The Special Fish Report by Albert W. Small (December 1944) reformatted by Tony Sale (c) March 2001

INTRODUCTION

The original document is held in The American National Archive (NARA) College Campus Washington.

NR 4628 SPECIAL FISH REPORT (BOX 1417)

This Report was written by Albert W. Small an American Cryptanalyst in the U. S. Signal Corps who was seconded to Bletchley Park and joined the Newmanry to work on breaking the German Lorenz cipher.

Formatted for HTML and PDF by Tony Sale (c) March 2001

Note: pages 30-54 and 85-92 are currently omitted. These are work sheets some of which are A3 size.
It is hoped to add these when a suitable method has been found for displaying them on the web.

This reproduction has been produced from photo copies of the original document. Unfortunately the original had been photographically copied, so was white on black. A challenge for scanning, image inversion and Optical Character Reading.(OCR). An attempt has been made to format each page as closely to the original as possible, including the use of double line spacing and some rather idiosyncratic use of spaces.

There is also some difficulty in representing the mathematical formulae appearing in the original document. Very often these have been written in by hand. After some discussion with colleagues a text representation has been agreed for the Greek symbols. This is explained on the "Current Notations" page 112 at the end of the document. The results are not entirely satisfactory but it may be possible to produce a more mathematically acceptable version in the future.

OCR is notoriously difficult to proof read and I am indebted to Frode Weierud, Andrew Hodges and many others for help in this. Reporting of any remaining "garbles" would be greatly appreciated.

This arduous task has been undertaken because of the interest in Colossus and the need to fully understand and debate the wartime use of Colossus to enable the completion of the Colossus rebuild by the Colossus Rebuild Project under the direction of Tony Sale.

Tony Sale, March 2001

CX/MSS TOP SECRET

1 December 1944

Subject: Special Fish Report

To: C.O., S.S.A., War Department
(Attention: SPSIB-3)

Herewith my notes on Fish problem, as per instructions.

(signed)
Albert W. Small
Cryptanalyst
U.S. Signal Corps

enclosure: 112 pages, numbered
consecutively.
Also pages 22 1/2, 27 1/2, 42 1/2,
60 1/2, 74 1/2, 83 1/2 94 1/2.

- -

SPECIAL FISH REPORT

* * * *

Daily solutions of Fish messages at G. C. & C. S. reflect a background of British mathematical genius, superb engineering ability, and solid common sense.

Each or these has been a necessary factor. Each could have been overemphasised or underemphasised to the detriment of the solutions; a remarkable fact is that the fusion of the elements has been apparently in perfect proportion. The result is an outstanding contribution to cryptanalytic science.

The original mathematics of the Fish problem was brought to us from England by Captain Seaman and Mr.Ferner,and kept scrupulously up-to-date by Capt. Fried. There is little difference therefore between the mathematics at Arlington

and that now existing at G.C. & C.S. We have all the elements of the story -- in fact, many more elements than are used in the story -- and we lack only the operational perspective to piece them together. It is a perspective that we can hardly attain unless we ourselves do operations on an "assembly line" basis.

(Tony Sale: Notes written in pencil)

The formula* for the Fish machine is:

(Plain + periodic key + aperiodic key = cipher)
P + X + PSI' = Z

and since arithmetic is modulo 2:

Z + X = PSI' + P

and this may be expressed by the parametric equations:

(cipher + periodic = pseudo plain)
(l) Z + X = D

(2) D + PSI' = P.
(pseudo + aperiodic = real plain)

G.C. & C.S. Fish sections are organized on the basis of these parametric equations. Mr. Newman must solve the former, obtaining pseudo plain text D from cipher text Z. Major Tester must solve the latter, obtaining plain text P from pseudo cipher text D.

Operations (1) and (2) are equally important, since no plain text can be read unless both are solved. However we can say that solution of equation (1) has taken the greatest amount of statistical science, and is of great primary interest to Arlington Hall. Equation (2) has taken much knowledge of language and is of great interest as an art.

- -

*There is appended a list of symbols and meanings.
To give a better understanding of G.C. & C.S., it is necessary to retell much that is already known to Arlington, and at the same time to assume a fair knowledge of the Fish on the part of the reader.

Characterisation of D

To solve (1) for pseudo plain text D, it is necessary to know the characteristics of D, as resulting from equation (2).

Individual letters of D are random in appearance, since D results from the enciphering of plain text P by the addition of PSI' key, the letters of which are fairly random when distributed singly. However the successive letters of PSI' text are non-random in their pairings, since PSI' is generated from PSI wheels which either all move with probability of "a," or all don't move with a probability of "l-a." A letter in the PSI' key will be repeated when the Mtotal= . ; or when the Mtotal = x and every one of the five signs of PSI, remains the same. Thus a double letter in the PSI' text has a probability of occurrence of (1-a) + a(l-b)^5. Similarly, a letter which is the exact inverse will follow a PSI' character with a probability of ab^5.

In teletype modulo-two arithmetic, if a letter is added to itself the result to always the teletype character "/" and if it is added to its inverse the result is always an "8". Thus by taking each letter of the PSI' text and adding it to a letter equal to the letter following it, this resultant text, called Delta-PSI' text, shows a slant for every double letter. Thus, / has a probability of occurrence of (1 - a) + a(l - b)^5. Similarly, 8 has a probability of occurrence in Delta-PSI' of ab^5. The probabilities of Delta-PSI' text characters are as shown in the following table:

Delta PSI' text:	Probability:
/	(1 - a) + a(l - b)^5
ET934	a(1 - b)^4 b
OARSD MILXZ	a(l - b}^3 b^2
BCFGJ MPUWY	a(1 - b)^2 b^3
5KQVX	a(l - b)b^4
8	ab^5
--------	----------

and as we are now dealing from necessity with delta PSI prime text, we must apply the delta process to our original parametric equations, expressing them now in the form:

(i)Delta-Z + Delta-X = Delta-D

(2)Delta-D + Delta-PSI' = Delta-P

so that Delta-D is now the important parameter to be solved "for" in equation(1) and solved "from" is equation(2).

We said that Delta-D was the sum of Delta-PSI prime and Delta plain, and gave the characteristics of Delta-PSI prime. The characteristics of Delta plain text, on the June Jellyfish circuit, were much as follows,

A	767	I	420	Q	498	Y	712
B	174	J	1049	R	577	Z	621
C	549	K	505	S	695	/	1156
D	594	L	490	T	432	3	1131
E	612	M	786	U	1472	4	479
F	991	N	451	V	380	5	2698
G	755	0	1001	W	504	8	1168
H	542	P	665	X	576	9	566

(Total 25,600 letters of plain text)

and we can see therefore that while D text was random in appearance,

Delta-D text will not be random, being the sum of Delta-PSI'(with definite characteristics)

and Delta-P (with definite characteristics).

Characteristics of Delta-D

Thus a set of Delta-D characteristics for late April and early May Jellyfish Berlin (20 motor dots) is as follows, based on 16 messages"normalised" to 3,200 letters length:

/	128	R .x.x.	92	A xx...	96	D x..x.	89
9 ..x..	110	C .xxx.	90	U xxx..	124	F x.xx.	100
H ..x.x	102	V .xxxx	94	Q xxx.x	101	X x.xxx	87
T x	99	G .x.xx	100	W xx..X	89	B x..xx	82
O ...xx	104	L .x..x	92	5 xx.xx	143	Z x...x	89
M ..xxx	100	P .xx.x	96	8 xxxxx	112	Y x.x.x	97
N ..xx.	100	I .xx..	96	K xxxx.	89	S x.x..	104
3 ...x.	113	4 .x...	90	J xx.x.	103	E x....	89

Total 3,200

I might add that the Delta-D characteristic frequencies are so important to the successful prosecution of the Fish problem, that practically every man, every Wren, in the Newmanry and in the Testery, knows them by heart, together with the variations to be expected on the different circuits and among the different correspondents. The solution of equation (1) is not considered complete unless the results match the expected Delta-D characteristics; and Maj. Tester's section utilises these to the fullest in the art of converting Delta-D into plain text.

From the above table of deltaD characteristics, the following important facts about the various impulses of Delta-D result:

Delta-D Impulse Combinations: Score:

1.2.	856
1x2x	857
(1p2).	1,713
1x2.	737
1.2x	750
(1p2)x	1,487
(4p5).	1,659
(1p5).	1,619
(1p4).	1,618
(2p5).	1,660
(2p4).	1,641
(3p4)x	1,656
(1p3).	1,632
3xlx2.	388
1.2.4.	439
1x2x4x	447
1=2=4	886
1.2.5.	451
1x2x5x	445
1=2=5	896
1.2.4.5.	238
1x2x4x5x	255
1=2=4=5	493
1x2x(3p4)x	471
(4p5).1x2x	475
(4p5).1.2.	442
(4p5).(1p2).	917
2x	1607
(3p4)x2x	855
1.2.(3p4)x	429
(1p2).(3p4)x	900
2.5.(1p4).	427
2.(1p4).	797
1.2.3.	444
2.(1p3).	832
(1p2p3).	1,639
(4p5).(3p4)x	863
1x2x3x	426
1=2=3	870

and from this very important table, we may estimate the following:

Delta-D Impulse Combinations :*	Expected Sigmage	Estimated Stand.Dev. of Sigmage	Actual Stand.Dev. of Sigmage
(1p2)./	4.0	+- 2.0	+- 1.6
4=5/=1=2	3.6	+- 2.2	+- 2.3
4x5x/1x2x	3.3	+- 2.0	
(4p5)./1x2x	3.2	+- l.6	+- 1.6
(3p4)x/1x2x	2.9	+- 2.2	+- 1.4
(4p5)./(1p2).	2.9	+- l.9	+- l.2
(3p4)x/2x	2.6	+- 1.9	
(5p4).(3p4)x usually run as (5p/4). (3p/4)x simultaneously	2.6	+- 1.6	+- 1.4
(2p5)./	2.1	+- 1.9	+- 1.7
(4p5)./	2.1	+- 2.0	+- 1.3
(3p4)x/(1p2).	2.1	+- 1.9	+- l.5
5./2.(1p4).	2.0	+- 1.7	+- 2.0
(3p4)x/	2.0	+- 1.7	+- 1.8
5=/1=2	1.9	+- 2.2	+- 1.8
4.5./1.2.	1.9	+- 1.9	
(3p/1).2.	1.8	+- 1.8	+- l.7
5./1.2.	1.6	+- 2.2	
(2p4)./	1.4	+- 1.9	+- 1.6
3x/1x2.	1.4	+- 1.4	+- l.0
4=/1=2	1.4	+- 2.2	+- l.6
(3p/1p2).	1.4	+- 1.9	+- l.7
4x/1x2x	1.3	+- 2.2	
(1p3)./	1.1	+- 1.7	+- 1.0
5x/1x2x	1.1	+- 2.2	
3x/1x2x	1.1	+- 1.6	+- 1.2
(4p5)./1.2.	1.0	+- 2.1	+- l.1
4./1.2.	0.8	+- 2.1	
(1p5)./	0.7	+- l.8	+- 1.5
3=/1=2	0.7	+- 2.0	+- 1.4
(1p4)./	0.6	+- l.9	+- 0.85
(3p4)x/1.2.	0.1	+- 1.4	+- 1.1
3x/1x2x	-0.2	+- 1.7	+- 1.3

and also for the 3rd impulse the following:

3x/1x4.	2.2	+- l.6	
(3p/2)x1x4x	2.1	+- 2.2	+- 1.5
3./1.2.5.	1.4	+- 2.6	
3./1.2.4.5.	1.2	+- 3.4	

------- - -- ------------ -----

*The slants have been placed in to show what "runs" are usually made on Colossus. Thus 4=5/=1=2 means that given 1 and 2, this is a good way to find 4 and 5 simultaneously. But 4=/5=1=2 and 5=/4=1=2 are just as plausible to run under proper circumstances.

In the preceding table, "Expected sigmage" is the excess of the average observed causal score over random score, divided by random sigma. The variance of such sigmage is assumed to be the sum of the variances of the sigmages of the letters involved. (These individual variances are in turn the mean square deviation-from-average-sigmage in the 16 messages.) The "estimated standard deviation" is the square root of the variance so computed. "Actual standard deviation of sigmage" is computed from sigmages recorded in runs.

It might almost be said that the whole story of the Newmanry is told in the preceding table, and for this reason such a table deserves close scrutiny. Different traffic networks would or course have different tables, and while such have not been computed for all cases where they are needed, they exist in the minds and experience of the Colossus operators.

A thorough knowledge of Delta-D characteristics is needed either for setting known X wheels or for breaking unknown wheels.

Setting known X's is of course easier. In an ideal world the basis for setting them would be to "run" all positions or Delta(X1, X2, X3, X4, X5) against Delta-Z text; and to match the resultant Delta-D text against the theoretical distribution. Where the match would be greatest, there would be the most probable answer. Since 41x31x29x26x23 settings would require a run of the Z tape 22,041,684 times through Colossus, this is obviously impossible on today's machinery. Compromises must be found. Accordingly the X's are set in smaller combinations rather than in toto; and this is the reason the table of Delta-D impulse combinations is so important. Any run involving the setting of two X wheels at once is called a "long run;" any involving the setting of only one X wheel is called a "short run". The average long run takes 8 minutes* as against 2 minutes for a short run.

*Utilising five counters in "multiple testing" to cut the time in one fifth.

From the table of Delta-D impulse characteristics, it is easy to see that (1p2)= . is an admirable characteristic to "shoot for" in wheel setting runs, since the average count of (lp2) in a 3,200 letter message set correctly on x1x2 is 113 dots more than the number of dots expected by random. This is 1600 + 4((1/4)*3200)^1/2 or "four sigma" above, and the odds against this being attained in the wrong place are 1/41 x 1/31 x 33,000 (from the normal distribution tables) or 25:1. In fact, the ease of setting lp2/ runs has been found to be a rough measure of the ease of solving the whole message.

Therefore in a X-wheel setting procedure, a message tape is placed on Colossus; X1, X2, X3, X4, and X5 patterns are plugged up; the machine is set to Delta-ize; and the counters are switched so as to count all the occurrences of (DeltaXl+DeltaZl)+(DeltaX2+DeltaZ2)= . throughout the message, given in turn each of the 1,271 possible initial settings of Xl and X2; and the electromatic typewriter to record all totals which exceed 2.5 sigma above random (which is the conventional "set total" for long runs.)

Below and following is a set of Colossus runs, made in an attempt to set message number SG 908. (These are carbon copies of the electromatic typewriter tape.) The first such run represents lp2/. Colossus #3 was used. Message date was October 22; filing time, 11:28. Total length of message, 2,683. Expected random score for lp2/ was 1341. Sigma of random score was 25.5 Set total was 1,404. Wheel settings are shown for XI and X2 ("K1 K2") together with the "counter letter" and the score. (Five counters are used at once in "multiple testing" to cut the time in one fifth).

COL. 3.

SG 908 22/10 1128 time

T 2685 a l341 £ 25.5 st 1404

1p2
k1 k2

14 12 c l417
18 14 d 1421
28 07 d 1407
27 12 e 1411
31 18 a 1416
40 08 b 1405
01 04 a 1465 ch 4.9 £ (tick)
01 19 a 1406
04 12 c 1409
14 12 e l417

In this 1p2/ run, the setting Xl = 0l and Xl = 04 gave a sigmage of 4.9. Thus the odds against getting the setting by random are approx. 1/41 x 1/3l x 2,000,000, or 150:1. It was therefore deemed correct. The "ch" beside the score means that Colossus was set at X1 = 01, X2 = 04, and the score was quickly checked.

