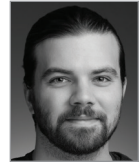




A gravitational lens in the Fornax constellation.
(Image from ESA/Hubble and NASA, S. Jha;
acknowledgment: L. Shatz/CC BY 4.0.)

Sebastian Fernandez-Mulligan is a PhD candidate in the program in the history of science and medicine at Yale University in New Haven, Connecticut. For his dissertation, he is examining how ideas from statistical physics influenced information theory, economics, and art. This feature is adapted from his article "From the model to the glance: How astronomers learned to see gravitational lenses, 1960–2020," *Historical Studies in the Natural Sciences*, volume 54, page 461, 2024.



Learning to see gravitational lenses

Sebastian Fernandez-Mulligan

In the 1970s and 1980s, iconoclastic astronomers used diagrams, computer models, and their own intuition to convince the community that they had observed celestial objects that noticeably bend background light.

Strong gravitational lenses are hard to find. Since the late 1970s, when the first one was observed, astronomers have discovered only a few hundred. But that is about to change. In the next decade, a new generation of astronomical sky surveys will probe the cosmos with unprecedented sensitivity. Scientists predict that the data from those surveys will contain more than 100 000 lenses. The first data release, coming from the space telescope *Euclid*, launched in July 2023, is slated to occur this month.

Gravitational lensing is a consequence of general relativity: Massive objects curve the space around them and bend the trajectories of passing photons. Sometimes the effect is minor; in what is termed weak gravitational lensing, the paths of photons are only slightly warped. Strong lensing occurs when the light from a background object is so severely deflected by a massive foreground object that astronomers observe it as two or more distinct images. Those images can appear distorted or magnified. By boosting the brightness of the images, strong lenses can allow astronomers to see extremely distant sources that would

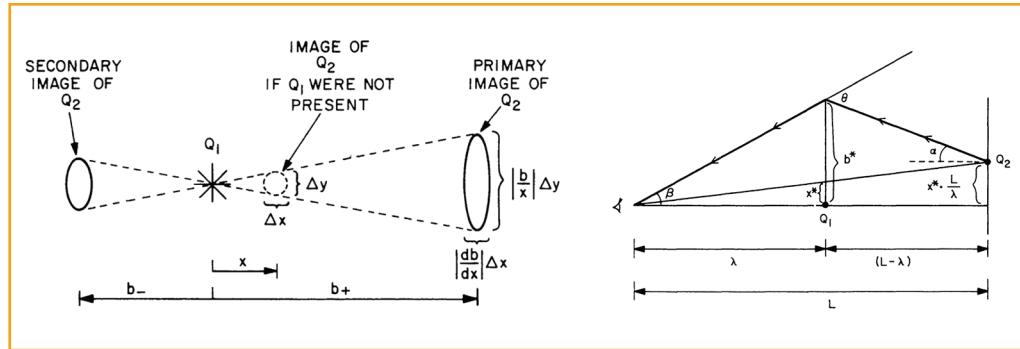
otherwise be too faint to observe. With a vast sample of lenses in hand, astronomers hope to conduct statistically robust studies of high-redshift galaxies.

Since the 1980s, astronomers have relied on learned visual intuition to find gravitational lenses. Certain signatures become visible only after years of work, certain shapes become important only after one has seen them many times, and certain faint objects can be spotted in an image field only by an expert. The seasoned astronomer becomes well versed in those tacit skills. With thousands of images to comb through, quick

GRAVITATIONAL LENSES

FIGURE 1. EARLY ARGUMENTS FOR THE EXISTENCE OF GRAVITATIONAL LENSES

relied heavily on geometric schemata based on simple optics models, such as these images from a 1974 paper by J. Richard Gott III and James Gunn. (Images from ref. 7.)



and intuitive visual analysis is a key aspect in the data-processing pipeline. In a glance, the trained eye sees things that amateurs cannot.

The oncoming deluge will overwhelm even the quick, intuitive glance: 100 000 lenses cannot be found by hand. Recent work has thus focused on developing and deploying algorithms that can automate the search for gravitational lenses. But the increasing sophistication of algorithms has not spelled the death of observational intuition. When the programs are tested, the control is often a human astronomer, who combs through the same simulated dataset and uses their visual intuition to discover gravitational lenses. The success of the model is predicated on how well it compares with the trained eye.

