCOMOTION SYSTEMS PROPERTIES FOR AUTOMATED PLANETARY ROVERS.

A planetary rover is a space apparatus (SA) made in the form of automated or remote-controlled transport mean (TM) of high cross-country ability delivered to the site of operation by means of a spacecraft and intended for movement and ensuring operation of scientific equipment placed on its board.

Two main structural components are noted in this formulation: service systems (locomotion, control, telecommunication, power supply, temperature control, navigation and machine vision) and payload as scientific equipment. The all-important structural component is the locomotion system (LS) realizing the transport function and consisting of subsystems which are traditional for a transport machine (Table 1).

Table 1

The planetary rover locomotion system
Running gear: - mover (propulsive device); - suspension; - mechanisms for wheels turning, unfolding from transport position, etc. (are specified when developing concept).
Electromechanical wheel drive: - tractive motors; - tractive reducers; - subsystem of ensuring serviceability and service durability.
Brake subsystem
Load-carrying structure (frame)
Control subsystem for drives and mechanisms
Safe subsystem of motion
Temperature-control subsystem
Cable system

The most important characteristics of the planetary rover as SA are minimum mass and transport overall dimensions, high reliability at minimum power inputs. The most important characteristics of the planetary rover as TM are cross-country ability, maneuverability, overturning stability, speed of motion, controllability, distance margin and resource. Combining qualities of SA and TM together, it is possible to offer the following definitions and quantitative indices of the planetary rover properties (Table 2).

One distinguishes obstacle-crossing ability and soil-crossing ability for TM. To characterize profile (relief) of terrain are usually used the following absolute quantities: a height of a ledge (scarp) — h_1; a height of a bench (counterscarp) — h_2; width of a ditch (cleft) — b; height of a separate stone — h_3, and an angle of grade (slope) — α°.

Soil-crossing ability is characterized by bearing strength of soil. Indices of soil-crossing ability are especially important for weak-coherent soils: lunar regolith; the Earth's loose and lightly deformed soils (some types of volcanic, quartz and other sands).

Table 2

	The most important characteristics of planetary rovers	
	Properties	Indexes (criteria)
1	*Cross-country ability:* ability to move around on the Moon surface without loss of mobility with the least deflection from a given route	Angle of the coherent and weak-coherent soil slope a, deg Height of a scarp and counterscarp surmounted, h_1, h_2, m Height of a stone under the bottom, h_3, m
2	*Maneuverability:* ability to change value or direction of travel speed under the predetermined law including a turn about geometric center, i.e. with the turning radius equal to naught	Minimum turning radius: R = 0 m Range of turning wheel angles: from 0 deg to 90 deg Equivalence of forward and reverse motion
3	*Economical efficiency:* ability to move by a route with the least power inputs.	Specific power inputs: power inputs of the planetary rover for motion related to the planetary rover mass and distance traveled, E_{sp}, W·h/km·kg
4	*Relative mass of the locomotion system*	Relation of a locomotion system mass to the planetary rover mass
5	*Service life/distance margin*	Total service life, day / distance traveled, km
6	*Reliability:* ability of fail-safe (unfailing) operation, ability to carry out functions of motion and scientific exploration even in the case of failure of separate elements.	Reliability index (0.999 for *Lunokhod-1, 2*)
7	*Stability:* ability to resist turning over in dynamic régimes of operation and when stopping	Maximum permissible speed of motion: maximum angle of slope of the surface α, deg
8	*Converting:* ability to ensure possibility of decreasing a volume occupied in transport position and unfolding into operating position	Ratio of volumes occupied in transport and operating position

Investigations made on the Moon surface [2, 5, 10] showed that bearing strength of the lunar soil decreases with increasing an angle of slope. Therefore the most important integrated index of soil-crossing ability and obstacle-crossing ability is an angle of a grade α° on weak-coherent soils. It is necessary to note that an angle of repose for fine-grained lunar and martian soils can be more than α° = 30 deg, and so it is necessary to create planetary rovers able to climb the grade with such steepness confidently. One can establish that this problem today has not yet been solved for planetary rovers that already have been realized.

Results of processing the *Lunokhod's-1* telemetric information showed that maximum angles of grades surmounted were of α° ≈ 22...26 deg, specific power inputs on different sections of the route were equal to E = (0.26-0.34) W·h/km·kg and wheel slippage δ dependent on an angle of terrain grades . Values of the parameters E and δ were determined by formulas:

$$\mathring{A} = \frac{U \cdot t\Sigma I_i}{m \cdot S_f} \quad (1)$$

$$\delta = 1 - S_f / S, \quad (2)$$

where *t* is duration of motion; *U* is supply voltage; I_i is a current strength for motor of a i-wheel; *m* is the rover mass (or a mock-up mass meeting the similarity criterion); S_f is actual distance traveled, *S* is a theoretical length of a route determined by a speed of wheel rotation and its radius *r*:

$$S = \omega r t \quad (3)$$

In conclusion of this section we note that a level of requirements on cross-country ability for automated vehicles is higher than for manned vehicles. In this connection we would remind that when the LRV rover of the *Apollo-15* mission lost mobility when climbing a grade, astronauts transferred the lunar electric car over dangerous section by hand and continued traveling.

2. OPERATING CONDITIONS OF VEHICLES ON THE MOON, MARS AND OTHER CELESTIAL BODIES OF SOLAR SYSTEM

In the late 70s – in the early 80s when the tide of success in exploration of the surface of the Moon, Mars and Venus by means of interplanetary automatic stations had place, VNIITransmash analysed physical conditions and made a prognosis of possibility and methods of traveling over the surface of other planets of the Solar system and their most interesting satellites. Both at that time and today the Moon, Venus, Mars and its satellite Phobos are of the most interest for investigation with contact methods.

Main factors defining conditions of using the Moon rovers are physical and climatic conditions on the lunar surface as well as structure, relief and physical and mechanical properties of upper layers of soil. Climatic conditions are defined by the temperature of the soil surface, absence of atmosphere, daily and seasonal changes of a flow of solar radiation. A free fall acceleration on the Moon is 1.62 m/s^2. Duration of the lunar day is almost 28 days. The distance from the Earth is about 384,000 km. The Moon is facing the Earth by one and the same side. The opposite side is not visible from the Earth and up till now planetary rovers were not used at this side.

One can choose two the most typical regions on the lunar surface: marine regions and continental regions. Marine regions are characterized by a rather darker colour and even surface and occupy about 15% of area. The continental regions have more light colour and complex relief. They occupy the opposite side of the Moon mainly. To present day the Soviet and the American Moon rovers operated only on the visible side of the Moon in marine regions and at the boundary of marine regions with the continental regions.

The most typical elements of the lunar relief in marine regions are craters formed mainly as a result of the meteoric impacts and having dimensions from several hundred kilometers to tens of centimeters (Fig. 2).

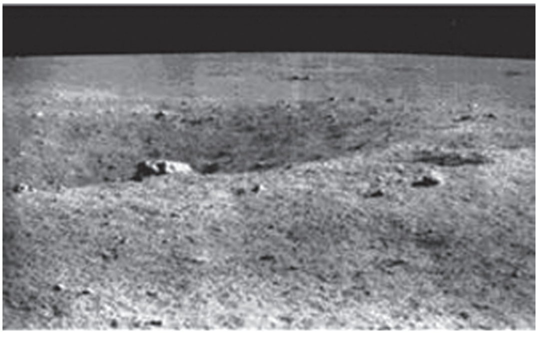


fig. 2

Crater distribution on the lunar surface depending on their dimensions is described with the dependence [2]:

$$N=10^{10.9} D^{-2} \qquad (4)$$

where *N* is an amount of craters with a diameter larger than *D* on the area of 10^6 km^2, *D* is a crater diameter, *m*.

This dependence is typical for craters with a diameter up to 100 m. Rather different dependence is taken place in the case of greater diameters. Craters are the most typical obstacles for the planetary rover motion.

Another typical element of relief on the lunar surface is stones and stony ridges. Stones are often met close to craters. At the same time the larger a crater the more stones around it. An amount of stones increases especially sharply near craters with a diameter more than 20 m. A considerable part of stones concentrates on the crater bank. Forms of stones are various. Presence of stones near craters is often explained by ejection of the bedrock when forming craters.