The next series of runs usually made in wheel setting, includes runs called:

"C-1," wherein count is made of the number of
times that 4 = / 1 = 2
In the deltaD texts.
"C-2" wherein 5 =/1 = 2
"C-3," wherein 4 = 5 = / l = 2
"C-4," wherein 3 = /l=2.

The reasons for these can be seen in the table of deltaD impulse characteristics

Accordingly, the next run in the above message was a "multiple test" on wheels X3, X4, and X5, run independently though simultaneously against the known Delta-D1 and Delta-D2. It is shown below, marked "Cl 2 & 4."

```
set k1  01 k2 04

01.2 & 4.

r 1465 a 732 £ 19.1 st 751

k3 k4 k5
   02        b 0761
      03     c 0785 <- likely for K5
05            a 0751
07            a 0796 ch 3.3 £  <- for K3
   07         b 0752
      07     a 0757
   13         b 0756
      13     c 0767
      14     c 0758
18            a 0759
   19         b 0761
      19     c 0769
      20     c 0766
22            a 0754
23            a 0769
   24         b 0752
      03     c 0785
28            a 0755
   02         b 0761
      07     c 0757
   07         b 0753
05            a 0751
```

Total positions looked at = 1465. Expected random score = 732. One sigma equals a count of 19.1. Set total for short runs (equals 1 sigma up = 751.

Nothing valuable on the above run showed for X4, and only 3.3 sigmage for X3 at 07.

Assuming X3 to be correct (odds = 1/29 x 1,400 = 48:1 if X1X2 were correct) the next try was to obtain X4 from (4p/3}x1x2x and this run is shown below, together with a C-3 run.

```
set k3  07

4p/3x,1x2x
r  730  a 369 £ 13.5 st 382

k4
02   a 0382
11   a 0383
13   a 0389
15   a 0389
19   a 0386              (nothing)
20   a 0389
24   a 0384
11   a 0383
13   a 0389
15   a 0389

03  (4=5=/1=2)
r  1465  a 366 £ 16.5 st 406

k4 k5
01 01 a 0386
16 03 a 0423  3.6 sigma  (this agrees for K5
16 20 a 0415             (with the C2 run
```

A new attack was next tried to set X3 properly, as below:

```
3x/1x2

r  566  a  283  £ 11.8 st 295

k3
07 a 0306
08 a 0296
13 a 0298
21 a 0326  -  3.8 £ ch  (this contradicts
24 a 0302                (the C-4 run, and
05 a 0295                (is temporarily accepted
07 a 0305  -  low
```

With this assumption for X3, the Colossus operator attempted to get more data on X4 and X5 as shown:

```
set k3  at 21

4p/3x,1x2x

st 382
k4
03  a 0383
06  a 0389
10  a 0391
14  a 0387                (hopeless, this makes
18  a 0392                (the C-3 run doubtful
22  a 0382
03  a 0383

5./1.2.

r  725 a 362  £ 13.5 st  375

k5
01  a 0385                (to check 5, but
03  a 0387                (not availing
07  a 0392
13  a 0390
19  a 0378
01  a 0385
03  a 0387
```

These runs being unsatisfactory, the run (4p5)./1X2X was made:

```
4p5/1x2x

r 759 a 369 £ 13.5 st 402
k4 k5
04 03 c 0407              (run in case operator
03 12 d 0405              (used less doubter
05 17 b 0404
09 21 c 0402
19 03 c 0408  2.8 £       (not too sure. confirms
23 17 d 0404              (the k5 .. vaguely,
03 17 e 0404              (and suggest k4 at 19
04 03 c 0407
03 01 d 0405
```

5./1.2.3.

r 350 a 175 £ 9.3 st 185

k5
02 a 0185
03 a 0195 2.2 £
07 a 0185
08 a 0187
13 a 0189
14 a 0186
15 a 0186
19 a 0185
20 a 0185
02 a 0185

try k5 at 03

Margin note:

Another attempt [in place of 5./1.2.] to set K5.

We now accept K5 at 03 because same sense as in C2 and elsewhere

Another attempt at X5 is shown here. Odds are 5:i in favour.

Accepting this setting for X5, and with the settings for X1, X2, X3, already obtained, the operator tried to set X4 by the highest count of 9, U and 5; then by the highest count of J and 0; then using /; then using G and Z.

Only the / run proved effective.

He next assumed the five wheels were set correctly and made the 32 letter count shown on the next page.

999uuu555
r 555 a 277 £ 1l.7 st 288
k4
02 a 0294
04 a 0294
13 a 0290
19 a 0296
23 a 0294
02 a 0294

Now running for K4 with K3 accepted from 3x/1x2

jjjj0000

r 344 a 172 £ 9.0 st 181
k4
03 a 0198
07 a 0190
14 a 0184
18 a 0181
03 a 0198
07 a 0190

//////

r 195 a 97 £ 6.9 st 104
k4
19 a 0122
25 a 0106
01 a 0104
04 a 0105
05 a 0107
10 a 0104
11 a 0l06
19 a 0122 <- 3.35 assume K4 set at 19

ggggzzzzz

r 282 a 141 £ 8.3 st 149
k4
06 a 0164

```
try k4 at 19

settings  01 04 21 19 03
32 lec                          (32 letter count)

/  0122
9  0102
h  0080
t  0079

o  0076  - too low
m  0103  - too high
n  0090  - too high
3  0073  - too low

r  0091
c  0076
v  0097  ) wrong way round
g  0092  )

l  0088  ) wrong way round
p  0077  )
i  0069
4  0062

a  0110  ) very much wrong way round
u  0076  )
q  0083     ) wrong way round
w  0093     )

5  0117
8  0105
k  0077
j  0078

d  0047  )
f  0073  )
x  0075  )
b  0058  )
         )  Fairly good
z  0044  )
y  0080  )
s  0098  )
e  0092  )
```

The operator looked at the above count, shook his head sadly, and made the conment which were copied down beside the letters.

It should be of interest here to note that the 32 letter count is usually the acid test. It is used in the light of experience.

From the preceding 32 letter count it was judged that X3 was possibly wrongly set, and therefore X4 since it was based on X3; runs were made at an alternate possibility for X3 on X4 to test this theory with success as shown below:

```
try K3 at 07              (See the C-4 run)

runs for k4

9999uuuu5555
r 567 a 283 £ 11.8 st 294
k4
19  a 0304   -   close runner up
20  a 0300
24  a 0302
02  a 0300
04  a 0297
k4
13  a 0308    -   not too significant
16  a 0296
19  a 0304
20  a 0300
24  a 0301
02  a 0299
04  a 0296
O8  a 0296
13  a 03o7   -  highest

///////

r 227 a 113 £ 7.4 st 120

k4
19  a 0151
01  a 0125
04  a 0123
05  a 0125
10  a 0125
15  a 0123
19  a 0151   ch 5 £    <-  this is it! K4 is set at 19
                              with K3 at 07

set k4 at 19
```

A new 32 letter count was needed at this stage, and here it is:

```
32 lec

settings 01 04 07 19 03

/  0151  )
9  0073  )
h  0083  )  ok
t  0076  )

o  0093
m  0086
n  0087
3  0076  - too low

r  0073
c  0094  )
v  0093  ) somewhat too high
g  0096

l  0080
p  0085
i  0067
4  0064

a  0065
u  0121
q  0083
w  0093 - still too high, but not bad

5  0110
8  0112  )
k  0083  ) not too high since strokes are high
j  0072

d  0060  )  too high
f  0060
x  0069
b  0064  )  too high

z  0060
y  0064
s  0113  )  too high
e  0077
```

This count looks much better, but seemed dubious in spots.

The operator tried to improve X3 on the following runs just to make certain,

but was unable to improve X3:

```
                    runs for k3        (another attempt to see
                    333333             (if could improve K3
                                       (counting 3
                    r 163 a 81 £ 6.3 st 87
                    k3
                    10  a 0088
                    22  a 0090
                    24  a 0088        ( nothing
                    03  a 0090

                    jjjjjjjx
                    ffffffff

                    r 119 a 59 £ 5.4 st 64
                                       ( counting F
                    k3
                    09  a 0064
                    12  a 0068
                    20  a 0070
                    21  a 0073         ( nothing
                    24  a 0068
                    25  a 0064
                    05  a 0067
                    09  a 0064

                    xxxxxxxx
                    r 135 a 66 £ 5.6 st 72
                                       ( counting X
                    k3
                    13  a 0082
                    21  a 0073
                    27  a 0074
                    29  a 0078         ( nothing
                    05  a 0073



                    settings    01 04 07 19 03

                    inevitable despite counts for E, D, S

                                       ( initials G T written on  )

                    Time - Almost 2 hours
```

Accordingly the settings 01, 04, 07, 19, and 03 were deemed correct, and the message, with the last 32-letter count, was sent to Testery.

(Normally, wheel setting is not allowed to go on longer than one hour per message. The message illustrated had such poor Delta-D characteristics that it had to be abandoned later in the Testery by the PSI "breakers," as I learned after this report was written.)

Immediately below I have enclosed another wheel setting run on another message, which "broke" in 20 minutes. See next 2 pages also. It should give something for those interested, to work out by themselves.

SB3115 23/10 col-3

t 9567 a 4788 £ 49 st 4912

1p2/

k1 k2
06 1l a 4921
02 12 e 4922
06 13 a 4948
06 15 a 4920
02 16 e 4977
06 17 a 4926
05 18 b 4926
02 18 e 4988
02 20 e 4954
06 21 a 5005 <- 4.4 sigma
05 22 b 4914
02 24 e 4919
03 25 d 4925
02 26 e 5015 <- 4.6 sigma
07 20 e 4926
19 26 c 4928
26 12 a 4921
25 19 b 4930
25 21 b 5038 <- 5.1 sigma
25 29 b 4929
29 18 c 4946
36 11 a 4964
36 13 a 5055
36 15 a 4995
36 17 a 4993
35 18 b 4926
36 19 a 5047
36 21 a 5384 <- 12.2 sigma ch k1 k2 !!
36 23 a 4975
36 25 a 4965
36 27 a 4989
36 29 a 5013
36 31 a 4964
38 08 d 4933
37 16 e 4937
37 22 e 4918
38 08 e 4933

```
set k1 36 k2 21

c1,2 & 4

r 5384 a 2692 £ 36.5 st 2728

k3 k4 k5
01          a 2938  <- ch.  6.8 rho ! k3 !
   01       b 2763
      01    c 2803
   02       b 2733
   04       b 2782
      04    c 3003  <- ch.  8.6 rho ! k5 !
05          a 2829
   05       b 2783
06          a 2740
   07       b 2802
      07    c 2750
   08       b 2774
   09       c 2811
10          a 2769
11          a 2751
   11       b 2742
      12    c 2759
   13       b 2823
      l4    c 2733
15          a 2750
16          a 2743
   16       b 2826
   19       b 3093  <- ch.  11.1 rho ! k4 !
      19    c 2759
20          a 2785
   21       b 2733
   22       b 2823
      22    c 2779
24          a 2740
      01    c 2803
   25       b 2796
26          a 2744
   01       b 2763
      04    c 3003
   02       b 2733
29          a 2776
01          a 2938
   04       b 2782
      07    c 2750


k3 01 k4 19 k5 04
```

Letter count on 36,21,01,19,04

0599 /
0324 9
0305 h
0265 t

0321 o
0255 m
0230 n
0318 3

0229 r
0229 c
0291 v
0299 g

0266 l
0311 p
0264 i
0225 4

0244 a
0379 u
0320 q
0224 w

0513 5
0475 8
0261 k
0351 j

0204 d
0297 f
0291 x
0232 b

0232 z
0290 y
0295 s
0236 e

All certain (G.T.) (initials of Colossus operator)

The run lp2/ does not always prove out so nicely. For instance the following messages:

Message #		Motor dots
GDBA	37	28
GDBY	273	18
GDBY	291	18
GDBY	277	18
JPY	23	28
JPA	254	26
JBA	98	27
JBY	192	28
JBY	191	28
JBY	188	28
JBY	176	25

could not be set with 1p2/ but had to be set with (3p4)x/ instead. The Delta-D totals (normalized to 3,200 and analyzed) showed the following counts: lp2: 1684 (as against 1600 random); 3p4x: 1733 as against 1600 random.

There is no rigid rule or procedure in wheel setting; each operator has his pet methods; the closest parallel to it in Arlington Hall to my knowledge is the manner in which we used to salve J-19. Page 19 of this report, listing runs of 4.4 sigmage, 4.6 sigmage, 5.1 sigmage, and 12.2 sigmage, as competitors, shows that actual sigmage is not in itself a measure or certainty of the setting, due in part to the self-match property of the X patterns. The important consideration is actually "how high is the score itself as against all the other ones?"

The next page shows some data gleaned from recent runs. All the messages involved were very favourable. Sigmages are not usually this high.

Wheel Setting Sigmage Data
(NOT normalised to length 3,200.)

Messages:	d	Length	1p2/	3p4x/	4p5/ 1p2	3p4x /1p2	(C1) 4=/5=2	(C2) 5=/1=2	3x /1x2	3/123	p/1.2
BR 7747	27	10,168	8.8	8.4	3.6	12.2	3.9	-.3	6.3	2.4	6.2
BR 2468	23	4,000	10.7	5.5	13.4	8.1	4.5	6.4	4.8	2.2	4.8
GDB 4116	26	10,172	10.7	7.0	9.6	1l.3	.5	2.4	4.9	.3	3.7
" " notx2=x	26	4,922	5.7	7.5	11.1	11.4	1.7	1.7	5.5	4.2	6.5
GDZ 8031 II	21	6,984	7.1	4.0	6.5	6.8	4.5	5.2	1.7	0.0	1.1
SBE 130 I	15	10,168	8.0	-.8	10.2	1l.5	5.5	5.1	3.1	6.9	7.2
SBE 389	20	10,172	10.7	3.2	12.7	4.1	3.6	1.6	2.7	3.4	4.4
SBE 427	23	10,168	10.4	.6	10.0	-.5	8.7	7.3	6.3	6.2	9.0
SBE 477	24	10,168	9.2	1.1	8.4	-1.3	7.3	3.4	4.8	6.6	8.3
BDX 805	16	10,172	10.7	-0.0	12.5	-6.8	6.5	8.4	3.8	7.5	8.4
BDX 1010	28	9,146	13.1	0.3	16.5	-0.1	10.0	9.6	6.0	7.8	9.9
GKB 3118	27	9,116	13.0	4.6	10.9	2.9	5.7	5.4	4.0	3.0	4.9
" " notX2=X		4,705	8.8	2.2	9.7	1.8	3.7	-	2.3	1.4	2.6
WB 1626	27	9,643	18.3	-7.2	20.5	-6.4	10.5	5.8	4.3	8.6	9.3
WB 1743	18	10,168	14.8	0.3	18.1	-0.3	9.4	9.9	3.1	8.2	8.4
WB 1847	28	10,172	14.4	-1.2	20.4	-3.7	12.4	10.7	5.6	7.4	9.4
WB 2003	25	9,066	16.4	2.3	19.7	12.4	5.2	7.3	3.0	8.0	8.2
CZ 1132	25	10,172	10.1	9.1	9.7	10.4	2.8	3.0	6.6	4.2	7.3

(Sigmage shown 1s of course for the correct setting of the current wheel.)

