Exploring the time-varying Universe

Richard Strom
ASTRON
Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
E-mail: strom@astron.nl

Lodie Voûte
ASTRON, Anton Pannekoek Inst. Of Astronomy, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands
E-mail: lvoute@gmail.com

Benjamin Stappers
School of Phys. & Astron., Alan Turing Bldg., University of Manchester, Oxford Road, Manchester M13 9PL, UK
E-mail: Ben.Stappers@manchester.ac.uk

Gemma Janssen
ASTRON
Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
E-mail: janssen@astron.nl

Jason Hessels
ASTRON
Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
E-mail: hessels@astron.nl

50 Years Westerbork Radio Observatory, A Continuing Journey to Discoveries and Innovations
Richard Strom, Arnold van Ardenne, Steve Torchinsky (eds)

© Published with permission of the Netherlands Institute for Radio Astronomy (ASTRON)
under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).
Introduction

The WSRT interferometrically measures Fourier components of the sky brightness distribution from a region set by the primary beam of the telescope elements, at a radio frequency determined by the receiver. This information is used to construct a two-dimensional image of radio emission from the piece of sky observed. Because it is an east-west interferometer array, the information obtained at any instant of time can only be used to construct a one-dimensional map (the telescope so synthesized has the response of a fan beam – narrow in one direction, but orthogonally very elongated). The information needed for a full map is obtained by observing for one or more 12-hour periods.

On short timescales the sky signal received carries no information useful for constructing (synthesizing) a map, so the data can be smoothed with a long time constant (typically 10 seconds). This limits the amount of data which has to be read out and stored, and the interval at which it must be done. But it also means that sources which vary rapidly will have their fluctuating signal smoothed out. For most natural radio emitters this is no problem: before 1967 few sources were known which flickered at a frequency of 1 Hz or more (solar active regions were one exception). But in August of that year, the first pulsing radio star was discovered with a large radio telescope at Cambridge, UK.

Pulsars, as they came to be known, are neutron stars formed during the huge explosion of a young, massive star which has almost exhausted its energy supply. The resulting supernova, SN, (explosion) also initiates a rapidly expanding hot bubble, a remnant (or SNR) which produces luminous radio and X-ray emission (see highlight from 1981 on the SN of 1572). As the lifetime of the extended remnant (<10^5 yr) is much shorter than the typical age of a pulsar (10^6-10^7 yr), only a few pulsars can be associated with SNRs: of these, the best known (and most extensively studied) is the Crab Nebula. It resulted from a SN seen in 1054.
Despite its shortcomings, the system, with some minor modifications, was used to observe several known pulsars. In collaboration with H. van Someren Gréve, Strom worked on a method for observing pulsed emission using the WSRT in synthesis mode. By restructuring the system for reading out the interferometer signal, it was possible to create eight temporal windows which run synchronously with the (known) pulsar frequency. Each window samples a different phase of the pulse period, and so one is able to create maps of pulsars both on and off the pulsed emission. Figure 3 shows an example of such a pair of maps. Of some 40 pulsars observed in this way with the WSRT, the positions of ten were improved by the method.

Full-fledged observing of pulsars only came with the inauguration of PuMa. See the following contributions for the evolution of the WSRT and Pulsars.

Around 1987, R. Strom became interested in the possibility of using the WSRT to search for pulsars. Specifically, the unusual SNR CTB80 (Figure 1) has a component with some of the characteristics of the Crab Nebula: it might harbor a pulsar. With the help of A. van Ardenne, the tied-array system (for VLBI) was adapted so it could be read out rapidly to observe pulsars. An example is an observation of PSR 1929+10 (Figure 2). Although the system clearly could observe pulsars, it was unable to detect pulsed emission from the presumed pulsar in CTB80 (which was detected shortly thereafter by the group in Berkeley). Its period (39 ms) was too short, and the emission probably too weak for detection with the improvised setup used by Strom and Van Ardenne.
At that time (1994) a new PhD student (albeit somewhat grey haired and the author of this section) was asked by his supervisor prof. Dr Jan van Paradijs (UvA, astronomical institute Anton Pannekoek) to familiarise himself with the intricacies of WSRT and to assist with the installation of the FFB at its site. With this new equipment it was hoped that the research for his thesis on the subject of so called giant pulses from some radio pulsars could be done. After a visit with Hans Weggemans (ASTRON, digital electronics) to Caltech, it was agreed to continue and complete the development of the second prototype in-house at the Westerbork observatory.

