

THE DISCOVERY OF THE BINARY PULSAR

Nobel Lecture, December 8, 1993

by

RUSSELL A. HULSE

Princeton University, Plasma Physics Laboratory, Princeton, NJ 08543,
USA

Exactly 20 years ago today, on December 8, 1973, I was at the Arecibo Observatory in Puerto Rico recording in my notebook the confirming observation of the first pulsar discovered by the search which formed the basis for my Ph.D. thesis. As excited as I am sure I was at that point in time, I certainly had no idea of what lay in store for me in the months ahead, a path which would ultimately lead me here today.

I would like to take you along on a scientific adventure, a story of intense preparation, long hours, serendipity, and a certain level of compulsive behavior that tries to make sense out of everything that one observes. The remarkable and unexpected result of this detective story was a discovery which is still yielding fascinating scientific results to this day, nearly 20 years later, as Professor Taylor will describe for you in his lecture. I hope that by sharing this story with you, you will be able to join me in reliving the challenges and excitement of this adventure, and that we will all be rewarded with some personal insights as to the process of scientific discovery and the nature of science as a human endeavor.

PULSARS

Pulsars were first discovered in 1967 by Antony Hewish and Jocelyn Bell at Cambridge University, work for which a Nobel Prize was awarded in 1974. At the time, they were engaged in a study of the rapid fluctuations of signals from astrophysical radio sources known as scintillations. They were certainly not expecting to discover an entirely new class of astrophysical objects, just as we were certainly not expecting to discover an astrophysical laboratory for testing general relativity when we started our pulsar search at Arecibo several years later. Pulsars have indeed proven to be remarkable objects, not the least for having yielded two exciting scientific stories which began with serendipity and ended with a Nobel Prize.

The underlying nature of pulsars was initially the subject of intense debate. We now know that these remarkable sources of regularly pulsed radio emission are in fact rapidly spinning neutron stars. A sketch of a

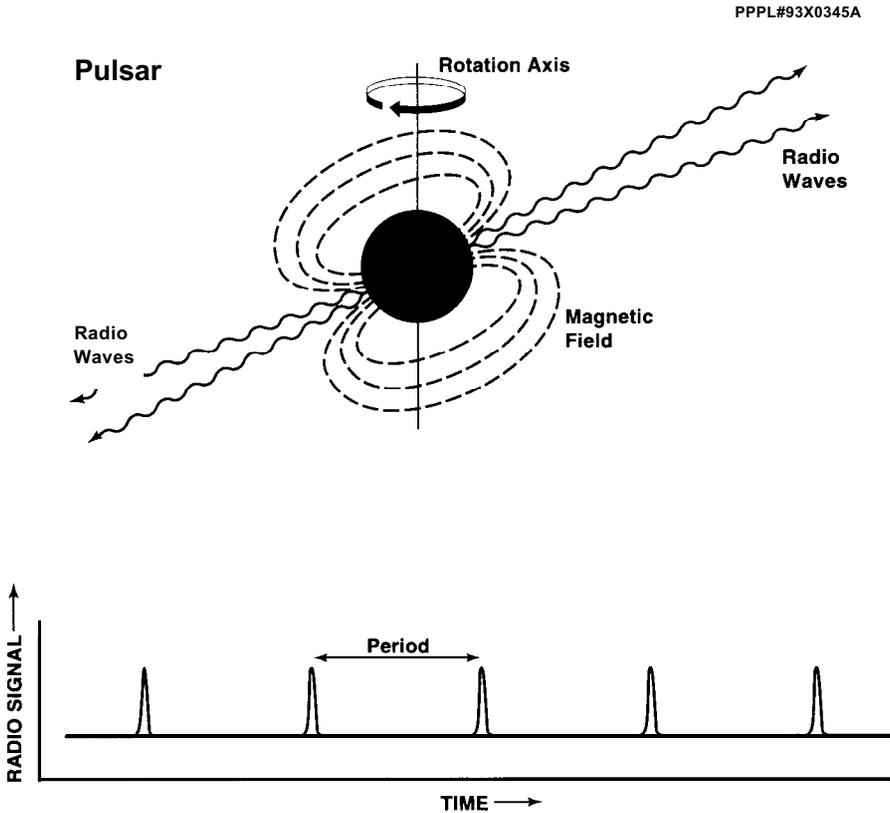


Fig. 1: A conceptual sketch of a pulsar, showing a rapidly rotating neutron star emitting narrow beams of radio waves from the polar regions of its embedded magnetic field. Also shown is a sketch of the periodic radio signal produced by a pulsar as seen by a radio telescope at the earth.

pulsar and its associated pulses as seen by a radio telescope is shown in Figure 1.

Neutron stars have roughly the mass of our sun compressed into an object only ~ 10 km in radius. Since the radius of our sun is about 7.0×10^5 km, one can immediately see that we are dealing with extraordinarily dense objects indeed. The intense gravitational field of a neutron star is sufficient to crush the electrons and nuclei of ordinary atoms into matter consisting primarily of neutrons. In a sense, a neutron star is a giant atomic nucleus, held together by gravitational rather than nuclear forces.

Neutron stars, and hence pulsars, are formed in supernovae, the spectacular final stage of stellar evolution for stars of many times the sun's mass. When the nuclear fuel of such a massive star is finally consumed, the resulting collapse of the stellar core produces a neutron star together with a cataclysmic explosion which expels the remaining stellar envelope.

The origin of pulsars in supernova explosions explains why in 1973 at the start of my thesis it was not considered surprising that all the pulsars detected to that point were solitary stars, despite the fact that a high

fraction of normal stars are found in orbiting multiple star systems. It seemed reasonable to expect that the immense mass loss of a star undergoing a supernova would tend to disrupt the orbit of any companion star at the time that a pulsar was formed. The discovery of the first binary pulsar produced an immediate addendum to this conventional wisdom, to the view that while disruption is quite likely, it is not quite obligatory.

The conservation of angular momentum and magnetic field during the stellar collapse which forms a neutron star leads to the fast rotation of pulsars along with an intense magnetic field on the order of $B \sim 10^{12}$ Gauss. This field is sketched in Figure 1 as predominantly dipolar, although actually it is inferred to possess some fine-scale structure as we will discuss in a moment. The combination of fast pulsar rotation with this high magnetic field gives rise to the very narrow beams of radio waves shown emitted from the region of the magnetic poles. If one (or, occasionally, both) of these beams happen to be aligned such that they sweep across the earth, we see them as the characteristic pulses from a pulsar. The effect is very similar to that of a lighthouse, where a continuously generated but rotating narrow beam of light is seen as a regular, pulsed flash by an outside observer each time the rotating beacon points in their direction.

Synchronously averaging together a few hundred individual pulsar pulses by folding the pulse time series shown in Figure 1 modulo the pulsar period produces an average pulse profile with a good signal-to-noise ratio. This average or integrated profile has a highly stable shape unique to each individual pulsar, just like a human fingerprint. They range in form from very simple single peaks to complex forms with double or multiple peaks, reflecting details of the magnetic field geometry of each individual pulsar. The widths of these profiles are characteristically narrow compared with the pulse period, having a duty cycle typically of the order of a few percent for most pulsars.