Wheel Breaking

The most interesting solution of Delta-Z + Delta-X = Delta-D occurs when Delta-X is known and Delta-D are unknown (except for general characteristics which depend upon the type of traffic.)

This is called wheel-breaking. It is an attempt, through the study of at least 5,000 characters of deltaZ, to formulate Delta-X and Delta-D simultaneously, thereby obtaining the X wheels as well as the Delta-D text.

We need mathematics in wheel-breaking. Wheel-breaking procedure involves establishing hypotheses of varying probabilities, and building further hypotheses on them. Odds in favour or the final outcome are most important to know. By the theory of probability, these final odds are the product of the prior odds and all the factors derived from the individual pieces of evidence. Because a product is involved, statements of odds and of factors are usually made in logarithms (to base 10) called "bans;" the actual working unit is for convenience a "deciban."

Note the following equations in probability:

$$P(E+A) = P(E)\,P(A/E) = P(A)\,P(E/A)$$
$$P(E+notA) = P(E)\,P(notA/E) = P(notA)\,P(E/notA)$$

If the center and right hand-member of the upper are divided by the center and right hand member of the lower, we obtain:

$$\text{Odds of A, given E} = (\text{Prior odds of A})\ \frac{P(E/A)}{P(E/notA)}$$

and the fraction on the right is the "factor" referred to above.

A "pip" is the factor in favour of a hypothesis, resulting from the occurrence of some standard event.

It is the custom in wheel-breaking to begin with the Delta-Z1 Delta-Z2 rectangle.

This rectangle is usually constructed on the Colossus, or on Garbo. The rectangle is converged crudely by hand, by methods already familiar to Arlington.

"Accurate convergence" is almost never used.

I have already submitted as part of this report the famous "Significance Test IV of the Black file." Patterns derived from the crude convergence of the Delta-Z1, Delta-Z2 rectangle, for Delta-X1, are checked for significance by means of this test. Actually the theta terms of the test are not computed, unless the answers comes out close to the borderline.

It has been found that the theta terms seldom sum to more than 300,000/N.

Therefore the standard for the X1 wheel is taken as:

$$(2.17\ X^2)/N > 219 - 300{,}000/N$$

or in simpler terms:

$$X > 10\ (N-1500)^{1/2}$$

In the above, X=Sum|Xi| wherein Xi are the excesses of dots over crosses in the resultant pattern.

Prom this formula in simplest terms is calculated a value for X given each length N, which must be equalled or exceeded by the Delta-X1 wheel before the rectangle is called significant.

The above formula is currently being "argued out;" some persons favour substituting 1271 for the value 1500 claiming the latter too high; others favour the older rule, X>(80N)^1/2. But the one given is the one now in use, and is referred to in one of the screeds sent by Capt. Fried as "the better rule." When the computors have converged a pattern to this point they then turn the rectangle over to the wheel breakers.

A table showing results obtained with significant rectangles follows. Jellyfish had not "come out" for six weeks, and Gurnard had "come out" well. The table to follow compares the two.

	Text length-	5,000 6,250	6,251 7,500	7,501 10,000	10,001 12,500	12,501 15,000	over 15,000	TOTAL
Jellyfish	# Rectangles	9	20	13	8	4	4	58
	# Significant	0	2	0	0	1	0	3
	# giving wheels out	0	0	0	0	0	0	0

	Text length-	5,000 6,250	6,251 7,500	7,501 10,000	10,001 12,500	12,501 15,000	over 15,000	TOTAL
Gurnard	# Rectangles	0	1	5	23	16	30	85
	# Significant	0	0	1	7	5	4	17
	# giving wheels out	0	0	1	3	0	4	8

Sgt. J. Levine at Arlington has gone into the accurate scoring of rectangles as far as anyone here, and all data compiled here on accurate scoring are already in Arlington's hands. Here they do not find accurate scoring in converging to be worthwhile on a production basis.

After the Delta-X1 & Delta-X2 patterns are deemed to be significant(and therefore the rectangle is called "significant,") the rectangle and the patterns for Delta-X1 & Delta-X2 are sent to the Colossus operator, together with the Z tape. The Colossus operator plugs the Delta-X1 & Delta-X2 patterns on the machine. Then he performs wheel-breaking runs.

For instance, using the deltaX1 pattern only, as against a Delta-X2 pattern of all dots, he makes a count of resultant Delta-D 1p2 = . positions. This is called the "Normal Score" ("NM"?) Then the last dot of the Delta-X2 pattern is changed to a cross, and another count is made on 1p2. positions. The last position of Delta-X2 is changed back to a dot and the second to last position of Delta-X2 pattern is changed to a cross, and a count made on 1p2. position. This is done for a cross in every position of Delta-X2.

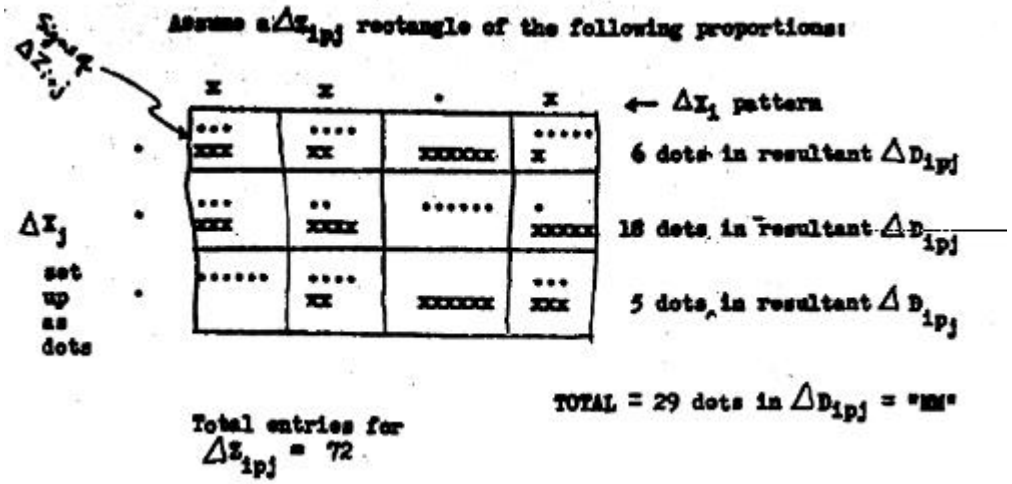
The excesses of these resultant scores over the normal score are exactly but opposite in sign
equivalent / to the excesses of dots over crosses that would be obtained

for the Delta-X2 pattern in a crude convergence of the given rectangle by

hand, with the given Delta-X1 pattern. A quick way of checking this

is as follows:

Assume a Delta-X ipj rectangle of the following proportions:



If we change the last sign of Delta-Xj to a cross we get:

6 dots in Delta-D ipj

18 dots in Delta-D ipj

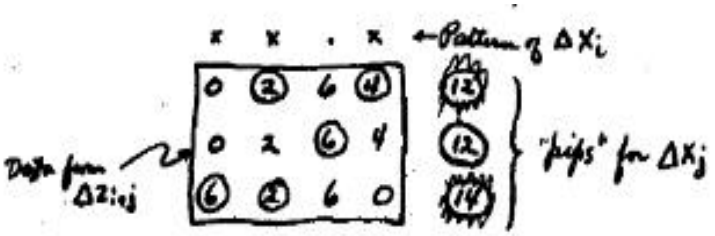
19 dots in Delta-D ipj

43 = TOTAL

with regard to sign of Delta-Xi
43 - 29 = 14 = value for excess of dots/ in last row opposite to what would

be given by crude convergence had we written the Delta-Z ipj rectangle as follows in

terms of excess of dots over crosses:



**Resultant Delta-D i+j crosses are not counted, nor is excess of dots considered, since Colossus is set to count only dots.

Thus a wheel-breaking run on Colossus is merely a means of converging into a pattern from any given data, by crude convergence, and using Colossus as a rapid adding machine.

It is necessary at this time to describe the decibanning of a wheel-breaking run.

Let: r = total number of values of Delta-Z ipj looked at in converging one given character of the unknown wheel.

r/2 + x/2 = number of dots counted.

r/2 - x/2 = number of crosses counted.

x = excess of dots over crosses.

Therefore "x" is the score given in the wheel-breaking run: the "pippage" for any given character of the unknown wheel.

Now the factor that the wheel was originally a dot in the character giving the score x, are estimated as:

$$\left(\frac{\frac{r}{2}+\frac{x}{2}}{\frac{r}{2}-\frac{x}{2}}\right)^x \quad \text{or} \left(\frac{r+x}{r-x}\right)^x \quad \text{and bans} = x \log_{10} \frac{r+x}{r-x}$$

Therefore (r+x)/(r-x) is the "value of a pip" (numerically) for that particular character of the wheel.

It is impossible to evaluate this for each character each time; so the average value may be considered as an estimate:

$$\frac{R+X}{R-X}$$

wherein R = Sum r1 or the total number of letters looked at in the whole run; end X = Sum |xi| if we assume the best possible score to be the true score.

It is a fact, however, that Sum |x1| is usually higher than Sum (x1 with regard to signs of the true wheel) - and therefore the more nearly accurate value of a pip (numerically) is:

$$\frac{R + qX}{R - qX}$$

wherein "q" is a factor for reducing X by the amout the maximum score may be expected to exceed the true score. This is getting involved and I hate to break up continuity here but it is really the best place to thrash it out.

What value shall be given "q"?

We said that x = Sum|xi|. The expected value of |xi| my be called Sum|xi|/W wherein W = wheel length, and therefore the expected value of |xi| (or E|xi|) = X/W.

Also, qX = correct value = Sum (xiei) (wherein ei = +1 ot -1 depending on the signs of the true wheel.) Let us call Sum (xiei)/W (the expected value of xiei) = x0 and let us call sigma xiei = sigma

Now, since |xiei| is a function of xiei(whose expected value is x0) we can derive the following directly from the normal function:

$$E\,|x_1 e_1| = \sigma\sqrt{2}\left(\frac{1}{\sqrt{\pi}}\,e^{\frac{-x_0^2}{2\sigma^2}} + \frac{x_0}{\sigma}\cdot\frac{1}{\sqrt{\pi}}\int_0^{\frac{x_0}{\sigma}} e^{\frac{-t^2}{2}}\,dt\right)$$

$$\text{But } E|x_1 e_1| = E|x_1| = \frac{X}{W}$$

$$\text{Therefore } \frac{X}{W\sigma} = \sqrt{2}\left(\frac{1}{\sqrt{\pi}}\,e^{\frac{-x_0^2}{2\sigma^2}} + \frac{x_0}{\sigma}\cdot\frac{1}{\sqrt{\pi}}\int_0^{\frac{x_0}{\sigma}} e^{\frac{-t^2}{2}}\,dt\right)$$

and we may graph x0 (which is actully Sigma xiei/W) against X/W Sigma and thereby solve for x0 and obtain Sum(xiei) which is qX.

The British put the formula (R + qX)/(R - qX) in the following form:

$$\text{Numeric value of pip} = \frac{(R/W)^{1/2} + qX/(R*W)^{1/2}}{(R/W)^{1/2} - qX/(R*W)^{1/2}}$$

(assumption here that
(sigma approx= (R/W)^1/2
(so that X/(W*sigma)app=X/(R*W)^1/2
{and we try to get a term
(in X/(R*W)^1/2
(so as to get a value for q

and have arranged the following table to obtain the value qX/(R*W)^1/2

X/(R*W)^1/2	qX/(R*W)^1/2 app = X/sigma
.9	.51
1.0	.72
1.1	.91
1.2	1.05
1.3	1.19
1.4	1.32
1.5	1.44
1.6	l.56
1.7	1.67
1.8	1.78
1.9	1.89
2.0	1.99
2.1	2.09
2.2	2.2

X/(R*W)^1/2 > 2.2, q = 1.

In actual practice, however, the table is not always used. The "crude decibanning" formula:

Numeric value of pip = (R + X)/(R - X)

is often considered of sufficient accuracy; if they feel there are a few characters of the wheel in doubt, they "knock a little off of X".

At this point I should state that the decibannage for a whole wheel is the sum of the decibans of the pips

For a wheel to be accepted as final: 1. It must integrate; 2. The lowest two pips whose signs may be changed and still give a legal wheel, must sum to more than 50 decibans:; 3. Individual characters of the wheels must be at least 20 decibans each, or if lower in spots they must be so marked and Testery notified for its information in solving delta psi prime; 4. The 32-letter count of the final Delta-D text must be satisfactory.

There is also available a brief significance test to use on a single wheel-breaking run:

Let R = total number of positions actually looked at in the run.

W = wheel length.

X = sum of the moduli resulting from the run (=Sum|x1|.)

The expected value of X at a wrong setting can be shown to be = .8(R*W)^1/2 from the normal distribution. And the variance of X at random an be shown to equal R(l - 2/Pi) or approximately .36R. So the standard deviation is approximately equal to .6(R)^1/2, and two standard deviations above expected random = .8(R*W)^1/2 + 1.2(R)^1/2. This is graphed for usable values of R and W. X of an run must exceed this if the run is to be called significant.

With these miscellaneous facts gathered together, we are prepared to study wheel-breaking.

Pages 30-34 inclusive are the actual work-sheets used during the runs shown in Pages 35 to 54.

The first line of the Delta-X1 worksheet (Page 30) and the first line of the Delta-X2 worksheet (Page 31) shows the pips obtained from converging a rectangle made on Garbo. Message involved, SB 3037 part l.(Original rectangle not available.) Other messages used during the run: SB 3037 pt 2; SG 917 Roll I pts 2 and 3.

A "part" is usually 10,168 characters (8x41x31); the last part is what is left.

(Tony Sale Note: pages 30-54 are at present ommitted (some ar A3 size) until a way can be found to put these on the web at a reasonable Kbyte size)

The 14th character of Delta-X3 was considered most likely to be doubtful. To check delta.t, a 32-letter count was made on message SB 3037 part 1, considering the doubtful character to be a dot; and then another considering it to be a cross. (I.E., the regular wheel-breaking type of run.) The result was to indicate character 14 as "20. db. down," that is a "cross," whereas in the same message earlier, run 10, this character 14 had been shown to be "3 db.up." The operator therefore removed the three db. up and added in the 20 db. down, or in effect added in 23 db. down, resulting in the final score of 1l down as shown clearly On the Delta-X3 worksheet.

The final "good" 32-letter count used in the above determination was then sent to Testery along with the dechi, as per custom. Therefore no exhibit is available to show the final step in clearing up Delta-X3.

Probably the reader has "skipped through" the wheel-breaking exhibits. Yet they are very important, because they demonstrate almost all the practical elements of decibanning, and answer so many questions that exist in the cryptanalysts' minds at Arlington. Take for example the problem of combining patterns from different rectangles. It isn't done here. They take pattern from one rectangle, strengthen the wheels, and then set them in other messages. If they can't seem to get anywhere, they use another rectangle and eventually try its wheels back in the first message. The effect is almost the same. They have found out by experience that several rectangle patterns can't be slid and merged anyway unless at least one rectangle is significant; and if it is significant the Colossus methods used make it unnecessary to add all the other ones to it at the beginning.

Thus many problems at Arlington do not exist here in the same form, and if Arlington wants them answered it must develop the desired experience.

How accurate is the pippage-deciban procedure in rectangling? To check this, twenty rectangles (18 Jellyfish, 2 Bream) of texts 4,800 to 24,531 letters long, were analysed.