To describe distribution stones on the surface the following relationship can be used:

$$N' = k \cdot d^{\gamma} \qquad (5)$$

where *N'* is quantity of stones with dimension more than *d*, *d* is a cross-section dimension of a stone, *k* and γ are constants depending on geomorphologycal situation ($k>0$; $\gamma<0$) [2].

Sometimes formations similar to stone are met. They are outwardly like stones but are large clods formed as a result of sticking together of fine-grained material, which are lightly destroyed by the Moon rover wheels.

More complex elements of relief are met on the lunar surface too: combination of slopes of large length complicated with craters and stones, terraces, ridges, taluses on steep slopes, etc.

A mechanism of impact crater forming leads to forming of a specific soil structure, i.e. fine-crushed material named as the regolith. The latter represents particles of initial material (basalt or other deposits) with an extremely wide granulemetric spectrum. However, particles of a too small dimension (less than 0.01—0.001 mm) are not formed with such method of soil destruction. Therefore the regolith is enough homogeneous class of soil, namely powdered sand with marked quantity of stones and stony boulders, which can be considered as separate projected obstacle from the point of view of movement. Impact crater forming leads not only to crushing soil but also to its very intensive transfer and redistribution on the lunar surface. Transfer of soil when the impact crater forming under conditions of absence of atmosphere takes place by the ballistic trajectory without marked differentiation of material rejected into space that promotes to averaging of soil properties on the considerable part of the surface.

Mars has the atmosphere, in which the carbon dioxide is the main component. In spite of the fact that this atmosphere is rarefied as compared with the Earth's, process of impact crater forming has lesser importance than on the Moon. Therefore physical and mechanical properties of the surface cover are more various. In particular, stone ridges with high strength characteristics of the material take place here. It is necessary also to take into account problems arising at the time of sufficiently frequent dusty storm on Mars.

The paper also considers peculiarities of operating conditions for automatons on the surface of Venus and the Phobos.

3. ANALYSIS OF TECHNICAL DECISIONS ON A LOCOMOTION SYSTEM DESIGN

The *Lunokhod-1* locomotion system (Fig. 1, p. 155) has characteristics given in Table 3 (p. 262).

It is necessary to note that in a period of creating this vehicle, a choice between wheeled and caterpillar methods of motion was greatly dramatic. The fact is that VNIITransmash was the leading research tank institute in the USSR. Therefore development of two versions carried out parallel practically up to beginning of manufacturing the flight units. While choosing the wheel arrangement designers took into account that the eight-wheeled mover can be changed into the two-caterpillar mover by means of replacement of the extreme wheels by driving sprockets and idlers, and the middle wheels by rollers.

VNIITransmash created the first in the world a mobile robotic rover PrOP-M that reached the surface of Mars during the Soviet *Mars-2* mission (1971). The rover moved using skis set on either side. The VNIITransmash's penetrometers for study of the strength properties of Venusian soil operated successfully as part of four landers of space stations *Venera* and *Vega*. Then the institute created the running mock-up of the Venus rover using the *Lunokhod-1* wheels as a component of the running gear.

As a result of investigations in the time interval after *Lunokhod-1* VNIITransmash's specialists created running mock-ups of the Moon rovers with the six-wheeled movers, determined experimentally effectiveness of application

Table 3

Lunokhod-1 mass, kg	756
Locomotion system mass, kg	105
Wheel arrangement	8×8
Wheel base, mm/wheel track, mm	1,705/1,600
Wheel diameter, mm/Wheel width, mm	510/200
Clearance, mm	380
Turning method	Side (as for a tractor)
Transmission	Electromechanical with individual drives of wheels
Brake system	Electrodynamics delay mechanisms and mechanical one-disk brakes with electromagnetic control
Suspension	Independent, torsion-bar with swinging of levers in longitudinal plane
Travel speed, km/h	0.8 or 2.0
Turning radius on a center of the supporting quadrangle, m: when moving/on the spot	2.7/0
Angles of static stability, deg: longitudinal/transversal	43/45
Typical obstacles surmounted (calculated): ledge, m/bench, m loose soil grade, deg	0.35/0.4 20

of swivel wheels. They investigated influence of elasticity of metal-mesh wheels, a height and disposition of grousers, design and characteristics of elastic and non-elastic lever-articulated (centerpoint) suspensions.

Biomechanics of a ‘walking’ mode of motion was studied and physical models and mock-ups of vehicles with walking movers were created. In the end this

Main technical characteristics:	
Mock-up mass, kg	320
Wheel base, m	1.7-2.2
Travel speed, km/h: -wheeled mode -wheel-walking mode	0.9 0.15
Loose soil slope surmounted, deg: -wheeled mode -wheel-walking mode	18 34

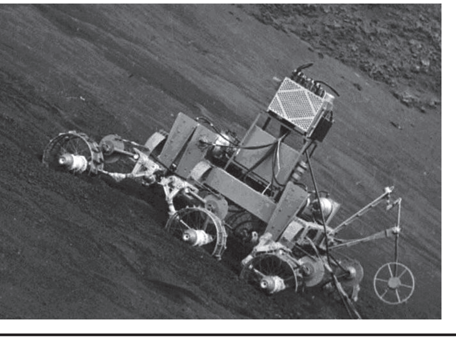


fig. 3

allowed to synthesize some successful schemes of a combined wheel-walking mover (WWM). Fig. 3 shows the first running mock-up of the Moon rover with WWM that is able to realize uninterrupted gaits of different types.

A design of the three-section self-propelled chassis with wide-profile cylinder-conical wheels received wide recognition in Russia and abroad (Fig. 4). A suspension is at all absent in this design, and joints providing two degrees of freedom for each of three wheel axles are immediately built in a frame. Notion 'clearance' is superfluous here, as the immovable bottom is absent. Besides enumerated merits this mock-up is able to realize the interrupted gait simultaneously by two wheels of each of three sections. It is no coincidence this innovative concept of a planetary rover received the silver medal of World Salon of Inventions (Brussels, *Heurica-95*) and was taken by Lavochkin Association (NPOL) (customer of locomotion systems for planetary rovers) in the capacity of a base for creating the Mars rover for the Soviet Martian program of the 80-90s of last century.

Main technical characteristics	
Mock-up mass, kg	200
Diameter of the wheel cylindrical part, m	0.51
Wheel base, m	1.4-2.5
Loose soil slope surmounted in wheel-walking mode, deg	34
Height of surmounted scarp, m	1

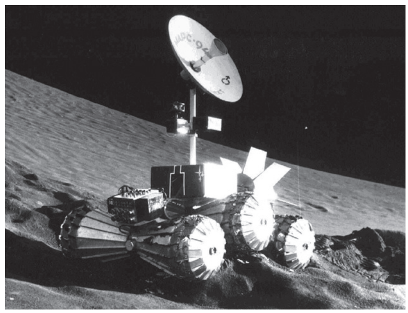


fig. 4

No entering in the theory of a wheel-walking mode it should be noted that mock-ups with WWM move as usual wheeled machines on good-quality sections of a route. The mode of wheel walking is used on complex sections, for example, on steep grades, weak-coherent soils when one, two or some driving wheels are a footstep of the walking mechanism. At that, other part of wheels is in a phase of repulsion and ensures a stop when applying pushing force to the axle of walking wheels.

While going over of Russia transfer to market economics a main factor of further development of planetary rover technologies in Russia is international co-operation, which is not limited by study of Soviet experience but is directed to development and realization of innovative projects. Especially close creative relations of VNIITransmash specialists and small enterprises Rover Co. Ltd., Actron Co. Ltd organized by these specialists were made up with the scientific and technical centers CNES (France), ESA/ESTEC (the Netherlands), DLR and Max Plank Institute (Germany).

The *Lunar Rover Mock-up Chassis* (LRMC) and its on-board manipulator, which has interfaces matched with the chassis and very effectively supplements it, were created under co-operation of VNIITransmash with ESA/ESTEC and with the assistance of Aktron Co. Ltd (Fig. 5, 6) [14]. The running gear has a synchronizing mechanism of position for wheels of the opposite sides decreasing angles of inclination of the frame when moving over uneven relief and a hill-side.