For the purposes of the pulsar observations which we will be discussing here, it is this average pulse that we will be referring to as the fundamental unit of pulsar observation. As an aside, however, it is interesting to note that the reproducibility of the average pulse shape for a given pulsar contrasts with a rich variety of quasi-random and regular variation in the individual pulses; they are not simply invariant copies of the average profile. In this sense, Figure 1 is an oversimplification in tending to show all the individual pulses as being the same. This stability of an average pulse contrasting with the pulse-to-pulse variation is understood by relating the average pulse shape (e.g., pulsar beam shape) to the fixed magnetic field geometry, while local smaller scale emission events moving within the constraints of this magnetic field structure generate the individual pulses.

Precision measurements of the intrinsic pulsation periods of pulsars as defined by the arrival times of these average pulses show that pulsars are extraordinarily precise clocks. This stability is readily understood as arising from the identification of the observed pulsar period with the highly stable rotation of a rapidly spinning, extremely compact ~ 1 solar mass object.

That pulsars are such precise clocks is in fact a central part of the binary pulsar story. In his lecture, Professor Taylor will describe pulsar timing methods in some detail. Suffice it for me to note for now that for PSR 1913 + 16, the binary pulsar we are going to focus on today, measurements of the pulsar period are now carried to 14 significant places, rivaling the accuracy of the most accurate atomic clocks. So you can certainly use a pulsar to set your watch without any hesitation!

SEARCHING FOR PULSARS

When I was approached by Joe Taylor at the University of Massachusetts to see if I was interested in doing a pulsar search for my thesis, it did not take too long for me to agree. Such a project combined physics, radio astronomy, and computers - a perfect combination of three different subjects all of which I found particularly interesting.

By 1973, pulsar searching *per se* was not by any means a novel project. Since the first discovery of pulsars in 1967, there had been many previous searches, and about 100 pulsars were already known at the time I started my thesis work. However, the telescopes and analysis methods used for these earlier searches had varied quite widely, and while many were successful, there seemed to be room for a new, high sensitivity search for new pulsars.

The motivations for such a new search emphasizing high sensitivity and a minimum of selection effects were several. First, a large sample of new pulsars detected by such a search would provide more complete statistics on pulsar periods, period derivatives, pulse characteristics, distribution in our galaxy and the like, together with correlations between these properties. Also, as pulsar radio signals travel to us, they are slowed, scattered and have their polarization changed in measurable ways which provide unique information on the properties of interstellar space in our galaxy. Thus finding a larger sample of new pulsars, especially ones further away in our galaxy, would be very valuable for these types of studies as well. In addition to statistical studies of the pulsar population, the search might also hope to uncover individual pulsars of unique interest, such as ones with very short periods. Indeed, since short period pulsars and very distant pulsars were precisely the types of pulsars which were the most difficult to detect for technical reasons, these most interesting of pulsars were precisely those the most strongly discriminated against in previous work. We saw an opportunity to improve on that situation with a comprehensive new computer-based attack on the pulsar search problem.

(And, yes, Joe's proposal to the National Science Foundation seeking funding for a computer for this search indeed pointed out, amongst these other motivations, that the discovery of "even one example of a pulsar in a binary system" would be quite valuable as it "could yield the pulsar mass". But we certainly had not been exactly counting on such a discovery, nor could we possibly have envisioned all that such a discovery might lead to!)

THE ARECIBO TELESCOPE

Since pulsars are relatively weak radio sources which must be observed at high time resolution in order to resolve their individual pulses, large radio telescopes are needed to collect as much signal as possible. Pulsars also have very steep radio spectra, meaning that they are strongest at low radio frequencies and are most often observed in the 100MHz to 1000MHz range, frequencies similar to those used for television broadcasts. Hence, the first requirement for an ambitious new high-sensitivity pulsar search was to use the largest radio telescope available for low-frequency pulsar work. The answer to that requirement was simple: Arecibo.

At 1000' in diameter, the Arecibo radio telescope is the largest single-element radio telescope in the world. An aerial view of the Arecibo telescope is shown in Figure 2. In order to support such a large reflector, the telescope was built in a natural bowl-shaped valley in the northwest corner of the island of Puerto Rico, several miles inland from the coastal town from which it takes its name.

Despite its fixed reflector pointing directly overhead, the Arecibo telescope is steerable to 20 degrees away from the zenith via the use of movable feed antennas. These antennas are suspended from a platform constructed of steel girders suspended by cables some 426 feet above the surface of the reflector. Using the 96 foot long feed antenna operating at 430MHz together with the Observatory's excellent low-noise receivers, Arecibo provided

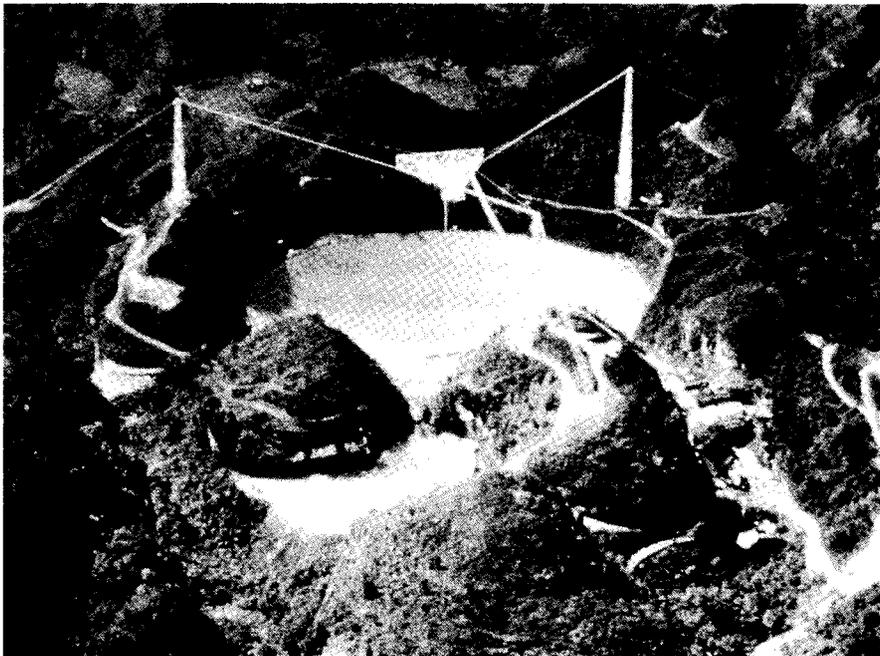


Fig. 2: The 1000' diameter Arecibo radio telescope, operated by Cornell University for the National Science Foundation as part of the National Astronomy and Ionospheric Center (NAIC).

just the capabilities needed for our pulsar search. The limited steering ability of the antenna was an acceptable tradeoff for obtaining this extremely high sensitivity, although it did have some interesting consequences in terms of the binary pulsar discovery, as we shall see later.