Patterns were crudely decibanned (i.e., each pip evaluated at 10 log10 (N+X)/(N-X) dbs., approx 8.7X/N dbs.) (which they have called "first approximation decibans in their tables) and the rectangles were arbitrarily divided into 4 classes:

Bad: X^2 /N < = 70.

Poor: 70 < X^2 /N < 80.

Moderate: 80 < = X^2 /N < = 90.

Good: 90 < X^2 /N

The tables that follow were then compiled from the actual results.

Delta-X1

	Classes of "first approximation dbs." into which pips fall:											
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	>=50	TOTAL:
BAD # of pips:	76	80	79	62	42	18	9	-	1	2	-	369
RECT- %right:	.51	.51	.66	.55	.84	.83	.89	-	1.0	1.0	-	.61
ANGLES (9 messages)												
POOR # of pips:	11	25	18	29	13	16	8	2	1	-	-	123
RECT- % right:	.46	.76	.61	.69	.85	.81	.86	1.0	1.0	-	-	.75
ANGLES (3 messages)												
MODERATE#of pips:	21	21	24	26	25	20	14	8	3	2	-	164
RECT- % right:	.62	.67	.71	.90	1.0	.95	1.0	1.0	1.0	1.0	-	.85
ANGLES (4 messages)												
GOOD # of pips:	8	1l	13	10	15	19	17	14	10	12	35	164
RECT- % right:	.50	.64	.78	.80	.94	1.0	1.0	1.0	1.0	1.0	1.0	.92
ANGLES (4 message.)												
TOTAL # pip.:	116	137	134	127	95	73	48	24	15	16	35	820
(20 msgs)%right:	.53	.59	.67	.68	.90	.91	.96	1.0	1.0	1.0	1.0	.74

Delta-X2

	Classes of "first approximation dbs." into which pips fall:											
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	>=50	TOTAL
BAD # of pips:	34	52	57	44	33	28	15	8	3	1	4	279
RECT- % right:	.50	.50	.53	.75	.70	.60	.60	.88	1.0	1.0	1.0	.61
ANGLES												
POOR #of pips:	6	7	16	18	12	6	12	7	4	2	3	93
RECT- % right:	.50	.71	.75	.72	.75	.67	.75	.57	.75	1.0	1.0	.72
ANGLES												
MODERATE#of pips:	6	8	15	20	18	15	15	10	8	4	5	124
RECT- % right:	.17	.50	.53	.55	.83	.93	.93	1.0	1.0	1.0	1.0	.75
ANGLES												
GOOD # of pipe:	1	4	5	7	9	10	9	17	10	7	45	124
RECT- % right:	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.00
ANGLES												
TOTAL # or pipe:	47	71	93	89	72	59	51	42	25	14	57	620
% right:	.47	.54	.59	.73	.80	.76	.90	.90	.96	1.0	1.0	.74

*("%" is to be read as: "proportion of")

The problems of wheel-setting and wheel-breaking have been presented without regard to motor limitations, because in general such limitations are ignored. It is a fact however (see Fried #96) that most of today's circuits use limitations of some form or other. If the limitation is simply X2 it can be used to advantage in the "stickier" Colossus runs: for instances in wheel-setting, operators may "look at" only those positions wherein X2oneback = x, and thereby gain approximately 20% in expected causal sigmage (more or less depending upon motor dots) the bulge occurs mostly on X2oneback = x positions. This fast is also of value in "wheel-breaking" the Delta-X2 patterns. High pippages in the Delta-X2 wheel can be expected at points wherein X2oneback = x and low pippages at X2oneback = . positions. Thus a preliminary pattern can often be completed by inspection of the worksheet, utilising this knowledge as a guide to marking in the X2oneback = x positions, and working this back and forth with Delta-X2 values already in.

Motor and Psi Runs

Last spring when Chi and Psi pattern remained permanent throughout each calendar month, the setting of Psi's might possibly have become a bottleneck. At that time the Newmanry used to make Motor and PSI runs to help the Testery. But this is done only about twice a week today, because the pressure to get out Chi's daily is too great.

However P5twoback limitations is beginning to disappear from the networks. The desirability of resuming motor and psi runs on a larger scale has therefore increased.

The principle of motor runs is simple. Bulges in the Delta-D result

mostly from Mb = . positions (ignoring limitation effects.) We can find the positions of Mb which give the greatest bulges opposite Mb = . positions, end these are the M settings. With Ms set, the Delta-PSI' job is easier.

Below is a sample motor run.

```
32 L.C.                  SBU  8  24/9            WD  24/9
                         ==============          COL 1
/  0214                  T 3652
9  0098                  R 214  A 133 £ 7.3  ST 156
h  0105                  E.S. 174
t  0101                  K1 16 K2 25 K3 29 K4 24 K5 01
o  0109                  M1  M2
m  0100
n  0099                  53  06  0160
3  0131                  53  07  0158
r  0091                  50  19  0157
c  0090                  54   1  0156
v  0106                  50   1  0156
g  0125                  51  07  0163
l  0111                  52  12  0156
p  0109                  12  03  0159
i  0105                  13  02  0161
4  0098                  54  07  0173 CH  4.8£
a  0088                  57  25  0159
u  0156                  39  37  O156
q  0119                  39  24  0157
w  0092                  50   1  0156
5  0172                  52  06  0161
8  0172                  55  25  0156
k  0127                  60  07  0158
j  0108                  13  02  0161
d  0100                  26  23  0156
f  0102                  54   l  0156
x  0106                  53  25  0160
b  0111                  09  04  0159
z  0086                  36  13  0157
y  0099                  50  08  0157
s  0117                  12  03  0159
e  0104                  53  06  0160
                         53  07  0158
Total    3651
```

SETTINGS M37 07 M61 54.

J.B.M.

Certain M R S

In the motor runs shown, "slants" were counted because they showed the greatest bulge from the 32-letter count.

Total letters in message: 3,652. Total slants, 214. Average score expected at wrong settings, 138. Sigma, 7.3 Set total, 156. Expected score right position, 174.

This last deserves explanation. If there are "d" dots in M37 then "a'" = (37-d)/37 and total number of Mb = x positions in a message "T" long = T(37-d)/37. Average score for any given character run opposite Mb = x positions (assuming Delta-D characters to be perfectly random at such positions) = (T/32)*(37-d)/37 . Assume that the character being run has a total of "r" as shown in the 32 letter count. Since r = sum of counts at Mb = x and Mb = . position (at right settings), we have Mb = . position count should equal r - ((T/32)*(37-d)/37), provided the motor is set correctly. If the motor is set incorrectly, Mb = . positions should give a score of r*(d)/37.

Thus, in the example given, if d = 24, then a' = 13/37, r*d/37 = 138, and r - (T/32 * 13/37) = 174.

Since average wrong score = r*d/37, Sigma = (r(d/37)*37/(37-d))^1/2 and excess of right over wrong = (r - T/32)((37-d)/37)

and sigmage = (r - T/32)(37-d)^1/2 /(r*d)^1/2 .

Psi runs give terrifically high bulges. Motor runs are done first; and then the chi's and motor are both known. The undeltaed Z text is then run on Colossus against psi1 and psi2, with the chi's and motor being added in correctly, so that at the right settings of psi1 and psi2 the resultant text is P(1 plus 2). P(1 plus 2) has a great advantage in scoring: out of 2,000 letters a score of 1400 or more can be expected.

Setting PSIs by "Hand"

Setting psi's involves the equation D + PSI' = P, and is therefore a problem belonging to the Testery rather than the Newmanry. It is done by hand just as at Arlington; but they do it much "smoother" here. Experience, and I believe experience alone, has made these capable lads artists at their work. The time usually required to set all 5 psi's varies: given known patterns, and high motor dottage, they can be set in 20 minutes. Adverse conditions can require 8 hours or more. Day in and day out, the average time is 4 hours counting the unbroken ones in that average; one hour is the most usual time.

The Newmanry sends the Testery a dechi, written out on width 31, plus their 32-letter count of Delta-D from Colossus. The Testery writes the X2 pattern across the top line of the dechi page,if a limitation involving PSI'1oneback is known or suspected; or the X2oneback pattern if any other limitation is known or suspected. The 32-letter count is then studied to determine the P-text form of "stop"; if the 32-letter count shows U, A, and 5 high, "5M89" is indicated; If U, O, 5, then "5M98;" but since O Is usually much higher in frequency than A anyway, it must be quite high before "5M98" is indicated; high slants in the Delta-D 32-letter count indicate doubling in the "stop" such as "55M889" or "5M889."

The Testery also studies the 32-letter count to see if, in their judgement, some X wheel could be wrongly set. (If X3 is correct, 3 should be higher than N, U than A, J than K, X than B, G than V, O than M; if X5 is set correctly, / than T, F than X, U than Q.) Testery finds Newman's section fairly accurate, especially on wheel-breaking; but Testery has never forgotten the time a Colossus operator sent them a dechi marked "all certain," which was set on wheels of the wrong date. In fact, X3 wheel setting is quite

Psi one and psi two therefore set with a colossal bang.

When the limitation is a P5twoback limitation, a combined motor and psi run is made. Naturally the best setting in the motor run will result when psi-five is set correctly, and so psi-five and motor can be set simultaneously. Then psi one and two follow as before. The difficulty with P5twoback limitation is of course the tendency for errors due to garbles, to upset the statistics. This has prevented widespread psi setting (by machines) of P5twoback limitation traffic. Whenever attempts are made, however, the "spanning device" is used - this is a device on Colossus that permits the operator to set up a beginning position and an ending position on his "spanning dials," and Colossus will look at the material between these two positions only. Spans for runs described are usually 800 letters in length.

The spanning device was put on Colossus for just these runs. But it has been used constantly ever since for the other runs (chi setting, etc.) and P5twoback runs are seldom made.

Crib runs and key-breaking runs are also done in Newmanry, but they will be discussed later.

likely to give trouble on any date. (The job done by Newmanry today is miraculous; 95% of those marked all certain by Newmanry are "broken on psi's" by Testery.)

As soon as the P text value is decided upon for "stop," the D is looked at for likely positions of this value. Likely positions are those giving repeats in PSI' or anti-repeats in PSI', all being compatible with the limitation if known. For such discovered positions, suspected abbreviations are tried in the vicinity, and the resultant PSI' text is written out. This is condensed where possible, and the crosses, and dots of S1,S2,S3, and S5 jotted dawn on cross-section paper where they are studied to see if actual wheel patterns can be matched. If not, the PSI' pattern is extended each way, with either a repeat or an anti-repeat, and the resultant plain-text studied, new guesses written in, and further PSI' obtained. From this material, multiple possibilities for setting the PSI wheels are considered, and the equation D + PSI'= P worked back and forth between possible PSI' and possible P, until the PSI is set in at least one level. When it is set uniquely on at least one impulse, it is then possible to project forward or backward to another stretch of D where text has been guessed, using the known motor dottage as an aid, and try to establish the distance between the two stretches on the known PSI wheel. Since this is the same for all PSI wheels, the others can be strengthened with the increased data from both plain-text-guesses, and possibly they can also be set.

Since in general the guessed-in cribs should be as close together as possible, the problem is to guess in as much as can be done, even if some is wrong. Thus they are constantly guessing in and testing assumptions, and the busier they keep the sooner the psi's are set. A quiet, unhurried speed is the essential of a good worker, who writes little, and does much mentally.

Powerful aids to psi setting exist. The D worksheet has been marked

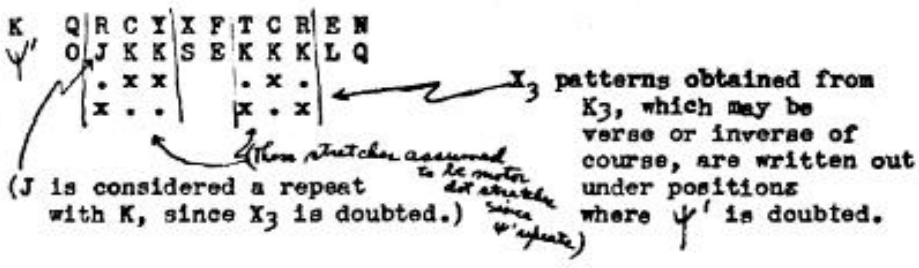
by the "dechi registrar" (a Testery man) with "Hand," "Auto," and "Pause," and all other useful operators' information that can be gleaned from the "red form" (message intercept from Knockholt.) The Testery knows from experience that after each pause or hand transmission, the resumed automatic sending is likely to be a repeat of 30 to 150 letters of plain text (or on rare occasions is often a new message and when not, then it \ states address, call signs, etc.). Such a P text repeat is called a "go-back." If a stretch of P text has been fitted in near the beginning or end of one Auto transmission, then an attempt is made to find a possible go-back area at the end or beginning of another such auto-transmission. To do this, Delta-D texts are written out by hand for about 150 letters from the D texts involved, and slid against each other, counting "clicks." For if both stretches are correctly superimposed, every place there is a motor dot in both texts simultaneously, the resultant Delta-D texts will be the same. These positions can be further checked by any limitation if it exists; since if the Delta-Ds are correctly set, then the P5twoback is the same in each Delta-D, and therefore the X2oneback signs must agree or else one out of the supposed pair of motor dots wasn't a dot.

If accurate scoring is desired, tables exist to show what characteristics the sum of Delta-D(1st transmission) plus Delta-D(2nd transmission) should take on, since this sum is simply the sum of the first PSI'stream and the second PSI' stream, and the PSI's always have definite characteristics depending on motor dottage. I shall give these tables a few pages later on. Go-backs are also an excellent way to detect whether the Xs are off in one impulse; the clicks from superimposing Delta-D texts will be good on only 4 impulses if a X wheel is off, (unless the distance between the go-backs is a multiple of the length of the wheel) and this method determine very easily which X wheel is misset, as well as permitting it to be corrected without obtaining plain text first.

In setting PSIs, quite often some PSI refuses to set, so that the corresponding X wheel setting is suspect. In such a case the PSI' text being worked

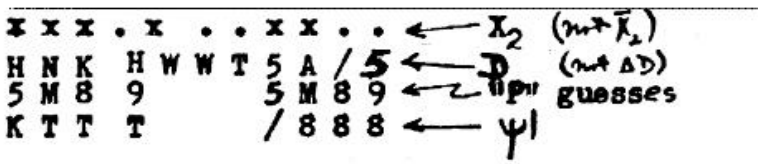
is examined for double letters, and those that would be double or triple or more except for the doubted impulse, are considered as being really doubled or tripled. Then the Z text from the "red form" is written out, the P crib added to it, and a believed-to-be-correct-Key is obtained. (This can also be obtained by adding the Newmanry X to the phoney? PSI':) This is written down, the PSI is written under it, and the Key patterns for the doubted impulse are written below that in the positions where there are double or triple letters in the PSI' text only.

Thus, it X3 is doubted we might have:



The possible combinations of patterns are now matched to the X wheel, finding new possible settings. The deX is modified by the X wheel's new settings, to obtain new PSI's, to re attempt to set PSI.

"Alphabet rods" are also found useful by some PSI setters, though others shun them. In a case like the to flowing:



it is possible (given simple X2 limitation) for there to be an extension of PSI at the position WT in the Z text. If so, W and T will both be enciphered by the same letter. By juxtaposing the W and T rods, and looking down them, all possible digraphs from such encipherment can be seen at a glance, end if none is likely then there was probably a motor

cross. In such case, the distance between the two corresponding stretches of recovered psi is known, and it may be possible from the two stretches together to fit in all the psi wheels.

