Parabola Volume 40, Issue 2 (2004)

The Breaking of the JN25 Series of Ciphers 1939 – 1945

Peter Donovan[1]

1. The National Security Agency.

The American National Security Agency, situated between Washington and Baltimore, is said to employ more mathematicians than any other organisation. This is a story of one of its predecessors, a joint facility of the United States Navy and the Australian Navy, that operated in Melbourne from 1942 to 1944. It continued as an Australian operation and was thus a predecessor of the Defence Signals Directorate. The Fleet Radio Unit, Melbourne, (FRUMEL) co-operated with other allied naval cryptology units. Its achievements were noted in the following message from the US Secretary of the Navy to the Australian Commonwealth Government (16 April 1947):

> The work of FRUMEL, was not highly publicized because of its secret nature, but the results obtained were of immeasurable importance in the successful prosecution of the late war. The successful accomplishment of its mission was in no small measure due to the unfailing devotion to duty of the 318 personnel of the Armed Forces of Australia who were attached to that unit and who worked side by side with the United States personnel.

2. The Modern Era.

The modern era in codes dates from August 1977, when the widely read journal *Scientific American* published details of a new encryption system that 'would take millions of years to break'. This depended upon availability of computing capacity, and could be rendered obsolete if quantum computers ever become a reality. Old-time cryptography ended at this time.

The crux of this innovation is a relatively simple trick. Given two largish primes p and q an encryption process can be defined that needs only the product pq, but the corresponding decryption process needs knowledge of the factorisation. In practice we can:

- Choose two 150-digit numbers by some random process. This could be done by throwing dice 300 times. It would not matter if the digits were restricted to 1, 2, 3, 4, 5, and 6.
- Calculate the next primes greater than or equal to each of these numbers and take these to be p and q.

[1]Peter Donovan is a retired Senior Lecturer in the School of Mathematics, University of New South Wales.

- Follow the processes described in *Scientific American*, noting that:
    * checking 150-digit numbers for being prime is quite easy using modern electronics, but
    * factorising 300-digit numbers that are in fact the product of two such unknown primes is not feasible.

3. The Use of Radio.

The introduction of radio and Morse code made large scale use of communication possible in warfare, but made ciphers essential. Note here the aim of the exercise is to delay interpretation of the message by the other side until it has no further relevance. Thus local tactical codes may be much simpler than the major operational code of a navy.

In the context of ciphers used late in the Second World War, there were certain other practices and pieces of technology that need to be mentioned:

- Radio Direction Finding (called RDF). The direction that a signal was coming from could be detected and thus if two different interception stations measured this together the location of the transmitter could be deduced.
- Radio Fingerprinting (called RFP) enabled an individual transmitter to be identified by idiosyncrasies in its output. At the time this British invention was kept very secret.
- The tabulator was a device for processing data stored on punched cards. With ingenuity this could be made to carry out a variety of quite specialised tasks to help cryptologists (code-breakers). It could also be used for tasks that would now be called information technology.
- Even if the equipment seems very primitive to modern eyes there were available – once the situation became desperate enough – immense human resources to carry out specialised clerical tasks.
- As radio interception had been practiced before 1939, some facilities were already in place.
- Some information could be derived from uninterpreted messages.

4. Code Book and Additive.

The following coding system is fairly obvious and was used quite widely. For present purposes we define a *group* to be a 5-digit number, perhaps with initial 0's. Thus there are $100\,000$ groups. A *code book* is prepared, containing in one column an alphabetical list of words and phrases to be used. Typically 'Message number' is such a phrase and 08540 may be the group assigned to it. Different groups are assigned to different

words and printed in the adjacent column. Some common words may be assigned two or three different groups. A reverse code book is also prepared, giving the groups used (the *book groups*) in numerical order with the corresponding words or phrases alongside.

Sending the sequences of code groups corresponding to the messages over the air would work for a while but eventually the other side would begin to work out the meaning of commonly occurring code groups. At this stage the code book would need replacement. So the task of the cryptologists on the other side is made much harder by bringing in an *additive table*. This would consist of (say) 100 pages each with 100 randomly selected groups as given in the following example.