PARAMETER SPACE, THE FINAL FRONTIER

In order to detect a signal of any kind with optimal sensitivity, we must completely characterize its properties and then devise a “matched filter” which will take full advantage of these known properties. From this point of view, the three critical parameters of a pulsar’s signal for a search are its dispersion, period, and pulse width, as illustrated in Figure 3. Dispersion refers to the fact that due to its propagation through the free electron density in the interstellar medium, the pulses arrive at the earth successively delayed at lower frequencies. The existence of a characteristic period for

PPPL#93X0374

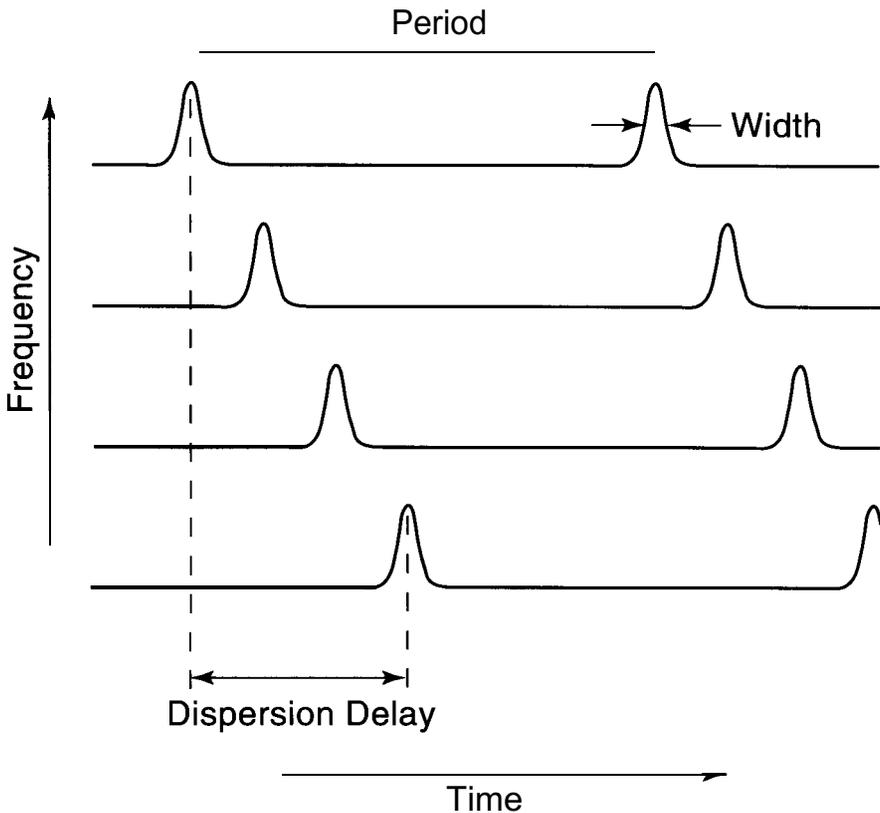


Fig. 3: A depiction of a received pulsar signal as a function of frequency and time, showing the three parameters employed in the pulsar search analysis: dispersion, pulse period, and pulse width.

each pulsar has already been discussed, while the pulse width has been adopted as the essential parameter describing the average pulse profile shape.

Since for a search we cannot know the pulsar's dispersion, pulse period, and pulse width in advance, in order to carry out an optimized detection we need to sweep across a wide range of possible values of these three parameters. This signal analysis also has to be done at every point in the sky. Adding the two stellar coordinate position parameters to the three pulsar signal parameters yields, in an overall sense, a 5-parameter search space.

Thus, the essence of obtaining the maximum possible sensitivity out of the new pulsar search was the ability to efficiently carry out an intensive and comprehensive search through the raw telescope data over this 3-dimensional parameter space of dispersion, period and pulse width. In turn, this meant that the essential new feature of this pulsar search relative to what had been done before lay in the computer analysis, not the telescope or other hardware.

DE-DISPERSION

Radio signals from pulsars suffer dispersion during their passage to the earth through the free electron density in interstellar space. This effect causes the pulses to arrive first at higher radio frequencies, and then successively later at lower frequencies, as illustrated in Figure 3. The magnitude of this effect is proportional to the integrated free electron density along the path between the pulsar and the earth. This column density is referred to as the dispersion measure (DM), given in units of cm^3pc .

Dispersion produces a significant special problem for pulsar observations. We always wish to use a large receiver bandwidth when observing a radio source so as to improve the sensitivity of the observations. The magnitude of this dispersion effect, at the frequencies commonly used for pulsar observations, is such that the arrival time of a pulse is significantly delayed across even a typical bandwidth of a few megahertz. Hence without special processing, the pulse would be smeared out in time when detected by the receiver. For example, when observing at 430MHz at Arecibo, the differential time delay across the available 8MHz bandwidth for a distant pulsar with $\text{DM} \sim 200 \text{ cm}^3\text{pc}$ would be 170 ms. Since this dispersion across the receiver bandwidth is much larger than a typical pulsar pulse width or even the entire period of short-period pulsars, "de-dispersion" is usually part of any pulsar observation. This is accomplished by first using a multi-channel receiver to split the wide total observed bandwidth into narrower individual adjacent channels such that the dispersion smearing within each channel is now acceptable for the pulsar study in question. Successively longer time delays are then added to the data streams from each higher

frequency channel in order to restore the proper time relationship between the pulses before summing all of the channels together.

While de-dispersion is a relatively straightforward process when observing a known pulsar at a given dispersion measure, a pulsar search must generate many de-dispersed data streams simultaneously, covering all possible dispersion measures of interest. This requires that the incoming data streams from the multichannel receiver must be added together not just with one set of delays added to successive channels, but with a whole range of sets of delays. An algorithm called TREE (by analogy with its computational structure) was used to carry out the required delays and summations on the incoming data as efficiently as possible by eliminating redundancies between the calculations required for different sets of delays.

This search de-dispersion processing is seen as the first step in the pulsar search data flow diagram shown in Figure 4. With the Arecibo telescope observing the sky at 430MHz, the multichannel receiver divided the 8MHz observing bandwidth into 32 adjacent 250kHz bandwidths. Detected signals from each of these channels were then sampled, de-dispersed, and written to magnetic tape in real time by the program ZBTREE running on the Modcomp II/25 minicomputer as the telescope scanned the sky.

It is interesting to note that a TREE algorithm can be used for a pulsar search all by itself, by simply scanning the resulting data streams for the characteristically dispersed pulses from a pulsar. An excess of strong pulses or an increase in the fluctuating signal power produced by weaker pulses occurring at some non-zero dispersion measure serves as the detection criterion. Such an approach, which relies only on the dispersed property of a pulsar signal, is much less sensitive than the search described here which includes a full periodicity analysis. But, in fact, it was just a search of this much simpler kind which was originally contemplated for Arecibo, and for which Joe Taylor had focused his funding proposal to the National Science Foundation for support to purchase a new computer. Once this new minicomputer arrived, however, it provided sufficient processing power to make a full de-dispersion and periodicity search a tantalizing possibility. The decision to attempt this much more ambitious goal was a critical factor in the ultimate success of the Arecibo search.

A vestige of the original dispersed pulse search was in fact retained in the REPACK program. This intermediate processing step on the Arecibo Observatory's CDC 3300 computer was needed since the Modcomp had no disk drive, and hence did not have enough space to re-organize ZBTREE data into the 136.5 second long data blocks required for the periodicity analysis program. REPACK displayed histograms of the signal strength distributions in each of the de-dispersed data channels in order to keep open the possibility of discovering unusually erratic pulsars or other types of objects which might generate sporadic non-periodic pulses. In practice, no discoveries were made this way, although the histograms did prove useful for monitoring the de-dispersion processing and indicating when severe terrestrial interference (both man-made as well as from lightning storms) was present in the data.