As soon as all five psi's are set uniquely, the worksheets are sent into the "setters' room"* where they compute the approximate position of the psi's at the beginning of the message, type out a PSI' stream on a Tunny machine, and drag this stream through the beginning of the message until they have established the original starting point of the psi's. Then they decipher about 120 letters of text by anagramming, with which to set the motor wheels.

On page 63 a table for (DeltaPSI')A plus (DeltaPSI')B characteristics was discussed.
The mathematics leading to the tables, and the tables themselves, follow:

It is known that: (Assuming M37=20 dots.)

P(0 crosses in DeltaPSI') = 1-a + a(1-b)^5 = .273 = A0

1/5*P(l crosses in DeltaPSI') = ab(l-b)^4 = .005 = A1

1/10*P(2 crosses in DeltaPSI') = a(b^2)(l-b)^3 = .011 = A2

1/10*P3 = .023 = A3

1/5*P4 = .050 = A4

P5 = .110 = A5

And therefore, since we are after the sum of two delta-psi-prime stream:

---

*The room where the original psi setting had just been found for a stretch of guessed-in plain-text, is called the "breakers' room." This is probably due to the fact that even when psi patterns are known, it is necessary to "break" a stretch of psi prime, before psi wheels can be matched to it. But "setters" have only to set the psi wheels at the beginning of the message, and then set the motor wheels.

P(0 cross in DeltaPSI'A) = A0^2 + 5A1^2 + 10A2^2 + 10A3^2 + 5A4^2 + A5^2
(plus DeltaPSI'B) = .106

1/5 * P(1 cross in DeltaPSI'A) = 2A0A1 + 8A1A2 + 12A2A3 + 8A3A4 + 2A4A5
(plus DeltaPSI'B) = .027

1/10 * P(2 in sum of DeltaPSI's)= .024

1/10 * P(3 in sum of DeltaPSI's)= .026

1/5 * P(4 in sum of DeltaPSI's)= .038

P(5 in sum of DeltaPSI's)= .067

If wrongly overlapped (Delata-D)A + (Delta-D)B can be called flat in its characteristics (and it is assumed so for the tables) then:

(Delta-D)A + (Delta-D)B or (DeltaPSI')A + (DeltaPSI')B Given:	Factor in favour of correct setting of go-back.	D.bans:
0 crosses	.106/(1/32)	5.3
1 cross	.027/(1/32)	-0.7
2 crosses	.0241(1/32)	-1.2
3 crosses	.026/(1/32)	-0.7
4 crosses	.038/(1/32)	0.8
5 crosses	.067/(1/32)	3.4

If a go-back is set properly, it is sometimes desirable to attempt to split the \ Delta psi prime streams back into its original components. sum of the

If its original components are more likely to arise from a pair of crosses than a pair of dots or a dot and a cross, we have gained valuable knowledge, and can write in some of the motor and thereupon make better guesses.

Let us define "Li" as a letter having "i" crosses in it. Then:

If((DeltaPSI')A + (DeltaPSI')B)=	The sum most likely come from:	
L0	L0 L0	70%
	L4 L4	12%
	L5 L5	11%
L1	L4 L5	42%
	L3 L4	35%
	L2 L3	11%
	L1 L2	10%
L2	L0 L2	24%
	L3 L5	21%
	L4 L4	21%
	L3 L3	14%
	L2 L4	14%
L3	L0 L3	48%
	L3 L4	27%
	L2 L3	13%
	L2 L5	9%
L4	L0 L4	76%
	L2 L4	12%
	L3 L3	9%
L5	L0 L5	89%
	L2 L3	7%
	L1 L4	4%

Another useful table assumes that only 4 impulses are satisfactory, the remaining one having been wrongly set on the chi wheel; it also gives the probability of crosses in the sum of the delta psi primes.

It is known that:

Given only 4 impulses:

P(0 crosses in delta psi prime) = 1-a +a(1-b)^4 = .277 = A0

1/4 * P1 = ab(l-b)^3 = .016 = A1

A2 = .034

A3 = .074

A4 = .161

and therefore:

P(0 cross in the delta psi prime sum) = A0^2 + 4A1^2 + 6A2^2 + 4A3^2 + A4^2 = .133

1/4 *P(1 in the sum) = 2(A0A1 + 3A1A2 + 3A2A3 + A3A4) = .051

1/6 *P(2 in the sum) = .050

1/4 *P(3 in the sum) = .064

P(4 in the sum) = .105

(All the above is also based on a motor dottage or 20.) Assuming the sum of (Delta-D)A plus (Delta-D)B to be random if the two are wrongly overlapped:

(Delta-D)A + (Delta-D)B or(DeltaPSI)A + (DeltaPSI)B	Decibans in favor of correct setting of go-back
Given (out of 4 impulses only):	
L0	3.3
L1	-0.8
L2	-0.9
L3	0.1
L4	2.3

It is also possible to make up tables taking into account the limitation. Thus:

(Assuming limitation = x and M37 has 20 dote):

P(0 crosses in DeltaPSI1) = 1-a' + a'(l-b)^5 = .542 = B0

1/5 *P1 = .003 = B1

1/10*P2 = .007 = B2

1/10*P3 = •015 = B3

1/5 *P4 = .032 = B4

P5 = .069 = B5

(Assuming limitation = .):	P(0 cross in)	= (l-b)^5 = .003 = C0
	1/5 *P1	= .007 = C1
	1/1O*P2	= .015 = C2
	1/10*P3	= .032 = C3
	1/5 *P4	= .069 = C4
	P5	= .151 = C5

Now, if limitation in 1st position = limitation in second position

P(0 crosses in the delta psi prime sum) = (1/2)(B0^2 + C0^2 + 5B1^2 + 5C1^2 +10B2^2 + 10C2^2 + 10B3^2 + 1OC3^2 + 5B4^2 + 5C4^2 + B5^2 + C5^2) = .183

P(1 cross in the delta psi' sum) = .029

P(2 crosses in the delta psi' sum)= .024

P(3 crosses in the delta psi' sum)= .024

P(4 crosses in the delta psi' sum)= .029

P(5 crosses in the delta psi' sum)= .047

So that with limitations alike:

(Delta-D)A + (Delta-D)B or(DeltaPSI)A + (DeltaPSI)B	Decibans in favor of correct setting of go-back
L0	7.7
L1	-0.3
L2	-1.1
L3	-1.2
L4	-0.3
L5	1.8

and expected score in right position : 855 decibans on a length of 1,000;

and expected score in wrong position = -516 decibans on a length of 1,000.

Again given that the limitations are alike, the sum of the delta psi prime streams breaks down into component parts as shown on the next page:

If((DeltaPSI')A + (DeltaPSI')B)=	The sum most likely came from:	
L0	L0 L0 x x	80%
	L4 L4 . .	7%
	L5 L5 . .	6%
L1	L4 L5 . .	36%
	L3 L4 . .	31%
	L2 L3 . .	10%
	L4 L5 x x	8%
	L3 L4 x x	6%
	L0 Ll x x	6%
L2	L4 L4 . .	20%
	L3 L5 . .	20%
	L0 L2 x x	15%
	L2 L4 . .	13%
	L3 L3 . .	13%
	L4 L4 x x	4%
	L3 L5 x x	4%
L3	L0 L3 x x	34%
	L3 L4 . .	28%
	L2 L3 . .	12%
	L2 L5 . .	9%
	L3 L4 x x	6%
L4	L0 L4 x x	59%
	L2 L4 . .	14%
	L3 L3 . .	10%
	L1 L5 . .	3%
L5	L0 L5 x x	81%
	L2 L3 . .	10%
	L1 L4 . .	5%

If limitation in first position not = limitation in second position, we have:

P(0 crosses in the delta psi prime sum) = B0C0 + 5B1C1 + 10B2C2

+ 10B3C3 + 5B4C4 + B5C5 = .029

l/5*P(l cross in the delta psi prime sum) = .024
1/l0*P(2 crosses in the delta psi prime sum) = .023
1/10*P(3 crosses in the delta psi prime sum) = .029
1/5 *P(4 crosses in the delta psi prime sum) = .046
P(5 crosses in the delta psi prime sum) = .088

and the deciban table is as follows, when the limitations are different:

(Delta-D)A + (Delta-D)B	Decibans in favor of correct setting of go-back:	
L0	-0.35	
L1	-1.1	
L2	-1.2	Expected score
L3	-0.3	right position
L4	1.7	length l,O00...282 db.
L5	4.5	Wrong:.....-250 db.

and (given limitations different) if the sum	It most likely comes from:	
L0	L4 L4	38%
	L5 L5	36%
	L3 L3	16%
	L0 L0	6%
	L2 L2	3%
L1	L4 L5	39%
	L3 L4	34%
	L0 L1 x .	15%
	L2 L3	10%
L2	L0 L2 x .	34%
	L3 L5	19%
	L4 L4	19%
	L2 L4	10%
	L3 L3	10%
L3	L0 L3 x .	59%
	L3 L4	21%
	L2 L3	9%
	L2 L5	7%
L4	L0 L4 x .	81%
	L2 L4	8%
	L3 L3	6%
L5	L0 L5 x .	93%
	L2 L3	5%

When setting or breaking psi's, by anagramming to develop the psi prime streams, a "breaker" may work out long stretches of plausible texts which are actually wrong. The following "curios" are noteworthy:

```
De X:        J H F T T H Y Z 9 I E 5 A V I B
Phoney P:          8 8 9 G E H E I M 5 5 A A
Phoney PSI':       K K T K T T U U U O S S G
Real P:      9 5 5 5 5 V 8 9 Z W E I 9 A 9 F
Real PSI':   K K P J J R R Y Y Y / X U X 4 H

De X:           Z Q C I / K 3 Q K 5
Phoney P   (1)     G E H E I M 9
phoney PSI'(l)     H H H C C J J
Phoney P   (2)     G E H 5 M A 9 8
Phoney PSI'(2)     H H H H H H J 9
Real P          5 X 9 8 R O E M 9 8
Real PSI'      R R R R R O D J J 9


De X:        F 4 Y 9 T J I C F P X N X V O 8 L I 4
Phoney P:    A N 9 G E N 5 M 8 9 K D O 9 5 M 8 9
Phoney PSI':C C Z V Z U X L L L L S S G A A F 4
Real P:      5 A A 9 8 9 H S I X N 9 5 Q U P 9 Q W
Real PSI'    P E P / K K L J J J Z 3 I D 8 D P Z Z
```

Psi Breaking by Hand

Psi breaking is obviously not different in its initial stages from psi setting: for in psi setting, it is necessary to "break" stretches of psi prime to match them against the known psi wheels. Psi breaking merely requires that more and longer stretches of psi prime be broken, since no aid can come from being given known psi patterns. Stretches of psi less than sixteen characters can not usually be projected forward or backward by multiples of the psi wheels without losing by the resultant stagger.

Of all the "de-chi days" broken by Newmanry since February, Testery has failed to break the psi's on only four.

On high-motor-dot traffic, "55M889" is a wonderful help since plain text can usually be guessed between any close-together-"stops" whenever the nature of the message is foretellable from the log. Go-backs are a help because they (usually) combine psi prime streams at just the right interval to project results successfully back and forth.

Perhaps the greatest occasional aid in breaking a set of psi's is the discovery and utilization of "stretches of psi in the de-chi." These stretches occur when the German message tape runs out of the tape reader in such a way as to cause blanks to be enciphered until the operator comes over and stops the machine. Since Testery is working with a de-chi, it discovers these stretches as pure psi prime. At the end of such a stretch a pause in the radio-transmission usually occurs (while the German operator fixes his machine.) Said pause helps the Testery locate and believe said stretches. Also, if the German tape gets stuck, some one letter may be enciphered repeatedly until the operator pulls the tape through (usually with the same sort of pause). Thirty two possibilities for the letter that was repeatedly enciphered at this point must be tried (in order of preference) when obtaining the resultant psi patterns.

It is sometimes possible to utilize recovered stretches of broken pattern which appear to be pattern repeats. If such repeats are visible on two impulses, the exact psi distance may be calculated by solving a diophantine equation (e.g., 43n + 17 = 47m + 12 where PSI'1 and PSI'2 contain repeats.) Patterns can then be played back and forth as developed.

Dragon

Little can be said about Dragon that is not already known by Arlington. It is indeed proving itself to be quite worth while. The British are not using it to break psi's; and they are not using it to set psi's on days of high motor dottage, since this can be done quickly by hand.

But when they want to set psi's on days when motor dots are low, Dragon gets answers in a truly astounding fashion. For this reason it is kept busy on low-motor-dottage traffic. It solves an average of 4 or 5 transmissions a day which might not have been set otherwise. Captain Fried has been forwarding the weekly Dragon reports, and also sent along the information that Dragon was modified to permit leaving out any impulse which might be in doubt in the de-chi. When it operates on only four impulses, the full 10-letter crib allowed by the board must be used, and the opinion here is that future Dragons really should make provision for 12-letter cribs, although good cribs of over 9 letters are hard to find.

Good cribs fond in a study or P text area

Gurnard (Berlin end)

Crib:	Frequency per 100,000	Length
89ROEM95	65	8
M89ARMEE	64	8
ANGRIFFE9	37	9
9DER9FEIND	34	10
5M89GR5M89	29	10
M89PZ5M89	28	9
TAETIGKEIT	23	10
9ANGRIFF9	28	9
5M89DIV5M	20	9
DG9HOSF95	16	9

Jellyfish (Berlin end)

89R0EM95	61	8
5M89KDO5M8	38	10
9DG9HOSF95	29	10
5M89R0EM95	27	10
ABSCHNITT	23	9
9ANGRIFF	31	8
5M89D5M89	20	9
5M89A5M89	20	9
5M89GR5M89	16	10
5M89WEST	34	8
TAETIGKEIT	15	10
AUFKLAERUN	14	10
9NACHR5M89	15	10

Stickelback (Berlin end)

889ROEM955	53	10
55M889SUED	24	10
9ROEM955	83	8
955LL8899	51	8
88ARMEE55	26	9
9LM9RAUM9	27	9
SUEDUKRAIN	21	10
GRUPPE9SUE	20	10
999PZ55M88	20	10
ANGRIFF9	32	8
HEERESGRUP	17	10
TAETIGKEIT	16	10

Whitting (Riga end)

95L5L595M5	19	10
95K5K58	21	7
89A5M89K5M	16	10
HEERESGRUP	13	10
95AA95M5A8	11	10
95M5A899	27	8
89ROEM95	15	8
M5(QP)5M5R5R	9	10
WIRTSCHAFT	8	10

Breaking Motor Wheels

Motor wheels are broken as at Arlington (with slight differences.) The breaker does not work back to the front of the message as he would if only motor setting, but uses the stretch of text already placed. First he types out about 1,000 letters of psi stream on a Tunny machine, on a width of 61. Then he writes out 600 letters of dechi on a width 61. He anagrams, marking down " - " for places where the limitation compels a psi change; " x " where a psi change exists without being compelled; " . " where psi's are extended. The " x . " pattern shows through the limitation, and is recorded in a rectangle 61 wide. (See worksheet.) Columns between which a change has occurred are headed with a cross over the left-hand-member-of-the-pair. Now a column can be chosen and matched to another column at a distance to the right of it of 37 plus the number of motor dots in between. If several columns occur together under a small stretch of XXXXX's, this block can be slid to the right and will match a similar block if it is also under XXXX's. (continued page 76.)