Main technical characteristics	
Chassis mass, kg	60
Payload mass, kg	60
Overall dimensions, mm:	
- length	1,200
- width	900
Type of frame	Rigid
Type of mover	Wheeled, 4×4×4
Maximum travel speed, km/h	0.35
Turning method	Cinematic, side

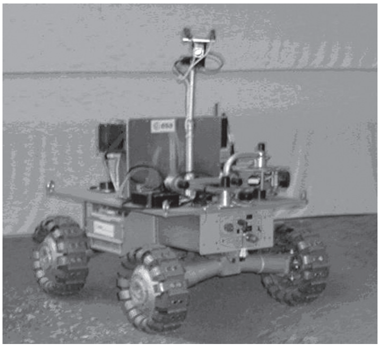


fig. 5

The on-board manipulator has a swivel head on which compact scientific instruments are placed (Fig. 6) [13]. It ensures conducting experiments at one and the same part of the surface using instruments one after another without their change that increases reliability. Besides, the manipulator has arrangements which can take and load minor stones and loose soil samples on the rover board. Having own mass of about 4 kg the manipulator is able to operate (under condition of the Earth's gravity) with payload having a mass up to 1.5 kg when its length is maximum (more than 0.8 m) without using mechanisms for partial weight compensation (imitation of natural low gravitation).

Important innovations have been putted in the *ExoMaDer* scale model (1:2) design developed at Rover Co. Ltd. under co-operation with ESA/ESTEC within the framework of a stage of choosing the concept of a rover for European *ExoMars* project. Kinematics of the lever-articulated (balanced) suspension provides leveling of normal reactions of soil applied to wheels when surmounting obstacles that allow the rover to increase essentially obstacle-crossing ability (Fig. 7) [15].

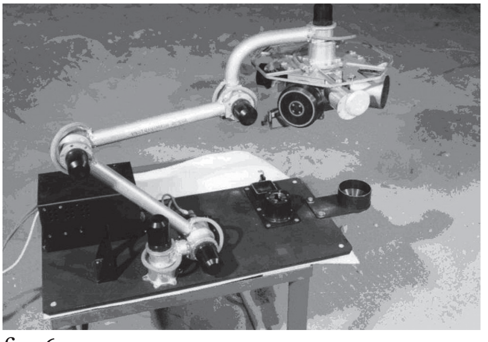


fig. 6

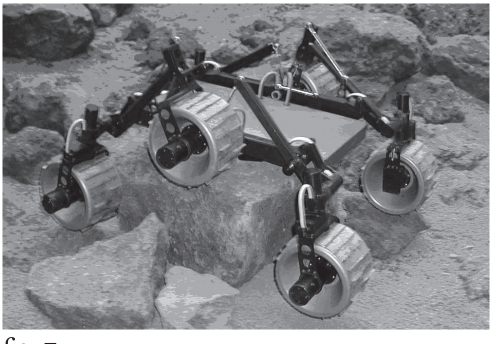


fig. 7

Fig. 8 shows a general view and technical characteristics of the *Chassis Marsochod* Autonomy (CMA) demonstrator created at VNIITransmash in co-operation with CNES and with the assistance of Rover Co. Ltd. [12]. Upper picture shows transport position, and lower one shows operating position. New technical decisions in this case are as follows: decisions ensuring a good chassis conversion and its unfolding into operating position by means of own drives; possibility to realize a wheel-walking mode of motion; possibility of cinematic turn of the chassis when drives for wheel turning are absent; possibility to regulate normal reactions of soil to wheels over the wide range (up to breaking off a pair wheels from the surface and its lifting). All this ensures surmounting obstacles with a height exceeding the wheel diameter and excellent maneuverability as well.

Main technical characteristics:	
Mock-up mass, kg	21
Payload mass, kg	30
Overall dimensions in transport position, mm:	
- length	900
- width	750
Type of frame	Articulated
Type of mover	Wheel-walking, 6×6×1+2
Maximum travel speed, km/h	0.2
Turning method	Side and cinematic

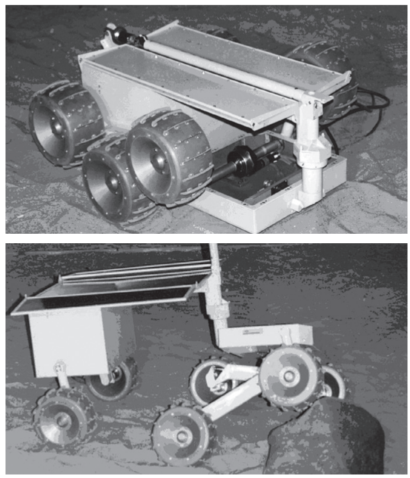


fig. 8

In July of 2006 the INTAS-CNES project No. 4063 'Innovative Mars exploration rover using inflatable or unfolding wheels' had been completed. The initiator of the project is CNES. Participants of the project from Russia are VNIITransmash (substantiation of characteristics and development of the concept of the Mars rover with inflatable and unfolding wheels), Keldysh Institute of Applied Mathematics of Russian Academy of Sciences (modelling of landing and motion of the Mars rover) and specialists of Lavochkin Association (materials and technology of manufacturing inflatable wheels).

Fig. 9 shows two concepts of the innovative Mars rover. A diameter of wheels is accordingly 1.4 m and more than 2 m. Such Mars rovers can climb the 30-degree loose soil grades even in the wheeled mode [16]. However, these newest technologies require experimental confirmation.

During the last years co-operation of Russian VNIITransmash and Chinese NOVERI was quickly being developed. In 2005 specialists of these enterprises carried out joint running tests of the planetary rover demonstrator with a mass of 150 kg and wheel diameter of 0.38 m in a soil canal (on a stand 'Soil canal')

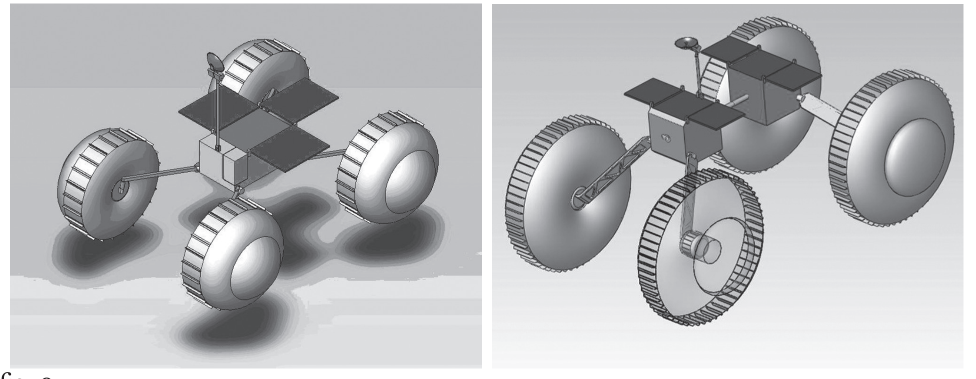


fig. 9

at VNIITransmash and at NOVERI (Fig. 10) [17]. Tests showed convincingly advantages of the wheel-walking mode of motion which are a new one for the Moon rovers. Tests were conducted using dry quartz sand. The demonstrator could not climb a grade of more than 20 deg when moved in the wheeled mode. When moving in the wheel-walking mode the demonstrator climbed the grade equal to angle of repose for the dry sand (~30 deg).

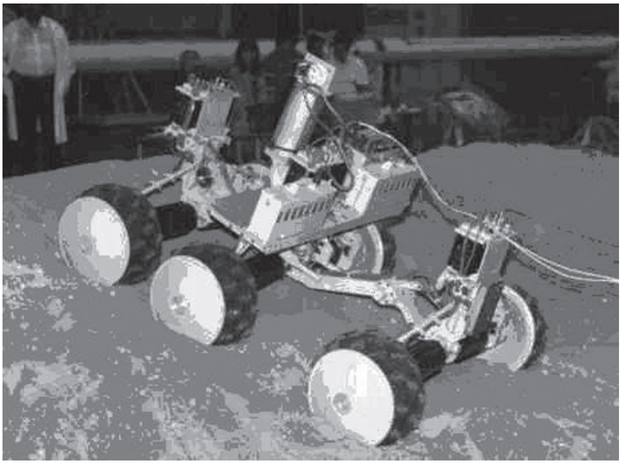


fig. 10

4. METHODS OF MOTION OVER THE SURFACE OF CELESTIAL BODIES

A wheeled mode of motion is optimum for the Moon rover of the class considered (purpose is transportation of the cosmonauts and equipment; a mass is less than 1,000 kg). A caterpillar mover will be possibly claimed just as on the Earth for the heavy Moon rover of another purpose, for example, intended for processing of soil.

