GENERAL REPORT ON TUNNY

With Emphasis on Statistical Methods.

TABLE OF CONTENTS

GENERAL REPORT ON TUNNY

With Emphasis on Statistical Methods.

TABLE OF CONTENTS

Part 6

61 Raw Materials and Production with Plans of Tunny Links

Part 7. REFERENCE

71 Glossary and Index
72 Notation
73 Bibliography
74 Chronology

Part 8

81 Conclusions

Part 9 APPENDICES

91 5202
92 Motor Rectangles
93 Thrasher
94 QKP Research
95 Mechanical Flags

91 THE 5202 MACHINE

91A PRINCIPLE OF THE 5202

(a) Introduction.

The 5202 is a photographic machine designed for setting messages enciphered on the Tunny machine. It is based on counting the number of coincidences between two sequences of (teleprinter) letters, one derived from the cipher and one from the chi wheels.

The machine was produced just too late for the European war, but it was decided to experiment with it for two months.

The theory of the method is simple. It depends on the equation

$$\Delta\chi = \Delta Z + \Delta D$$

(b) Example: 1x2.3x

Suppose that we are dealing only with the first three impulses of ΔD and that the letter (from an eight letter alphabet), $\Delta D_1 = x$, $\Delta D_2 = .$, $\Delta D_3 = x$, has a frequency above random. In order to count the number of times the above three-impulse letter occurs in ΔD, using the 5202, we record the data on 35mm film as a series of transparent spots on an opaque background.

A transparent spot in level	$\Delta\chi$ Film $\Delta\chi_1\ \Delta\chi_2\ \Delta\chi_3$	ΔZ Film $\Delta z_1\ \Delta z_2\ \Delta z_3$
1	. . .	x . x
2	x . .	. . x
3	. x .	x x x
4	x x .	. x x
5	. . x	x . .
6	x . x	. x .
7	. x x	x x .
8	x x x	. x .

It will be seen on inspection that the spots will appear on the same level in both films, and thus give a coincidence when superimposed, only if $\Delta D_1 = x$, $\Delta D_2 = .$, $\Delta D_3 = x$.

(c) Example: 1 = 2 = 4

Although we are dealing with an eight-letter alphabet, we can put the data into four levels of the film as follows:

Level		$\Delta\chi_1$	$\Delta\chi_2$	$\Delta\chi_3$	ΔZ_1	ΔZ_2	Δ_3
		•	•	•	•	•	•
1	or	x	x	x	x	x	x
2		x	•	•	x	•	•
	or	•	x	x	•	x	x
3		•	x	•	•	x	•
	or	x	•	x	x	•	x
4		x	x	•	x	x	•
	or	•	•	x	•	•	x

In this case a spot will appear at a given level in the $\Delta\chi$ and ΔZ films if one or the other of the two three-impulse letters shown for that level is generated. This will be seen to be equivalent to working with two generlized impulses, 1 + 2, and 1 + 4.

(d) Generalised statement of principles.

The fundamental equation

$$\Delta\chi = \Delta Z + \Delta D \qquad (A1)$$

can be generalised by generalising the idea of an impulse.

Definition. A generalized impulse is any (non-null) sum of ordinary impulses. For example Z_{34}, i.e. $Z_3 + Z_4$, is a generalised impulse of Z. We refer to an impulse in this sense as an 'impulse' (in inverted commas).

If $\Delta\chi$ is considered as a letter in an n-'impulse' alphabet we write it as $(\Delta\chi)$. Similarly for (ΔZ) and (ΔD). Then (A1) can be written in the generalised form

$$(\Delta\chi) = (\Delta Z) + (\Delta D) \qquad (A2)$$

Let $L_1, L_2, \ldots\ldots L_r$ be letters in ΔD which have a frequency above random. Then

$$(\Delta\chi) \longrightarrow (\Delta Z) + (L_i) \qquad (A3)$$

These can be represented on 2^r levels of film. On one film we can record $(\Delta\chi)$ (a single spot) and on the other $(\Delta Z + L_i)$ (r spots for the r alternative letters L_i). (Or we can interchange $\Delta\chi$ and ΔZ and obtain an equivalent result.) Then if the two films are superimposed the number of coincidences of transparent spots gives us the number of occurrences of one or other of $L_1 \ldots L_r$ in ΔD.

