The night sky in southern Estonia, with the Milky Way visible at center. (Courtesy of Martin Mark, CC BY-SA 4.0.)

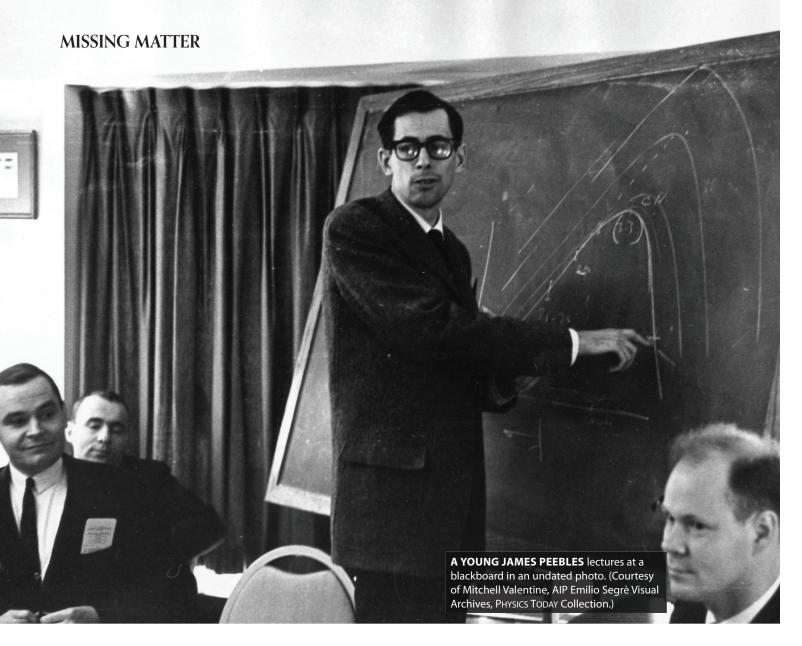
Jaco de Swart is an AIP Robert H. G. Helleman Memorial Postdoctoral Fellow at MIT in Cambridge, Massachusetts. A trained physicist and historian, he is currently completing a book manuscript on the history of dark matter.



Jaco de Swart

Groups on both sides of the Iron
Curtain hypothesized in the 1970s that
a yet-unknown material makes up
most of the mass in the universe.

wo papers that appeared in 1974 changed the face of the universe. Independently authored by separate collaborations, one in the US and the other in Estonia, they argued that galaxies are 10 times as massive and extensive than had previously been thought. Both groups combined various astronomical observations to show that most of the universe's mass is hidden in invisible clouds around galaxies. The universe itself, too, they illustrated, is heavier by a factor of 10 than had previously been believed, potentially changing human understanding of the fate of the cosmos. Their arguments marked "a watershed in our understanding of galactic structure, galaxy formation, and cosmology," read a review in the 1999 centennial issue of the Astrophysical Journal.1 Five decades ago those papers proposed the existence of what we now know as dark matter.



Today dark matter is not only one of the pillars of modern cosmology but also one of its central conundrums. The existence of unseen matter distributed throughout the universe is key to understanding cosmic structure and evolution: It explains how galaxies move about and why they exist in the first place. But at the same time, after decades of dedicated research and experimentation, the exact nature of dark matter—what the stuff is actually made of—is still unknown. Currently several dozen massive international experimental efforts, including ones in underground mines and in space stations, are attempting to detect evidence of hypothesized dark-matter candidates. The two papers from 1974 formed the basis of that profound hypothesis and initiated an exhilarating new era in cosmic understanding.

Here I tell the story of how those two papers made dark matter come to, well, matter. That story is unlike usual dark-matter histories, which typically center on the roles of astronomers Fritz Zwicky and Vera Rubin. In the 1930s Zwicky found that galaxies in clusters are unstable without extra mass, and in the 1970s Rubin observed that galaxies

rotate faster than their luminous mass would imply. Astronomy textbooks normally cite those observations as evidence for the existence of dark matter.