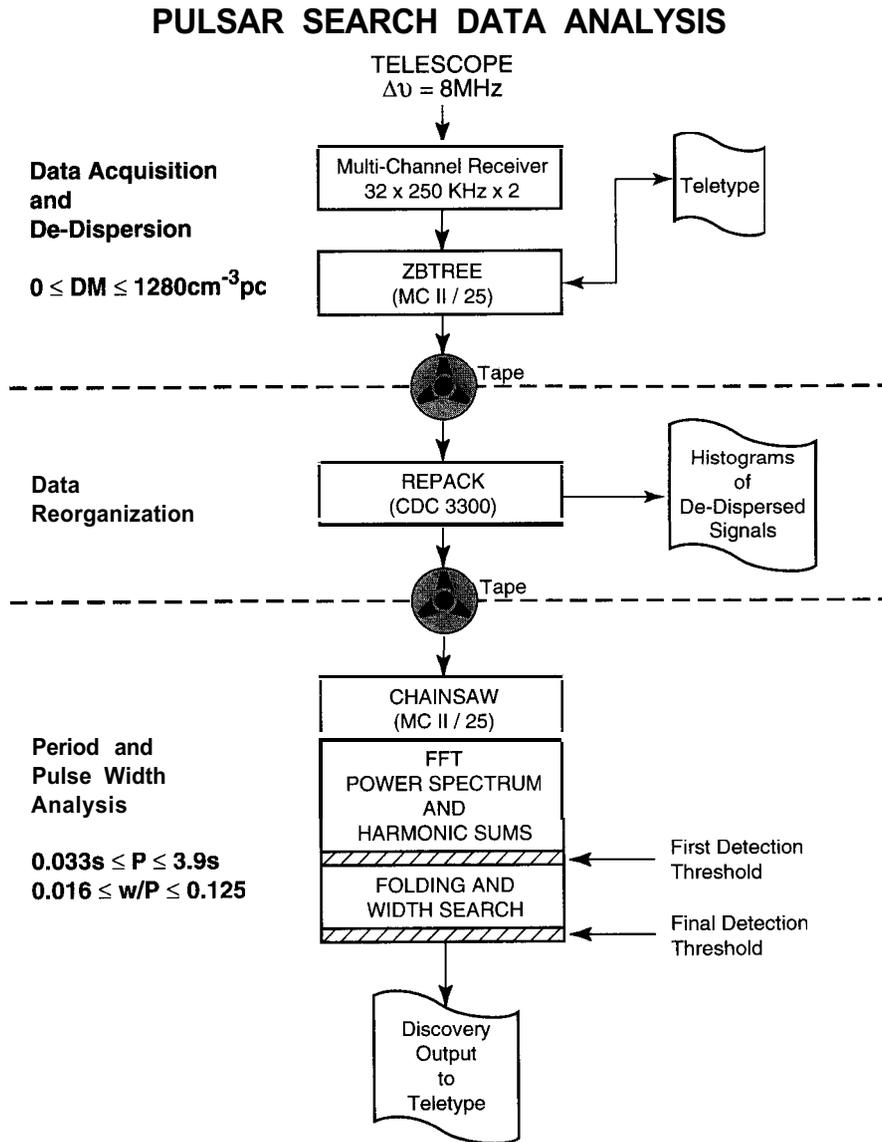


Fig. 4: A flow chart summarizing the pulsar search data analysis.

PERIOD AND PULSE WIDTH ANALYSIS

Again referring to Figure 4, we see that for the final periodicity and pulse width analysis the magnetic tapes written by REPACK were read back into the Modcomp II/25 minicomputer off-line between search observations using a program called CHAINSAW (since it reduced TREE data, of course - I think Joe Taylor gets the "credit" for coming up with this particular name).

This periodicity analysis was by far the most computationally intensive part of the overall search, and a great deal of effort was focused on finding an algorithm that would be efficient enough to meet the time constraints of the search while not sacrificing ultimate sensitivity.

The algorithm eventually developed for the period and pulse width search task was a hybrid approach, combining both Fourier analysis and folding of the data. In the Fourier analysis stage, a Fast Fourier Transform (FFT) was first used to generate a power spectrum from each data integration. This accomplished, the next task was to deal with the fact that the frequency spectrum of a periodic train of narrow pulses, such as that from a pulsar, presents the signal energy distributed across many harmonics of the fundamental pulsation frequency, with the number of harmonics present inversely proportional to the pulse width. Harmonics were thus summed within this power spectrum, with separate sums carried out over different possible numbers of harmonics for each fundamental frequency in order to cover the desired range of periods and pulse widths.

In the second phase of the analysis, the strongest signals revealed by the harmonic sums across the power spectrum were identified. This effectively formed a first detection threshold, which served to eliminate from further consideration the vast majority of the period and pulse width parameter space for each set of data which had no significant likelihood of containing a pulsar signal. The periods for each of these strongest "internal suspects" for each dispersion measure were then used to carry out a full folding and pulse width fitting analysis, with a signal-to-noise ratio calculated for each. If any of these signal-to-noise ratios exceeded a specified final threshold value, they were printed out to the teletype as possible pulsar discoveries. This final threshold was typically set at 7σ (a 7 standard deviation signal-to-noise ratio), which may seem to be a very high value. This high threshold was necessary, however, in order to keep the false detection rate reasonably low given the large parameter space being searched.

PUTTING IT ALL TOGETHER

In the end, the computerized search system shown in Figure 4, when combined with the Arecibo telescope, achieved a pulsar detection sensitivity over ten times better than that reached by any previous pulsar search¹. At each point in the sky scanned by the telescope, the search algorithm examined over 500,000 combinations of dispersion, period, and pulse width in the range of $0 < DM < 1280 \text{ cm}^{-3}\text{pc}$, $.033\text{s} < P < 3.9\text{s}$, and $.016 < w/P < 0.125$.

Having described what the pulsar search system looked like on paper, let me now show you what it looked like embodied in a real computer down at Arecibo. Figure 5 is a picture of me with the Modcomp computer in the Arecibo telescope control room. I will confess that looking at this picture does make me feel a little old. This reaction comes not so much when I look at myself in the picture, but rather when I look at this computer, which I

still remember thinking of as an impressively powerful machine. While it was certainly quite powerful as used for this purpose at that time, computer technology has certainly come a long way since I was working down at Arecibo.

You may note that the computer is housed in two rather crude looking wooden boxes. I made these up out of plywood at the University of Massachusetts to form a combination packing crate and equipment cabinet for shipping the computer to Arecibo, an example of the range of skills that graduate students often employ in doing their thesis work. You may also be able to see hash marks on the side of the crate near the top, which recorded a running tally of the pulsars discovered by the system.

The computer itself had a core memory of 16K 16 bit words. There was no floating point hardware, as all the calculations were done in scaled integer representations for speed. The teletype served for interactive input/output, while the tape drive provided the necessary data mass storage and the ability to communicate with other computers.

What you can't see in the picture is that this machine was programmed entirely in assembly language. Over 4000 statements on punch cards were eventually written for the analysis codes. Also, due to the limited hardware on the system (16K memory, no disk drive), I could not use the available operating systems provided by the manufacturer, which would have been too large and slow for the purposes of this application anyway. Hence, the task of programming the computer also involved writing my own custom drivers for the peripherals, servicing the interrupts, and the like. I do recall