The spots are made visible by shining a light through the film thus activating a photo-electric cell. A fundamental difference between the 5202 and Colossus or Robinson is that the 5202 measures the amount of light rather than the number of spots of light. If the amount of light exceeds a certain amount (cf. set total on Colossus) the position is indicated by a lamp. Later on, an exact count can be made of the number of spots of light in these positions by means of a special counter film (see below 91B(b).

(e) Example of conditional run.

Assume that χ_1 and χ_2 have been set and that we wish to make use of the evidence of ΔD_1 and ΔD_2 in doing a run to set the other three chis. The example considered is the run for the letters 5U/F in ΔD.

5	x x . x x
U	x x x . .
/	. . . . .
F	x . x x .

It is seen that a 5 or U can occur only if ΔD_1, ΔD_2, is xx, a stroke if ΔD_1, ΔD_2 is .. and an F only if ΔD_1, ΔD_2 is x. . The machine, which makes the film (the generator), has circuits available for each of the letters 5U/F which can be plugged to prevent any data appearing on film unless the required conditions of ΔD_1, ΔD_2 are satisfied

	ΔX Film			ΔZ FILM.										
	One spot in each group of 8 levels			When $\Delta D_1, \Delta D_2$ = XX, two spots in each group of 8 levels.						When $\Delta D_1, \Delta D_2$ = •• one spot			When $\Delta D_1, \Delta D_2$ = X• one spot			When $\Delta D_1 \Delta D_2$ = No spot
				S			U			I			F			
LEVEL	ΔX_3	ΔX_4	ΔX_5	ΔZ_3	ΔZ_4	ΔZ_5	ΔZ_3	ΔZ_4	ΔZ_5	ΔZ_3	ΔZ_4	ΔZ_5	ΔZ_3	ΔZ_4	ΔZ_5	
1	•	•	•	•	X	X	X	•	•	•	•	•	X	X	•	
2	X	•	•	X	X	X	•	•	•	X	•	•	•	X	•	
3	•	X	•	•	•	X	X	X	•	•	X	•	X	•	•	
4	X	X	•	X	•	X	•	X	•	X	X	•	•	•	•	
5	•	•	X	•	X	•	X	•	X	•	•	X	X	X	X	
6	X	•	X	X	X	•	•	•	X	X	•	X	•	X	X	
7	•	X	X	•	•	•	X	X	X	•	X	X	X	•	X	
8	X	X	X	X	•	•	•	X	X	X	X	X	•	•	X	

91B TECHNICAL ASPECTS

(a) The Film.

The actual width of the transparent spots is .006 inch. Hence if we only had sufficient levels on the film for one letter per column, a message of 5000 letters would require 2ft 6 ins of film whereas, as will be seen later, only three inches of film can be scanned at one time. Hence the film has 80 levels, divided by an opaque strip into two groups of 40, which if desired, can be used to record two different alphabets derived from the same stream. Thus with a 3-'impulse' (8 letter) alphabet, using the whole film we could record 10 consecutive letters of $\Delta\chi$ or ΔZ on one column of film, or using the top half of the film for one alphabet and the bottom half for another, we could record 5 consecutive letters. For a 4-'impulse' (16 letter) alphabet, the whole film would be necessary to record five letters. It would seem that with a 2-'impulse' alphabet, using only 4 levels, 20 letters could be recorded. In practice this is not so. The reason is that it is necessary to expose the film a column at a time. Hence a method of storing data is necessary, in order that all the data for one column should be transferred to film simultaneously. This storage circuit is only capable of storing data from 5 or 10 consecutive letters.

To summarise we are concerned with 3 cases.

Case (i) Two impulse run. Uses 4 levels. Two alphabets can be run, one on each half of the film. 10 positions recorded in each column.

Case (ii) Three impulse run. Uses 8 levels. Either 10 positions recorded using whole film or if we have two alphabets, one on each half of the film, then 5 positions recorded.

Case (iii) Four impulse run. 16 levels required for each position. 5 letters recorded in each column.

There is one further point to consider about the film. In preparing the ΔZ film, the initial letters will always be placed in the top set of levels. As it would be very inconvenient to have to give the films two dimensional relative motion in comparing them, it is necessary for all possible combination of chi settings to occur in the top level in order to try all possible initial chi settings. Hence it is necessary to repeat the complete $\Delta\chi$ stream, 10 times over (or if only 5 letters are recorded in each column 5 times over). If we record 10 letters in a column we strike a snag if $\Delta\chi_4$ is involved, since 26, the length of the wheel is not prime to 10. In this case it is necessary, half way through making the film, to move all the χ wheels one place in order that both odd and even positions of $\Delta\chi_4$ should occur in the top levels.