So a little piece of celluloid is used to mark down these columns that form a good starting block, and the celluloid is slid back and forth to different matches and partial matches, and the process is repeated until the 61-wheel pattern has been obtained. The 37-wheel pattern then of course "falls out."

Messages in depth

The fundamental fish equations:

$$P + PSI' + X = Z$$

may also be expressed by the parametric equations:

$$(3) \quad K = P + Z$$

$$(4) \quad PSI' + X = K$$

both of which must usually be solved to read entirely any given message. Newmanry and Testery both, apply themselves to their solution.

Solution of (3) may be done by a process called "anagramming messages in depth;" or by a process called "cribbing."

Solution of (4) is called "key breaking."

Recovery of key by anagramming messages in depth is well understood by everyone. While the British are unconvinced of the ease of anagramming two messages in depth in the Sturgeon machine, which Dr. Levine indicates can be done without too much difficulty, in his paper on band transposition, (wait till he gets over here!) they do it all the time on Tunny. Such a worksheet follows. Procedure is to write out:

```
            Message 1 - - - - - - - -
            Message 2 - - - - - - - -
            Difference  - - - - - - -
                                                N               9
If the difference is a 3, it probably came from 9 (or of course N) and if
          8                      E
a 5, from 9, and if an F, from N. In any case, they search through the
```

difference row for familiar sequences, dragging probable words against this difference row to see where good text #2 results. They believe hand methods to be quicker than machine, and are unimpressed by our I.B.M methods. In some ways I think them right. Facility gained at 32-letter arithmetic and visual recognition of the various forms of P text, stands them in good stead later when it comes to breaking psi wheels. And it is not too great a mental burden for them, because they are reading messages that mean something to the war effort. Our failure at Arlington to work operationally destroys the huge incentive that makes the British daily mental gymnastics not only bearable but also pleasurable.

When bits of plain texts in scattered stretches are read, there is a difficulty in the association of these plain texts with their correct messages. P1 + Z1 = K1 it is true, but Pi + Z2 = ?? Aids in the correct association of plain and cipher texts are:

1. The log book gives the type of plain texts to be expected.
2. Hand transmissions tie in stretches, and also result in characteristic plain text.
3. Go-backs are self-identifying.
4. When a X2oneback limitation exists, the Delta Key can be distinguished from the Delta "rubbish."

The last process is done as follows:

Delta "believed-to-be-key" is written on a width of 31. The

of Delta-K2 so written

excess of dots (or crosses) in any column \ should have a proportional bulge of Beta, because of the X2oneback limitation. Let the modular sum of these 31 columns excesses be called X. (i.e., X = Sum|X1|.) Then a quick significance test to determine if the text is Delta-K2 with a X2oneback limitation is that shown on page 29, i.e., X >= .8(RW)^1/2 + 1.2 (R)^1/2

In the anagramming example shown, the anagrammist had those clues to aid him which he has marked on the worksheet.

Note that both K and Delta K are written out on the worksheet. This worksheet was merely a convenient place to Delta the K for the "key breaking."

Breaking Key from Anagramming

Key obtained from the preceding anagram worksheets was broken by the British in the following steps:

1. They made out a "Turingery" Delta K sheet. Time: 15 min.

2. They constructed Delta K1,5 Delta K 2,5 Delta K3,5 and Delta K4.5 "rectangles" and "flags of rectangles" Time: 2 men 2 1/2 hours each. (these flags obtained by cross products of the rows? for Delta K5 patterns)

3. They combined the flags into one master flag by straight addition and crudely converged the master flag. (If necessary, they could have flagged the master flag.) Time: 1 man 1 hr.

4. They took the resultant Delta X5 wheel through the 4 rectangles obtaining embryonic wheels for Delta X1 Delta X2 Delta X3 and Delta X4, by crude convergence. Time: 1 man 1 hr.

5. They transferred the embryonic wheels to the Delta Key "Turingery" sheet, dechi-ing on 4 impulses, and writing down the resultant delta psi prime. Time: 1 man 3/4 hr.

6. They "counted out the Delta X wheels" (to be described). Time: 1 man 1 1/2 hours. One counting was enough to get legal, good looking X wheels which were immediately sent to Newmanry for setting other messages, before the psi's had been recovered.

7. They broke the psi wheels. Time: 1 man 1 1/2 hrs.

TOTAL TIME: 1 man, 1 shift; plus 1 assistant 2 hrs. (And I was bothering them)

All worksheets except for the 4 flags are given in Pages 84 through 92. The combined flag made from the four flags is included however (Page 85.)

The "Turingery" sheet is necessarily shown in its final stage. Originally it contained only Delta K. While little apparent effort was put on it, actually a great deal or work has been erased in the process of completing it. The part of the key I have encircled (in red) is incorrect, a result of incorrect anagramming. This was discovered only after the X's were out and psi's were being broken, by the psi's not fitting in that area.

DeltaK15 DeltaK25 DeltaK35 DeltaK45 rectangles were first flagged each (by standard cross product procedure)/on a width of 23, i.e. in reference to the fifth impulse. These flags were laid over each other and combined by simple summing. The total number of entries in any such combined flag is dependent upon the total number of comparisons in the 4 rectangles, which may be (length of the generating text) considered the "N" value / for the combined flag. The combined flag was then crudely converged. The random sigma of the score obtained by crudely converging any "excesses" / flag is (N)^1/2 and therefore for the combined flag, sigma = square root of the sum of the number of comparisons from the 4 rectangles. But number of comparisons equals .0648(N)^2 - 2N + 8 wherein N = length of Delta K text.* Sigmage may therefore be measured as:

$$\frac{\text{.}\qquad\text{X}\qquad\text{.}}{2(.0648(N^2) - 2N + 8)^{1/2}}$$

and in the flag given turns out to be 15.5 which is significant indeed.

- -

*Proof: Let N = kw + j, wherein w = width of rectangle, and 0 <= j < w.
There are either k entries or k+1 entries in any column.
Total comparisons therefore = ((k+1)k)/2 *j + (k(k-1))/2 *(w-j)

$$= (N^2)/2W - N/2 + (1-j/w)/2$$

Now for four rectangles wherein w = 26, 29, 31, 41, and the expression is summed:

GRAND TOTAL = 1/2 N^2 Sum-w 1/w - 2N + Sum-w (j/2)(l-j/w) This last term is given an approximate value of 8. Thus the answer:

GRAND TOTAL = .0648(N^2) - 2N + 8.

For this reason, the results may be believed. The British doubted one weak character in Delta X5 shown on the flag sheet, tried it as a dot (so there would be 12 crosses) and tried to integrate to a X5 wheel. This failed. (See flag sheet, Page 85.) So they tried the doubted character as a cross, as the score actually indicated it to be, and reversed the Delta X5 pattern to get 12 crosses. This integrated.

The Delta X5 pattern was then pushed through each rectangle. It was pushed through in inversed form, because they hadn't versed Delta X5, as yet (Pages 86-89.) This gave embryonic delta X wheels on all impulses. These embryonic wheels were then scribbled down on the wheels sheet shown, in true versed form. See wheels sheet, Page 90. In doing this they doubted any character which shoved less than 3 pips, and did not record it. Since lOLog10 of (1 + Beta)/(l-Beta) equals 3 decibans if b = 2/3, the average rectangle pip is worth three decibans, and they therefore were doubting any character of less than 9 decibans.

The embryonic wheels of the wheel sheet were then held up to the Turingery sheet, and the delta key on every impulse excepting the second wan de-delta-chied. (The known X2oneback limitation would give a good X2 by other means, so it was not used) Thus the resultant was deltaPSI' 1,3,4,5.

The second impulse of delta psi prime was then "counted out." This was done by examining the other delta psi prime impulses at each character and if they were dots, to assume the motor key was a dot and therefore the unknown delta-psi-prime impulse had to be a dot. (See the count sheets and also the Turingery sheet.) Scoring was done as follows:

If other deltaPSI' impulses gave:	Then the deltaPSI' impulse in question was scored:
4 dots 0 crosses	16 db., favor dot.
3 dots 0 crosses	12 db. in favor of a dot
2 dots 0 crosses	7 db. in favor of a dot.
1 dot 0 crosses	3 db. in favor of a dot.

This scoring is quite clearly derived: If ther is one dot, the odds that the unknown is also a dot are: $P(..)/p(x.) = \frac{(1-a) + a(1-b)^2}{ab(1-b)}$ which equals 2 if b is 2/3, and thus 10Log10 = 3. Similarly if there are two dots, the odds in favour of the unknown being a dot are $P(...)/P(x..) = \frac{(1-a) + a(1-b)^3}{}$ = approximately 5, so decibans = 7. And so on for three and four dots, given no crosses.

In the example shown, the score values are given for crosses without dots, as are given for dots without crosses. This is justified in practice where it is not known whether or not the dots are really dots, and the crosses crosses, there being a possibility that the signs are reversed, but actually it is not correct. A complete and correct table for scoring, assuming b = 2/3, follows:

Probability that signs are the right way round:	SCORES IN DECIBANS FAVORING HYPOTHESIS THAT THE UNKNOWN IMPULSE IS A DOT (Relatively a dot)										
	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
GIVEN:						SCORES:					
....	3	4	4	5	6	7	8	9	10	12	16
...	3	3	4	5	5	6	7	8	10	10	11
..	3	3	4	4	4	5	5	6	6	6	7
.	3	3	3	3	3	3	3	3	3	3	3
xxxx	-16	-12	-10	-9	-8	-7	-6	-5	-4	-4	-3
xxx	-11	-10	-10	-8	-7	-6	-5	-5	-4	-3	-3
xx	-7	-6	-6	-6	-5	-5	-4	-4	-4	-3	-3
x	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
any mixture of dots and crosses	3	2	2	1	1	0	-1	-1	-2	-2	-3

In the sample given, the ???? ".5" should have been employed if ignorance was assumed.

The above table is derived from the following formula:

Let "p" equal probability that our signs are the right way round, then the probability of another dot given "n" dots in a ??? of Delta-PSI' is: $P = (1-p)b + \frac{\text{Probability (n+1 dots)}}{\text{Probability (n dots)}} = (1-p)b + p\ \frac{1-a+a(1-b)^{n+1}}{1-a+a(1-b)^n}$

Odds in favor = P/(1-P)

And in the case of the "impure character" (mixed crosses and dots) the probability of another dot is:

P' = p(l-b) + (l-p)b Odds = P'/(1-P')

Back to the Turingery: The delta psi prime sign indicated in the scoring (from "counting out") was combined with the delta key sign for that impulse, and the result recorded on the count sheet in terms of the Delta X.

Thus there had now been obtained the Delta X2 count sheet shown page 91. It was then reasoned that, if there was a X2oneback limitation, scoring would be heavy in columns immediately following positions where X2 = x. Therefore X2 = x was written in immediately to the left of each heavily scoring columns. Immediately below this were recorded the signs of delta X2 obtained from the columns themselves. These two checked, and resulted in a X2 wheel which was legal, and turned out to be correct.

In the same fashion, the X3 wheel was obtained, using the "counted out" X2 wheel plus the others. Then putting in this X3 wheel, X4 was "counted out. Now X1 was tackled, with fear that the data would be insufficient since delta X1 needed data for 41 positions. It so happened that there was only one position left vacant (see count sheet.) This character was assumed to be a cross since a string of 4 dots in delta X1 is not too likely. The X1 obtained was used with the others to give a final check on X5, and the job of breaking the X's was through.

These wheels, even though not "proved" (by obtaining psi wheels with the de-chi) were immediately sent to the Newmanry to use in setting other messages.

The psi's were broken as psi's are always broken: by inspection of psi prime streams, and the creation of more psi prime elements by anagramming where necessary, though such is not usually necessary since the psi's usually come out easily. For this reason, the psi breaking worksheets are not included.

When the psi's were finally out, the Newmanry was telephoned and told that the X's had been proved. The set of psi's and X's and the messages were given over to the "setting room" for the motor wheels to be broken, as described previously.

There are or course variations in technique here as in Newmanry. Often when the first X has been broken through "counting out," the corresponding psi prime is obtained, and by numbering the resultant groups the more certain motor dots can be inferred; whenever we have motor dots, delta Key is delta X and we can thereby obtain another X wheel or two and other psi's. This method short-circuits much of the need for "counting out" as described above.

Also it is not usually necessary to obtain every character of a X wheel, since "fiddling" with the unknown characters enables the breaker to assume some missing ones on the basis of known X wheel characteristics.

Other aids to key solution exist. The placing of rectangles end-to-end as follow.:

| 1-5 | 2-5 | 3-5 | 4-5 |

to use data from wheels 1, 2, 3, and 4 simultaneously in getting out the pattern for 5, has been found an aid. Results can be flagged in a 23x23 flag if desired, or the long rectangle can be converged by itself.

A "15O-square," which contains every possible delta Key i+j rectangle:

1-5	2-5	3-5	4-5	.....
1-4	2-4	3-4	.....	5-4
1-3	2-3	.....	4-3	5-3
1-2	.....	3-2	4-2	5-2
.....	2-1	3-1	4-1	5-1

is an even more powerful aid, but very seldom used. It is at least 3 feet wide and takes a prodigious amount of time to construct and converge. A significance test proposed when converging it is the famous "Significance Test IV:"

$$\frac{2.51\ (X/2)^2}{10N} + \text{sigma*theta} - 452\ \text{db}$$

wherein the 452 db. is an empirical constant (2^(l50-l)).

One half of X is used instead of X since the square counts everything twice. Such a square when used gives a good start in Turingery, which Turingery usually itself needs; but Turingery/makes use of all the data appearing in the l50 square and in addition uses motor dot information, so Turingery is actually more powerful although slower to start.

Getting a start in Turingery when 300 letters or more of key are available (and motor dottage is not too low) can also be done by the "5x5 square" method. Steps involved are:

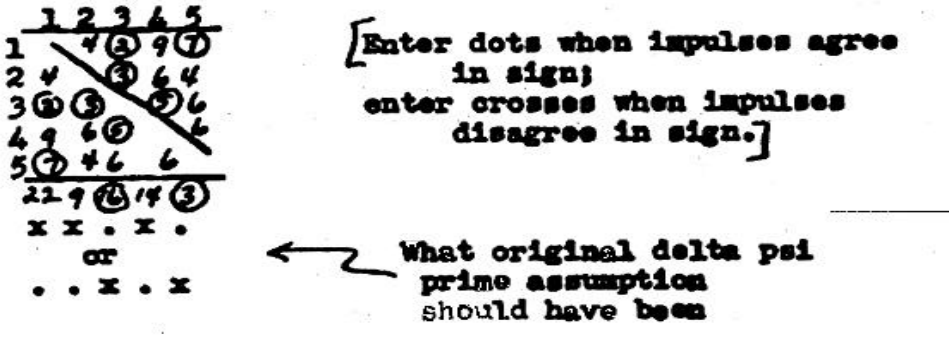
1. Write out delta Key on worksheet. (Regular Turingery procedure.)

2. Assume a character for delta psi prime in the first position of delta key. This is to be changed later if found wrong. It is actually assumed to be a slant.

3. Obtain the resultant delta X signs, and carry them forward throughout the worksheet to their "multiple" positions. At each of these points assume either a repeat or a complete anti-repeat in delta-psi-prime, obtaining thereby the delta X for the whole character. A complete delta X character so obtained is called a "child."