A year later after modifications and additions, a change in its nickname from The Beast to Beauty and the gaining of a lot of insight by its crew into the art of “how to observe pulsars” – the filterbank was up and running. However, in order to be at the forefront of pulsar research we realised that for observing these fast, shortlived transients, for measuring the four Stokes Parameters of (giant) pulses, to partake in pulsar timing and for detecting highDM pulsars at lower sky frequencies, a more comprehensive backend was required.

Up until then, pulsar backends at the various observatories were either fully analogue or hybrid filterbanks, and correlators. The advent of fast Digital Signal Processors (DSPs) gave Dr Ir Paul van Haren (then UU, Instrumentele Groep Fysica (IGF)) the idea to consider the possibility to go “all digital”. In itself quite a challenging exercise as this hadn’t been done before. After a feasibility study in 1996 by him and the author, a proposal for a new pulsar machine, nicknamed PuMa, was presented to ASTRON.

For maximum flexibility the operation of this new machine should be entirely under software control. For instance, when acting as a filterbank, the proposed backend should have a variable number of channels. Each of these channels should have (unlike the FFB) such a bandwidth that the full bandwidth range of the telescopes could be covered for an optimal signal to noise ratio. In this mode the power of the signals of each polarisation is recorded in each channel whereby the phase information is lost; this is called incoherent dedispersion. At the time, WSRT could observe a sky frequency range of eight bands of 10 MHz each. PuMa would split each band, with a choice of 16 to 1024 channels, into bands of 625 to 10 kHz. The user then had a choice to observe the two polarisations or to record all four Stokes-parameters.

In another mode PuMa could be used as a baseband sampler, allowing for off line coherent dedispersion later on. It would then sample the X and Y polarisation of two 10 MHz bands at 50 ns each or more bands at a lower rate.

At 2
will Deitch (Caltech/ASTRON, FFB software) and the author

1  To reduce the ‘smearing’ of the signal due to dispersion in the interstellar medium.


3  Will Deitch (Caltech/ASTRON, FFB software), Hans Weggemans (ASTRON, digital electronics), Harm Jan Siepel (ASTRON, analogue electronics), Marco Kouwenhoven (UU, Astronomical institute) and the author.
In each of these modes the user could specify the number of bits for each sample, ranging from 1 to 8.

For its realisation a steering committee was set up in 1996. This team had representatives from ASTRON, IGF and the astronomical institutes of the universities of Amsterdam and Utrecht. The actual design and construction were done at IGF with Paul van Haren as the project leader and the author as its project scientist.

On its input side PuMa had sixteen A/D converters for the X and Y polarisation of the eight analogue WSRT bands. Together with the general timing unit, allowing for a (then exceptional) few nanoseconds timing accuracy, these were designed by Theo Beijaard. They were followed by eight clusters of four, homemade, digital processing boards using in total 192 SHARC DSPs as designed by Jaap Langerak. Quite an achievement, as its first design turned out to be spot on.

The software for the DSPs, covering all the modes of operation of PuMa, was written by Dennis Driesens. The processed data from these boards were assembled by two in-build workstations (programmed by the author) and written to eight, then very fast, Cheetah hard disks, prior to final storage on DAT tapes. The whole setup allowed observation runs of up to several hours. With a PuMa, SHARCus and Cheetahs our team sometimes felt like Zoo keepers, under the most appreciated and inspired leadership of ‘tamer’ Paul.

An essential aspect of the reliability of the hardware and software was to validate all modes of its operation. To this end Marco Kouwenhoven (UU, then also a PhD student) wrote independent software that simulated pulsar signals from the telescopes and the corresponding output that should be produced by PuMa. To our great relief all results were well within established parameters.\(^4\)

Two years later, in 1998, the completed PuMa backend was shipped from IGF to WSRT and installed successfully\(^5\). Its completion almost within the estimated time and well within budget, which is, also in the scientific community, rather exceptional. At last the author could start with the observations for his PhD thesis\(^6\).