But facts and observations themselves do not tell a history (see the article by Matt Stanley, Physics Today, July 2016, page 38). To understand the origin of the case for dark matter, we need to know how prior observations made by Zwicky, Rubin, and others were interpreted to be evidence for its existence. In what context were they used to show that the universe had preponderous amounts of missing matter? Who started to care, and why? That happened independently 50 years ago on both sides of the Iron Curtain.

A search for two numbers?

Half the story starts with a prolific young astrophysicist named Jeremiah Ostriker. An expert on stars, Ostriker received a BA in physics and chemistry from Harvard University in 1959 before pursuing a PhD at the University of Chicago with Subrahmanyan Chandrasekhar, who was famed for groundbreaking work on stellar astrophysics that would

earn him the 1983 Nobel Prize in Physics (see the article by Freeman Dyson, Physics Today, December 2010, page 44). Under Chandrasekhar, Ostriker started a career in the physics of stars and their rotation. His PhD research was devoted to showing that there is a hard limit to how fast stars can rotate before they disintegrate. After a brief stint at the University of Cambridge, Ostriker received an assistant professorship in 1965 at Princeton University, where he continued his influential work on stellar physics.²

Ostriker's area of research—the properties and evolution of stars—had dominated astronomy since the 1930s. But the field's focal point began rapidly shifting in the 1960s. With the aid of Cold War—era technological developments, astronomers began opening new windows to the universe with observations across the electromagnetic spectrum. Interest resurged in Einstein's theory of general relativity. The

astronomical workforce increased dramatically, and a new generation of researchers began observing novel phenomena on the galactic scale and beyond, including quasars, pulsars, and the cosmic microwave background. Cosmology became a new focus for young astrophysicists.³

That new generation, Ostriker included, began to part ways with Edwin Hubble's classical cosmological endeavor, which was termed in the headline of a February 1970 Physics Today article by Allan Sandage "a search for two numbers": the Hubble constant, measuring the expansion of the universe, and the deceleration parameter, quantifying the rate at which the expansion is slowing. In the late 1960s New Zealand astronomer Beatrice Tinsley and colleagues had shown that the brightness of galaxies changes as they age. As Ostriker recalled in an interview, that work aroused "suspicion" of the traditional enterprise. "All of a sudden we realized galaxies have to evolve," he said.4 With that revelation, some of Hubble's classic cosmological tests were deemed unreliable. Researchers instead began using observations to work out the many physical processes that govern galaxies and the universe.

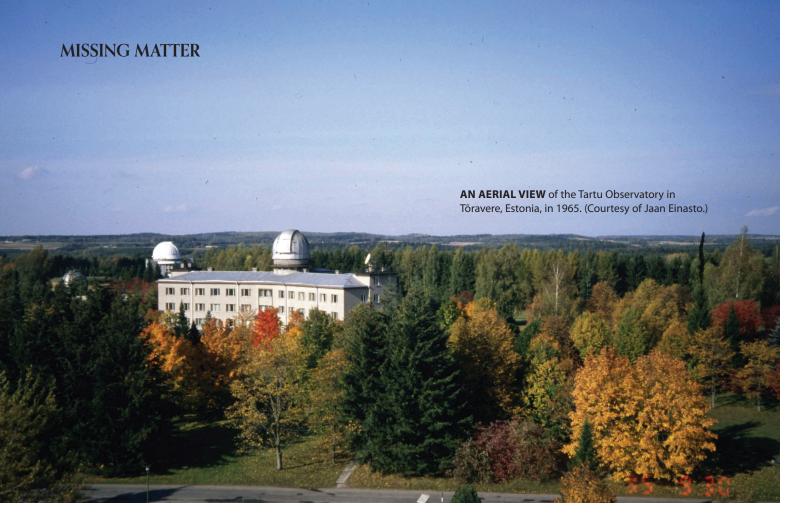
By 1971, when Ostriker was promoted to full professor at Princeton, he had shifted his focus from stars to galaxies. He used his expertise in

the evolution of stars to show how stellar processes could influence the total luminosity of galaxies during their lifetime. Another question Ostriker delved into was inspired by his graduate work on rotating stars: How do rotating galaxies maintain stability during their lifetime? Answering that question, however, required modeling galaxies with computers—a practice that Ostriker was not familiar with. He turned to his colleague James Peebles for help.