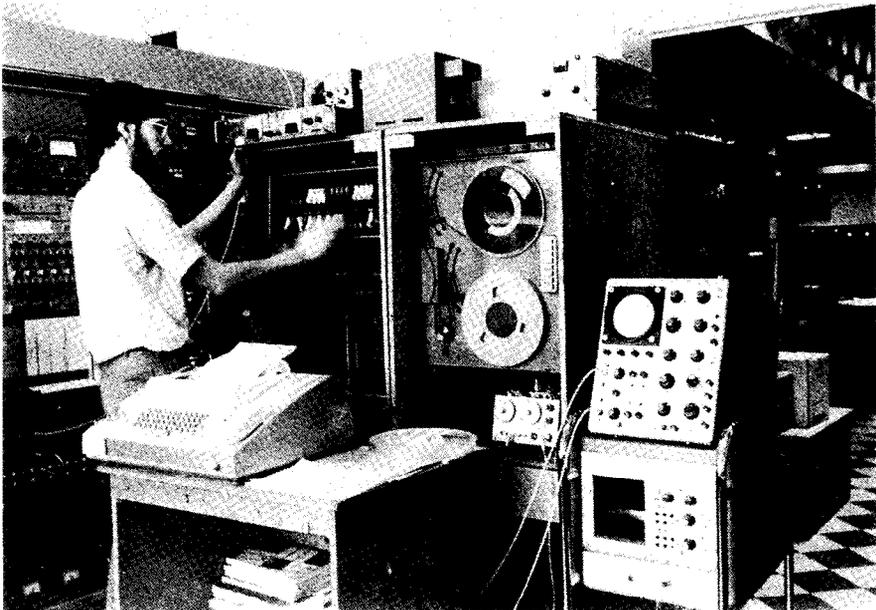


Fig. 5: A picture of myself with the Modcomp II/25 minicomputer used for the pulsar search at Arecibo, taken in the Arecibo control room shortly after the discovery of the binary pulsar.

a certain sense of pride that every bit set and every action carried out by that machine was explicitly controlled by code that I had written. I also recall saying when it was done that while it had been quite a valuable and interesting experience to understand and program a computer so extensively at such a fundamental level, once in a lifetime for this type of experience was enough!

SCANNING THE SKIES WITH COMPUTER AND TELESCOPE

As it happened, the Arecibo telescope was undergoing a major upgrade at the time we were proposing to use it for our search. Fortunately, while many other types of observations were made impossible during the construction work that this entailed, pulsar searching was not severely limited by the degraded telescope capabilities during this time. Also, as observing time had to be fit in between construction activities, a graduate student who could spend long periods living on site could take advantage of any available time on an ad hoc basis. Hence we had the opportunity to spend many more hours pulsar searching than I think we ever expected at the beginning of the search. At one point, I remember someone asking how much longer I would be staying down at Arecibo, to which Joe quipped "he's been down there for just one more month for the last year". In fact, I ended up at the observatory carrying out the pulsar search intermittently from December 1973 to January 1975, with time at Arecibo interspersed with periodic trips back to Massachusetts when observing time was not available.

Of course, my long days at Arecibo yielded a wide range of experiences, some of which are of the sort which are funny now but which weren't so funny then - such as the various occasions on which equipment failed, interference ruined entire observing sessions, and the like.

Coping with interference is a central fact of life for a pulsar astronomer since the relatively low radio frequencies used for these observations are plagued with interference of every type. For example, I can recall trying to eliminate one persistent long period "pulsar", which was ultimately traced to an arcing aircraft warning light on one of the towers which support the telescope platform structure. Lightning storms often led to the wholesale loss of entire observing runs, and when the U.S. Navy decided to hold exercises off of the coast, there was no point even trying to take data - I just sat in the control room watching signals from the naval radars (or whatever) jump around on the observatory spectrum analyzer.

As I alluded to before, successfully carrying out one's thesis often requires one to employ skills which weren't quite part of the original job description. In my case, this included the necessity of becoming an amateur computer field service engineer, when minicomputer components including the power supply, tape drive controller, and analog-to-digital converter multiplexer failed at different inopportune moments.

Just using the Arecibo telescope was quite a memorable personal experience in and of itself. I particularly liked the fact that I could set up the

telescope controls myself when doing a run, watching out the picture window in the control room as the huge feed structure moved in response to my commands to start an observation. It contributed to a feeling of direct involvement in the observations, of gazing into the heavens with some immense extension of my own eyes. At some telescopes, the control room is situated such that you can't even see the telescope during the observations. While this is hardly a practical problem for the observations, it is somewhat more romantic (and also reassuring to an astronomer's confidence that all is well) to be able to see the telescope move as you observe with it!

In the end, the pulsar search system performed beautifully. At the conclusion of the search, 40 new pulsars' had been discovered in the roughly 140 square degrees of sky covered in the primary search region near 19^h right ascension where a part of the plane of our galaxy is observable with the Arecibo telescope. These 40 were accompanied by the detection of 10 previously known pulsars, giving the search a 4-to-1 success rate in multiplying the pulsars known in this region. While for me this would have been quite a satisfying achievement and formed the basis for a nice thesis, it was of course eclipsed by the discovery of what was to become by far the most remarkable of these 40 new pulsars, PSR 1913 + 16.

MAKING DATA MAKE SENSE: THE BINARY PULSAR

By July of 1974 the pulsar search had settled down to something of a routine. I had even by then typed up and generated blank copies of various "forms" to be filled out to catalog and organize the discovery, confirmation, and subsequent period improvement observations for each new pulsar. Figure 6 shows the teletype output for the discovery observation of one such new pulsar, from data taken July 2, 1974. This output from the search system presented a succinct summary of vital information, including the pulsar position. The coordinates of this particular discovery correspond to those of PSR 1913 + 16, which we now also know as the binary pulsar. In this case two discovery outputs were generated, showing a signal with the same periodicity appearing in both dispersion channels 8 and 10. The pulsar period is shown on the output as about 53 ms, but my handwritten notation added to the output at the time corrects this to the true period of 59 ms. This standard correction factor of 1.1112 for the search output

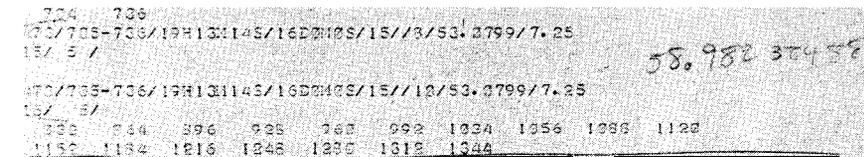


Fig. 6: The teletype output from the off-line analysis program showing the discovery of PSR 1913 + 16 using data taken on July 2, 1974. Note that due to a shift of the sampling rate of the search to avoid 60Hz power line interference, the 53 ms pulsar period typed by the analysis program needed to be adjusted to arrive at the correct 59 ms value I noted at the time by hand on the printout.

reflected a shift of the system sampling rate from 15ms to 16.668 ms in order to avoid interference generated by the 60Hz power line frequency.

The number 7.25 on the right of the output line is of particular note, as it shows that the initial detection of the binary pulsar occurred at only 7.25σ . Since the final search detection threshold was set at 7.0σ for this as most other observations, had the pulsar signal been only slightly weaker on this day it would never have even been output by the computer. The initial discovery of the binary pulsar was a very close call indeed ! In retrospect, all that work focused on getting every last bit of possible sensitivity out of the pulsar search algorithms had indeed been worth the effort.

The discovery form from my notebook for PSR 1913 + 16 is shown in Figure 7. It shows the original 7.25σ detection on scan 473, along with subsequent confirming re-observations on scans 526 and 535 which resulted in much more comfortable signal-to-noise ratios. The "fantastic" comment at the bottom refers to the 59 ms period of this pulsar, which at the time made it the second fastest pulsar known, the only faster one being the famous 33 ms pulsar in the Crab Nebula. At that point in time, I had no idea what lay ahead for this pulsar. However, a clue as to forthcoming events is evident on this form. You will note that as later observations of this pulsar proceeded, I kept going back to this discovery form changing the entry for the period of this pulsar, eventually crossing them all out in frustration.