PuMa’s high sampling rate of 50 ns made it possible to study accurately the variations in the arrival times of giant pulses, as shown below for the “millisecond pulsar” 1937+21.

These two panels above show detections of giant pulses of the millisecond pulsar observed at 1365 MHz. The two sets of three vertical lines in the top panel mark the peak and edges of the main pulse (left/cyan) and interpulse (right/magenta). The lower panel is a zoom-in, just right of the main pulse, its sample time resolution is 50 ns.

---

\(^4\) For more information on this subject see https://www.imdb.com/title/tt0708727/quotes

\(^5\) Shortly after PuMa’s installation at WSRT, Jan van Paradijs passed away. Prof. Dr Richard Strom took over as the author’s supervisor.

\(^6\) The author wishes to express his profound gratitude to Ir Jan (ASTRON) and Johanna Noordam for looking after his social needs during all of his ‘WSRT’ years.
As these graphs show, there are some minor variations in the arrival times and all observed giant pulses arrived shortly after the main pulse. This is in contrast to the giant pulses of the “Crab pulsar” 0531+21, which were observed to arrive almost anywhere during the main pulse as shown below.

Another aspect of study was whether there was a relation between the power of giant pulses and the time between the previous or following giant pulse. Such a relation would, like the “earth quake” model, suggest that an underlying, power generating, mechanism for giant pulses would be existing. However, an observation run with some 1400 giant pulses of the Crab pulsar showed no such relation. This seems to indicate that the mechanism that generates giant pulses has no memory of its history, their generations seem truly random.

Due to its flexible setup PuMa was also suitable for concurrent multi-frequency observations, whereby subsets of the fourteen WSRT telescopes were tuned to different sky frequencies, for instance 382, 825 and 1405 MHz. This feature allowed us to study various aspects of the giant pulses in detail.

For instance we did indeed find that the average flux of giant pulses, as a function of the sky frequency, satisfies a power law (spectral index), which suggests that its emission mechanism is indeed broad band. However we found that there are considerable deviations from this law for individual giant pulses. In one Crab pulsar observation, with over a thousand giant pulses, only less than three hundred could be observed significantly at all three sky frequencies.

It was quite exceptional that in one of the early multi-frequency observations we found a giant pulse from the Crab pulsar which had a signal to noise ratio (SNR) of approximately a 500,000 times stronger than the SNR of an average normal pulse. Up until then these giant pulses were only detected at several hundreds of times stronger.

The very strong giant pulse from the Crab pulsar as shown on the cover of the author’s thesis, together with his dog Puma.

Variations in longitude of the various giant pulses of the Crab pulsar, all coinciding with the main pulse. The three vertical lines show the start, peak and tail of that main pulse.

Variations in longitude of the various giant pulses of the Crab pulsar, all coinciding with the main pulse. The three vertical lines show the start, peak and tail of that main pulse.

Variations in longitude of the various giant pulses of the Crab pulsar, all coinciding with the main pulse. The three vertical lines show the start, peak and tail of that main pulse.

The very strong giant pulse from the Crab pulsar as shown on the cover of the author’s thesis, together with his dog Puma. (This was then a WSRT novelty, made possible by Hans van Someren Gréve MSc (ASTRON).)
Chapter 5.3 Taking the Pulse of Neutron Stars with WSRT

With the imminent arrival of a top of the range pulsar machine (see previous section) at the WSRT the author was hired by the University of Amsterdam to work with Ramachandran (ASTRON and UvA) and the team working on building the machine to help develop the data analysis software and to start undertaking the science.

The excitement of being part of something new, pulsar research using the WSRT with cutting edge instrumentation, combined with the strong pull of the “home” country of the Netherlands for the children (or their partners) of Dutch immigrants not only attracted the author (New Zealand) but also the very talented post-doctoral fellows: Edwards (Australia), van Straten (Canada) and Hessel (Canada) who have all helped shape pulsar research in the Netherlands.