Can galaxies survive?

Born and raised in Winnipeg, Manitoba, Canada, Peebles got a BS from the University of Manitoba before moving to Princeton to study physics in 1958. Although he started in particle physics, Peebles eventually became charmed by the work of Robert Dicke, a gravitational physicist who by 1960 was well known for his unique approach, which involved





experimentally testing different gravitational theories.⁵ Dicke asked Peebles in 1964 to consider the consequences of a potential remnant from the universe's hot and dense early history: an observable cosmic background of microwaves in the sky. When the cosmic microwave background was observed in 1965, Peebles was standing at the cradle of the modern Big Bang theory. One of the pioneers of a new, "physical" cosmology, he worked on such topics as the synthesis of nuclear elements in the Big Bang and the formation of galaxies and the cosmic structure—work for which he received a share of the 2019 Nobel Prize in Physics.

When Ostriker knocked on his door, Peebles had just come back from a visit to the famed nuclear research facility in Los Alamos, New Mexico, where he was invited to help make sense of a highly energetic flash of gamma radiation detected by a satellite in 1969. (Although detected by a satellite meant to monitor nuclear weapons tests, that flash was later recognized as the first detection of a gamma-ray burst.) Peebles made good use of his visit: He used the facility's supercomputers—normally used to model nuclear weapons and explosions—to create the first simulations of galaxy clustering in the universe. It showed how a homogenous soup of mass would increasingly grow clumpy under gravity.

Following Ostriker's suggestion when he arrived back in Princeton, Peebles used his computer-punch-card dexterity and "banged out some n-bodies." To be precise, it was 500 bodies: His model simulated the stability of galaxies by using starlike particles moving under gravity in a disk. Ostriker

and Peebles quickly found that rotating galaxies in their model were "rapidly and grossly unstable." In their computation, disks of stars disintegrated after a single rotation.⁷

Something was off. It was well known that the Milky Way had existed for much longer than a single rotation. So how did it survive in the real universe? One explanation, they argued, involved rethinking the distribution of galactic mass. Instead of positing that all the mass was located in the bright disk, they proposed that more mass was in the spherical bulge of the galaxy, which they termed a "halo." That halo would help stabilize a galaxy as it rotated.

Few galactic dynamicists were enthusiastic about Ostriker and Peebles's new idea. Among other reasons for skepticism, it was unclear if their analysis would hold for all galaxies. Would it account for galaxies that are not rotationally symmetrical? If so, is a massive halo the only way to prevent instability? Many astronomers agreed with MIT astronomer Alar Toomre, who called the idea "a real migraine" a few months after Ostriker and Peebles's paper came out in 1973.8

The mass of the universe

Although Ostriker and Peebles's stability argument remained controversial, it managed to inspire Amos Yahil, a lecturer at Tel Aviv University in Israel, on leave at the time. Unsatisfied with the field of particle physics—in which he obtained his PhD at Caltech in 1970—Yahil had recently shifted to cosmology. While a postdoc at the Institute for Advanced Study in Princeton, New Jersey, in 1971–73, he began studying the

distribution of mass in the universe, which was understood to be one of the key parameters to understanding the fate of an expanding cosmos. Was the universe "open," meaning that it would expand forever? Or was it "closed," meaning that it would have so much mass that gravity would cause it to collapse together again sometime far in the future?