The history of gravitational lensing provides insight into how that intuition was formed and how it became accepted. Rather than take that skill for granted, the astronomy community should acknowledge its historical development. Visual markers that seem obvious today—for example, a doubly imaged quasar or giant lensed arcs—were not initially viewed as clear signs of lensing. Their path to clarity was marked by befuddlement and contestation. By looking at the historical development of intuition, astronomers can

ask a question about the present: What role does intuition play in today's computational age?

Visualizing what cannot be seen

Gravitational lensing was an active area of theoretical research during the 1920s and 1930s, when scientists were clamoring to confirm or contest the conclusions of general relativity. Arthur Eddington proposed gravitational lensing in 1920; his ideas were independently echoed by Orest Khvolson in 1924. Albert Einstein himself privately toyed with the concept in 1912 before publishing a short paper on lensing in 1936.¹ The following year, maverick astronomer Fritz Zwicky made one of the earliest arguments that gravitational lensing could be used to measure the mass of intervening galaxies.

But by the 1940s, as more astronomical observations solidified the credibility of relativity, work had all but ceased. Although lenses had offered an observable example of space-time curvature, their predicted rarity made them unappealing research topics for observational astronomers. Even Einstein shared such pessimism, concluding his paper with the proclamation that “there is no great chance of observing this phenomenon.”²

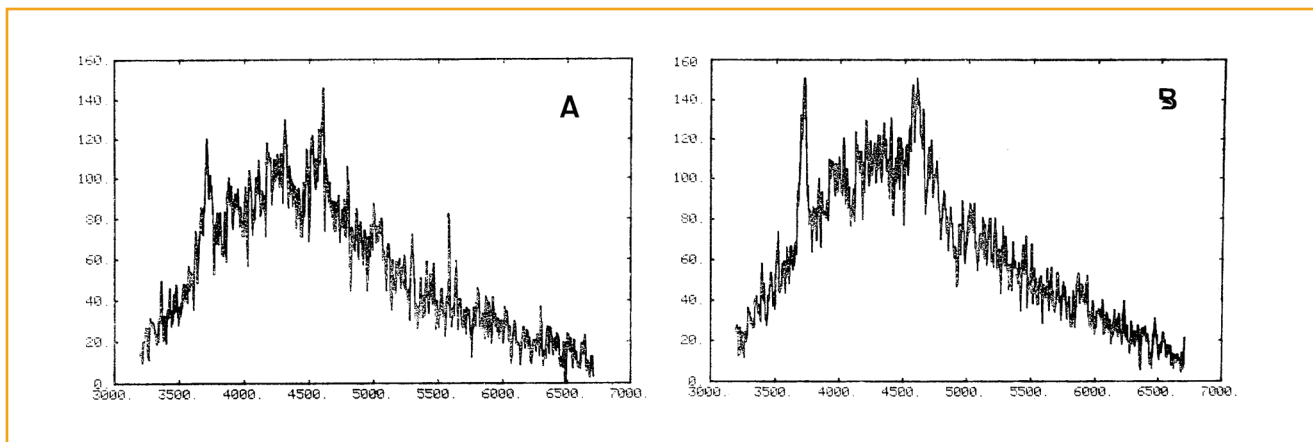


FIGURE 2. TWO QUASAR SPECTRA taken by Robert Carswell and Dennis Walsh on 29 March 1979, which they later identified as evidence of a gravitational lens. The first spectrum is on the left; the second, at right, was taken just a few moments later. (Images from ref. 8.)

As long as astronomers believed that gravitational lenses were impossible to observe, work on them remained sporadic. That pessimism remained until Maarten Schmidt's 1963 discovery of the first quasar (see the article by Hong-Yee Chiu, *PHYSICS TODAY*, May 1964, page 21) sparked renewed interest in gravitational lensing. The newly found objects were puzzlingly bright—so bright that some astronomers argued that they might be the result of magnification from gravitational lensing. Married collaborators Jenő Barnothy and Madeleine Barnothy Forró were the most radical proponents of that theory, arguing that quasars were simply lensed galaxies. They predicted that there were hundreds of lenses across the sky.