We will now consider the equipment of the 5202 in more detail. This is not meant to be a technical description, as one already exists in the American "reference Manual for 5202 Equipment". This account, which is only meant as a basis for explaining the cryptographic use of 5202, is based on the reference manual.

The equipment fallsinto three parts.

(1) The camera and target for preparing the film.

(2) The generating unit for controlling the target.

(3) The comparator for examining the two superimposed films in all possible relative positions.

(b) The Comparator.

This is the machine which examines the film in all possible relative positions and measures the number of coincidences between the film.

This is done as follows. The message film, of which 3 inches or 500 columns can be examined at one time, is kept stationary. The $\Delta\chi$ film is moved over it. Light is directed down on the two films in such a way that the light intensity is uniform over the whole message film. Then the light shining through the two films is proportional to the number of coincidences of transparent spots. This light is focussed on to a photocell. The photo-cell controls a light, which flashes on when the amount of light focussed on the photo cell exceeds a certain amount. The actual amount of light necessary to flash the lamp can be varied by means of two dials on the machine. It is not possible to correlate the number of coincidences on the film with the readings on these dials, as the photo-cells tend to vary. Hence it is only possible to cut down the stops to a pre-determined number and then record these. When we have cut the number of stops down to say three or four we can set the films to these settings as follows. We throw the automatic stop switch. When the $\Delta\chi$ film reaches a places where the lamp is flashed, the $\Delta\chi$ film is automatically stopped and reversed. It then runs back slowly until it reaches the hit when it stops again. Due to the momentum of the film it will not set exactly on the correct place. It is necessary to do the final setting by hand, using the flashing of the lamp as a guide. When the film is correctly set the lamp should remain on.

We can then throw onto a screen a magnified image of the identification strip (described in more detail in para (c)). This enables us to read the settings by seeing which letters on the $\Delta\chi$ film come against the arrow at the beginning of the message film. Because there is an identification letter only for every other column, it is important to see whether the arrow points to an identification or between two of them.

We can count the score by means of a special counter film: this has transparent spots 4 levels high, one and only one in each set of 4 levels i.e. one in levels 1-4, one in levels 4-8 and so on. These spots are spaced at a distance of 3 inches, so that only one of them is viewed at a time. The rest of the film is opaque but at one end a section is cut away except for a thin strip in the middle. This strip is in the same position on the film as the opaque sections (between lamps 40 and 41) of other films, making it possible to keep the counter film in the machine without interfering with setting operations. When we wish to count, this film is moved along under the other two films and light shines through all three only when the counter film spot is beneath a coincidence on the other two. The number of these coincidences can be counted on an electronic counter and this total gives us the score for the run.

(c) The camera and target.

The camera affords a means of exposing the film and synchronising the exposures with the generating unit. The film is moved on .006 inches after each exposure so as to be in a position to expose the next column in the next exposure. The camera has two rates of exposure: 1 per sec. or 10 per sec.

The "target" consists of a row of 80 lamps, divided into two groups of 40, one corresponding to each level of the film. These lamps are controlled by the generating unit, which selects which of the 80 lamps should light up. After exposure and development, the levels corresponding to lamps which have lit, have transparent spots produced on them.

In addition it is necessary to have some sort of identification on the films in order that it should be possible to read off the settings when running the film in the comparator. This is done on the chi film by photographing the chi settings onto the film, at the top of the column in five levels used for this purpose. In order to make the settings readable this is done for every other column. The settings are recorded on five counter wheels which are illuminated internally by a lamp flashing in phase with the chi wheels. The counter wheels move in step with the chi wheels themselves. In order that it should be possible to read this identification

through the message film another lamp photographs a transparent channel onto the message film in the same position as the identification on the chi film.

In addition another level on the message film is also transparent, but has a small arrow photographed on it at the beginning of the message and if required, at preassigned intervals throughout the message. The chi film has a clear channel in order that the arrow may be seen.