It was immediately clear that the top quality instrumentation combined with the excellent sensitivity of the WSRT when compared to the other telescopes undertaking pulsar research that it would be excellent for high precision pulsar timing (see Section 5.4). This was particularly true due to the range of frequencies, and the speed with which the receivers could be changed, with the MFFEs. However building up a significant time baseline to be competitive with other instruments takes time. The distributed nature of the dishes in an interferometer also mean that, without significant compute resources, they aren’t particularly good for wide-area pulsar searches. So it was recognised very early on that we could make significant contributions to the understanding of the radio emission from pulsars using the combination of high-quality WSRT data and sophisticated new analysis techniques.

The majority of these initial projects evolved around understanding the single pulse emission, which again benefited from the high sensitivity of the array. These formed an important component of the PhD theses of Voute, Kouwenhoven, van Leeuwen and Smits, and the Masters thesis of Janssen.

The pulses of some pulsars are made up of subpulses which appear to drift through the pulse window, defined by the average of thousands of pulses, in an organised fashion and this proved to be a rich vein of research with the WSRT. A particular favourite object was one of the first pulsars found, PSR B0809+74, which was known to have drifting subpulses but also was known to null, i.e. there are entire rotations, or series thereof, when no radio emission is detected at all. In a pair of papers led by van Leeuwen (2002, 2003) it was shown that the nulls affect the rate of drifting and tells us about where the emission is coming from in the pulsar magnetosphere. In 2003 Edwards and the author came up with a new method to study the drifting subpulses called the two-dimensional fluctuation spectrum, which we applied to PSR B0809+74 initially and then to a series of other pulsars (Edwards et al 2003a,b). The latter work took advantage of the frequency agility of the WSRT to study the frequency dependence of these effects.

The development of the new technique allowed one to search for drifting subpulses and other pulse modulation mechanisms in weaker pulsars and it was recognised that understanding the phenomenon required a much larger sample than the handful of well-studied bright sources. The main topic of the thesis of Weltevrede was therefore to undertake a survey of as many sources as possible with the WSRT at two relatively widely separated frequencies of 1400 and 350 MHz. The results were published in two influential publications in 2006 and 2007. This work showed that many more pulsars than had previously thought exhibited subpulse drifting and suggested that the phenomenon may be ubiquitous amongst pulsars and that it may be geometry which governs whether they are visible. There was also some evidence that the phenomenon is more often seen at the lower radio frequencies. See Figure 1.
The new technique also allowed us to look for pulse-to-pulse modulation properties in millisecond pulsars. These are the pulsars that are spun-up to spin periods of just a few milliseconds and which are famed for their high precision clock-like stability. They are being used for experiments to test theories of gravity and also to try and detect gravitational waves (see Section 5.4). Understanding their use as clocks requires us to understand how they pulse and in a paper led by Edwards we showed that a number of millisecond pulsars showed strong modulation and there was also evidence that they exhibit drifting subpulses, so despite their very different spin period and magnetic field strengths they show properties analogous to their young brethren.

The flexibility of the WSRT and the PuMa backend also allowed us to undertake some unique experiments that are only now being implemented at other arrays where the greater number of dishes makes them more suitable for science. In preparation for the ballet that would use live pulsars sounds, which unfortunately never happened, we implemented a mode which enabled sub-arraying (the modern term) the WSRT dishes and point them simultaneously at 3 different pulsars and process the data in real time. Also using the sub-arraying concept we split the array to observe some pulsars simultaneously at multiple frequencies. The most successful of these were the observations of the Crab pulsar which are part of Lodie Voute's thesis.

The single pulse work with the WSRT and PuMa will be a lasting legacy of the contribution they have made to pulsar astrophysics.

Pulsars are highly magnetised, rapidly rotating neutron stars. They emit radio beams along their magnetic axes, acting like cosmic lighthouses. One of the most important features of pulsars is their highly stable rotation. Using pulsars as distant clocks through high-precision timing measurements gives a unique tool to probe various astrophysical processes. Besides monitoring the intrinsic spin-down behaviour of pulsars, pulsar timing can be used to measure orbital parameters of pulsars in binary systems, and in specific cases test Einstein’s theory of gravity by measuring relativistic effects in the orbits (see Section 5.5). When having a large observing bandwidth, or multiple observing frequencies available, effects of the ionised interstellar medium on the pulse arrival times can be studied and monitored.