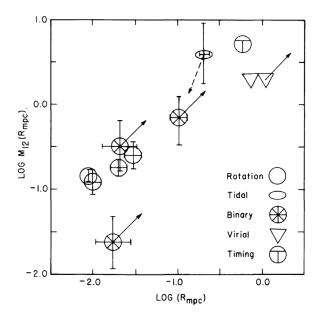
To study the universe's mass distribution, Yahil began modeling clusters of galaxies. Although other researchers had modeled galaxy clusters by using clear spatial boundaries, Yahil observed something curious: Galaxy counts suggested that there was no clear end to a cluster. In other words, a cluster's density appeared to have no sharp spatial cutoff. After hearing Ostriker talk in 1972 about the possibility of massive galactic halos, Yahil started to wonder whether galaxies, like clusters, also lacked a clear point where their mass suddenly ends. He sent a draft manuscript on the subject to Ostriker, initiating their fruitful collaboration.

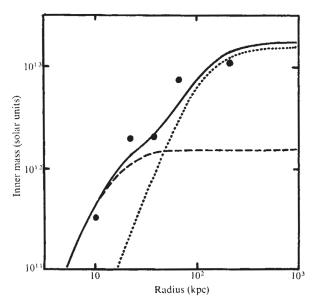
Ostriker himself had been similarly inspired by the consequences of his halo idea. "If the disk doesn't dominate the [mass of a galaxy's] interior, then maybe it doesn't dominate the exterior either," he recounted during an interview. Ostriker

and Yahil began collaborating to see whether their idea indeed held. What if a spherical dark component increasingly dominates the mass of a galaxy in its outer edges? The duo soon asked Peebles, who had just written a textbook on cosmology and the mass distribution in the universe, to join their collaboration.

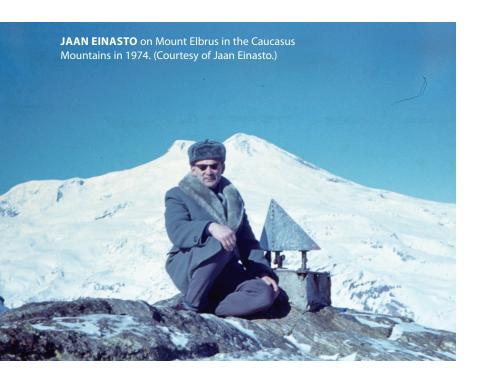
The team gathered dynamical measurements of galaxy masses from various astronomical subdisciplines. Those determinations were based on the gravity needed to explain the movement of cosmic objects. They included galaxy pairs, small groups and large clusters of galaxies, and rotating galaxies. The three men found that those dynamical measures of galaxy mass kept increasing with a galaxy's radius: In other words, the farther out the mass of a galaxy is measured, the higher the mass of the system—even if it was measured outside of the luminous disk of the galaxy. "The masses of ordinary galaxies may have been underestimated by a factor of 10 or more," they concluded in their paper, which was published in October 1974. Galaxies had no clear boundary but were surrounded by extended, invisible massive halos, possibly of "faint stars." Added together, those hidden halos

Demonstrating the existence of dark matter





Diagrams from the landmark 1974 papers by the Princeton University and University of Tartu groups that demonstrate the existence of large halos of unseen mass surrounding galaxies. The diagrams plot different masses of galaxies as measured within a distance *R* from their centers. They show how the mass of a galaxy does not stop at a fixed point but instead keeps increasing linearly with the radius far beyond the bright visible disk of a galaxy. At left is a plot from the article by Jeremiah Ostriker, James Peebles, and Amos Yahil. The data points signify observed mass as determined by different methods, including ones based on galaxy rotation, galaxy pairs, and cluster dynamics (labeled "virial" in the diagram after the statistical mechanics theorem used to determine the total mass of galaxies in a cluster). At right is a plot from the article by Jaan Einasto, Ants Kaasik, and Enn Saar. The dots represent the observed values obtained from five groups of galaxy pairs, the dashed line is the mass function of known stellar populations, the dotted line is the implied mass distribution of the "dark" corona, and the solid line is the total mass distribution of the galaxy, including the corona. Mass is given in units of solar mass *M*_o. (Left diagram from ref. 9; right diagram from ref. 15, *Nature* citation.)



significantly increased the mass of the universe, which suggested that the universe might be closed.

