Profile: David Harper

Once an astronomer, now a genome researcher, David Harper talks about the fascination of celestial mechanics, computing by means of cassette tapes – and life after professional astronomy.

Like most kids in the late 1960s, I was fascinated by the Apollo Moon landings. My parents bought me a telescope for my sixth birthday, soon after Apollo 11, and I remember bringing it home on the bus from Liverpool. It was just a little telescope, too small for real stargazing, but it was exciting to look at the Moon. I learnt to recognize the constellations in the sky thanks to a wonderful book called *The Stars: A New Way to See Them* by HA Rey. The new way was to join the stars up in constellations to make realistic patterns – so that Leo actually looked like a lion, for instance. The book is still in print and I'd recommend it to anyone interested in stargazing.

I was good at maths and sciences at school, although we didn't really do astronomy – we went outside to see a partial eclipse of the Sun in 1976, but that was about it. When I applied to university I wanted to study astronomy, but also use my strength in maths. University College London offered a joint honours degree in maths and astronomy, and that's where I went. Our tutor was Derek McNally, who warned us early on, "The Maths department will expect 75% of your time, and so will Physics and Astronomy", and that was pretty accurate. It was hard work but fun. Most of my friends were on the physics and astronomy side and I think I pushed the limits of tolerance of the maths department because I insisted on taking practical astronomy classes.

In my first year, I spent one evening a week at the University of London Observatory at Mill Hill. On one of the few clear nights we had – north London is not ideal for observing – we were sent out into the grounds with theodolites to find our latitude and longitude. We had to measure the altitudes of stars, start a stopwatch, then telephone the speaking clock and count back to get a fix on the time. Then we determined our location by plotting position lines on graph paper. I was a bit of a smart alec and thought I'd do it using least squares instead, so I was relieved to get a position in central London – one of my classmates reckoned he was in the middle of the North Sea!

I'd always had a strong practical bent, so I took the astrophysics lab class in the second year, and did an astronomy project in my final year, comparing methods of determining orbits of comets. I wrote one program to generate simulated observations of comets, and several other programs to analyse them by different methods.

I wanted to do a PhD and in my final year looked around for something in celestial mechanics – not the most fashionable subject, but one whose mathematical elegance appealed. Jim Message at Liverpool University gave me the most positive response, and when I met him we hit it off straight away. We talked for about six hours straight, about all sorts: my final year project, the work he was doing and trains. Looking back, many of my conversations with Jim involved trains, to a greater or lesser extent.

Saturn's satellites

My PhD at Liverpool concerned the orbits of satellites of Saturn, reworking dynamical theories, and fitting them to observations. There was a wealth of astrometric data from the 1870s onwards, but by 1930 celestial mechanics had got a bit passé. In the early 1960s, when NASA planned to send spacecraft to the outer solar system, they discovered that existing satellite ephemerides were not accurate enough, because no-one had made any observations for almost 40 years. They took steps to get more data from then on.

In the summer of 1985, at the end of my first year in Liverpool, I spent three months at the Royal Greenwich Observatory. I worked with Andrew Sinclair and Don Taylor in the Nautical Almanac Office. Don had collected all of the available astrometric observations of Saturn's satellites, scouring dozens of old journals and annals of foreign observatories in the RGO Library. Andrew had written a program to numerically integrate the equations of motion of the satellites. They generously shared their work with me, and the main part of my PhD was a detailed analysis of the orbits of Titan, Hyperion and Iapetus using the pre-1930 observations.

Computing, old-style

After my PhD, I had hoped for a postdoc, but my application wasn't approved. Instead, I found a job in the Computer Laboratory at Liverpool University. I've always worked with computers. My school had a Z80 microcomputer which was used for computing O-level. By today's standards, it was a very primitive machine. It had just 16 kilobytes of memory, and the BASIC interpreter had to be loaded from a very temperamental tape recorder. It was incredibly sensitive to the volume that you played the tape at, and if you got it wrong you had to start again. It was kept in a store-room at the back of a classroom. I was given permission to use it during my free periods, but I was trapped in the storeroom for the whole hour so that I wouldn't disturb the class next door! When I went to UCL, undergraduates were allowed to use the central mainframe and I taught myself Fortran. At Liverpool, I discovered computer algebra systems – programming languages that enable you to perform algebraic calculations on a computer.

After my PhD, my job was a three-year post to promote and support the use of computer algebra systems in the UK academic community. It was fascinating – I visited pretty much every university in the UK, giving seminars and live software demonstrations across the JANET network. This was the pre-Internet version of JANET, and it was always touch and go whether the link would stay up. It was a very enjoyable job, and I loved Liverpool, but I still harboured ambitions to be an academic. I applied for a job at Queen Mary and Westfield College (QMW, as it was then), where Carl Murray had been awarded a five-year research fellowship, and I was appointed to cover his teaching and administrative duties.

I soon discovered I had not been realistic

about what academic life entailed, especially the amount of teaching and administration involved, but that was what I was employed to do. I got back to my celestial mechanics research, and Don Taylor and I decided to do the last word in analysis of the orbits of Saturn's satellites. I also finally got my hands on a proper astronomical telescope, through a series of observing runs on the Jacobus Kapteyn Telescope on La Palma. This was in part to support British involvement in the Cassini Mission. The JKT is the smallest of the telescopes up there and, unlike the bigger instruments, you get to drive it yourself, which appealed once again to my practical side. Observing time on the JKT was allocated in seven-day chunks, which was ideal for the type of observations we needed to make, especially as we could work on bright nights – the full Moon was not a problem. I was saddened when I heard that the JKT would no longer be funded by the UK. There are now no telescopes available to UK observers for astrometry of planetary satellites.