A pulsar timing programme has been running at WSRT using the pulsar machines PumaI from 1999 to 2010, and using PuMaII from 2006 until the shutdown of the MFFE system in 2015. The timing programme using PumaI started observing about 10 millisecond pulsars (MSPs) and 10 slow pulsars. More pulsars were added every year, with the observing time also gradually increasing from 18 hours per month at the start to more than 48 hours per month in the later years allowing the coverage of more than 50 (mostly millisecond) pulsars at two or three observing frequencies. A revised timing programme will be started with the ARTS backend this year (see Chapter 15).
The next-generation Pulsar Backend: PuMaII

With the PuMaI backend already delivering good results for single pulse work and pulsar timing, a new step was taken in the early 2000s towards the development of a new pulsar instrument. PuMaII, as this was called, was designed to be capable to handle the full observing bandwidth that was covered by the MFFE frontends (160 MHz > 1 GHz, 80 MHz < 1 GHz). PuMaII was designed to be simultaneously capable of storing recorded data from each band, on 8 storage nodes, and had 32 independent processing nodes available for handling the data after observing. Most importantly, this new PuMaII cluster was capable of coherently dedispersing the data, which significantly improved the timing precision for millisecond pulsars due to the fact that inter-channel smearing due to dispersion was mitigated (Karuppusamy, Stappers and Van Straten 2008, PASP 120, 191). Also, with largely extended storage space, baseband data could be recorded up to 12 hours and temporarily stored for later processing (which was in particular used by the LEAP project).

Using the MFFE capabilities of observing at different frequencies, PuMaII was routinely used to probe and monitor the effects of the interstellar medium on pulsar arrival times. Overall, the multi-frequency observing availability, in combination with the excellent timing resolution of the data and the large storage and processing capability, the PuMaII-WSRT pulsar timing data has generated a long-term data set on over 50 millisecond pulsars that has been extremely valuable for various pulsar timing experiments.

Using millisecond pulsars in a pulsar timing array

In 2005, a team of European pulsar astronomers including ASTRON’s Ben Stappers were awarded the EU Descartes Prize for their collaborative research in pulsar astronomy. Building from this collaboration, in 2006 the European Pulsar Timing Array (EPTA) collaboration was founded. The EPTA is a collaboration using the five large radio telescopes in Europe: besides the WSRT, the EPTA combines data from the Lovell telescope at Jodrell bank, and the Effelsberg, Nançay, and Sardinia radio telescopes. In all pulsar timing experiments, both the total baseline (in time) of observations, as well as the quality and the cadence of the observations are relevant to the quality of the timing solutions for each pulsar. Since the total observing time at any telescope is limited, it directly pays off to collaborate with other teams, and create larger data sets by combining the pulse Time of Arrival (TOA) from each telescope in a combined timing solution. Besides improving timing for many individual millisecond pulsars, and studying the many aspects of the various pulsar systems themselves, the overarching goal of the EPTA is to detect gravitational waves in the nanoHertz (nHz) regime. Gravitational waves (GWs) are small disturbances in space-time that are caused by the acceleration of masses. They originate from a variety of sources, and are expected to exist over a wide range of frequencies. A pulsar timing array (PTA) uses an array of MSPs as the endpoints of a Galactic-scale GW detector. It is sensitive to GWs with frequencies in the nHz regime, being complementary to the frequency bands addressed by other experiments like LIGO/VIRGO and LISA. The most important sources of GWs in the nHz regime are supermassive binary black hole systems (SMBHBs) formed in merging galaxies in the early universe. Understanding the formation and evolution of SMBHBs on cosmic timescales provides unique information on galaxy evolution and star formation processes in general, and studying them through GWs will open up a completely new era of astronomy. In this way, the PTA experiment will complement the recent detections of stellar-mass black hole mergers in the higher-frequency LIGO/VIRGO GW band.
A PTA will detect GWs by measuring correlations in the timing residuals of pulsars that are distributed across the sky. To reach this goal, the prime requirement is to observe a large number of millisecond pulsars and measure the arrival times of their pulses to the best precision possible. Since observing GWs in the nHz regime means that the signals are stretched out for years, a total observing campaign over at least a decade is required to make a detection. Recently, the EPTA has used a combination of WSRT observations with data of the other European radio telescopes of the 6 best timed millisecond pulsars to set an upper limit on a nHz gravitational wave background generated by SMBHs (Lentati et al. 2015, MNRAS 453, 2576), which is already excluding some of the most conservative scenarios for SMBHB evolution.