Main advantages of a wheel are well known. They are simplicity and the lesser structural mass, high efficiency and reliability. The new advantage has been revealed by Russian scientists and designers. This is possibility to use a combined wheel-walking mode (WWM) of motion over the very complex sections of a route. Application of the WWM, of course, complicates the running gear design, increases amount of drives and sensors that brings to some complication of apparatus and program parts of a control system.

However these shortcomings are not so serious that possibility of quality improvement of soil-crossing ability and obstacle-crossing ability even with not great diameter of wheels (less than 0.5 m). The fact is that mock-ups with such wheels climb the loose soil grades with angle of repose when moving in wheel-walking mode.

As to other modes of movement — walking (with breaking off the foot from the surface), jumping, rotor-spiral (screw), somersaulting — in our opinion, only walking mode has perspectives to compete with a wheel. Problems of control in flight are not decided for jumping mode. The rotor-spiral and somersaulting modes require considerable specific power inputs non-compared with the specific power inputs for wheeled mover.

5. CONCEPTS OF COMPONENTS FOR PLANETARY ROVERS' LOCOMOTION SYSTEMS

High effectiveness of locomotion systems are ensured in the stage of designing, firstly, by correct choice of the common concept for the planetary rover, which most fully meets the problems of the specific space mission and, secondly, by means of quality of technical decisions on a design of main components of a locomotion system and their harmonious coupling as a part of a new mobile space rover. Especially important problem is designing of a running gear which consists of the mover, suspension, bearing structure and electromechanical drives of locomotion system. Just these assemblies in many respects define a technical appearance not only of the locomotion system but the planetary rover as a whole. Some key creating technologies of the running gear approbated by VNIITransmash specialists.

- A multiwheel mover. The wheel arrangements are 6×6, 6×6×4 (the last figure corresponds to the quantity of swivel wheels) and 6×6×6.

Advantages of wheel arrangement 6x6 as compared with 8x8 are possibility to increase a wheel diameter; a lesser LS mass at approximately equal cross-country ability and turning ability.

Advantages of the wheel arrangement 6×6 as compared with 4×4 are more high reliability of fulfillment of side turning with R_t=0; surmounting the more high ledge (scarp) and bench (counterscarp) (when diameters of wheels are equal); decrease of pressure in the wheel contact spot with soil.

- Realization of side (by means of wheels reverse of opposite sides) turning with R_t=0 for the planetary rover with different versions of the wheel arrangement.

Advantages of wheel arrangement 6×6 is a lesser LS mass. Limitation is a wheel track should be approximately equal to the wheel base (L≈B).

Advantage of the wheel arrangement 6x6x4 are more high reliability of fulfillment of side turning with R_t =0, decrease of limitations on ratio of a wheel track and a wheel base.

Advantages of the wheel arrangement 6×6×6 are preservation of all advantages of the previous wheel arrangement and, besides, ensuring of possibility to change direction of the planetary rover motion by means of preliminary swivel of all wheels at the same angle (without turning of the planetary rover).

- Application of metallic wheels of two types:

- rigid wheels (including wide-profile wheels) for the automated distance-controlled planetary rovers with minor (less than 2 km/h) travel speed;
- metal-elastic wheels with a net tyre and tyre-deflection stops.

Advantage of rigid wheels are high axial stiffness necessary to realize side turning with R_t =0; possibility to organize a developed system of high grousers increasing soil-crossing activity on weak-coherent soils; simplicity of manufacturing.

Advantage of metal-elastic wheels are: increasing of the wheel contact spot with soil that increases soil-crossing activity on weak-coherent soils; additional cushioning.

- All-wheel drive scheme of the electromechanical transmission; drives and brakes installed into each wheel.

Advantage of the all-wheel drive scheme is possibility to increase cross-country ability and turning ability.

Advantages of the transmission with the individual electromechanical wheel drive are simplicity and convenience of splitting energy of the on-board system of power supply by means of cables.

- Application of lever-articulated suspensions of two types:

- independent lever-articulated suspensions with longitudinal or transversal swinging of levers and elastic elements in the form of bar spring of torsion (torsions);
- balance-articulated interdependent cinematically non-elastic suspensions.

A field of applying suspensions of the first type is the manned and remote-controlled planetary rovers with a travel speed more than 0.5...1 km/h. Their advantages are increase of smoothness of motion, decrease of overload level for equipment when moving; ensuring a contact for all wheels with soil when surmounting complex obstacles. Shortcoming is significant irregularity of loading wheels when moving on a grade and over complex relief.

A field of application of suspensions of the second type is the autonomous or remotely controlled planetary rovers with not great (less than 0.5...1 km/h) travel speed. Their advantages are possibility of optimization of kinematics of balance-articulated suspension mechanisms of all wheels to exclude their longitudinal movement and to increase obstacle-crossing ability by means of levelling wheel loads.

- Application of load-carrying frames of two types:

- rigid frame including the frame in the form of a bottom of the pressurized container;
- articulated sectional frames.

Advantages of rigid frames are convenience of organization of the common pressurized container or common platform for all on-board equipment; absence of additional joints; convenience of tracing cables; convenience of organization of solar-battery panels.

Advantages of sectional frames are increase of soil-crossing ability and obstacle-crossing ability by means of smoothing wheel loads and possibility of application of wide-profile wheels to ensure the effect 'without clearance' machine; high static and dynamic overturning stability.

- Application of the combined wheel-walking mover, which up to now has been debugged as applied to planetary rovers at VNIITransmash only.

Advantage is qualitative increase of soil-crossing and obstacle-crossing ability of locomotion systems having even not great 'traditional' wheel diameters.

Application delay of a new type of the mover for planetary rover in world practice is explained, in our opinion, by the fact that designers do not have methods of synthesis of the wheel-walking mover stated in detail only in the Russian-language literature [7,8].

REFERENCES

1. Mobile laboratory on the Moon – Lunokhod-1 / Ed. by Vinogradov A.P. V . 2. Moscow: Nauka, 1971.
2. Mobile laboratory on the Moon – Lunokhod-1 / Ed. by Barsukov V.L. V. 2. Moscow: Nauka, 1978.
3. The first panoramas of the lunar surface. Moscow: Nauka, 1960—1969. V. 1—2.
4. Lunar Roving Vehicle – Apollo XVII, 1972, http://www.hg.nasa.gov
5. Kemurdjian A.l., Gromov V.V., Cherkassov I.I., Shvarev V.V. Automated stations for studying the surface cover of the Moon. Moscow, Mashinostroenie, 1976.
6. Dynamics of planetary rovers / Avotin E.V., Bolkhovitinov I.S. Kemurdjian A.L., Malenkov M.I., Shpak F.P. Moscow: Nauka, 1979.
7. Movement on soils of the Moon and planets / Gromov V.V., Zabavnikov N.A., Kemurdjian A.L., Kadzukalo I.F., Malenkov M.I., Naumov V.N., Nazarenko B.P., Rozdenstvensky Yu.L., Moscow: Mashinostroenie, 1986.
8. Planetokhody (Planetary rovers). 2nd revised and added edition / Kemurdjian A.L., Gromov V.V., Kadjukalo I.F., Malenkov M.I., Mishkinjuk V.K., Petriga V.N., Rozentsveig I.I. Moscow: Mashinostroenie, 1993.
9. Kemurdjian A. From the Moon Rover to the Mars rover // The Planetary Report. V. 10. № 4. 1990. P. 4—11.
10. Kemurdjian A.L., Gromov V.V., Shvarev V.V. Investigation of physical and mechanical properties of extraterrestrial soils // Successes of Soviet Union in

investigation of space. Moscow: Nauka, 1978.