After a group of new pulsar discoveries had been accumulated by the search, an observing session would be devoted to obtaining more accurate values for the pulsar periods. The first attempt to obtain a more accurate period for PSR 1913 + 16 occurred on August 25, 1974, nearly two months

PSR 1913+16*				
RA	DEC	PERIOD	DM	
19 ^h 13 ^m 14 ^s	16°00'	.05910332	150	
191313 ±4 ^s	16°00'24" ±60"	59 ms 0591030	167±15	
SCAN	DUMPS	DM CHNL	FREQ/tree	S/N
473	705-736	8,10	430/28	7.25
526	1-128	8,9	430/28	13.0
535	1-384	8	430/28	15.0
$a = 49.95$ $b = 2.11$ <i>fantastic!</i>				

Fig. 7: The "discovery form" from my notebook recording the detection and confirmation of PSR 1913+16.

after the initial discovery. The standard procedure involved making two separate 5 minute to 15 minute observations for each pulsar, one near the beginning and the other near the end of the approximately 2 hour observing window that Arecibo provided for a given object each day. The data was then folded separately to arrive at a period and absolute pulse arrival time within each of these two independent observations. After correcting for the effect of the Doppler shift due to the earth's motion, the periods determined within these two integrations were accurate enough to then connect pulse phase between the observations using the two measured arrival times, providing an accurate period determination across this one or two hour baseline.

As you can see in Figure 8, my first attempt to carry out this by then standard procedure for PSR 19 13 + 16 on August 25 produced a completely perplexing result. I always routinely compared the periods found within each of the two different short observations before proceeding with the remainder of the calculations, just to double check that nothing was amiss. As shown in my notes, something was indeed amiss for this data - and quite seriously so! Instead of the two Doppler-corrected periods being the same to within some small expected experimental error, they differed by 27 microseconds, an enormous amount. My reaction, of course, was not "Eu-

PERIOD REFINEMENT FOR PSR 1913+16 E

DATE 28/8/94

RA 19 13 14 DEC 16 00

CDC INTERFACE SAMPLES EVERY 5 MSEC FROM 10 MSEC FREE, CHNL 11

SCAN	AST START	# RECORDS	2209 PERIOD	2209 ZEROLAG	DOPPLER
567	20 46 30	180	.059 028 056	.0196	.999 950 152
574	22 11 10	180	.059 000 696	.0539	.999 949 530

FWHM ~ 15 msec (badly skewed - 2001 has ~~period~~ problems)

Check agreement between 567 and 574 $dp \sim \frac{(-.015)(.06)}{900} \sim .000001$

$P_{K567} = .059 025 113 5$

$P_{K574} = .058 997 710 2$

$22202 \quad .000 027$ - is already 2209 is not able to get a good period due to structure nature of pulse

DP $\frac{1}{2.4}$ $\frac{1}{25160}$ ~ 2.4 at integration not possible

DN $\sim (-.000 027)$ $\frac{1}{25160}$ $.059$

Fig. 8: A "period refinement form" from my notebook, showing the failure of my first attempt to obtain an improved period for PSR 1913 + 16.

reka - its a discovery" but instead a rather annoyed "Nuts - what's wrong now?" After a second attempt to carry out the same observation two days later resulted in even worse disagreement, I determined that I was going to get to the bottom of this problem, whatever it was, and finally get a good period for this one recalcitrant pulsar.

While I really could not imagine what specific instrumental or analysis problem could be producing such an error, there was the fact that at 59 ms, PSR 1913 + 16 was by far the fastest pulsar I had observed. Its pulses were thus not very well resolved by the 10 ms data sampling I was using for these observations, so I speculated that perhaps this had something to do with the problem. The difficulty with going to higher time resolution was that I was using the ZBTREE search de-dispersion algorithm as part of the period refinement observations, and the search computer simply could not execute ZBTREE any faster than the 10 ms rate I was already using. So getting higher time resolution meant setting up a special observation for this one pulsar, and writing a special computer program for the Arecibo CDC 3300 computer to de-disperse the new data and format it for analysis.

Data was first taken using the new observing system on September 1 and 2, about a week after the problems with the period determination became evident. As I worked on the new system, I tried to convince myself that as soon as the new, higher time resolution observations were available, the problems would just go away. The new data, of course, yielded just the opposite result-the problem was now even worse, since the period was still changing, but now poor time resolution could no longer be blamed. Figure 9 shows a plot taken from my notebook of this first period data taken with the new system. PSR 1913 + 16 was observed for as long as possible on each of these two days, and the observed period of the pulsar is shown versus time for the roughly 2 hours the pulsar was observable each day. I had color-coded the data from the two days in my notebook, which does not

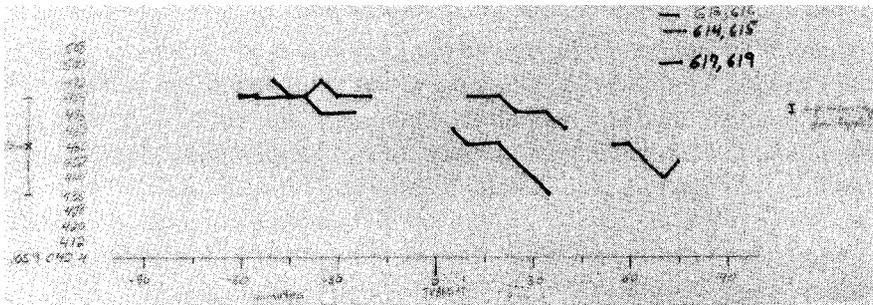


Fig. 9: The first two observations of PSR 1913 + 16 using the improved system specially designed to resolve the difficulties with determining the pulsation period for this pulsar. The observed pulsation period in successive 5-minute integrations is plotted versus time before and after transit. A calculation showing the magnitude of the change in the earth's Doppler shift is also seen on the right. Looking at this plot of data from September 1 and September 2, I realized that by shifting the second of these curves by 45 minutes the two curves would overlap. This was a key moment in deciphering the binary nature of PSR 1913 + 16.

reproduce well here in the figure. The data from the first day forms the broken upper curve, while data from the second day forms the uniformly lower broken second curve. The breaks in the curves were due to unavoidable gaps in the data. Both curves show the same behavior: the period starts out high, and then drifts to progressively lower values. The magnitude of the change of the earth's Doppler shift during this time is indicated by the small error bar on the right of the curve; so even if this correction had been done completely incorrectly, it could not really be a factor in what I was seeing here. Such a consistent drift during the course of an observation or experiment is usually quite suspiciously indicative of an instrumental problem, but, by this point, I had no good ideas left as to what could produce this effect. Certainly the observatory's precision time standards could not be suspected of drifting this much! But then, looking at this plot, it struck me that the two curves, while exhibiting the same overall downward trend, were really not identical, but that they would indeed match if shifted 45 minutes with respect to each other. While I clearly recall this as a crucial insight into the reality of the period variations, I do not recall whether or not the entire picture of PSR 1913 + 16 being in a binary system was also immediately clear to me at the same instant. But certainly within a short period of time I was sure that the period variations that I had been seeing were in fact due to Doppler shifts of the pulsar period produced by its orbital velocity around a companion star.

I immediately focused my observing schedule on getting as much more data on PSR 1913 + 16 as I possibly could, as fast as I possibly could. But a certain caution still remained; this was quite a remarkable assertion to be making, and I wanted to be absolutely confident of my conclusions before I announced my results. I thus set myself a strict criterion. If the pulsar was indeed in a binary, at some point the period would have to stop its consistent downward trend during each day's observations, reach a minimum, and start increasing again. I decided to wait for this prediction to come true to assure myself beyond all doubt that I correctly understood what I was seeing.