### The Large European Array for Pulsars

Even with large-bandwidth instruments and processing techniques like coherent dedispersion, the achievable timing precision that is required for GW detection is at the limit of what is possible with single telescopes today. To increase overall sensitivity, the Large European Array for Pulsars was developed since 2009, where in a VLBI-like way (but in the time domain) the EPTA telescopes were coherently combined to produce a 200-m equivalent dish (Bassa et al 2016, MNRAS 460, 2207). After simultaneously observing a set of millisecond pulsars, raw voltage data of each telescopes were transferred to Jodrell Bank where those were coherently added with specifically designed software for beamforming and correlating pulsar observations (Smits et al. 2017, A&C 19, 66). Regular observations for the LEAP project have been running since 2010, and the resulting data is now routinely used to improve the EPTA GW detection efforts.

### Combining PuMa and PuMaII to discover new pulsars: the 8gr8 survey

Although in present times it has become common standard to use radio interferometers as search instruments, at the time that the pulsar machines for WSRT were developed, interferometers were not deemed useful for large area pulsar (or transient) surveys due to the small tied-array beams. A first attempt at creating multiple beams to solve this problem was started in the early 2000s (by Stappers, van Straten, Edwards and Braun). The new method called 8gr8 used the fact that the first 12 WSRT telescopes (an East-West linear array), could be placed at fixed and equal separation. When creating a tied-array beam, their equal separation results in a grating response on the sky consisting of highly elliptical fan beams. WSRT had eight independent signal chains, which were usually covering adjacent frequency bands to maximise the observing bandwidth for pulsar observations. However, when tuning all eight bands to the same centre frequency, but slightly shifting their phasing respective to each other, the combined mapping of the grating patterns of each band provides full coverage of the total primary beam of a single dish.

During a 2 hour observation, PumaI was used to record the signals of the eight bands separately. The rotation of the Earth leads to rotating of the fan beam patterns, and a complicated offline-processing weighting scheme was devised using the new PuMaII computing cluster to form small subbeams, formed by combining different elements of each grating pattern over the observation. For each observation, this effectively results in many small subbeams on the sky that could be searched for pulsar signals.

The survey was mostly sensitive to relatively long-period pulsars due to data-rate limitations. A promising part of the sky was selected (Cygnus area) where stars can be expected to be formed, and therefore young, slow-rotating pulsars should be present. The initial observations were completed in 2004 and 2005, and over 200000 subbeams were searched on the PuMaII cluster. Candidates that were present in both initial and confirmation observations were later followed-up in a timing programme. Eventually, three pulsars were found (Janssen et al. 2009, A&A 498, 223) of which one turns out to be a mode-switching pulsar.

### Scattering features and variability in the Crab pulsar

For pulsar timing, observations above 1 GHz generally give the most precise results. However, low-frequency data play a key role in the efforts to optimally measure and correct the data for interstellar weather influences, since those...
The discovery of radio pulsars just over 50 years ago gave the first direct evidence for the existence of neutron stars, as well as a handy tool to study their properties. The radio pulsations provide a clock-like tick that we can use to measure a pulsar’s spin rate with fantastic precision. Using a large radio telescope like the WSRT, along with a high-time-resolution recording backend like PuMaII, astronomers register up to billions of pulses from a pulsar and then model these to distill important physical insights. These “rotational ephemerides” enable us to model all the recorded pulse times of arrival at the telescope and then infer exactly how many rotations the pulsar has made between observing sessions.