11. Malenkov M., Gromov V., Koutcherenko V., Manukyan A. Highly Effective Locomotion System of Planet Rovers for Exploring and Mastering the Moon and Planets // Proc. of the Int. Conf. «Systems and technologies of the future learning and mastering of an outer space», Moscow, ISTC, 2003.
12. Maurette M., Lamboley M., Koutcherenko V., Malenkov M. Development of the Demonstrator of the Self-propelled Chassis for the Mars Rover on the INTAS-CNES Project // Proc. of the «ASTRA-2002», ESTEC, Noordwijck, The Netherlands, 2002.
13. Malenkov M., Putz P., Planet Rover Manipulator for Investigation of Moon's and Planet Surface. Proc. of the 4th Workshop ASTRA'96, ESTEC, Noordwijk, The Netherlands, 1996.
14. Winnendael van M., Malenkov M., Koutcherenko V.. The Lunar Rover Mockup of ESTEC's Planetary Utilization Testbed // Proc. of the ASTRA'98, ESTEC, Noordwijk, The Netherlands, 1 — 3 Dec. 1998.
15. Koutcherenko V., Bogatchev A., Development of the concept of Mars Rover's locomotion chassis.Europe project's *Exo Mars* // The Report of Int. Workshop Planet Rovers and Earth-based robots, St.-Petersburg, VNIITransmash, 2004.
16. Koutcherenko V., Malenkov M., Vladykin S., Kozlov O. The concepts of locomotion system of planetary rovers with inflatable and unfolding wheels // The Report of the Int. Workshop "Planet Rovers and Earth-based robots", St.-Petersburg, VNIITransmash, 2005.
17. Malenkov M.I., SuBo, Gromov V.V., Koutcherenko V.I., Kuznetsov I.S., Stepanov V.V., GuoXuan Lu, Wenming Wang Key Technologies of the Moon Exploration — realization and perspectives of creation of highly effective Locomotion systems for the Moon rovers // The Proceedings of 8th ILEWG Conference on Exploration and Utilization of the Moon, Journal of Astronautics, Chinese Society of Astronautics. 2007. V. 28 Sup. P. 105—114.

INTERNATIONAL COOPERATION IN SPACE EXPLORATION

R.Z. SAGDEEV,
University of Maryland,
USA
S. EISENHOWER,
The Eisenhower Institute,
USA

UNITED STATES—SOVIET SPACE COOPERATION DURING THE COLD WAR[1]

The Space Age spawned two outstanding space programs as a result of the hot competition between the United States and the Soviet Union. Both countries gave primary emphasis in their space efforts to a combination of national security and foreign policy objectives, turning space into an area of active competition for political and military advantage. At first, this charged political environment accommodated nothing more than symbolic gestures of collaboration. Only in the late 1980s, with warming political relations, did momentum for major space cooperation begin to build. As the Soviet Union neared collapse, with its ideological underpinnings evaporating, the impetus for the arms race and competition in space declined, allowing both countries to seriously pursue strategic partnerships in space.

Throughout the years between 1957 and 1991 the USA and the USSR both increased areas of cooperation, including space, as a symbol of warmer relations while cutting cooperation off when ties worsened.

The birth of the Space Age following the Soviet launch of Sputnik came out of the confluence of two seemingly incompatible developments. From the end of World War II, the Soviets made rockets their most important military asset. By the mid-1950s, they were ready to test their first intercontinental ballistic missile (ICBM). In 1957, the International Geophysical Year was launched, a multinational effort to study Earth on a comprehensive, coordinated basis. To highlight the effort, organizers had urged the United States and the Soviet Union to consider launching a scientific satellite. On Oct. 4, 1957, a seemingly routine test launch of a Soviet ICBM (now known as the R-7 rocket) carried the first artificial satellite to orbit.

Sputnik's launch had dramatic repercussions for the Cold War rivals. After reaping the first political dividends from military rocket technology, the Soviets continued to pursue a highly classified military-industrial approach in developing its space program. Conversely, the US government decided to make NASA a purely civilian enterprise, while focusing its military space efforts in the Pentagon and intelligence community.

Early on, President Dwight D. Eisenhower pursued US-Soviet cooperative space initiatives through a series of letters he sent in 1957 and 1958 to the Soviet leadership, first to Prime Minister Nikolai Bulganin and then to Premier Nikita Khrushchev. Eisenhower suggested creating a process to secure space

1 Abridged version of the article first published in *50 Years of Exploration and Discovery* edited by NASA. We are grateful to NASA's administration for the permission to include the article into this book.

for peaceful uses. Khrushchev, however, rejected the offer and demanded the United States eliminate its forward-based nuclear weapons in places like Turkey as a precondition for any space agreement. This would be the first of many times when space was linked with nuclear disarmament and other political issues.

Meanwhile, the United States energetically proceeded with its multinational initiative under the umbrella of the United Nations to develop a legal framework for peaceful space activities. This eventually led to the Outer Space Treaty and creation of the United Nations Committee on the Peaceful Uses of Outer Space, which the Soviet Union eventually joined.

In the scientific community, the role of an international space science union was assumed by the Committee on Space Research, with its unusual charter giving a mandate to both superpowers to appoint vice presidents. This arrangement opened an opportunity for dialogue and informal contacts between American and Soviet space officials. Academician Anatoli Blagonravov, the Soviet Union's representative for negotiating multilateral space science cooperation agreements, became the group's first appointed vice president. However, nothing could happen in the body without Kremlin approval.

The civilian nature of NASA, legislated in the 1958 Space Act, made it possible for the American researchers to collaborate on and disseminate scientific advances, an opportunity envied by many of us Soviet scientists. The actual work and industrial efforts for the Soviet space program were run under the classified umbrella of the Ministry of General Machine Building, with its enormous and rapidly expanding network of design bureaus and production facilities. The military was its principal client. The military also owned and operated every launch site and the network of ground control centers. The ministry had to report to the Communist Party's Central Committee and the Commission on Military-Industrial Issues of the Council of Ministers. Work beyond defense contracts was given secondary priority.

As a result of this critical dependence on the military, the Soviet aerospace industry relied entirely on domestic hardware, all the way down to the tiniest individual microcomponents. This resulted in an internationally isolated technological culture that would have created enormous barriers of incompatibility for any joint endeavor.

In April 1960, in advance of a planned Eisenhower-Khrushchev summit meeting, the leadership of Moscow's scientific community was anticipating a chance for major breakthroughs in bilateral cooperation, perhaps including the space area, following Eisenhower's *Atoms for Peace* initiative. However, the much expected summit was cancelled in the aftermath of the May 1 downing of a U-2 spy plane over the Soviet Union.

Early in his presidency, John F. Kennedy made repeated attempts to engage the Soviet Union in space cooperation. In his inaugural address, Kennedy said, 'Let both sides seek to invoke the wonders of science instead of its terrors. Together let us explore the stars'. Khrushchev, still persuaded of the eternal supremacy of Soviet rocketry, was not moved. Less than three months after Kennedy's inauguration, on April 12, 1961, Soviet cosmonaut Yuri Gagarin became the first human to escape Earth's gravity. In the aftermath of his brief flight, the piloted component of the Soviet space program rapidly grew to become indisputably dominant over any other type of space activity. Official Soviet propaganda was obsessed with everything that happened in orbit, including elaborate descriptions of the cosmonauts' menu at their last breakfast and all of the details of their physical exercise program. At the same time, the Soviets were left far behind in other key areas of space technology. Their first geostationary telecommunication satellite was launched 11 years after its American counterpart. In the case of getting meteorological data from a geostationary location, the gap was even bigger.

Despite the continued space competition between the United States and the USSR, Khrushchev sent Kennedy a letter raising the possibility of space cooperation on a modest level after John Glenn became the first American to orbit Earth on Feb. 20, 1962. That led to two rounds of discussions between NASA's Deputy Administrator Hugh Dryden and Soviet academician Blagonravov. An agreement led to the opening of cooperation in three areas: 1) the exchange of weather data from satellites and the eventual coordinated launching of meteorological satellites; 2) a joint effort to map the geomagnetic field of Earth; and 3) cooperation in the experimental relay of communications. This link became a primary forum for subsequent US — USSR interaction on space.