Other scientists used the attention of the quasar discovery to highlight additional potential lensing applications. Astrophysicist Sjur Refsdal, for example, rigorously defined how lenses could allow astronomers to infer the mass of intervening galaxies or to measure the Hubble constant through a lensed supernova flash. As he and coauthor Jean Surdej later wrote, his and others' work was received as "particularly promising because of the recent discovery of quasars by Schmidt."³ Lenses had transitioned from mathematical oddities to observational possibilities.

But how could they be found? No prior observations existed. There was no standard practice to replicate, no routine data to collect, and no agreed-on logic to follow. Although astronomers predicted that double images could occur, they had no empirical example to search for in practice. Using existing tools, scientists had to develop techniques that would make lenses visible both to themselves and to their colleagues.

Nigel Sanitt, a graduate student at Cambridge University in the early 1970s, sought to turn possibilities into observations. Roger Blandford, Sanitt's office mate at the time, remembered "berating him for working on a phenomenon that was unlikely ever to be observed."⁴ Despite those apprehensions, Sanitt forged on with his thesis work, and he isolated five candidates for gravitational lensing from a catalog of radio sources. Of the five, he argued that one, 3C 268.4, exhibited high potential for lensing because a secondary image was present near the source.

That interpretation of 3C 268.4 was contested. What Sanitt argued was a "faint ... companion image 2.5 arcsec away,"⁵ other astronomers such as Jerome Kristian had previously identified as a "closer galaxy about [2.5 arcsec] to the south of the quasar."⁶ Sanitt used the radio position data and the mathematical theory of gravitational lensing to argue that the faint image was indeed a lensed image and not a distinct galaxy.

Because the analysis of the telescope image was disputed, Sanitt's publication relied little on visual data. Instead, he used geometric schematics to logically buttress his reading of existing data. That style of argumentation was peppered throughout several papers in the 1970s. Further studies, such as the work of J. Richard Gott III and James Gunn, relied on



FIGURE 3. RAY WEYMANN, pictured in 1970. In collaboration with Robert Carswell and Dennis Walsh, he coauthored the 1979 paper announcing the observation of the first gravitational lens. (Image courtesy of the AIP Emilio Segrè Visual Archives, John Irwin Slide Collection.)

theoretical drawings to make arguments about the possibility of observing lenses⁷ (see figure 1).

Those papers achieved mixed success. Stick-figure schematics did not convince the astronomy community that gravitational lensing had been observed. But astronomers nevertheless welcomed the geometric drawings: They became the standard visualizations for a phenomenon that had not yet been observed.

Is seeing believing?

In 1979, possibilities became observations. At 2:00am on 29 March of that year, atop Kitt Peak National Observatory in Arizona, Robert Carswell and Dennis Walsh were midway

GRAVITATIONAL LENSES

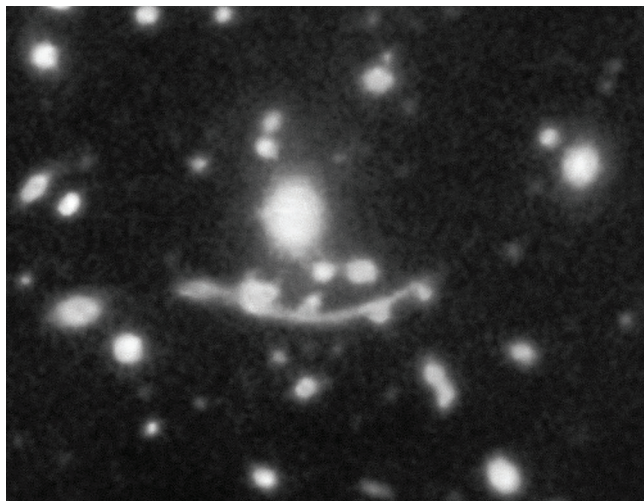


FIGURE 4. THE GIANT ARC found in the Abell 370 galaxy cluster, imaged in the visible spectrum. Arcs are now one of the telltale signs astronomers look for when searching for gravitational lenses. (Image adapted from NOIRLab/NSF/AURA/R. Lynds, V. Petrosian/CC BY 4.0.)

through an observation run to survey quasars. Having already slogged through a long list of objects, they plugged in the next series of pointing coordinates. The telescope heaved toward its programmed position. Two bright blue dots appeared on the viewfinder: the double object 0957+561. Two years earlier, Carswell and Walsh's collaborator Anne Cohen had measured the accurate optical position of that strange pair—seemingly two quasars that were very blue, very bright, and only six arcseconds apart.