4. Enter on a width of X1 each induced Delta X1 value (from the children) together with a subscript showing the basic impulse from which it was induced. (I.e., in the X1 case, subscripts representing impulses 2, 3, 4, or 5.) Values down the columns tend to be alike, if the originally assumed delta psi prime character ("slant") was correct. But, down a column, a cross with a subscript 2 for example and a dot with a subscript 3, would tend to indicate that the signs of the original delta psi prime assumption were opposite in impulses 2 and 3. A count is therefore made, in each column, of the number of agreements and disagreements between each possible pair of subscripts; these counts, for each pair of subscripts, are then totalled in terms of excesses of agreements or disagreements. The same procedure is carried out for X2 X3 X4 and X5 widths. The grand total count for each pair of subscripts, in terms of excesses of agreements or disagreements, is then entered in a 5x5 flag which is doubled about the diagonal into a 5x5 square.

This square is then crudely converged. The resultant pattern (or its inverse--whichever is "slant," or if not "slant," whichever has most crosses in it) usually gives what the original delta psi prime character first assumed should have been. Sample square shown:



5. The corrected version of the original delta psi prime assumption is now entered, all work corrected accordingly, and the Turingery process carried forward as at Arlington, by carrying the delta X signs from the children forward and backward to their "multiple" positions, "counting out" wheel patterns as described at the bottom of page 81, etc.

Colossus breaking of key is also possible, and is similar to Colossus wheel-breaking (only much faster.) The machine attack on parametric equations (3) and (4), called "cribbing" and "key breaking on Colossus" are next to be described.

"Cribbing"

Solution of parametric equation (3) K = P + Z, can be done by Robinson "crib runs," and key obtained used to solve equation (4), PSI' + X' = K.

With X2oneback limitation, or X2oneback + P5twoback limitation, cribs used should be at least 500 letters in length. Without limitation, or with X2oneback + PSI1oneback or with X2oneback + PSI1oneback + P5twoback limitation, cribs should be over 1,000 letters. A usable crib arises when the Germans use the same message tape in the transmission of the (same) message on different keys. Messages deciphered by Testery and suspected to be circular, are checked by "Sixta" (traffic analysis people) against log receipts on other circuits. Sixta also looks for messages, the last two digits of whose serial numbers are the same. (Serial numbers' last two digits are sent in clear by the German radio operators in receipting procedure. These clicks of course may be accidental.

An average of two recoveries a week results from "cribbing."

It is shown in Fried #F46 P4 that for X2oneback + P5twoback limitation transmission:

$$? \left[\Delta_{31 \cdot k} (\Delta P_2 + \blacksquare P_5) = \Delta_{31 \cdot k} (\Delta Z_2) \right] = b^2 + (1 - b)^2$$

and the crib run for X2oneback + P5twoback limitation consists therefore of running the first function in the square brackets against the second function, on Robinson counting dots.

The value of "k" is usually 3,4, 9, and 11. (See P.2 of #F 75 report by Fried.)

A Delta-P2 + P5twoback tape is therefore constructed for Robinson as follows:

Level 1(DeltaP2 + P5twoback) differenced at 93

Level 2(DeltaP2 + P5twoback) differenced at 124

Level 3(DeltaP2 + P5twoback) differenced at 279

Level 4(DeltaP2 + P5twoback) differenced at 341

Level 5 Control: . x . x . x . x

and a Z2 tape made similarly (NOTDeltaZ2. Since DeltaZ2 can be obtained on Robinson from a Z2 tape it is also possible to form Delta/31.k DeltaZ2 on Robinson from a Delta/31.k Z2 tape) The Z tape looks as follows:

Level 1 Z2 differenced at 93

Level 2 Z2 differenced at 124

Level 3 Z2 differenced at 279

Level 4 Z2 differenced at 341

Level 5 Control: . x . x . x . x

Control levels are placed in tapes to permit elimination of portions of the tape not desired, from the count. Also, since Robinson must have one tape in which "start" and "stop" impulses are punched, and since "stop" must follow "start" by at least 2,000 impulses to give the relays time to operate, shorter texts can be placed on a tape of over 2,000 positions, and the control punched "out of phase" in the blank area, causing the blank area not to be counted at all, rather than to be counted as slants as would have been done otherwise. This has been described in writeups sent by Fried.

In view of the control being so constructed it is necessary

to make each run in two parts, looking at even offsets of one tape with reference to the other on the first run (by plugging Robinson to count only where Level 5 of tape 1 (control level) plus level five of tape 2 (control level) = dot), and looking at odd offsets of one tape with reference to the other on the second run (by plugging Robinson to count only where level 5 of tape 1 plus level 5 of tape 2 = cross.) This does not take twice as long, because the tapes are in effect slid 2 positions at a time with reference to each other in each run, and the over all time is the same as if only me one run were made and the tapes slid one position at a time.

The two tapes are placed on Robinson and the Robinson set to record all positions wherein simultaneously:

P tape level 1 + DELTA of (Z tape level 1) = dot.

P tape level 2 + DELTA of (Z tape level 2) = dot.

P tape level 3 + DELTA of (Z tape level 3) = dot.

P tape level 4 + DELTA of (Z tape level 4) = dot.

P tape level 5 + Z tape level 5 = dot for even runs
or cross for odd runs.

In other words, Robinson is plugged up to record positions wherein:
Delta/93 (DeltaP2+P5twoback+DeltaZ2) = Delta/124 (DeltaP2+P5twoback+DeltaZ2) = Delta/279 (DeltaP2+P5twoback+DeltaZ2) = Delta/341 (DeltaP2+P5twoback+DeltaZ2) = dot

and the causal probability for such is b^5 + (1-b)^5 whereas the random probability is 1/16. Such a run is much stronger than a run on a single impulse of the special tapes, wherein causal probability is b^2 + (1-b)^2 as against a random of 1/2.

A sample Robinson crib run as described above is shown next page, and described in the pages to follow.

page 98 (still to be formatted)

Positions of tapes relative to each other are shown on the run as the first four digits dawn the columns, and the scores at such positions are shown as the last four digits. The P tape was 2,510 long. The Z tape was 2,512 long, of which only 1,222 characters were due to the Z text and 1,290 were blank. (That is, the crib as 1,288 longer than the cipher.) Start and stop impulses were placed on the Z tape. The Z tape was made two positions longer than the P tape, so that as the two revolved side by side in the Robinson "bedstead" (tape rack), the first position of the Z tape changed relative to the P tape by two notches each revolution. For this reason the setting "1052" shown on the accompanying Robinson run really means position 2 x 1052 or 2104. (This was an even run. On odd runs the position is one less than twice the recorded position.)

In this particular crib run it should be noted that if the first position of the Z tape was set along 2,104 positions on the P tape, then the first 406 characters of the Z tape matched the last 406 of the P tape. The remaining 816 characters of the text (non-blank) part of the Z tape have been matched against the beginning of the P tape (and are therefore random) This is considered better procedure however than to let the Z tape "run off the end" of the P tape, in which case average random scores would have to be re-figured at each of such positions.

The random score in the sample is obviously 1/16 times 1,222 = 76.4. Sigma is (1,222*(1/16)*(15/16))^1/2 = 8.46. A score of 114 is therefor 4.5 sigma. This was attained in spite of the fact that only 406 characters counted toward the causal score, whereas 816 were random!

These 406 characters actually represent 406+341 or 747 letters of

plain text cipher text coincidence, since the 4th level of the P tape came from delta-ing the actual P text at interval 341, and therefore the tape is that much shorter than the real text.

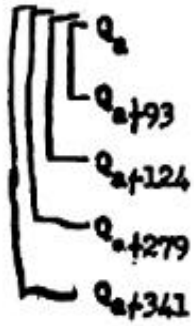
Consider for a moment the results of adding the two tapes as described page 97 and as done in the run shown. Let us call

$\Delta P_2 + \bar{P}_5 + \Delta Z_2 = Q$. Then we have in effect:

(A)

as the sums of the P and Z tapes at any given moment. These 4 levels of delta/31.k *Q actually represent the result of summing:

(B)



But there are really 5*4/2 or 10 sums possible in (B) of which we have shown in (A) only 4. It is desirable to use all 10 sums in any test proposed.

Now if there are "t" dots in the 4 levels of any Delta/31.k *Q character from(A) that means that in (B), sign Qa agrees with "t" of the signs Qa+93, Qa+124, Qa+279, Qa+341. There are therefor in (B) really "t+1" signs alike and like Qa and "4-t" signs alike and unlike Qa.
The number of dots in the 10 possible combinations of (B) is

therefore: 1/2(t+1)(t+1-1) + 1/2(4-t)(4-t -1) = T. The following table can be made from this formula:

t	T .
0	6
1	4
2	4
3	6
4	10

If at any given juxtaposition of the two tapes, we now make a count not only of the 4-dot characters (as done in the crib run shown, for all juxtapositions) but also of 3 dot 1 cross, 2 dot 2 cross, l dot 3 cross, and 0 dot 4 cross, characters ("t" = 4, 3, 2, 1, 0) we can multiply such counts by the proper T scores from the above table, and obtain the best possible scoring for that particular given juxtaposition. This is easily done by setting the tapes to the chosen juxtaposition, and making a "sixteen letter count," and then grouping these letters into categories of "t," applying the T values.

Since there are (1/2)*5*4*N or 10 N comparisons possible, the number of dots expected at random is 5 N, and sigma is of course (2.5N)^1/2. Thus after a crib run has been made as shown on page 98, a 16 letter count is usually made at the juxtaposition suggested by the first run, the T score computed, and resultant sigmage computed. If the resultant sigmage is above that obtained from the original run, and is high (usually 5 sigma or better) then the juxtaposition is accepted, and the key recovered from adding crib and cipher together at that juxtaposition. In the present example the sigmage from the 16 letter count (this count not shown) was 6.1 sigma, so the setting was deemed proper.

If the limitation in the sample run given had been X2oneback only, and not X2oneback + P5twoback, tapes would have been made the same as those described on page 96 except that the P5twoback impulse would have been omitted entirely. Thus the P tape would have been simply:

Level 1 (DeltaP2) differenced at 93

(etc.)

If there had been no limitation whatsoever, or a X2oneback plus psi-one-prime-bar limitation (and therefore unpredictable at this stage) the original tape would have had to have been different. The mathematics leading to the differencing of P and Z at small multiples of the X2 wheel only, would have been of no avail. X wheels would have to have been considered in pairs, and differencing \ done at products of their cycles. The of P and Z P tape then would have looked like:

Level 1 (DeltaP1&5) differenced at 943

Level 2 (DeltaP4&5) differenced at 598

Level 3 (DeltaP3&5) differenced at 667

Level 4 (DeltaP2&5) differenced at 713

Level 5 Control . x . x .x . x

and the Z tape would have been the same with "Z" substituted for "P." The same type of runs would have been made, and the same tests employed, since probabilities of getting dots are the same. The above tape however requires much longer texts because the differencing interval is so much greater.

Key obtained from Robinson cribbing may now be solved as described under "Key Breaking" (Page 79) or by key breaking runs on Colossus, next to be described.

Key Breaking on Colossus

Key-breaking runs are identical to wheel-breaking runs, already described. Decibanning is done the same way. Wheels come out much quicker, however.

In general, a DeltaK45 rectangle is made from the Delta Key. This rectangle is usually flagged (cross product) and the flag converged by crude convergence. A pattern is thus obtained for DeltaX5 and this may be pushed through the rectangle (crude scoring) to obtain DeltaX4. If necessary the DeltaKl5 DeltaK25 DeltaK35 rectangles are also made as described on Page 80, but this is not usually necessary on key obtained from cribs since such key is sufficiently long to give good results with less work. As soon as patterns X4 and X5 have been obtained, a DeltaK tape is placed on Colossus and the following runs made (wheel-breaking runs):

4p/5 as a test of score.

3=/5 (or if it is known that the 5 pattern is correct as to direction of signs, and not inversed, then the run 3./5. is made instead.)

2=/5, (or 2./5. if there is no possibility that 5 pattern signs have been inversed)

1=/5 (or 1./5. if as before.)

Note that an attempt is made to build up a large number of stronger partial wheels, rather than to make a \ run such as 2=/4=5 which would have less text for study, and might therefore not have resulted in any stronger pattern anyway at this stage.

The next wheel-breaking run is often 4=/3=5 which should "finish off" wheel X4. Then comes 3=/4=5 which usually finishes X3. Then 2=/3=4=5. Finally comes l=/2=3=4=5 and in half-an-hour after starting on Colossus, "the wheels are out". (On short messages with low values of b, it is or course not as easy as appears here.)

As described at the bottom of Page 77, De1ta-K2 may be written out in a width of 31 if there is a X2oneback limitation, or else Delta-K2 plus P5twoback may be written out in a width of 31 if there is a X2oneback + P5twoback limitation, and the excesses of dots over crosses in the columns then indicate the pattern of Delta-X2 + X2oneback + cross (or "chi-two-cap-inverse.") From such a pattern, if complete and accurate, the X2 wheel can be obtained and the direction of signs will be proper (i.e. verse instead of inverse.)

An attempt to get such a pattern is therefore often made with key obtained from cribbing. Results are used in conjunction with the DeltaK35 rectangle, or with Colossus wheel-breaking runs, as a "head start" to X recovery.

* * * * * *

Description of Machines Used in "Newmanry"

TUNNY: Newman's section has three tunny machines: the first two use standard telephone relays (type P.0. 3,000) to generate the chi and psi and motor patterns. The third and latest model uses telephone relays and will perform the delta operation in addition.

Each Tunny is primarily used for deciphering messages; it is also used for de-chi-ing; it can work backwards as well as forwards. Each Tunny has available 2 reperforators, 1 teletypewriter printer, and 1 tape reader ("transmitter.")

On model three, any tape can be sent through and a delta tape generated; or wheel patterns my be delta-ed. The control panels and pattern generators are on relay racks with total width of 8 feet, and height of eight feet. The tape transmitter, reperforators, and typewriter are on a desk nearby. Any impulse can be included or excluded.

JUNIOR: (Three available.) Consists of 1 reader, 1 electromatic typewriter, 1 plugboard. Used to print de-chis.

GARBO: Used primarily to make rectangles. Model I can make delta-de-chi tapes or prints, although models II and III do not have memory apparatus needed for delta Garbo consists of 2 I.B.M. tape readers, 1 I.B.M. reqenerating typewriter,

and 1 perforator; the assembly is on two tables and there is a small rack for switches. In Garbo I the memory relays are arranged in 4 small banks on the switch rack. Delta process can be done by using two tapes, one tape plus memory, and in case of delta-de-chi, by one tape for Z text, memory, and one delta-X tape. Results can be printed on typewriter or punched on tape. Garbo can transfer impulses to different levels, etc. Any letter on the tape my be made to operate any letter on the electromatic typewriter.

Rectangling was originally done on Garbo by setting the typewriter to the width of the smaller of the two wheels involved, and recording the Z tape (or K tape, or etc.) line after line until completing one product cycle of the two wheels; then setting the paper back in the typewriter to just below the original starting point and recording the second cycle position by position just below the equivalent parts of the first cycle. A tally sheet resulted. Results of this tally sheet, in excesses of crosses or dots, had to be written diagonally on a rectangle sheet in accordance with standard procedure for forming rectangles.