(d) Generating unit: χ wheels

The chi films are made with the camera working at an exposure rate of about 10 per second. Since in most types of run used each exposure records 10 positions of the chi wheels, the chi wheels must operate at the rate of 100 characters per second. The generator consists of 5 wheels geared to a shaft with ratios 41:10, 31:10, 29:10, 26:10, 23:10 i.e. in ratios proportional to the wheels lengths. In addition there is another wheel termed the master storage strip cam which is geared 1:1 to the shaft. Each χ wheel carries two wire brushes, one making contact with a metal disc connected to -310 volts, the other contacting one of a set of equally placed metal segments equal in number to the wheel length and corresponding to a position on the χ wheel. These segments are wired to the corresponding position on the χ pattern board I. This board is so plugged that a connection is made when there is a x in the $\Delta\chi$ wheel. Thus we get a signal of -310 volts if there is a cross in the $\Delta\chi$ wheel, and if there is a dot there is no connection through the plugboard, in which case the voltage of the line is -330 volts. Whatever wheels we require are connected through plugboard V to a set of five double triodes which have the effect of putting potentials of -175 volts on one line of a pair, and -300 on the other, if a dot signal is received. We will term a line active if it is at -175 and inactive if it is at -300. We will term the line of a pair which is active when a dot signal is received the dot-line and the other the cross-line. Then we have 5 pairs of lines, each consisting of a dot-line and a cross-line, which can be made to correspond to the five impulses of the $\Delta\chi$ stream. These 5 pairs of lines come out on plugboard III at the plug holes $P_1^-, P_1^+, P_2^-, P_2^+ \ldots P_5^-, P_5^+$. The $-$ corresponds to a dot-line and the $+$ to a cross line. There are 7 plug-holes for each line.

(e) Translating Circuit.

This translates a letter in teleprinter form (a pattern of r dots and crosses) into a "single spot in one of 2^r levels."

The 5 pairs of χ lines can be joined across to 4 pairs of lines which form part of the translating circuit. These 4 lines come out at the plug-holes $T_1^-, T_1^+, T_2^-, T_2^+ \ldots T_4^-, T_4^+$. We can therefore join whatever impulses we are interested in across to these lines. These lines constitute a resistor matrix which controls a set of 16 valves $T_1 \ldots T_{16}$ termed the translator tubes. These valves are each connected to one and only one line of each pair, this being possible in 16 ways. A translator tube operates if all the lines to which it is joined are active. Hence each possible 4 impulse character (here we mean teleprinter impulse) will operate one valve. The actual scheme is as follows.

T_1	. . . .	T_9	. . . x
T_2	x . . .	T_{10}	x . . x
T_3	. x . .	T_{11}	. x . x
T_4	x x . .	T_{12}	x x . x
T_5	. . x .	T_{13}	. . x x
T_6	x . x .	T_{14}	x . x x
T_7	. x x .	T_{15}	. x x x
T_8	x x x .	T_{16}	x x x x

If on the other hand we are only interested in 3 impulses then we can make both T_4- and T_4+ active by connecting both valves to -175v on Plugboard III. In this case each possible three impulse character operates two valves.

(f) Combining tubes.

The 16 translating tubes control the inputs to 16 further tubes termed

the combining tubes. The connection between the translator tubes and combining tubes is not fixed but is done on Plugboard II. Each of the inputs controlled by $T_1 \ldots T_{16}$ has plug-holes. The 16 combining tubes $C_1 \ldots C_{16}$ have 4 plug-holes. Those for $C_1 \ldots C_8$ are numbered $C_{1-i}, C_{2-i}, C_{3-i}, C_{4-i}$ $(i = 1-8)$ and those for $C_9 \ldots C_{16}$ are numbered $C_{5-i}, C_{6-i}, C_{7-i}, C_{8-i}$ $(i = 9-16)$. A combining tube operates unless one of the translator tubes to which it is connected is operated. The combining tubes act as part of the control of a bank of 80 pairs of thyratrons. These are arranged in 16 columns of 5 each, each column being associated with one combining tube. These thyratron pairs are associated in groups of 4, which receive a pulse from the master storage trip cam through a connection made on plug-board VIII. This pulse goes through a phasing switch which has 10 positions and comes out at the 10 plug-holes $P_1 \ldots P_{10}$. The groups of 4 mentioned above come out to the plug-holes $L_1 \ldots L_{20}$, on the same plugboard. The actual arrangement of these groups of 4 is shown on diagram III. These thyratron pairs correspond to the 80 lamps. The groups $L_1 \ldots L_{10}$ correspond to lamps 1-40 in that order and $L_{11} \ldots L_{20}$ to the lamps 41-80.