For instance, measuring the spin rate and how it changes with time gives properties like the dipole magnetic field strength of the pulsar, its age, and how much rotational energy it is losing. Some pulsars are in binary systems with another star. In this case, we observe the pulsar’s apparent spin rate being Doppler-shifted as it swings around in its orbit. This allows us to measure the pulsar’s orbit with great precision. As an example, the projected size of the orbit of PSR J1909-3744 – a particularly good “pulsar clock” – is known to be over a million kilometres, but the uncertainty in this determination is only about 20 metres.

Why would we care to measure pulsar orbits with such ridiculously high precision? Simply put, because pulsars are extreme objects that bend space-time strongly and challenge our understanding of the physics of dense matter. For some binary pulsars, we can measure how the pulses are delayed as they travel through the curved space-time around their companion star. This allows us to measure their mass precisely, and hence better understand the physics that describes dense matter. We can also compare the observed orbital effects to the predictions of Einstein’s theory of general relativity and use these systems as
natural laboratories for testing gravity. Famously, the first indirect detection of gravitational waves was made by measuring the orbital decay of PSR B1913+16, which is a binary pulsar whose stellar companion is also a neutron star. In this system, general relativity predicts how the orbit will shrink as gravitational waves are released; this prediction matches with the measurements perfectly, thereby providing an important test of Einstein’s theory. Russell Hulse and Joe Taylor were awarded the 1993 Nobel Prize in physics for the discovery of this Hulse-Taylor pulsar binary system.

More recently, our team has used the Green Bank Telescope (GBT), in West Virginia, to discover the first known pulsar in a stellar triple system (Ransom et al. 2014, Nature, 505, 520). This pulsar system goes by the name of PSR J0337+1715 (or just “J0337” for short). Discovered by PhD student Jason Boyles (West Virginia University), regular monitoring of this pulsar soon showed that it was behaving strangely. The pulsar was clearly in a 1.6-day orbit with another star (a white dwarf), but after correcting for the Doppler effect induced by this orbital motion the pulsar’s spin rate still appeared to be speeding up.

All radio pulsars slow down with time because their huge magnetic fields act as a brake. So how could this pulsar be spinning faster with time? Continued monitoring using the GBT, Arecibo and WSRT revealed the answer: the pulsar and its binary companion were actually orbiting yet another white dwarf star once every 327 days (Figures 1 and 2). This was causing the pulsar to move towards us, and thus further Doppler-shifting its period to make it appear like it was rotating faster and faster. Such an orbital configuration is known as a hierarchical triple system because it is effectively a binary system nested within another binary system. The orbits of the three stars can be as long as the inner binary separation is much smaller than the outer binary.

We had an exciting new system to study, but also a big challenge at hand: given the complicated way in which the pulsar and the two white dwarfs pull on each other, was it going to be possible to accurately model their motions?

High-cadence monitoring observations were key to understanding this system. Though GBT and Arecibo provided very sensitive measurements, WSRT was able to play a critical role because it could make near-daily observations during the early preparations for Apertif in 2012 and 2013. Even if only a subset of the dishes was available, and even if only for 30 minutes at a time, this still provided valuable pulsar timing data. Having observations each day meant that we could start with a rough model of the system, and then slowly refine this description each time a new measurement was added. However, this modeling was inaccurate because it assumed that the nested inner binary and outer binary do not interact with each other. In reality, however, the two orbits are constantly tugging on each other and changing the pulsar’s motion in subtle, but measurable ways.

Enter Anne Archibald, who fearlessly took on the ambitious task of accurately modeling all effects in the system. In her approach, Anne created a computer simulation of the orbital motions, tracking how the three stars dance around each other under the influence of gravity. She then tweaked the starting point of this simulation, the orbital properties, and the masses of the three stars, until the simulated motion matched with the observed arrival times of pulses from the pulsar.
From this modeling, nearly the full geometry of the orbits can be determined, as well as the individual masses of pulsar and two white dwarfs. We found that the pulsar has a mass 1.4 times that of the Sun, whereas the inner and outer white dwarfs have masses of 0.2 and 0.4 that of the Sun, respectively. Each of these masses could be determined with a precision of one part in 10 thousand.