There were large differences between the two negotiating partners. The Soviet Academy of Sciences did not run the space program, but rather served as an official front for a vast network of secret enterprises controlled by the military and Communist Party apparatus. An asymmetry existed also in the fact that while the Russians knew about the American planning process, everything about the Soviet space program was a classified secret. It was difficult to persuade our Soviet authorities, including the president of the Academy of Sciences, academician Mstislav Keldysh, that we should reciprocate. The Soviet system had a different culture and mentality.

Following the ouster of Khrushchev in October, 1964, the new Soviet leadership of Leonid Brezhnev and his colleagues took even a harder line toward overall US-Soviet relations. Brezhnev previously had served as the curator of the military industry on behalf of the Politburo. He knew well there was a 'missile gap' in favor of the United States, and he was about to embark on an unprecedented build up of deterrent forces. The negative atmosphere at higher levels was reflected in the Soviet academy's dealings with NASA. Soviet opposition to the US war in Vietnam led to more bitterness.

In December, 1968, only weeks after Richard Nixon's election, *Apollo-8* orbited the moon, followed by the lunar landing of *Apollo-11* in July, 1969. Meanwhile, the Soviet Union experienced a series of failures in its manned lunar program. The opportunity for using dramatic space cooperation efforts as a means of reducing the US-Soviet Cold War rivalry had passed. As painful as it was for the Soviet leadership, the time of their country's dominance in heavy rocket launching technology was over. Cooperation in space now would have to come at more modest levels. The triumph of the *Apollo* program signified a crucial benchmark in the superpower space race by ending Soviet leadership in space exploration. The Soviet Union was simply unable to match such largescale US efforts. Nor did the Soviets have an institutional structure like NASA that was capable of running a program like *Apollo* in an open and transparent way. While not ready to publicly admit their defeat, the Soviets argued that scientific work on the Moon could be better achieved robotically. Unmanned Soviet lunar missions, initially introduced as a shadow program with a much smaller budget than the manned version, occurred at the same time as the *Apollo* program. The *Lunokhod* moon rovers and sample return probes earned a great deal of admiration from international scientists. However, inside their close circle, the Soviet leaders, in a rude awakening, conceded that the era of Soviet dominance in space was gone forever.

The challenge for both sides was determining where to go next. While the Americans eventually pursued the development of the space shuttle, the Soviets embarked on a program to place crews in space for extended periods of time by building the *Salyut* series of orbital space stations.

In reality, that space station program was not the result of major brainstorming or serious debates about a new national vision for space exploration. It came from the spontaneous process of internal competition between rivals within the Soviet aerospace industry. The Soviet military initially supported the approach, which was reminiscent of the US Air Force *Manned Orbiting Laboratory* project, which was canceled in 1969 after a single, unmanned launch. Reflecting military priorities, the key instrument on early *Salyut* stations was a big optical Earth observation camera, the Soviet version of 'open skies' technology. Of course, official propaganda said this mission had nothing to do with military interests.

After this type of assignment was passed to unmanned spy satellites, the real motivation for expanding the *Salyut* program became the desire to undertake long-duration flight. Longevity records for humans in space became the benchmark for judging the success of these flights. In order to move in that direction, the *Salyut* program worked to excel in two important areas: achieving the safety of its manned flight hardware and developing a solid base in space medicine. Eventually, these would be two of the most important contributions the Russians would make to the International Space Station partnership.

In the early 1970s, the Nixon administration sought to reduce US-Soviet tensions, and launched a major effort to reach a strategic arms limitation breakthrough, as well as new cooperation in space. In 1970, during a meeting with Keldysh, US Academy of Sciences President Philip Handler mentioned an American movie starring Gregory Peck and Gene Hackman called *Marooned*, in which Soviet cosmonauts helped rescue three US astronauts stranded in Earth orbit. Handler suggested the United States and the USSR develop a mutually compatible docking system that would make possible such rescues, as well as non-emergency space dockings. This imaginary movie scenario touched a chord within space communities on both sides, which already had experienced emergency situations in real life. Talks led to the *Apollo-Soyuz* Test Project docking mission of 1975, which developed compatible rendezvous and docking systems still in use today, and the establishment of a few topical working groups in different space science and applications disciplines.

Implementation of *Apollo-Soyuz* cooperation was dictated by the political will of the two countries' political leadership. The cooperation presented a serious management challenge for both sides, given the overall lack of compatibility between the two space programs. NASA had to work with a counterpart that could not even be clearly identified. The Ministry of General Machine Building was still shrouded in secrecy and Soviet authorities instructed the Academy of Sciences to act as a cover for all activities during *Apollo-Soyuz*. Soviet industry experts had to introduce themselves as employees of the Space Research Institute and military officers from Soviet Space Command changed into civilian clothes while insisting that the Soviet academy administered the launch site in Baikonur, Kazakhstan.

Despite this artifice, the docking in orbit in July, 1975 was a rare and dramatic display of US-Soviet friendliness during the depths of the Cold War. Leonid Brezhnev and President Gerald Ford exchanged messages of friendship and congratulations. This was to be the last dramatic international handshake in space for years to come. Soon after the flight, both sides met to discuss potential follow-on space projects and agreed to establish a special bilateral working group. I chaired the Soviet group and worked with NASA's Charles Kennel on a scenario in which a specialized science module, a blend of Russian and American station designs, could be delivered to orbit by the US space shuttle. Unfortunately, politics intervened again. Incoming President Jimmy Carter was concerned by congressional charges that the Soviets had obtained valuable US technology during the *Apollo-Soyuz* Test Project. By late 1978, the Carter administration had ended discussions on additional cooperation with the Soviets. After the Soviets invaded Afghanistan in December, 1979, any hope of significant cooperation in space was gone. The United States pursued cooperation with Europe through projects such as a *Spacelab* module that could ride aboard the space shuttle, while the Soviets maintained their focus on flying the manned *Salyut* space stations.

On the planetary exploration front, we were quite impressed by the successes of the Mars *Viking* missions and the *Voyager* missions to Jupiter, Saturn, Uranus, Neptune and the outer limits of the Solar system. At the same time the principal Soviet robotic missions were repeatedly directed toward Venus.

Nevertheless, the Soviet robotic space program was successful in its own right. The Soviets learned also from US achievements. Anticipating the success of the US *Viking* mission, the Soviet Academy of Sciences decided to abandon Mars as a priority and see how the American program would develop. The open and predictable nature of the US space program gave Soviet scientists an opportunity to find their own niche with realistic projects that would have a scientific impact and avoid direct competition.

Our *Venera* program to Venus was quite successful. Following simplistic probes in the late 1960s, we managed to deliver sophisticated hardware to the planet's surface in 1975 and send back panoramic pictures. Because the United States and the USSR agreed to share the results of NASA's *Pioneer Venus* mission in 1978 and the Soviet *Venera* missions, scientists and space experts on both sides placed enormous symbolic and scientific value on the results of these joint efforts.

US-Soviet cooperation in life sciences and biomedical research also took root in the 1970s. In 1977, seven US biological experiments or medical devices flew aboard the Soviet *Cosmos-936* mission, which also carried experiments from France and a number of Soviet bloc countries. This mission investigated the impact of long-duration spaceflight on the human body. A later *Cosmos* mission, *Cosmos-1129* in 1979, carried 17 additional US experiments and devices. And on May 6, 1979, the United States and the USSR signed a treaty that provided for the deployment of an international system of emergency beacon receivers aboard satellites.

When Reagan was elected to the presidency in 1981, Cold War tensions were rising. The Soviet invasion of Afghanistan, imposition of martial law in Poland and NATO's placement of *Pershing* rockets and cruise missiles in Europe – which was countered by deployment of the Soviets' SS-20 medium-range nuclear missiles – characterized the tenor of the period. In the midst of the Poland martial law crisis, the Reagan administration announced on Dec. 29, 1981, that it would allow the US-Soviet space cooperation agreement, due for renewal in May, 1982, to lapse. Mutual suspicion grew to the point that the Soviets began attributing potentially aggressive intentions to the *Space Shuttle* Program. It would be another 10 years before the conditions finally were ripe again for cooperation.