(g) Illustrative example.

The actual control of the lamps is best shown by a simple example. We will consider the case of a $\Delta\chi$ film in which $\Delta\chi_1, \Delta\chi_2, \Delta\chi_4$ are involved and the alphabet used is the ordinary 3-impulse alphabet, the impulses being ordinary teleprinter impulses.

In this case we put $\Delta\chi_1$ signals on the pair of lines $T_1^- T_1^+$, the $\Delta\chi_2$ signals on to the second pair of lines and $\Delta\chi_4$ on to the third pair. The fourth pair of lines are both made active. In this case two of the translating tubes will be operated by each impulse, one being in the set $T_1 \ldots T_8$, the other in the set $T_9 \ldots T_{16}$, the tubes operated being similarly placed in each set.

This run clearly uses 8 levels. Hence we need 8 lamps for each character and consequently 8 thyratron pairs to control them. Accordingly we connect the sets L_1 to L_2 to P_1, L_3 to L_4 to $P_2 \ldots$, L_{11} to L_{12} to P_6 L_{19} to L_{20} to P_{10}. The connection between translating tube and combining tube is simply straight across i.e. T_1 to $G_1, \ldots, T_{16}$ to G_{16}.

Now suppose we start the run. In the first position the master storage trip cam sends a pulse to the 8 thyratron pairs in L_1L_2. This puts the first thyratron in each pair in a condition to fire. The actual one that fires is determined by which combining tube is not operated and hence in this case, by which translating tube is operated. In the next position the first thyratron in one of the pairs in L_3L_4 is fired and so forth up to the tenth position. When the 10th impulse is received the information on the first thratron of the 80 pairs is transferred to the second of each pair, the first thyratrons now being ready to receive the data from the next 10 positions. When the camera is operated, it clears the second thyratrons and lights the lamps corresponding to the thyratrons that have struck. In this way the data is transferred to the film.

There is one further case to consider, and that is the case when 16 levels are required. In this case it is necessary to transfer the information every 5 instead of every 10 positions. This is done by commoning P_1 to P_6 ..... P_5 to P_{10} so that the special 10th pulse is received every 5th position instead of every 10th. In this case of course, 4 groups of 4 thyratron pairs have to be plugged to each P plug-hole.

(h) Message film. The message is recorded on six impulse tape. This is read on an ordinary tape reader and is then transmitted via a special deltaing circuit; from then on the process is essentially the same as the recording of patterns. The only important difference is that the exposure rate of the camera is only 1 per second.

(i) Conditional de-chis and G-circuits.

The case sometimes arises that we have set two of the chis and we wish to use this information to set the other three wheels. In this case we can actually add the $\Delta\chi$'s to the impulses which are set and thus obtain ΔD on these two pairs of lines. To do this we have to reset the chi wheels to their known settings. Since the wheels cannot be moved independently, the time taken becomes prohibitive if we deal with more than two wheels.

To consider the actual operation of the G-circuits it is easier to consider a definite example. The example considered is one that we actually used, that is, running for 5U/F when $\Delta\chi_1$, $\Delta\chi_2$ are set and we are trying to set $\Delta\chi_3$, $\Delta\chi_4$, $\Delta\chi_5$ (see 91A(e)). In order to do this we require a straight 345 chi film i.e. one where the impulses are the ordinary 3rd, 4th and 5th, teleprinter alphabets. The message film however is modified by the information we obtain from the de-chi on impulses 1 and 2.

It will be remembered that each of the combining tubes had 4 plug-holes for its input which were numbered, in the case of the tube $G_1(G_6)G_{11}, G_{21}, G_{31}, G_{41}$. So far it has not been necessary to distinguish between these 4 plug-holes. However in the type of run now being considered, these 4 plug-holes serve slightly different functions. Between the plug-hole and the cathode of the combining tube to which it is connected, is a relay. Normally, that is, when the G-circuits are not in use, this relay is closed. If however, the relay operates and thus opens the circuit, no signal is received from the translating tube, and the combining tube, unless prevented from operating by another translating tube connected via another plug-hole, will be in its normal, i.e. operating, state. The relays in all the lines terminating in plug-holes with the same lower suffix are all operated when a thyratron connected to all of them strikes. This thyratron, in turn, strikes unless a controlling triode conducts. The grid of this triode is connected on plugboard III to one line of either one or both pairs of lines, carrying the ΔD_1, ΔD_2 signal, and if either line to which it is connected is inactive, prevents the triode from conducting.