I had about two weeks to wait before I was able to obtain the final bit of evidence that I needed. The next page in my notebook shown in Figure 10 shows the new data as it was accumulated over this time. As before, the observed pulsar period is plotted versus time during the observing run, with data from the first two days of Sept. 1 and Sept. 2 seen copied over near the top of the page. At the lower right is the curve from Sept. 16, finally showing what I had hoped to see, the period reaching a clear minimum and then reversing direction.

By Sept. 18, my analysis of the data taken Sept. 16 showing the period minimum was complete, and I wrote a letter to Joe back in Amherst to tell him the amazing news that the pulsar was in a high-velocity binary orbit with about an 8 hour period. (Even at that point I referred to him as "Joe", rather than "Professor Taylor" - an informality that marked his approach to his students as colleagues rather than subordinates which I greatly

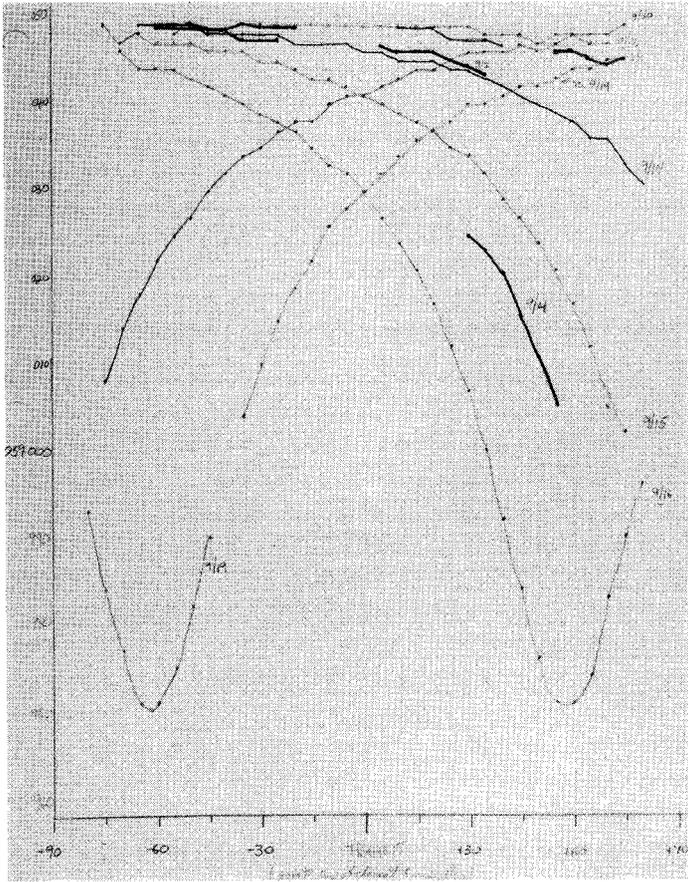


Fig. 10: After the realization that PSR 19 13 + 16 was probably in a binary system, data was taken as often as possible on succeeding days to fully confirm this hypothesis. The criterion that I established for myself was that I would need to see the period derivative change sign and start to increase. This final confirmation of the binary hypothesis was obtained with the September 16 data shown at the lower right.

appreciated.) But just sending a letter seemed rather inadequate for such news, and as telephone connections from Arecibo were somewhat problematical, I decided to try calling him using the observatory's shortwave radio link to Cornell University. This radio transceiver communicated with its counterpart in Ithaca, New York, where a secretary could then patch the radio over into the telephone and connect me to Joe at the University of Massachusetts in Amherst. I do not recall exactly what we said, beyond telling him the news, but as you can imagine Joe was on a plane to Arecibo in very short order. He arrived at Arecibo with a hardware de-disperser which allowed the continuing observations of the binary to proceed much more efficiently than had been possible using the ad hoc system I had put together for the first observations.

Upon his arrival, we compared our thoughts on the status and potential of the binary pulsar discovery and set up the new equipment. Soon Joe

started to program a more formal least squares fitting procedure for the orbital analysis, while I concentrated on accumulating and reducing more period data from the pulsar and finishing up the pulsar search work needed for my thesis. After Joe left to head back to Amherst, the new binary pulsar data was transferred between the pulsar data acquisition system (me) and the orbit analysis system (him) by my reading long series of numbers to him over the shortwave radio, a rather low-bandwidth but straightforward data link.

ANALYZING THE ORBIT WITH NEWTON AND EINSTEIN

The binary pulsar system revealed by the period variations which I have just shown is depicted in Figure 11. Instead of existing as an isolated star as was the the for all previously known pulsars, PSR 1913 + 16 is in a close elliptical orbit around an unseen companion star. In making this sketch, I have allowed myself the benefit of an important additional piece of information which was not immediately known at the time of the discovery. As Joe Taylor will describe, the pulsar and its companion are now known to have almost identical masses, and hence their two orbits are drawn as being closely comparable in size.

Combining the observed period variations shown in Figure 10 with subsequent other observations provided the first full velocity curve for the binary pulsar orbit. These measurements of the pulsar period converted to their equivalent radial velocities are shown as the data points in Figure 12, reproduced from the binary pulsar discovery paper³ published in January 1975. These velocities are shown plotted versus orbital phase, based on the $7^{\text{h}}45^{\text{m}}$ period of the pulsar's orbit.

That the binary's orbital period is almost exactly commensurate with 24 hours explains why the observed period variations had shown their tantaliz-

PPPL#93X0345B

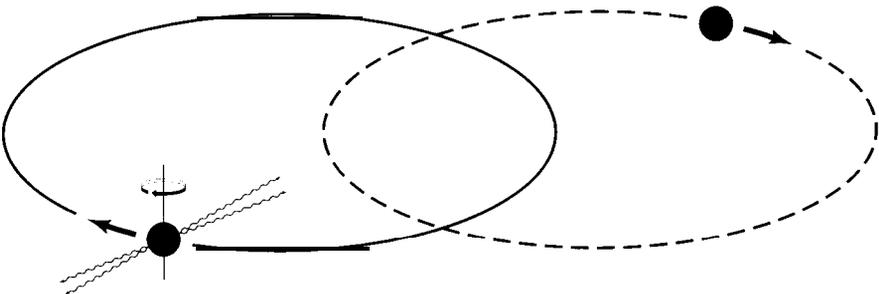


Fig. 11: A sketch of the pulsar in its elliptical orbit around a similarly massive companion star.

ing 45 minute shift each day. By pure chance the binary's orbital period was such that each time the pulsar was observable overhead on each successive day at Arecibo, it had also come around again to almost exactly the same place in its orbit. The result was that I would see almost exactly the same period variation all over again during the next day's observation. The critical 45 minute shift identified in Figure 9 thus was just the difference between three whole $7^{\text{h}}45^{\text{m}}$ orbits and 24 hours. Had Arecibo been a fully steerable telescope, such that I could have observed the pulsar for a longer time each a day, there would have been much less mystery and drama involved in the identification of the pulsar as a binary. But at Arecibo I did have two critical advantages which offset this problem. In the first place, the telescope's enormous sensitivity allowed me to first discover and then re-observe the pulsar in relatively short integrations, short enough so the pulsar could still be detected despite its varying period. A more subtle advantage was that at Arecibo I had the rare opportunity to use a "big science" instrument in a hands-on "small science" fashion. I had extensive access to telescope time, and I was able to quickly set up, repeat, and change my observations as I saw fit in pursuing the binary pulsar mystery.