The Soviet parallel

On the other side of the Iron Curtain, almost 6500 kilometers east of Princeton, a similar conclusion was drawn. As Yuri Gagarin launched into space in 1961, on land the Soviet Union was developing a strong workforce in cosmological physics. One of its main drivers was Yakov Zeldovich, a physicist who was famous for his work on the Soviet nuclear bomb project in the 1940s. He began gathering a group of bright physicists to tackle the problems of the cosmos. Working in parallel to Dicke's group in the US, Zeldovich and his team quickly became internationally renowned for their work on neutrinos, quasars, black holes, and the cosmic microwave background. Zeldovich's weekly two-hour seminar at Moscow State University's Sternberg Astronomical Institute became a central hub for anything cosmology and drew scientists from all around the Soviet Union. 10

In 1971 one of the Sternberg seminars was given by a 42-year-old astronomer named Jaan Einasto, who had traveled by train from Tõravere, a small town 20 kilometers outside of the city of Tartu in what was then the Estonian Soviet Socialist Republic. Tõravere housed the Tartu Observatory, which Einasto had helped set up when it was relocated outside the city in 1964. Estonia's astronomical heritage dates back to the early 19th century, when the famed Baltic German astronomer Friedrich Georg Wilhelm von Struve helped make a name for the Tartu Observatory with his observations of double stars. Einasto was part of a new postwar generation in Estonian astronomy.¹¹

Einasto was invited to Moscow to discuss new theoretical

models of galaxies-a subject he had worked on since his PhD work under Grigori Kuzmin. Working in parallel to Tinsley and other astronomers in the US, Einasto aimed to mathematically describe and understand the evolution of galaxies by modeling their luminosity and distribution of matter. Above all, he aimed to precisely model galaxies by using known populations of stars. Taking existing observations, he modeled how star populations were distributed in galactic components, such as the bulge, the core, and the disk. In Moscow, he gave a seminar on his recent model of the Andromeda galaxy, the nearest neighbor to the Milky Way. It piqued the interest of Zeldovich, who invited Einasto to present at the annual Soviet winter school for astrophysics in the Caucasus Mountains.

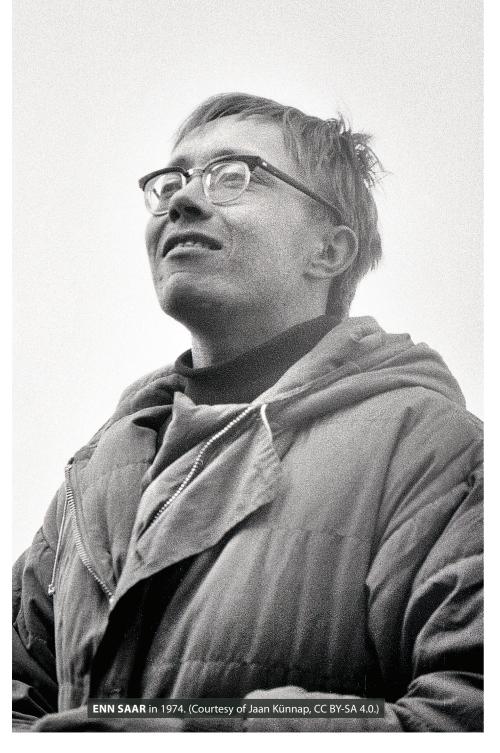
By 1972, after many conversations at the winter school about a model of Andromeda. Einasto had hit what he recalled in an

interview as a "deadlock." He had come across a paper from Morton Roberts, a radio astronomer at the National Radio Astronomy Observatory in Green Bank, West Virginia, that described using hydrogen clouds to model Andromeda's rotation. Roberts's data showed that those clouds were moving remarkably fast along the galactic edges—beyond the visible stars in the disk. What could be out there at the edges of the galaxy that explained this? "No combination of stellar populations was able to explain rotation data of galaxies," Einasto wrote in his autobiography. It was a problem Einasto was unable to address properly until he came to speak about it with a colleague at the observatory.