The present Garbo method results in a similar tally sheet, but the cycles have been combined in groups of five so that there are less tallies in each cell to contend with. Suppose a Delta Z1&2 rectangle is desired from 11,000 letters of Z. A Z1&2 tape is made on Mrs. Miles (a machine to be described) showing the 11,000 Zl&2 characters divided into 5 levels:

Level 1	1	2	. . . . .	1271	. . . . .	2542				
Level 2	2543	2544	. . . . .	3813	. . . . .	5084				
Level 3	5085	. . . . .	6355	. . . . .	7626	< - Z1&2 text				
Level 4	7627	. . . . .	8897	. . . . .	10168					
Level 5	10169	. . . . .	END(11,000)	.						

(5*1271) < (11,000 < 2*(5*1271) so the tape is 2*1271 long.

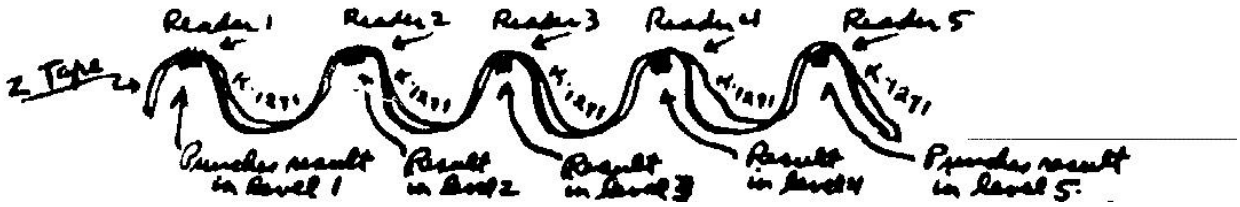
If the text length is expressed as k.5.1271, the tape will be k.1271 long.

The above tape is now passed through Garbo which deltas it as it goes through, and prints:

If Delta Z1&2=	Value printed:
/ (no crosses)	- zero
E,4,9,3,T (1 cross letter)	-l
two cross letters	-2
three cross letters	-3
4 cross letters	-4
8 (5 crosses)	-5

This printing is done by simple monoalphabetic substitution with variants, plugged up on the Garbo electromatic typewriter. There are now only k entries in depth in the cells of the tally sheet, instead of 5k entries as in the earlier method. The entries show the number of crosses. These are converted into excess of dots over crosses, and written in diagonally on a rectangle sheet as desired.

Mrs. MILES: Models I and II are each made up of 5 Western Union tape readers and two reperforators and two banks of Siemens high speed non-chatter relays. Each will combine up to five tapes in any manner not requiring memory, and punch out results in any arrangement of levels on one or two tapes Each sums, transfers, etc. Mrs. Miles III has approximately 70 vacuum tubes instead of relays, to do the summing job with less trouble. The distributors in the tape readers however give trouble to the vacuum tube counters, since the distributor brushes have a tendency to bounce, so Mrs Miles IV(under construction) will have a vacuum tube ring as distributor. Mrs. Miles makes up tapes for crib runs. Also she prepares tapes for Garbo rectangling. For this the Z tape is placed into 5 readers simultaneously as follows:



ROBINSON: The original machine has one "bedstead"(a large rack with wheels) which holds two tapes and photocell apparatus; 1 counter rack, and 1 control rack. The bedstead has sprocket drive for the two tapes to keep them in perfect step, since they both run at once. (as they do NOT do in Colossus.) Robinson has two counters: one operates while the other prints out. The controls available are: a. Set-total keys. b.Impulse comparison jacks and plugs.

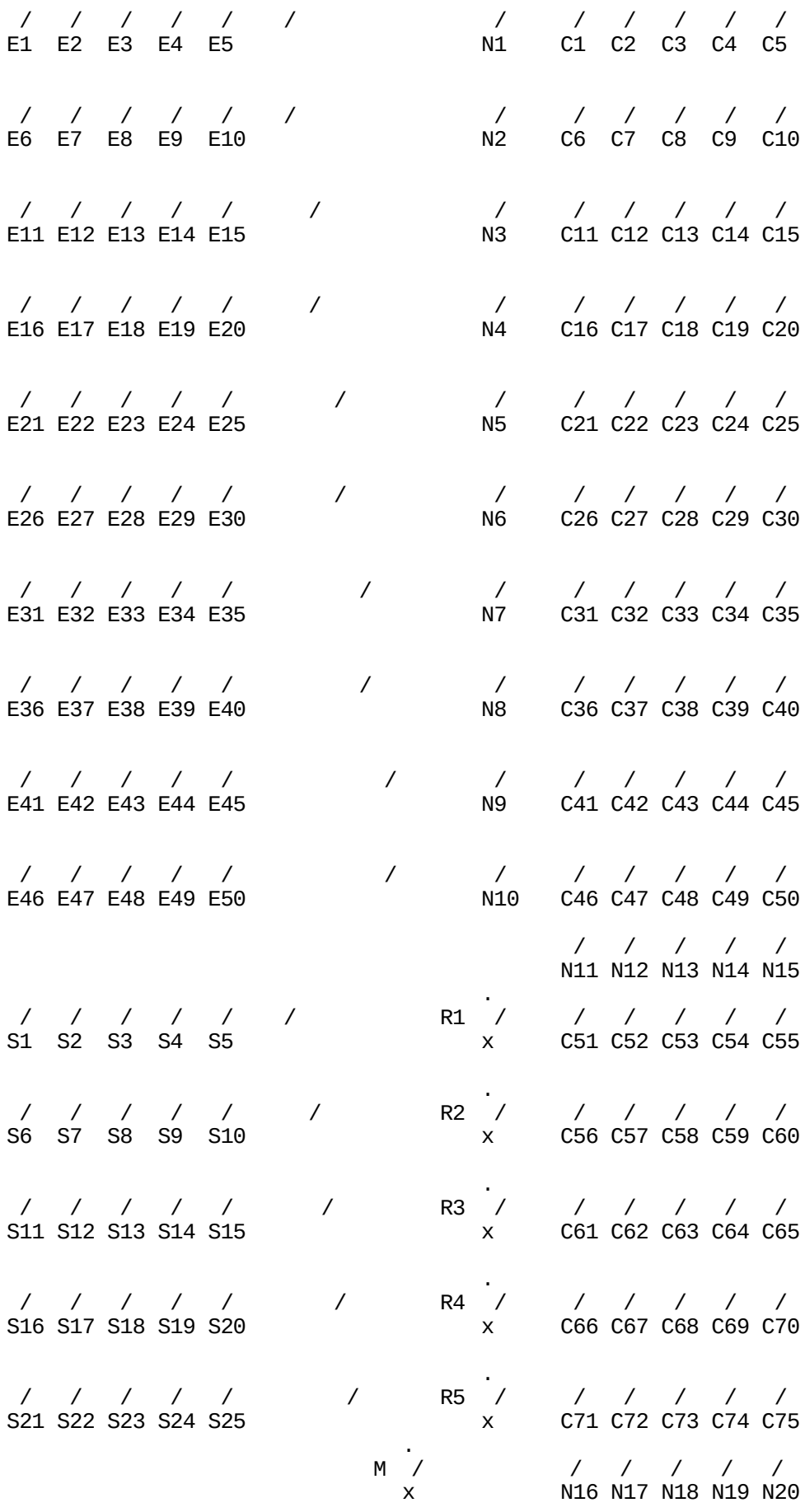
c. Counter jacks and plugs. Delta process is done by means of two sets of

photocells for each tape, so that a pair of adjacent levels may be read simultaneously and differenced. Thyratrons count units of 1, 2, 4, 8; high speed relays count units of 16, 32, 48, 64; nickel-iron relays count 80, 160, 240, 320, 400, 800, 1200, 1600; standard telephone relays ("P.O. 3,000") count 2,000, 4,000, 6,000, and 8,000. Tape lengths possible on the bedstead are from 2,000 to 1l,000. DOUBLE ROBINSON: (two built) Same as above but double Robinson has 2 bedsteads and photocells racks, so it can read from 4 tapes at once. NEW DOUBLE ROBINSON (under construction): same as the double Robinsons just mentioned, except that it will have vacuum tube counters, and will permit multiple testing. Other Robinson are being converted to this type later.

COLOSSUS: Bedstead has positions for two tapes, but only one runs at once. Each tape travels past only one set of photocells. Tape is driven by friction between tape and the wheels, without the use of sprocket-pull as in Robinson. Colossus has vacuum tube: chi pattern generator, psi pattern generator, and motor pattern generator; also has a vacuum tube memory so it can delta. "Triggers" (rotary telephone switches in this case) select any of 5 sets of wheel patterns for chi, psi, and motor wheels, each set of which my be plugged up at will and selected at will. There are five counters, which may all work at once if desired, and when counters are overloaded the machine runs idly (without stepping the wheel-patterns) until the counters print out and are clear again. "Span counters" permit the selection of a stretch or span of tape for study, eliminating all before and after the counting.

Colossus has a jack board for plugs, to be used in controlling it; and also a key board. The key board may be used instead of the plugs for most runs and is more convenient, or in conjunction with the plugs. The "keys" are switches whose levers have "center," "up" and "down" positions, or in some cases only "center" and "down." A schematic of the key board appears next page.

COLOSSUS KEY-CONTROL-BOARD



In the schematic, keys have been numbered by me for convenience in describing switching. They are not numbered on the real keyboard.

The "C" keys on the right operate the 5 counters. The "E" and "S" keys on the left refer to the five impulses.

Suppose it is desired to count all the x x x x x characters that result in delta D from running a Z tape against a wheel pattern. The wheel pattern is selected on another board and switched to "delta;" the Z tape switched to "delta;" and then on the keyboard shown (page 109) keys E1, E2, E3, E4, and E5 thrown "down," and key C1 thrown "down". That means that counter 1 will count all cases wherein impulses 1,2,3,4,and 5, equal cross. If it is desired to count all the cases of . . . . . then keys E1 E2 E3 E4 and E5 are thrown "up" and key C1 "down;" if it is desired to count all the cases of x . . . . then keys E1 and C1 are thrown "down," and keys E2, E3, E4, and E5 thrown "up." The counter keys are thus always thrown down, to make counters operative; and the E keys are thrown up to make an impulse dot or down to make an impulse cross.

Suppose a separate count is desired on x x . x x, x x x x x, ... and . . x . . then one way would be to throw keys: E1E2E4E5C1 down E3 up; E6 E7 E8 E9 E10 C7 down; E11 E12 E13 E14 E15 up C13 down; E16 E17 E19 E20 up E18 C19 down; and the first pattern x x . x x would be counted by counter 1, the second by counter two, the third by counter 3, the fourth by counter 4.

Suppose it is desired to count positions wherein 1 plus 2 equals dot. Keys S1 and S2 are thrown down (giving 1 plus 2) key R1 is thrown up (giving "equals dot") and key C51 thrown down (giving "counter one to count such places.") The S keys are thus used in summing. For (1p2)./(4p5). we would throw down keys S1 S2; up key R1, down key C51; down keys S9 S10; up key R2; down key C56. Thus counter 1 would have two conditions imposed upon it simultaneously: that 1 plus 2 equal dot, and that 4 plus 5 equal dot.

Now suppose it is desired to count in the same counter all the places wherein either . . x . . or . . . . . occurs. If we throw keys E1 E2 E4 E5 up and E3 and C1 down, and then up keys E6 E7 E8 E9 E10 and down C6, that will make counter 1 count places wherein . . x . . AND . . . . . SIMULTANEOUSLY occur. This is not only impossible, but it isn't what was wanted. But if we up keys E1 E2 E4 E5 and down keys E3 C1 and down negativing key N1 , so that counter 1 now counts all position that are NOT(. . x . .), and then in addition throw up keys E6 E7 E8 E9 E10 and down keys C6 and N2. counter 1 must now count all positions that are SIMULTANEOUSLY NOT(. . x . .)AND NOT(. . . . .) The negative of this condition may be counted in counter 1 however by throwing negativing key N16 down. If this is done counter 1 now counts all positions that are NOT: SIMULTANEOUSLY NOT(. . x . .)AND NOT(. . . . .) or in other words that are EITHER . . x . . OR . . . . . It is really quite simple.

The "M" key left alone does nothing; with it up Colossus counts only if the motor is a dot; or down, only when the motor is a cross.

The undesignated keys are used in multiple testing wherein five successive positions are tested at once and the wheel-patterns stepped five notches each time; in such cases these keys must be thrown down and other plugging done on the jack control board to provide for five positions stepping. Remaining controls are not of particular interest to Arlington.

Colossus can do rectangling, the run being much like wheel-breaking runs. X1 and X2 patterns are plugged up/with one cross and all the rest dots, These each patterns are run in all 1271 starting positions against delta Z; there is a "depth" switch that can be thrown, so that the answers will be doubled and have the depth subtracted from them, and therefore print out as excess of dots over crosses (positively or negatively.)

It is planned to have a total of 12 colossi.

APPENDIX: Current Notations

- - - -

a = average.
a = proportion of crosses in MT.
a' = proportion of crosses in MB.
AT = auto transmitter.
bi = P(DeltaPSIi = x).
d = actual number of dots in M37
D = dechi text, = Z + X.
f = factor in Bayes theorem.
HC = hand check.
i,j = (suffixes)
K = Key text= PSI' + x.
l = length of message tape.
MT = total motor = MB + effect of limitation.
MB = basic motor (resulting from M 61 driving M 37).
n = text length.
n' = Defined intrinsically: n'/n = P(X2oneback = x).
o = odds.
P = plain text = Z + K.
P5twoback = level 5 of P characters two back
P(.....) = Probability of event (.....) usually by cause.
PB(.....) = Proportional Bulge of event (.....) and defined intrinsically:
P = p(l+PB)
p = probability, usually by random.
R = position of wheel tape (Robinson).
r = number of letters looked at.
SD = standard deviation.
ST = set total.
TM = MT.
TP = teleprinter.
x1 = score in terms of excess of dots or crosses.
X = Sum|xi|.
Z = cipher text = P + K.

Now follows the text form used by Tony Sale, of the Greek symbols handwritten in the original text. The most used are the Greek capital Delta (triangle on its base) /\ for which the word "Delta" has been used. (The word is also used at some points by Albert Small). The lower case "delta" is used for the Greek lower case delta. Another often used symbol is the Greek capital psi, used for the aperiodic set of 5 wheels in the Lorenz cipher machine, also known as the S wheels. In the edited text PSI has been used for capital psi and PSI' for psi prime. Sum has been used for the Greek capital Sigma and sigam for the Greek lower case sigma. PI is used for the Greek Pi.

betai = PB(DeltaPSIi = x) Defined intrinsically as bi = 1/2(1 + betai)
deltaA = PB(DeltaD = "A") Defined intrinsically as P(DeltaD = "A")= 1/32(1+deltaA)
Delta means differenced at interval 1; unless otherwise specified, ie. Delta/31
PSIi means one of the aperiodit (PSI) wheels or the text generated by the wheel.
PSI' means the text generated by extending PSI by MT.
Xi means one of the periodic (X) wheels or the text generated by it.
X2oneback = level 2 on X one position back.
^X2 = X2hat = X2oneback + Delta-X2
^^X2 = X2hathat = X2hat + cross
M37 and M61 mean motor wheels of size 37 and 61.
PIij = PB(Delta-P(ij)= .) Defined intrinsically as: P(Delta-P(ij) = .)=1/2(1+PIij)
sigma = standard deviation, equals square root of variance.
Theta(ij) - excess of dots over crosses in a cell of a rectangle.