	35	**86**	**79**	**65**	**49**	**72**	**52**	**03**	**62**	**12**
87	57721	56649	01532	86060	65120	90082	40243	10421	59335	93992
92	35988	05767	23488	48677	26777	66467	09369	47063	29174	67495
26	14631	44724	98070	82480	96050	40144	86542	83622	41739	97644
55	92353	62535	00333	74293	73377	37673	94279	25952	58247	09491
59	60087	35203	94816	56708	53233	15177	66115	28621	19950	15079
53	84793	74508	57057	40029	92135	47861	46694	02960	43254	21519
66	05877	55352	67331	39925	40129	67420	51375	41395	49111	68510
96	28079	84234	87758	72050	38431	09399	73613	72553	06088	93312
28	67600	17247	95378	36759	27135	15772	26102	73492	91394	07984
17	30103	41777	17780	88154	95706	61075	01016	19166	33401	52278

Here the *page co-ordinates*, given in heavy type, are numbered in a quite erratic way. The 100 pages would be numbered by 3-digit numbers, but not consecutively. Thus this page might be numbered **638**. An index to page order would assist the user.

The cipher clerk first takes the plain-language text and finds the corresponding code groups. Thus if the message was:

Send today five hundred tons fuel oil Truk

the code book might well give the groups:

BOOK 14159 26535 89793 23846 26433 83279 50288

The clerk would then choose randomly an initial place in the table of additives, such as page **638**, row **66**, column **52**. Consecutive entries from the table would then be written directly underneath the book groups as follows:

BOOK	14159	26535	89793	23846	26433	83279	50288
ADDs	51375	41395	49111	68510	28079	84234	87758

The additions start from row 66 column 52 and proceed left to right in subsequent rows (as in normal text).

These would then be added *without carrying* (this was called false addition) to yield the *groups as transmitted*, then called GATs, displayed in the third row:

BOOK	14159	26535	89793	23846	26433	83279	50288
ADDs	51375	41395	49111	68510	28079	84234	87758
GATs	65424	67820	28804	81356	44402	67403	37936

The starting place **6386652** would then need to be placed in disguised form at the beginning of the message together with an indication of the identity of the intended recipient.

There is an extremely slow method of breaking into such cipher systems which ultimately is based on the fact that some words are used much more than others. Of course the other side knows this risk and should change the additive reasonably frequently and the code book, which is harder to put together, somewhat less frequently.

In fact such systems, with a few other complications introduced to make things harder, served the Japanese Army quite well until January 1944 when a box of cipher material was found in a swamp in New Guinea. Yet the principal Japanese Navy operational code, or rather series of codes, JN25, was broken much earlier.

The steps in breaking such a cipher would appear to be:

(a) Write all the messages obtained in alignment, that is so that GATs corresponding to the same additive are vertically underneath each other. [The erratic choice of page co-ordinates makes this harder.]

(b) Work out what the additive is for the groups in each column. Thus simple subtraction (without carrying) recovers the book groups.

(c) Then you are back to the quite non-trivial problem of working out what the more frequently occurring book groups mean.

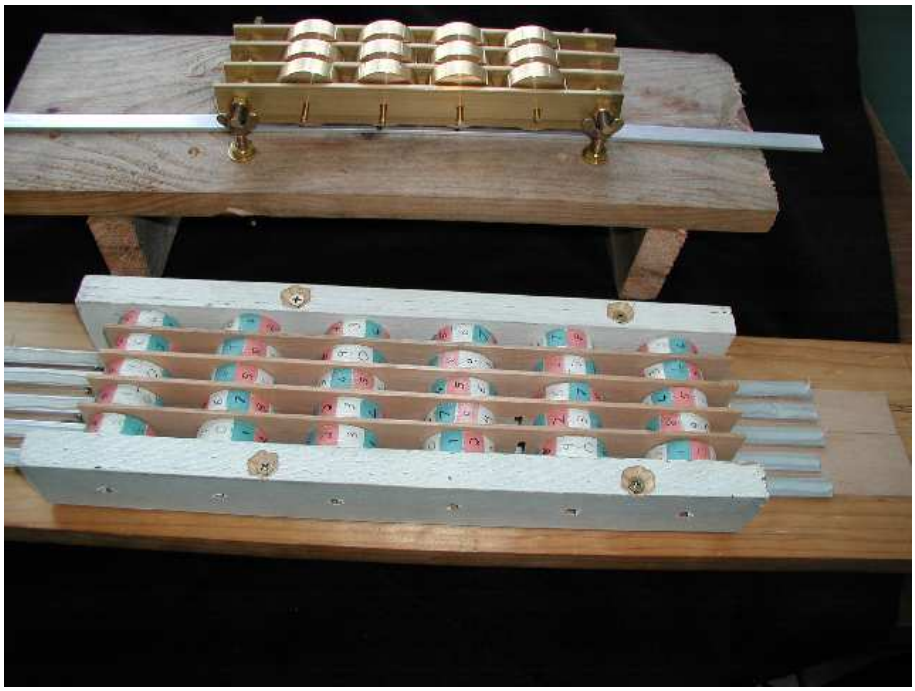
We will be examining the first two of these three steps for the JN25 series of ciphers. These had the special property that all the book groups when interpreted as natural numbers were *multiples of three*. Presumably the aim was to give a simple check for the deciphering clerk but *it turned out to be a quite remarkable blunder*.