Opportunity knocks

At the end of my time at QMW, I noticed an advert for a vacancy at the Nautical Almanac Office at the RGO. The job description could almost have been written for me and I joined the NAO in November 1996. I had got married the previous year, and in early 1997 we moved to Cambridge, feeling pretty optimistic. I had not known that the RGO was under threat of closure – that was my first surprise on starting the job. Throughout my first year, staff were trying to stop the closure, calling on the considerable support of colleagues around the world, but to no avail. PPARC announced the closure at the end of 1997. We had a pretty gloomy Christmas that year. While colleagues in the NAO arranged to transfer to the Rutherford Appleton Laboratory, I discovered that I wasn't eligible for any transfer expenses because I was on a fixed-term contract, nor was there any guarantee that my contract would be renewed. I chose not to go.

I spent two good years at the RGO, working with talented and dedicated people. We didn't just produce the Nautical Almanac, we were a source of official information. We acted as expert witnesses in legal cases where the time of sunset or twilight was a factor, and that led to me appearing in court at the Old Bailey! We also provided astronomical information to the general public, which is how I made myself unpopular in Lowestoft. It's the most easterly town in Britain, and as the year 2000 approached, the town's tourist office was claiming that it would be the first place to see the Sun rise on 1 January, 2000. BBC Radio Kent got wind of this, and invited me to explain to their listeners that the Sun rises in the southeast in early January, not in the east, so North Foreland, on the Kent coast near Dover would see sunrise about five minutes earlier than Lowestoft. Soon after, BBC Radio Suffolk asked me to appear on their afternoon show to break the bad news to the residents of Lowestoft, and I've never dared to visit the town since then.

Shortly before the RGO closed, my wife spotted an advert in the local paper for someone to run the website for the Sanger Institute. I applied, because it was local and I had some experience of setting up and developing the NAO website. I didn't get the job, but when I later saw an advert for software development at the same place, I thought I would try again. At the interview I persuaded them that they needed a mathematician. I got the job, and that's how I moved from astronomy to genomics. I've been there 11 years now.

From astronomy to genomics

The Sanger Institute is best known for its role in the international Human Genome Project. Back in 1999 it was at the forefront of DNA sequencing. A lot of pioneering work was being done and the learning curve was steep, especially as I hadn't studied biology since O-level. When I started it was like the first day at school: I didn't know anyone, I didn't know where to go and I had no shared history or vocabulary with anyone. I spent the first three months wondering if I had made the biggest mistake of my life in leaving astronomy. I had to read up about genomics, learn to recognize the buzz words and make sense of the fundamentals, but this was a very new area and there weren't really any textbooks. My new colleagues were happy to explain their work to me, and eventually it all began to make sense.

It's impossible to read the DNA code of an entire chromosome in one go, but we can read small pieces of DNA, a few hundred letters of the code at a time, so the trick is to shatter the chromosome into millions of short fragments, and read the code as strings of letters. Then we use computer software to find the overlaps between the strings, to reconstruct the original chromosome. It's like trying to complete a huge jigsaw, but you're working with several incomplete sets of the pieces, and there are gaps. I was in a group which was sequencing the genomes of pathogens, including the malaria parasite and the bacteria which cause plague, leprosy and tuberculosis. The hope is that if we can understand how these organisms work at the molecular level, we can design drugs to disrupt specific biological processes.

Genomics is very dependent on technology, in much the same way as astronomy. It's a very young science, but the pace of progress is astonishing. In astronomy, photographic plates took more than a century to become obsolete. The DNA sequencing machines which were cutting-edge technology when I joined the Sanger Institute are now outdated. They enabled us to sequence DNA fragments equivalent to eight human genomes per year. Now, ten years later, a new generation of machines can sequence 20 human genomes worth of DNA every day. The real challenge lies in handling the data. I recently heard a conference keynote talk about petabyte-scale scientific databases that will be generated by projects such as the Large Synoptic Survey Telescope and the Large Hadron Collider – and DNA sequencing! The computing facilities at the Sanger Institute are in the supercomputer league, and much of my work involves writing software in Java to manage and visualize large amounts of DNA data.

I still do astronomy – I think I'm lucky to be interested in a science you can do at home. My wife Lynne Stockman and I have a website which gives us the chance to connect with people around the world who share our passion for astronomy, genealogy and cats. She writes a monthly stargazing guide, and I write about blue moons. We launched the site in 1998 and the two blue moons in January and March 1999 gave us an early boost. I've also been involved in producing calendars linked to the phases of the Moon. I currently contribute Moon phase diagrams to each annual edition of "Stargazers' Almanac". And I'm still doing research in celestial mechanics. During my PhD I met a Chinese astronomer called Kaixian Shen, who was also working on satellite dynamics. We have maintained our friendship and our research collaboration for almost 25 years. Satellite astrometry is still a respectable subject in China and observatories there are collecting new data. Since I moved into genomics, Shen and I have published papers on the dynamics of the satellites of Saturn and Uranus, and we are now looking at the Neptune system. I can run orbit integrations and data analysis on my home computer, which is far more powerful than the IBM mainframe that I used for my PhD.

I don't regret changing my career. I enjoyed being a professional astronomer at QMW and the RGO, but switching to genomics has been tremendously rewarding. I've had the chance to work with great people in two completely different areas of science which both address fundamental questions: who are we, and where do we come from? I was proud to have a small role in the very successful Cassini mission and my sense of ownership of that science continues. Now I am playing a part in improving human health. In many ways, PPARC did me a favour, forcing me to change career from astronomy. If I had to do it again, I would. ●

David Harper works at the Wellcome Trust Sanger Institute, Hinxton, Cambridge, one of the world's leading genome research institutes.

More information
http://www.obliquity.com