A bomb in the Caucasus

It was coworker Enn Saar, a cosmologist, who helped Einasto by teaching him a valuable lesson in gravitational physics. The son of two Estonian fishers, Saar grew up walking to high school with matches and a newspaper to fend off wolves—and having a deep interest in the cosmos. Shortly after defending his PhD on Big Bang cosmology in 1972 at the University of Tartu under the supervision of cosmologist Arved Sapar, Saar began discussing with Einasto the problems in modeling the outer edges of galaxies. There was "sort of a difference in attitudes," Saar said in an interview. His experience in cosmology meant that he saw little problem with extended galaxies. He said that lacking borders was "a normal state of any gravitational body." Perhaps there was more than met the eye and galaxies could extend far beyond their luminous disk.

For Einasto, Saar's insight meant "abandoning the idea that only known stellar populations exist in galaxies." They might be surrounded by a new population of yet-unknown nature. Einasto and Saar named that invisible population the



"galactic corona." Einasto presented his ideas at the First European Astronomical Meeting in Athens in 1972. "Giant galaxies may be surrounded by massive coronae of very large dimensions," his abstract read. Einasto suggested that that "unknown matter" might exist in the form of "rarefied ionized gas." The response to his talk in Athens, Einasto recounted in an interview, was lukewarm at best. The half-hearted reaction made him determined to find more data to support his claim.

That was no small task in the Soviet Union. Iosif Shklovsky, the eminent Russian astronomer famous for his collaboration with Carl Sagan, is said to have once joked that all Soviet astronomical observations were done through the US-based *Astrophysical Journal* because Soviet astronomers were often

hampered by weather conditions and equipment problems. But even obtaining issues of US journals in the Soviet Union—especially in Estonia, far from scientific centers such as Leningrad or Moscow—was not trivial. Einasto was often forced to use foreign travel stipends to acquire his own copies of publications such as the *Astrophysical Journal*.

Through an arduous search of the international literature, Einasto discovered the long-standing "Zwicky problem": that galaxies in groups and clusters seem to move so fast that they must either be exploding or require large amounts of extra mass to stabilize them. Einasto found data on groups and pairs of galaxies that would complement Roberts's galaxy-rotation data and make a strong case that coronae of unseen mass must exist. He also learned about x-ray studies that showed that galaxies did not have enough ionized gas to account for all the mass in their coronae. That knowledge prompted him to hypothesize that the mass might be made up of something akin to a new population of stars. Einasto diligently worked out the calculations with Saar and local student Ants Kaasik.

They presented their work in January 1974 at Zeldovich's annual winter school in the Caucasus. Against the backdrop of Mount Elbrus, the tallest mountain in Europe, Einasto shared his idea about galactic coronae: There must be a still-unknown nonstellar population of stuff surrounding galaxies. As he wrote later, it was "as if a bomb had exploded." The avid young physicists in Zeldovich's group immediately started to do back-of-the-envelope calculations: Could

coronae consist of gas or neutrino clouds? As Einasto, Saar, and Kaasik began writing up their conclusions for a Soviet astronomical leaflet, the *Astronomicheskii Tsirkulyar (Astronomical Circular)*, their host intervened: "Zeldovich insisted this must be published in some really important journal," Einasto told me in an interview.

On Zeldovich's advice, Einasto and his group decided to translate their paper into English and send it to the renowned UK journal *Nature*. "It was a strange idea," Saar told me, "because we knew that it was practically impossible." Not only would it be their first-ever paper in such a prestigious English-language journal—in Tartu, it had long been standard to publish results in the local astronomical bulletin, the *Tartu Astrofüüsika Observatoorium Teated* (Notices of the Tartu

MISSING MATTER

Astrophysical Observatory)—but at the time, the KGB thoroughly examined every piece of outgoing international mail.