Understanding how such a triple system can exist at all is already an interesting puzzle in stellar evolution. Importantly, the pulsar spins with a period of only 2.7 milliseconds — i.e. it rotates hundreds of times a second. The paradigm for the origin of such rapidly rotating neutron stars is the recycling scenario, in which a low-mass companion star can steadily transfer matter and angular momentum onto a neutron star and thereby spin it up. Recycling is certainly part of PSR J0337+1715’s creation story, but the formation and evolution likely involved several stages of mass transfer, possibly also a stage where the outermost star was transferring mass onto the inner binary and creating a circumbinary disk. Though there must certainly be other pulsars in our Galaxy with two white dwarf companions, such systems are not easy to form and thus must be quite rare.

Fortunately, we can use PSR J0337+1715 as a unique natural laboratory, even if we don’t understand the details of how it was formed. As previously mentioned, pulsars have been used to perform important tests of gravitational theories. In the case of the Hulse-Taylor binary, that was by detecting orbital decay from gravitational waves. Other pulsar systems can test gravity in complementary ways.

Millisecond pulsars in very wide orbits with a white dwarf star have been used to test a fundamental aspect of general relativity: the universality of free fall. Universal free fall dictates that all objects, regardless of their mass or composition, will experience equal accelerations in an external gravitational field. Apollo astronaut David Scott demonstrated this principle in the near vacuum on the Moon by dropping a hammer and a feather. Sure enough, they hit the lunar surface at the same time. It is sometimes said that Galileo performed a similar demonstration from the leaning tower of Pisa, but it’s unclear whether he actually performed that experiment. Much more serious and precise tests have been done using lunar laser ranging, in which the exact distance to the Moon is determined by bouncing signals off a reflector left by the Apollo astronauts. In this way, we can see whether the Earth and Moon experience equal accelerations towards the Sun.

In the case of a pulsar and a white dwarf binary, the external gravitational field is provided by the Milky Way galaxy itself. Here we can see whether the pulsar and white dwarf also experience equal accelerations. If not, one would expect their orbit to become elliptical (oval shaped) in a way that points towards the centre of the Galaxy.

Such tests are important because, while general relativity perfectly describes the falling feather and hammer, maybe the extreme gravity of a pulsar will allow us to see where the theory starts to break down. Newton’s theory of gravity is an excellent practical description unless objects are moving very rapidly or bend space-time strongly, like a neutron star or black hole. So far, Einstein’s theory of gravity has shown that it provides an accurate description of gravity in the extreme conditions under which Newton’s theory breaks down. That said, maybe under even more extreme conditions, or with sufficient experimental precision, we’ll see that Einstein’s general relativity also has flaws. Testing the universality of free fall is key because it is a fundamental principle on which general relativity is based: gravity is purely a reflection of the geometry of space-time, and inertial and gravitational mass are equivalent.

The PSR J0337+1715 triple system allows us to test the universality of free fall like never before. Here the pulsar and inner white dwarf are falling freely in the gravitational field of the outer white dwarf. The accelerations that they are experiencing are many orders-of-magnitude stronger than those experienced by pulsar binaries in the gravitational field of the Galaxy alone. Hence, we can perform a much more powerful experiment compared to previous studies.

To do this, we collected over 1,200 hours of observations using WSRT, GBT and Arecibo. This provided over 27,000 measurements of the pulse arrival time, where each such value is derived by averaging close to a half million individual pulses. This is necessary to distinguish the pulsar signal from the noise of the telescope. What we found is that the pulsar and inner white dwarf experience identical accelerations towards the outer white dwarf, to within a few parts in a million (Archibald et al. 2018, *Nature*, 559, 73). This test is about 1,000 times more constraining than previous studies, and again Einstein’s general relativity has passed with flying colours.

Many alternative theories of gravity have been proposed, and our result enables us to place limits on how much these theories can deviate from the predictions of Einstein’s general relativity. At the same time, the quest to better understand the fundamental nature of gravity, dark matter, and dark energy continues. While the PSR J0337+1715 test will likely stand as the most stringent constraint on universal free fall for some time to come, observations of gravitational waves with LIGO/Virgo will also be testing gravity in other ways. With the upcoming Square Kilometre Array, we expect to discover a large fraction of all the pulsars in the Milky Way, and among this harvest will surely be other pulsar triple systems that may provide the next, even tougher exam for Einstein to pass.