Carswell and Walsh quickly measured a spectrum and estimated the redshift. When they looked at their results, they were shocked. Walsh recalled “two strong emission lines, the same two emission lines. Same redshift. Clearly, we’d made a mistake.” Assuming they had accidentally measured the same object twice, the duo repeated their observations. The second measurement rolled in, and the two spectra remained identical⁸ (see figure 2). For the blue quasar pair, the similarity in both categories meant, in the words of Carswell and Walsh, that “the initial conditions, age and environment influencing the development of the [sources] have been so similar that they have evolved nearly identically.”⁹

Confused by their results, Carswell and Walsh reached out to Ray Weymann (see figure 3), a colleague working at the University of Arizona's Steward Observatory. Whereas Carswell and Walsh had been looking at emission lines—namely, sharp peaks in the spectrum—Weymann studied the absorption features, or discontinuous dips in the spectrum. Intervening objects, such as clouds of interstellar gas, are opaque to photons at certain wavelengths. When light reaches a telescope after passing through gas clouds, portions of the spectrum become attenuated, much like how sunglasses block UV light before it reaches our eyes. Weymann measured the spectrum of each source and found, once again, that

the objects had the same redshift. More striking was that the objects had the same absorption features. And the two ostensible quasars were far enough apart that an intervening cloud of gas would need to be unprecedentedly large to cover both.

Weymann was the first to propose a gravitational lensing explanation. Having recently been asked by a colleague to examine the claim that quasars were gravitationally lensed galaxies, Weymann was well versed in the theoretical developments of the 1960s. If the two blue dots were images of a single gravitationally lensed source, he argued, then their spectra would be similar. And if a gas cloud sat in front of the original source, it would not need to be extremely large to explain the absorption line similarities. Carswell, Walsh, and Weymann published a paper in May 1979 arguing that their results were the first observation of a gravitational lens.⁹

But seeing is a tricky thing in astronomy. Observations occur at a wide range of wavelengths, and objects that look one way in the visible spectrum often look quite different in UV or radio wavelengths. Other colleagues scrambled to get multiwavelength data on the sources. “I remember well the mixed reaction [the paper] received,” Walsh recalled.¹⁰ Some astronomers noted that the shape of the two quasars did not look nearly as identical in radio images as they did in optical images. They showed that one of the two objects seemed to have an extended trail in the radio regime and argued that the dual objects were actually distinct. For months, debates raged over whether the sources were truly identical.¹¹

Criticisms petered out as more lensing candidates were identified from observational data. Just a year later, Weymann and a group of collaborators published results arguing that the triple quasar PG1115+08 was three lensed images.¹² As the results rolled in, the practice of observing a gravitational lens stabilized: Astronomers needed to demonstrate that the sources in question had identical spectral signatures. As he told me in a 2022 interview, Weymann recalled that moment as an inflection point: “The notion that the gravitational lens really exists and that we can actually observe it triggered the realization of the reality of looking for instances of it.” That was the thorny knot of discovery: To search for lenses, astronomers had to believe that they could be observed. To do so in the 1970s was to search for double, or even triple, quasars.

Modeling mysteries

In 1987, a new anomaly electrified the attendees of the American Astronomical Society annual meeting. Roger Lynds and Vahé Petrosian announced “the existence of a hitherto unknown type of spatially coherent extragalactic structure having . . . narrow arc-like shape, [and] enormous length.”¹³ Stretching over 100 kiloparsecs, the arc (see figure 4) puzzled astronomers. What was its origin? Some thought it was a shock wave from galactic explosions, others saw it as evidence for galaxy cannibalism, and still others asked whether it was the deformed images of a gravitational lens.