Nevertheless, in the absence of a formal intergovernmental agreement, the White House authorized low-profile cooperation on a case-by-case basis. Among the activities that continued were the satellite-based search and rescue efforts, which was based on the coordinated use of the US-Canadian-French SARSAT and the Soviet COSPAS satellites to locate airplanes or ships in distress. By the mid-1980s, the effort had helped save more than 400 people. NASA also was allowed to continue working with the Soviet Union in space biology and medicine. As part of that effort, four US medical devices were used in experiments on the 1983 *Cosmos-1514* mission, which was devoted to primate research. That tacit format of interaction led by Soviet academicians Oleg Gazenko and Anatoly Grigoriev and NASA's Dr. Arnauld Nicogossian, later would serve as an example for future cooperation between the Russian space station *Mir* and *Space Shuttle* programs and on the International Space Station. Meanwhile, exchanges of planetary data continued, but discussions of future cooperation in planetary exploration were cancelled.

The US side was pragmatic about keeping up its contacts with Soviet scientists during these times times of political tensions. Regular consultations on space science-related issues, for example, were carried out through a channel between the US National Academy of Sciences and the Soviet Academy of Sciences. Americans were keenly interested in learning about the effects of long-duration flights on the human body – an area where the USSR enjoyed a monopoly during NASA's six-year hiatus in human spaceflight from 1975 — 1981.

In addition to these cooperative activities, Soviet and American space scientists regularly met at Committee on Space Research sessions. Aerospace engineers and officials from industry also maintained a similar engagement under the umbrella of the International Astronautical Federation.

During what many would consider the coldest period of bilateral relations in the early 1980s, these contacts produced a very special cooperative project that sought to explore Halley's comet. The United States and the USSR both participated in the Inter Agency Consultative Group, which was set up in 1981 to bring together space and groundbased studies of the comet during its 1986 passage through the inner Solar system.

After deciding not to send a spacecraft to view the comet, the United States agreed to play a supporting role, which involved providing ground-based observation data on the comet. This data was used to support the parallel Soviet *Vega-1*, *Vega-2* and European Space Agency *Giotto* missions. The success of the encounter with the comet was to be critically dependent on precise navigation. Scientists from NASA's Jet Propulsion Laboratory in Pasadena, California, suggested a brilliant technical scenario for the *Vega* and *Giotto* spacecraft to use at the approach to the comet. This had to be done a few days prior to the arrival of *Giotto* in order to help it home in on the celestial whereabouts of the ultimate target: the comet's elusive nucleus. The whole procedure required close cooperation in real time. NASA's Deep Space Network was given all of the necessary parameters from the Soviet spacecraft communications systems, then both sides performed pre-flight calibration tests of the hardware. This helped *Giotto* navigate much closer to the comet's nucleus, providing scientists with outstanding data and producing some of the most awe-inspiring video footage ever taken in space. Several months before, when the Soviet *Vega* spacecraft had to release meteorological ballons in the atmosphere of Venus, the Deep Space Network played a crucial role in getting the first direct signals from these balloons and continued to track them as they were buffeted by Venus' unusual atmospheric circulation.

Ironically, such successes were achieved despite continued chilly relations between the two governments. Several private groups, however, worked to keep US-Soviet space ties alive. Among them was the new Planetary Society, which was created in 1979 by well-known astronomer Carl Sagan, NASA Jet Propulsion Laboratory Director Bruce Murray, and their associate, physicist Louis Friedman. After its founding in 1985, the Association of Space Explorers, composed of people who had flown in space, also became an important forum for discussions on the benefits of US-Soviet cooperation in human spaceflight. These efforts would provide a powerful impetus for getting stalled US-Soviet space cooperation back on track.

Not long after Reagan was elected president, NASA urged him to approve a space station to rival the Soviet station program. In his January 1984 State of the Union address, Reagan announced he was directing NASA to 'develop a permanently manned space station...within the decade' and 'invite other countries to participate'. Peggy Finarelli, a senior official in NASA's international office at the time, recalled that Reagan's approval of what became known as Space Station *Freedom* was 'a leadership issue very much in the context of the Cold War. We were challenging the Soviets in the high ground of space. We had to say that *Freedom* would be bigger and better than the Soviet space station'.The original estimate was that *Freedom* would cost about $8 billion. It was envisioned to be in orbit by 1992 in order to celebrate the 500th anniversary of Christopher Columbus' discovery of America.

While the Soviets were not invited to join the *Freedom* project, the Reagan administration indicated its willingness to resume space cooperation with the USSR prior to the 1984 State of the Union address. Only days before the speech, the administration privately suggested to Moscow a simulated space rescue demonstration mission in which US astronauts from the space shuttle would assist Soviet cosmonauts aboard a *Salyut* station. Both privately and publicly, the Soviet response was cool, because of the perceived asymmetry of a mission in which the Soviet crew was in trouble and the US crew would act as rescuers.

The Soviet government also revived the notion from the Khrushchev era that space cooperation would be possible only if there were progress in space arms control. The primary point of contention was the Reagan administration's proposed Strategic Defense Initiative, which had been announced in March, 1983. From the start of the Reagan administration, however, pressure for cooperation in space had been mounting.

The US Senate issued a more formal call for renewal of US-Soviet space cooperation with passage of Joint Resolution 236 on Oct. 10, 1984. President Reagan signed the resolution on Oct. 30, noting US readiness 'to work with the Soviets on cooperation in space in programs which are mutually beneficial and productive'.

When Mikhail Gorbachev emerged as the Soviet leader in 1985, Reagan thought he had found a willing partner. Gorbachev was interested in reducing the Soviet defense budget, and with the so-called Euromissile issue still unresolved, his government quickly signaled its readiness for a new round of arms control negotiations with the United States. When Reagan and Gorbachev met in Geneva that November to discuss arms control, they also signed an agreement on scientific cooperation. Once again, cooperation was symbolic of a thaw in the Cold War. However, Gorbachev still expressed strong Soviet opposition to the Strategic Defense Initiative and space was not included in the agreement. The Soviets had linked space cooperation to a demand that the United States abandon its plans for the initiative altogether.

Only three months after the Geneva summit, a tragedy occurred that would set the US space program back several years – the space shuttle *Challenger* disaster. Little noticed at the time was a diplomatic breakthrough that occurred only a few weeks after the *Challenger* accident. On Feb. 20, 1986, the Soviets launched the first of six modules that eventually would comprise the *Mir* space station, and in the wake of the *Challenger* accident and the launch of *Mir*, the Kremlin finally agreed to decouple non-military space issues from the Strategic Defense Initiative. The United States and the Soviet Union subsequently signed a five-year agreement on space cooperation in April, 1987. A number of joint scientific projects were agreed to, although there was no mention of cooperation in human spaceflight. More importantly, in an exchange of letters between Gorbachev and Reagan the previous summer, the link between arms control progress and renewed space cooperation was dropped. This paved the way for both sides to take meaningful steps toward actual cooperation.

J. BLAMONT
Centre National
d`Etudes Spaciales (CNES),
France
jacques.blamont@cnes.fr
G. HAERENDEL
Max-Planck Institute
for Extraterrestrial Physics,
Garching, Germany
hae@mpe.mpg.de

AN INITIATIVE TOWARDS INTERNATIONAL COOPERATION IN SPACE EXPLORATION

1. PRECEDENT AND RECENT INITIATIVES

The Human Space Exploration program is widely conceived to stem from the initiative taken by President Georges W. Bush in January, 2004. Its essence can be condensed in two ideas:

1. The Space Exploration initiative aims at sending men to the Moon and to Mars in the next half century, under an agenda to be defined.
2. The program will be placed under American leadership, but international cooperation is needed.

President George W. Bush characterized the Exploration Vision as ‘a journey, not a race’, recognizing that it will take place over many decades and will involve numerous elements, many of which will be large and complex programs in their own right. He also stated that: ‘We will invite other nations to share the challenge and the opportunities of this new era of discovery’ and then called ‘on other nations to join us on this journey, in a spirit of cooperation and friendship’.

However, long before, i.e. in November, 2001, ESA issued a call for ideas in human and robotic exploration missions. The majority of proposals received concerned in situ exobiology, Mars Sample Return and human exploration. Other main targets were the Moon, Europa, and asteroids. To be consistent with the long-term goal of human exploration, the program under the name *Aurora* was defined as follows: ‘The objective of the *Aurora Programme* is first to formulate, and then to implement a European long-term plan for the robotic and human exploration of the Solar system bodies holding promise for traces of life’.