Also shown in Figure 12 from the discovery paper is a curve fitted to the velocity data points using a Keplerian orbit, e.g., an orbit based on Newtonian physics. The orbital analysis used to arrive at this first formal fit to the

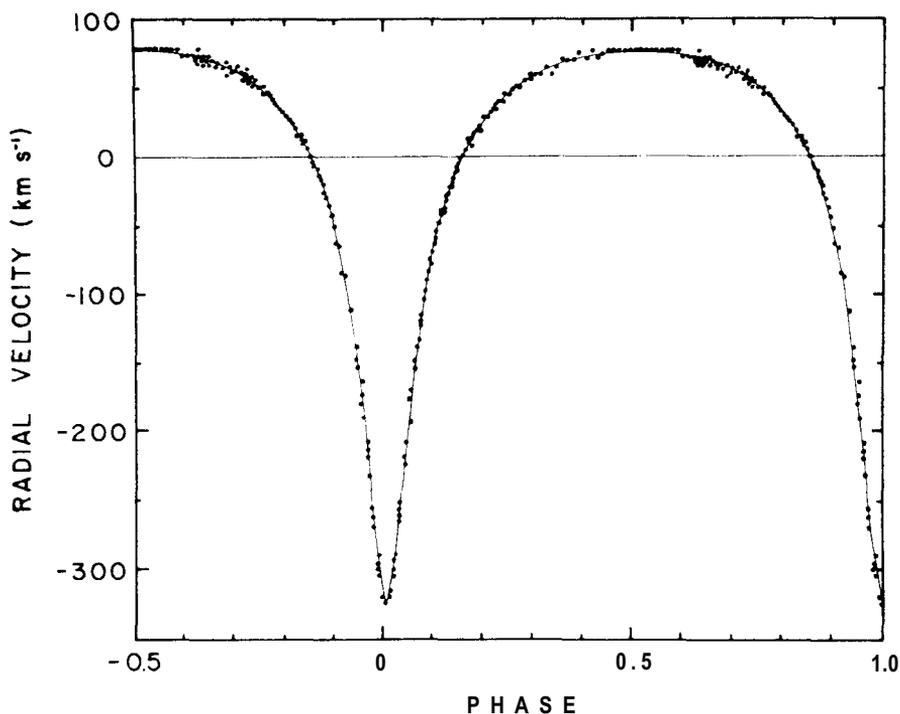


Fig. 12: The complete velocity curve for PSR 1913 + 16 from the discovery paper, fitted with a Keplerian orbital solution. The orbital phase is the fraction of a binary orbital period of $7^{\text{h}}45^{\text{m}}$ (from R.A. Hulse and J.H. Taylor, *Astrophys. J.*, 195: L51 -L53, 1975).

binary pulsar data was exactly the same as that historically applied by optical astronomers to the study of systems known as single line spectroscopic binaries. The only difference in the case of the binary pulsar was that instead of observing a periodic shift in wavelength of a line in the visible spectrum of one member of a binary star system, we were using the periodic shift of the spectral line formed by the pulsation period of the pulsar clock. Indeed, the first rough solution obtained for the orbital elements from this velocity curve involved the use of a graphical technique found in Aitken's classic book "The Binary Stars"⁴, first published in 1935! This hand analysis of the data was carried out at Arecibo using a velocity curve plotted on overlapping pieces of transparent graph paper held together with paper clips. The results from this crude fit were later used as the first approximation to the orbital elements needed to initialize the more accurate least squares numerical fit shown in the paper.

This initial purely Keplerian orbit analysis has now long since been superseded by a fully relativistic solution for the orbit, determined with exquisitely high accuracy using almost 20 years of observations. The orbital elements presented in the discovery paper based on this velocity curve, however, served to dramatically quantify why we knew that relativistic effects would be important in this system immediately after the binary nature of PSR 1913 + 16 was identified. While the Keplerian analysis did not by itself allow a full determination of the system parameters, it certainly sufficed to yield a breathtaking overall picture of the nature of this binary system: an orbital velocity on the order of ~ 0.001 the velocity of light, an orbital size on the order of the radius of our sun, and masses of the members of the system on the order of a solar mass. A succinct summary of the implications of these results was given in the discovery paper:

"This. . . will allow a number of interesting gravitational and relativistic phenomena to be studied. The binary configuration provides a nearly ideal relativity laboratory including an accurate clock in a high-speed, eccentric orbit and a strong gravitational field. We note, for example, that the changes in v^2/c^2 and GM/c^2r during the orbit are sufficient to cause changes in the observed period of several parts in 10^6 . Therefore, both the relativistic Doppler shift and the gravitational redshift will be easily measurable. Furthermore, the general relativistic advance of periastron should amount to about 4 deg per year, which will be detectable in a short time. The measurement of these effects, not usually observable in spectroscopic binaries, would allow the orbit inclination and the individual masses to be obtained".

As we now know, even this ambitious expectation was an understatement of the relativistic studies that would ultimately be made possible by this system. But at the time, excitement over the possibility of observing these potentially large relativistic effects was tempered by concerns over whether the

system would be sufficiently clean, that is, sufficiently free of other complicating effects, to allow these measurements to be made. But by early 1975, as I was starting to write my thesis, the first quantitative evidence arrived that the system might just indeed become the amazing relativistic laboratory that it has since proved to be.

This evidence came in the form of a measurement of the advance of periastron of the orbit using a more precise pulse arrival time method to analyze the binary pulsar data. This general relativistic effect involves the rotation of the elliptical orbit itself in space. Einstein's success in explaining the observed excess 43" of arc per century advance of perihelion of Mercury's orbit in our solar system by this relativistic effect constituted one of the earliest triumphs of his theory of general relativity. Initial results now showed what certainly seemed to be this same effect in the binary pulsar orbit, but now at the incredible rate of 4 degrees per year, just as predicted in the discovery paper! Thus in the 100 years it would take Mercury's orbit to rotate by a mere 0.01 degrees, the binary pulsar orbit would rotate more than 360 degrees, turning more than once completely around! This was a dramatic confirmation indeed of the role of relativistic effects in the new binary pulsar system. In fact, it should be noted that since the binary discovery almost 20 years ago, the orbit has already rotated nearly 90 degrees due to this effect, and if the binary pulsar were to be discovered today, the shape of its velocity curve would look nothing like that shown in Figure 12 from the original discovery paper.

I will end my story of the binary pulsar at this point, historically corresponding to the end of my doctoral thesis at the University of Massachusetts. In closing, first let me thank all those at the University of Massachusetts and at Arecibo who helped make my thesis and all that has followed from it possible. I would also like to observe that the long-term study and analysis of the binary pulsar was and is an exacting task that required a maximum of patience, insight, and scientific rigor. I have always felt that there was no one more suited to this study than the person with whom I have the honor of sharing this prize. I am therefore very pleased to now turn the binary pulsar story over to someone I have admired both as a person and as a scientist for over 20 years, Professor Joseph Taylor.

REFERENCES

1. R. A. Hulse and J. H. Taylor. A high sensitivity pulsar survey. *Astrophys. J. (Letters)*, 191:L59-L61, 1974.
2. R. A. Hulse and J. H. Taylor. A deep sample of new pulsars and their spatial extent in the galaxy. *Astrophys. J. (Letters)*, 201: L55 - L59, 1975.
3. R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *Astrophys. J.*, 195: L51 -L53, 1975.
4. Robert G. Aitken, *The Binary Stars*, Dover Publications, New York, 1964.