In fact the book groups could and should have been selected in a patternless way. The NSW State Lottery of that era used a barrel with $100\,000$ numbered wooden balls with diameter about 1cm to determine the prizewinners. Such a device could be used to allocate different book numbers to the words in the code book.

One key document in understanding this matter surfaced in 2000 in the US National Archives. A photographic copy of the original may be obtained on the web by asking a search engine to locate 'Turing Dayton'. The date was December 1942.

The author, Alan Turing, was the principal scientist at the British military cryptologic station in World War II. He was the key figure in the effort to break into the German submarine Enigma traffic in what became known as the Battle of the Atlantic. Indeed, the bulk of the report deals with the special purpose computing devices then called 'bombes' used for submarine Enigma. The following sentences refer to something quite different.

> SUBTRACTOR MACHINE. At Dayton [Ohio, USA] we also saw a machine for aiding one in the recovery of subtractor groups when messages have been set in depth. It enables one to set up all the cipher groups in a column of the material, and to add subtractor groups to them all simultaneously. By having the digits coloured white, red or blue according to the remainders they leave on division by 3 it is possible to check quickly whether the resulting book groups have digits adding up to a multiple of 3 as they should with the cipher to which they will apply it most. A rather similar machine was made by Letchworth for us in early 1940, and, although not nearly so convenient as this model, has been used quite a lot I believe.



The photograph shows the author's working 'mock-up' of the
the SUBTRACTOR MACHINE described in the text
(The assistance of Cameron Verrills of the UNSW Facilities Department is
acknowledged).

Turing seems to have used the words 'code' and 'cipher' indifferently. The phrase 'in depth' was the contemporary jargon for 'in alignment'. He considered subtractors rather than additives but there is no fundamental difference.

The blunder in the JN25 series of ciphers had been detected by the distinguished British cryptologist John Tiltman. The team of mathematicians at Bletchley Park (about half way between Oxford and Cambridge in the UK) found a way of exploiting it. It consists of doing a bit of task **(a)**, then a bit of task **(b)**, then more of **(a)** and so on. The output was not quite certain but the odds were strongly on its side. The process was eventually self-verifying.

5. The Battle for Australia.

This phrase is used to describe a sequence of major events which turned around the Japanese advance following the Pearl Harbor raid of 7 December 1941. These would include the Battle of the Coral Sea, the Battle of Midway, the Kokoda Trail fighting and the actions on Guadalcanal (Solomon Islands). All were heavily influenced by naval signals intelligence emanating from work on JN25. The short series B5555 of key FRUMEL material in the National Archives clarifies this assertion.

Series B5555 is stored in Melbourne and, most regretably, is not available via the world wide web.

6. Example of use of the Subtractor Machine for task (b).

Here nine messages have been put in alignment and so it is known that the groups as transmitted 32815, 30763, 35252, 43535, 27363, 64350, 59362, 57761 and 21289 were all obtained from multiples of 3 by (false) addition of the same random 5-digit group. These groups are placed as the first column each of the displays for TEST 1, TEST 2 and TEST 3 below. Throughout, multiples of three are marked with a $\surd$ and non-multiples by a $\times$. These nine entries are set up in the machine. Note that examining all $100\,000$ possible additives one by one would be prohibitively tedious on a large scale. So suppose that previous work has thrown up five common code groups, and these are the underlined groups in the display below. The machine is used to try all five of these in the top entry and so with the other eight groups changed too. In each case the relevant part of the display gives some groups marked with a $\times$ and so are not multiples of 3. The process continues with trying out each of the five in the second row but again without success. However, when it is tried in the third row success is achieved. One has thus stripped the additive (here 60646) from one group in each of nine messages and also found one entry, also 60646, in the table of random additives.

We now consider this process in greater detail. Consider the entries in TEST 1. The first column represents the nine messages.

The second column in TEST 1 has been obtained by positioning the first of our common code groups (24648) at the top and evaluating the false additive difference between the first of the code groups in column one (32815) and this common group. The false additive difference is 18277. We now subtract this amount from each of the code groups in column one to obtain the remaining entries in column two. A similar approach has been used to obtain the entries in columns three, four, five and six.