Shortly after the speech of President Bush, on February 16, 2004, the Director General of ESA, Jean-Jacques Dordain and the EU Commissioner, Philippe Busquin, issued a joint reaction to the Bush-Initiative with the spirit that a coherent European Space policy would not make sense if not seen in the wider global context: ‘Unlike the days of the Cold War, getting to the Moon and Mars is not about proving one's superiority over a political enemy. It is about all of us, around the world, working together for a common goal’.

How then should international, or rather a global cooperation be approached?

Emerging U.S. and European space exploration programs exhibit a number of similarities. Both are based on an open ended vision for a sustained long-term effort starting with the Moon, then Mars. Furthermore both foresee programs developed through incremental, adaptive decision-making with no major new funding available. Canada, China, India, Japan and Russia also have Moon/Mars exploration plans. With capable and focused space exploration programs in Europe and elsewhere, the United States has a number of potential partners in the pursuit of its vision.

A successful cooperation between space agencies was organized during the years 1980 and 1990, called IACG (Inter-Agency Consultative Group). It started as a light structure coordinating a handful of missions addressing one single object, Comet Halley. That was Phase 1. Founders were the Agencies providing a major contribution to the programme (ESA, InterKosmos, ISAS and NASA). Subsequently IACG widened its field of action to the Solar-Terrestrial theatre with ISTP and its more than 20 projects. This was Phase 2. Membership remained the same, reflecting the predominant role of the four founding Agencies in the programme. In Phase 3, from 1999 to date, IACG tried to widen its field of action even more by encompassing the entire classical space science domain, and include in its agenda the advanced identification and harmonization of the Member Agencies initiative in particular domains of joint interest. This effort met no support from the Member Agencies, and IACG was terminated in early 2006. However, some Working Groups remain active such as the present ones on ILWS (International *Living With the Star*, NASA led), on search for extra-solar terrestrial planets (ESA led), and the somewhat atypical IMEWG (The International Mars Exploration Working Group) and ILEWG (The International Lunar Exploration Working Group). While these working groups have some limited usefulness, they do not constitute the type of international cooperation that is needed and no progress has been made in the direction of concrete measures. The working group are satisfied by exchanging information... which is available on the Net.

The topic of international cooperation was addressed at the American Institute for Aeronautics and Astronautics 7th Workshop on International Space Cooperation, which was held in Anchorage, Alaska, May 3—6, 2004. A working group on 'International Co-operation in the Context of a Space Exploration Vision' had as mandate to discuss the modalities of defining and implementing an exploration vision as a coordinated international endeavour.

The workshop arrived at an interesting approach, considering that all national space exploration activities, taken together, comprise an inherently global enterprise-in effect, a Virtual Program of Programs.

Rather than trying to develop a cooperative concept for exploration as a whole, the Virtual Program would be composed of a coordinated set of individual activities, with each activity employing the most sensible international arrangement, as determined by the specific partners involved. Not all partners would be involved in all activities, and all activities would not necessarily be cooperative.

The coordination of the national programs would be provided by some 'International Space Council', which would only be an informal forum, at least in the initial stages.

In the words of Peggy Finarelli and Ian Pryke .

'The council membership would be open to any party that is actively involved in space exploration activities, committed to a synergistic, long-term virtual program of programs and willing to share information on its national program plans to that end. Although the process of the council need to be developed by its stake-holders, regular meetings at agency leadership levels would definitely be required. The council would operate flexibly over the long term, providing its members a forum for communication, consultation and coordination, leading ideally to an alignment of national exploration programs. The results of council deliberations would not be binding or controlling, but guiding. The council would also ideally promote the sustainability and continuity of the long-term global vision, so as to endure changes in various national commitments and possible transformational events in space programs in general'.

The July 12, 2004 issue of *Space News* also contained the editorial 'Cooperation and Competition' which expressed a number of the ideas the working group espoused. It noted that 'Different combinations of the current players — and new ones sure to come — should be working on a variety of exploration programs' and that 'if pursued in an independent but coordinated fashion, such programs could provide the balance between cooperation and competition necessary to keep things moving quickly instead of bogging down'. The working group's proposed paradigm shift is fully consistent with these ideas.

2. FOCUS AREA FOR COOPERATION

It is clear that what is needed is some equivalent to the IACG, but reshaped in order to meet the situation of space exploration during the 21st century, characterized by :

1. The appearance of new players,
2. The large size of NASA's programme compared to others.

We will take as a working hypothesis that the scope of exploration encompasses all aspects of planetary research including manned missions, occupation of the Moon and Mars and utilization of planetary grounds for science. All Space Agencies are not engaged in this domain with the same resolution, but the preceding considerations lead to the conclusion that, if the present attitudes remain unchanged, a major exploration initiative should be a co-operative venture among equals. Here lies the main hurdle, since nobody is equal in space exploration to the United States. This is the reason why the approach through an 'International Space Council' of 'separate but equal' partners seems to be sensible, and should constitute a first step.

We will not retain the word 'Council' but propose the term "Forum" for the structure we believe the Agencies should create, since we want to stress the fact that it would be a non compelling structure.

In order to integrate from the outset the already existing organizations of stakeholders, and with the aim of providing a more unified view of the programming of relevant missions, it is suggested that the two committees ILEWG and IMEWG should become a part of the International Space Exploration Forum.

However, exchange of information is not sufficient. If the idea of some ‘International Space Forum’ is certainly commendable, it falls short of providing enough fuel for a vigorous cooperation.

Under the umbrella of the Forum could exist structures providing eventually the possibility of joint actions. Following the idea of ‘equality’, a number of partners could cooperate in the following domains.

1. R&D: A number of specific products have to be continuously evolved, as technology advances, for the needs of exploration: components, instruments, systems, the burden of which could be shared between partners outside national and industrial restrictions.
2. Operations: Agreements could be passed between the various Deep Space Networks, not only for the availability of all existing facilities to the partner's missions, but also for the development of new ground capacities, including software.
3. Assets: A major goal of cooperation could be the joint development of ‘planetary facilities’ under integrated management as:

- Geostationary telecommunication network and navigation/localization network (GPS style) around Mars
- International automatic stations on the Moon or Mars
- International mobile laboratories on the Moon and Mars
- Planetary Internet.

4. Partnerships in missions: Possible joint ventures as sample return missions could be agreed upon with distribution of major tasks or systems among the participating partners
5. Access to space: global agreements with the industries possessing launch capacities could lower rocket prices.

This not exhaustive list of suggestions should be taken as illustrations of the potential contents of discussions among eventual partners of the International Space Exploration Forum.

3. PROCESS AND PROPOSAL

In proceeding along the here formulated lines one has to respect the very diverse levels of acceptance of the human element in the exploration program among the various nations and their people. This is manifested clearly, for instance, in the evolution of ESA's *Aurora* program from the Ministerial Conference in Edinburgh in 2001 to that in Berlin in 2005. While the number of participating European nations has increased, the word ‘human’ has been carefully avoided in the last program outlines, and the time horizon has been kept as close as consistent with the next robotic mission, *ExoMars*. By contrast, NASA is at present reformulating its science program with the goal to redirect freed funds towards the Human Exploration Program, with the Moon as its first target.

Taking account of this situation an avenue could be the formation of two groups of partners:

- a group interested in manned missions, in pursuance of the Bush initiative and following the original vision of ESA's *Aurora* program,
- a group interested in a more modest endeavour using only automatic systems.

The two groups would closely cooperate, at least at the programmatic level, but also on some technological developments, although each of them would be involved in two parallel strands of missions,

It is suggested that at the forthcoming COSPAR Scientific Assembly in Beijing, a panel would meet for creating (between Space Agencies) an International Space Exploration Forum, with the goal of paving the way towards a future global cooperation on space exploration by:

- integrating ILEWG and IMEWG inside this new and more permanent structure;
- analyzing the possibilities of joint actions for the elaboration of an international road map in planetary exploration;
- or even initiating some cooperative ventures through interested Agencies.