The process was tedious: Simple scientific words such as "atom" needed to be avoided because secrecy-obsessed KGB censors could associate them with nuclear weapons. One of the group's corrections to the proofs was never added because their return letter to the UK was held up by the censorship process and arrived after *Nature* went to press. Nevertheless, the publication was successful. It came out in July 1974, a few months before that of the Princeton group. "Evidence is presented," they wrote, "that galaxies are surrounded by massive coronas exceeding the masses of known stars by one order of magnitude." Their evidence also showed that the corona hypothesis would mean that the total mass of the universe was larger than previously thought by

a factor of 10, which implied that the unknown dark matter made up the majority of matter in the universe.

From heresy to orthodoxy

On either side of the globe, Einasto's and Ostriker's groups independently demonstrated the existence of dark matter. Despite working in vastly different political contexts, both groups involved collaborations between young astrophysicists and cosmologists studying galaxies. The evidence they presented was neither a simple proof nor a single observation, like that of Zwicky or Rubin, but an inference using a combination of different arguments. As Peebles stated when I interviewed him, "What was the best argument? None of them. This is a case of no one argument being compelling, but so many arguments pointing in the same direction."

The two papers were exemplars of the nascent field of physical cosmology and its interdisciplinary teamwork and methodology: combining data and arguments from different scales—from stars and galaxies to clusters—to form a consistent physical picture of the cosmos.

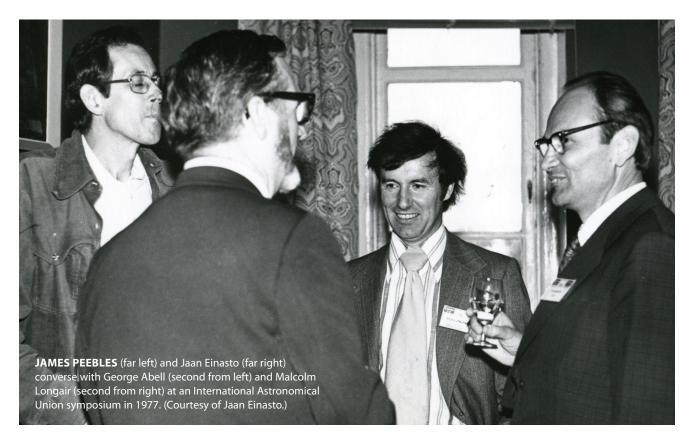
Despite their historical significance, the papers were not immediately received with open arms. "People thought it was just crazy," Ostriker told me. Saar recalled that "most astronomers and physicists didn't like this thing at all." Some astronomers disputed parts of the data that the groups used. Astrophysicist Geoffrey Burbidge was vocal about his disgust with the idea and authored a scathing response in the Astrophysical Journal a few months after the publication of the two papers. "Contrary to the results obtained by Einasto et al. and Ostriker et al.," Burbidge's paper reads, "we show that there is no unambiguous dynamical evidence which demonstrates that galaxies have very massive halos."16 He was particularly critical of the assumption that one could measure the mass of galaxies simply by observing their dynamics.

It was only later in the 1970s that the hypothesis of missing matter—the idea that galaxies are surrounded by coronae or halos of invisible mass—became a staple in astronomical and cosmological thinking. Both the Princeton and Tartu collaborations worked hard to gain acceptance for their proposal. In 1975 Einasto organized a conference in Tallinn, Estonia, to discuss the possible

YAKOV ZELDOVICH AND JEREMIAH OSTRIKER (left

to right) in Moscow in 1979. (Courtesy of the AIP Emilio

Segrè Visual Archives, gift of Jeremiah Ostriker.)



nature of the invisible corona with Zeldovich and his students, and that same year he set up a dedicated session on "missing mass" at the Third European Astronomical Meeting in Tbilisi, Georgia. Ostriker defended his ideas in a lecture at the National Academy of Sciences in 1976, arguing that "most of the mass is not in ordinary (solar) type stars, but some other dark form." By then the reception from most of the community had flipped 180 degrees: "Within two years, we went from heresy to orthodoxy," Yahil told me.