The arc was found in Abell 370, one of about 2700 galaxy clusters included in a well-known catalog compiled by



FIGURE 5. THE VERA C. RUBIN OBSERVATORY, which will carry out one of the next-generation sky surveys, under construction in Chile in 2021. (Image from the Rubin Observatory/NSF/AURA/O. Rivera/CC BY 4.0.)

George Abell in 1958. Although astronomers had been observing objects from that catalog for years, they had discarded the arc as an observational artifact—perhaps a scratch on the glass plate used to record the image. But Lynds and Petrosian took electronic photographs. With no glass to scratch, the arc became an astronomical anomaly.

When Geneviève Soucail returned with a spectrum, things got stranger. Not only was the redshift the same across the entire arc, but it was estimated that the object was twice as far from Earth as any other galaxy in the Abell 370 cluster. Along with the redshift, the emission lines were also the same across the entire object; moreover, the spectrum had a break at about 4000 angstroms, which is characteristic of galaxies.¹⁴

Of the gamut of explanations, Soucail's team argued that the arc was the signature of a gravitational lens. But the researchers were faced with a challenge: There was only one arc. Unlike in the case of the double quasar, the astronomers could not simply compare spectra to prove the lensing origin. Instead, they turned to models. The increasingly powerful computational resources available in the 1980s allowed Soucail and her team to generate a simulated schematic of the lensing system, which they published in the article next to an image of the system. Side by side with the

observational evidence, the model gave meaning to the arc. Lynds and Petrosian rapidly followed up with their own lensing models.

By making sense of arcs such as the one in Abell 370, the schematics transformed them into signatures of gravitational lensing. Arcs quickly became a key part of astronomers' intuition. Up late on an observation run in 1988, Patrick Henry and one of his graduate students pointed the telescope at Abell 963. A huge arc appeared on the screen. In a 2022 interview, Henry recalled immediately turning to his graduate student and joking, "Let's jump on it. A quick paper and we will . . . become rich and famous." When I asked Henry if he had taken a spectrum of the arc, he replied, "I'm not sure anyone ever got a spectrum of 963." Painstaking spectroscopy and analysis gave way to an intuitive assessment of the image.

As the coterie of astronomers studying gravitational lenses expanded in the 1990s, funding was found for large-scale search programs. The first of those, the MIT search program for gravitational lenses led by Bernard Burke, identified five gravitational lenses, the largest sample to date.¹⁵

The procedures of those search programs highlight how important intuition had become. The surveys began with an automated program that directed a radio telescope to map

GRAVITATIONAL LENSES

the positions of more than 6000 sources. Burke and his team then manually identified sources that had multiple, visually similar objects in close proximity, and those sources would be optically imaged at the 4-meter telescope on Kitt Peak and the 5.1-meter telescope at Palomar Observatory in California. With optical images in hand, manual analysis became even more important. Astronomers combed through the images and selected 40 candidates for intense spectroscopic study. They subsequently chose four for further examination. Throughout the process, visual analysis and intuitive skills were the grease between the gears of the data pipeline. Only at the end of the analysis pipeline did astronomers deploy their models.

Detection had become intuitive. As the number of known lenses increased rapidly in the 1990s, detection depended heavily on the visual examination of thousands of images. That kind of analysis continues today. At the University of Chicago, Michael Gladders trains the next generation of scientists in a hands-on astronomy course. As he told me in a 2022 interview, he entered a classroom in 2020 with 120 000 images of the sky. Dividing the portfolio among the students, he told them to be “fairly reflexive. If you’re looking at them one every two seconds ... you’re done in an afternoon of work!” Just as the professionals analyze their datasets, the students powered through thousands of images to find just a few lenses, building their intuition as they went.

Whither intuition?