Note that none of the trials in TEST 1 were successful because there is no column containing only entries that are divisible by three.

TEST 1.

$32815\times$		$\underline{24648}Y_\eta$		$\underline{18690}Y_\eta$		$\underline{95898}Y_\eta$		$\underline{35403}Y_\eta$		$\underline{83382}Y_\eta$
$30763\times$		$22596Y_\eta$		$16548Y_\eta$		$93746\times$		$33351Y_\eta$		$81230\times$
$35252\times$		$27085\times$		$11037Y_\eta$		$98235Y_\eta$		$38840\times$		$86729\times$
$43535\times$		$35368\times$		$29310Y_\eta$		$06518\times$		$46123\times$		$94002Y_\eta$
$27363Y_\eta$	$\rightarrow$	$19196\times$	$\rightarrow$	$03148\times$	$\rightarrow$	$80346Y_\eta$	$\rightarrow$	$20951\times$	$\rightarrow$	$78830\times$
$64350Y_\eta$		$56183\times$		$40135\times$		$27333Y_\eta$		$67948\times$		$15827\times$
$59362\times$		$41195\times$		$35147\times$		$12345Y_\eta$		$52950Y_\eta$		$00839\times$
$57761\times$		$49594\times$		$33546Y_\eta$		$10744\times$		$50359\times$		$08238Y_\eta$
$21289\times$		$13012\times$		$07064\times$		$84262\times$		$24877\times$		$72756Y_\eta$

In TEST 2, the additive difference has been obtained by comparing the common code groups with the second code group in the starting alignment. Again none of the columns have entries that are all divisible by three after this addition decoding.

TEST 2.

$32815\times$		$26790Y_\eta$		$10742\times$		$97040\times$		$37555\times$		$85434Y_\eta$
$30763\times$		$\underline{24648}Y_\eta$		$\underline{18690}Y_\eta$		$\underline{95898}Y_\eta$		$\underline{35403}Y_\eta$		$\underline{83382}Y_\eta$
$35252\times$		$29137\times$		$13189\times$		$90387Y_\eta$		$30992\times$		$88871\times$
$43535\times$		$37410Y_\eta$		$21462Y_\eta$		$08660\times$		$48275\times$		$96154\times$
$27363Y_\eta$	$\rightarrow$	$11248\times$	$\rightarrow$	$05290\times$	$\rightarrow$	$82498Y_\eta$	$\rightarrow$	$22003\times$	$\rightarrow$	$70982\times$
$64350Y_\eta$		$58235\times$		$42287\times$		$29485\times$		$69090Y_\eta$		$17979Y_\eta$
$59362\times$		$43247\times$		$37299Y_\eta$		$14497\times$		$54002\times$		$02981\times$
$57761\times$		$41646Y_\eta$		$35698\times$		$12896\times$		$52401Y_\eta$		$00380\times$
$21289\times$		$15164\times$		$09116\times$		$86314\times$		$26929\times$		$74808Y_\eta$

In TEST 3, the additive difference has been obtained by comparing the common code groups with the third code group in the starting alignment. This time we have success. All entries in the third column are multiples of three!

TEST 3.

$$\begin{array}{cccccccccc}
32815 \times & & 21201Y_\eta & & 15253 \times & & 92451Y_\eta & & 32066 \times & & 80945 \times \\
30763 \times & & 29159 \times & & 13101Y_\eta & & 90309Y_\eta & & 30914 \times & & 88893Y_\eta \\
35252 \times & & \underline{24648}Y_\eta & & \underline{18690}Y_\eta & & \underline{95898}Y_\eta & & \underline{35403}Y_\eta & & \underline{83382}Y_\eta \\
43535 \times & & 32921 \times & & 26973Y_\eta & & 03171Y_\eta & & 43786 \times & & 91665Y_\eta \\
27363Y_\eta & \rightarrow & 16759 \times & \rightarrow & 00701 \times & \rightarrow & 87909Y_\eta & \rightarrow & 27514 \times & \rightarrow & 75493 \times \\
64350Y_\eta & & 53746 \times & & 47798 \times & & 24996Y_\eta & & 64501 \times & & 12480Y_\eta \\
59362 \times & & 48758 \times & & 32700Y_\eta & & 19908Y_\eta & & 59513 \times & & 07492 \times \\
57761 \times & & 46157 \times & & 30109 \times & & 17307Y_\eta & & 57912Y_\eta & & 05891 \times \\
21289 \times & & 10675 \times & & 04627 \times & & 81825Y_\eta & & 21430 \times & & 79319 \times
\end{array}$$

Calculation of the last two columns here is really a waste of time: the almost certain correct solution has already been found.