More support for their arguments appeared after 1977. Optical and radio astronomers published new galaxyrotation data that showed more signs of unseen mass. Cosmologists began theorizing that missing matter affected galaxy formation, and particle physicists connected the mysterious substance to a potential background of neutrinos in the universe. In both instances, theorists accepted the evidence for missing mass and used the idea as a central thesis to underpin theories of cosmic particles and structure formation. In other words, what researchers soon began to call dark matter was now the basis on which theories of the universe were constructed. By the end of the 1970s, its reality appeared inescapable. Astronomers Sandra Faber and John Gallagher wrote in a 1979 review paper, "We think it likely that the discovery of invisible matter will endure as one of the major conclusions of modern astronomy."18 Indeed it has.

REFERENCES

- 1. S. Tremaine, Astrophys. J. 525 (centennial issue), 1223 (1999).
- 2. J. P. Ostriker, interview by C. Smeenk (14 May 2002), Niels Bohr

- Library & Archives, American Institute of Physics, College Park, MD.
- 3. J. de Swart, in *The Renaissance of General Relativity in Context*, A. S. Blum, R. Lalli, J. Renn, eds., Birkhäuser (2020), p. 257.
- 4. Unless stated otherwise, all quotations in this article come from interviews I conducted between 2014 and 2020 with Jeremiah Ostriker, James Peebles, Amos Yahil, Jaan Einasto, Enn Saar, and Alar Toomre. The transcripts will eventually be made publicly available at the American Institute of Physics's Center for History of Physics.
- P. J. E. Peebles, Cosmology's Century: An Inside History of Our Modern Understanding of the Universe, Princeton U. Press (2020).
- J. Peebles, interview by M. Harwit (27 September 1984), Niels Bohr Library & Archives, American Institute of Physics, College Park, MD.
- 7. J. P. Ostriker, P. J. E. Peebles, Astrophys. J. 186, 467 (1973).
- 8. A. Toomre, in *Highlights of Astronomy, Volume 3: As Presented at the XVth General Assembly and the Extraordinary General Assembly of the I.A.U. 1973*, G. Contopoulos, ed., D. Reidel (1974), p. 457.
- 9. J. P. Ostriker, P. J. E. Peebles, A. Yahil, Astrophys. J. 193, L1 (1974).
- R. A. Sunyaev, in Zeldovich: Reminiscences, R. A. Sunyaev, ed., Chapman & Hall (2004), p. 233.
- 11. J. Einasto, Dark Matter and Cosmic Web Story, World Scientific (2014).
- 12. Ref. 11, p. 81.
- 13. J. Einasto, in *Stars and the Milky Way System*, L. N. Mavridis, ed., Springer (1974), p. 291.
- 14. Ref. 11, p. 90.
- 15. J. Einasto, A. Kaasik, E. Saar, *Nature*, **250**, 309 (1974); "Dynamical evidence of massive coronas of galaxies," *Tartu Astrofüüs. Obs. Publ.* preprint (1974).
- 16. G. Burbidge, Astrophys. J. 196, L7 (1975).
- Ref. 11, p. 93; E. K. Kharadze, ed., Proceedings of the Third European Astronomy Meeting, Tbilisi, 1–5 July 1975: Stars and Galaxies from Observational Points of View, Academy of Sciences of the Georgian SSR (1976); J. P. Ostriker, Proc. Natl. Acad. Sci. USA 74, 1767 (1977), p. 1773.
- S. M. Faber, J. S. Gallagher, Annu. Rev. Astron. Astrophys. 17, 135 (1979), p. 182.