The scale of gravitational lens astronomy is shifting. Since the 1970s, astronomers have identified several hundred lenses through visual identification. With the next generation of sky surveys, scientists expect that they will observe more lenses than ever before. The *Euclid* survey is expected to ultimately find more than 150 000 galaxy–galaxy lenses. Later this year, the Vera C. Rubin Observatory (see figure 5) is expected to see first light. Its survey is predicted to observe thousands of lensed quasars and more than 100 000 galaxy–galaxy lenses. Data from those projects and others that are planned or under construction, such as the Square Kilometre Array Observatory, will give researchers unprecedented surveys in the optical, near-IR, and radio wavelengths. Astronomers will soon be working with a few hundred thousand lenses.¹⁶

For the first time, scientists will have a massive sample of gravitational lenses from across the cosmos. But they will be forced to work differently: To process the incoming datasets, astronomers will increasingly rely on mechanized algorithms, rather than visual identification, to find lenses. Some of those automated methods have been designed to look for explicit shapes, such as arcs or rings; others rely on machine-learning algorithms that have been trained on simulated datasets of gravitational lenses. Each of the methods promises labor-saving efficiency over the visual inspection of images.

But those techniques have not and will not erase the importance of the human eye. The swell of AI tools, alongside older algorithmic procedures, is often accompanied by claims

of human obsolescence.¹⁷ But the onset of mechanization has not made tacit skills irrelevant. As astronomers search for more accurate and more efficient methods, they consistently benchmark new algorithms against the visual examination by their colleagues. Although algorithms are faster than manual inspection, they often miss subtle cases of gravitational lensing, such as wispy arcs or complex visual deformations. In a recent comparison using data from the Kilo-Degree Survey, algorithms proved less accurate than human observers at identifying lenses. All the automated routines missed the “jackpot lens,” an extremely rare case where the lensed images formed two full rings of light from two different background sources.¹⁸

On the eve of a data deluge, intuition thus serves a new purpose—as an ideal. Rather than doing away with the importance of astronomical intuition, algorithmic tools have merely shifted its role in the process of detection. The history of gravitational lens observations highlights that such intuition is constantly under reevaluation. Before spectroscopic experiments convinced the astronomy community that two objects could be one, double quasars were not an obvious instance of gravitational lensing. Not until models accurately replicated the mysterious arcs did they turn into clear markers. Observational intuition is constantly being reevaluated as a product of past experiments, theories, and models. The successes of computational algorithms only become legible through all-too-human standards. As astronomers continue to develop models, it is important that they continue to develop their eyes.

Think you can spot a gravitational lens?

Go to physicstoday.org/lenses.

REFERENCES

1. J. Renn, T. Sauer, J. Stachel, *Science* **275**, 184 (1997).
2. A. Einstein, *Science* **84**, 506 (1936), p. 507.
3. S. Refsdal, J. Surdej, *Rep. Prog. Phys.* **57**, 117 (1994), p. 121.
4. R. D. Blandford, *Q. J. R. Astron. Soc.* **31**, 305 (1990), p. 306.
5. N. Sanitt, *Mon. Not. R. Astron. Soc.* **174**, 91 (1976), p. 96.
6. J. Kristian, *Astrophys. J.* **179**, L61 (1973), p. L63.
7. J. R. Gott III, J. E. Gunn, *Astrophys. J.* **190**, L105 (1974).
8. D. Walsh, in *Gravitational Lenses: Proceedings of a Conference Held at the Massachusetts Institute of Technology, Cambridge, Massachusetts, in Honour of Bernard F. Burke's 60th Birthday, June 20, 1988*, J. M. Moran, J. N. Hewitt, K. Y. Lo, eds., Springer (1989), p. 11.
9. D. Walsh, R. F. Carswell, R. J. Weymann, *Nature* **279**, 381 (1979), p. 383.
10. Ref. 8, p. 21.
11. D. H. Roberts, P. E. Greenfield, B. F. Burke, *Science* **205**, 894 (1979).
12. R. J. Weymann et al., *Nature* **285**, 641 (1980).
13. Quoted in B. Paczyński, *Nature* **325**, 572 (1987).
14. G. Soucail et al., *Astron. Astrophys.* **184**, L7 (1987).
15. B. F. Burke, in *Gravitational Lensing: Proceedings of a Workshop Held in Toulouse, France, September 13–15, 1989*, Y. Mellier, B. Fort, G. Soucail, eds., Springer (1990), p. 127.
16. T. E. Collett, *Astrophys. J.* **811**, 20 (2015).
17. L. Messeri, M. J. Crockett, *Nature* **627**, 49 (2024).
18. R. B. Metcalf et al., *Astron. Astrophys.* **625**, A119 (2019).

PT