7. Example of aligning two JN25 messages for task (a).

In this example we have two intercepted messages each with 24 GATs ('groups as transmitted'). As before $\surd$ indicates divisibility by 3. The first is

98319	26876	84038	31341	92269	11803	39045	52147
55734	67835	64288	97135	58304	32312	90596	60820
68894	95971	95186	09126	12977	76054	75396	02608

We are also given a string of six consecutive groups from the additive table.

40422	35613	73995	03881	36886	24108

We subtract these groups (called ADDs in the tables below) from runs of six consecutive GATs from the first message to obtain differences (DIFF) as shown below. As these differences are obtained it is checked whether they are multiples of 3. The first calculations yield:

GATs	98319	26876	84038	31341	92269	11803
ADDs	40422	35613	73995	03881	36886	24208
DIFF	58997 $\times$					

Once a difference which is not a multiple of 3 has been found, we move on to the next round of calculations to obtain:

GATs	26876	84038	31341	92269	11803	39045
ADDs	40422	35613	73995	03881	36886	24208
DIFF	$86454Y_\eta$	59425 $\times$				

The third calculations yield:

GATs	84038	31341	92269	11803	39045	52147
ADDs	40422	35613	73995	03881	36886	24208
DIFF	44616	06738	29374 $\times$			

The fourth calculations yield:

GATs	31341	92269	11803	39045	52147	55734
ADDs	40422	35613	73995	03881	36886	24208
DIFF	$91929Y_\eta$	$67656Y_\eta$	$48918Y_\eta$	$36264Y_\eta$	$26361Y_\eta$	$31536Y_\eta$

Thus there is a very high probability that we have aligned the intercepted message with the given run of six additives. The second intercepted message is:

17492	37961	68705	88286	00912	80906	62797	23270
49455	72754	85734	72351	05633	19011	71430	49311
06121	22548	10154	50865	81704	08147	63599	73309

Eventually one discovers that this time the very likely alignment is:

GATs	85734	72351	05633	19011	71430	49311
ADDs	40422	35613	73995	03881	36886	24208
DIFF	$45312Y_\eta$	$47748Y_\eta$	$32748Y_\eta$	$16230Y_\eta$	$45654Y_\eta$	$25113Y_\eta$

The two intercepted messages may now be put in (highly probable) alignment.

8. Notes on section 2, 3 and 4.

- In 2003 a notable advance was made in computer programs that determine whether a given large integer is prime. Thus the process of finding the smallest prime greater than a given 150 digit integer is not particularly formidable.
- Occasionally icosahedral dice are available in specialised shops. The shape is that of the regular icosahedron and the faces are numbered from 1 to 20. These could be used to produce random sequences of digits much more naturally than conventional cubical dice.
- It appears that only one detailed account was written down about the use of tabulators in code-breaking in the Second World War. It is file HW 25/22 in the Public Records Office in Kew, a suburb of London. Regretably it is not accessible by the web. The author remembered Tiltman attributing his success with JN25 to a hunch.
- The Australian Archives item A5954 530/2 is available from the web and contains an item from the Advisory War Council minutes of 17 June 1942. 'As it was, we had the great advantage in each action of knowing the Japanese plan beforehand.' This refers to the crucial Battle of Midway and a relatively minor diversionary action around the Aleutians, a chain of islands off Alaska. Access to this item is obtained from `http://www.naa.gov.au` by clicking on 'Recordsearch'.

- The Australian Archives item A425 C1947/514 documents the presence of a card-processing device called the NC-4 at FRUMEL. This was used in the alignment part of handling JN25 intercepts, probably from mid-1943 onwards.
- The Australian Archives items B5436 part A and B5436 part G are useful reference material for this article. They, and also the previous item, may be accessed from the web.
- The American National Cryptologic Museum has the transcript of an interview with Rudolph Fabian, the administrative head of the American side of FRUMEL for nearly two years. This confirms that a small British team had been working on JN25 in Singapore in 1940. The United States Navy was allowed to send an observer to this British unit early in 1941. Fabian was then responsible for switching the American cryptologic team he commanded in the Philippines over to working on JN25. More resources were put into breaking JN25 after the Pearl Harbour raid of December 1941. Fabian's team was evacuated to Melbourne in February, March and April 1942.
- This article reveals nothing about how to get started on a cipher of JN25 type. Hints on this aspect may be found in archive B5436 part b.