We can now consider in detail the plugging which effected the result shown in the example given. The letters 5U/F are controlled by G_1G_6, G_2G_7, G_3G_8, and G_4G_9 respectively by means of the ΔD_1, ΔD_2 signals through plugboard III. To obtain the ΔD_1, ΔD_2 signals we synchronize the chi wheels with the tape reader, preset $\Delta\chi_1$ and $\Delta\chi_2$ to the known settings and bring out the known ΔD_1 and ΔD_2 impulses to the pairs of lines terminating in P_4^+ P_4^- and P_5^+ P_5^- in plugboard III. The reason for bringing them out on these pairs of lines is that we can then put the 3rd, 4th and 5th impulses on to lines P_1^+, P_1^-, P_2^+ P_2^-, P_3^+ P_3^- and plug them straight across the top three lines of the resistor matrix, which simplifies the plugging on plugboard II from translating tubes to combining tubes.

On plugboard III all the plug-holes not marked correspond to the G symbol at the end of the row. Since 5 and U are controlled by G_1G_6 and G_4G_9 respectively, we join G_1, G_6 and G_4G_9 to the lines P_4^+ and P_5^+ by means of bottle plugs. This ensures that the relay will open unless the de-chi signal is xx and so spots will be put on the film only if $\Delta D_1\Delta D_2$ is xx. Similarly G_3 and G_8 are joined to P_4^- P_5^- and G_4G_9 to P_4^+ P_5^-. If now the signal received on lines P_4^+ P_4^-, P_5^+ P_5^- is .x, all relays operate and no spot appears.

Remembering that G_1 and G_5 control the letter 5, we plug from the T plug-hole to the G plugholes in the following manner to get the arrangement shown in the example given previously.

$T_1 \longrightarrow C_{1,7}$

$T_2 \longrightarrow C_{1,8}$

$T_3 \longrightarrow C_{1,5}$

$T_4 \longrightarrow C_{1,6}$

$T_5 \longrightarrow C_{1,3}$

$T_6 \longrightarrow C_{1,4}$

$T_7 \longrightarrow C_{1,1}$

$T_8 \longrightarrow C_{1,2}$

$T_9 \longrightarrow C_{5,15}$

$T_{10} \longrightarrow C_{15,16}$

$T_{11} \longrightarrow C_{16,13}$

$T_{12} \longrightarrow C_{15,14}$

$T_{13} \longrightarrow C_{15,11}$

$T_{14} \longrightarrow C_{5,12}$

$T_{15} \longrightarrow C_{5,9}$

$T_{16} \longrightarrow C_{5,10}$

In the resistor matrix the wiring of the first three pairs of lines of $T_9 - T_{16}$ is the same as for $T_1 - T_8$ so the plugging of $T_9 - T_{16}$ to $C_{5,9-16}$ simply puts data on the lower half of the film in the same way that the plugging $T_{1-8} \longrightarrow C_{1,1-8}$ does in the upper half. We are thus able to record data for ten letters in each column, five groups of eight levels in the top half and five more groups in the lower half. The plugging for $U_{\chi'}$ and F is simply a variation of the above.

(j) $\overline{\chi}_2$ limitation.

It is convenient to be able to consider only characters occuring against $\overline{\chi}_2$ x's. This is done by modifying the $\Delta\chi$ film when it involves $\Delta\chi_2$. The χ_2 pattern is set up reversed (i.e. interchanging dot and cross on plugboard I). This is read one back and a special suppressor circuit causes no light to light up if a cross signal appears i.e. $\overline{\chi}_2 =$.

91C TIMES AND ROUTINES

(a) Colossus time.

In order to understand the reasons for the uses we made of the 5202 machine it is necessary to consider how the machine compares with the other method of breaking Fish messages, i.e. by means of Colossus.

When a message is run on Colossus the following processes have to be gone through. The message is obtained in the form of 5-impulse tape and this has to be copied and stuck into a loop. This process, including checking takes an average time of 40 minutes for a message 5000 letters. The actual time for setting 5 chis on Colossus on a reasonably easy message is about 30 minutes. 4 minutes for a 1+2 break-in, 12 minutes for runs for the next two wheels, and 15 minutes or so, for the last wheel, letter counts, and checks.

(b) Time for 5202 processes.

For comparison, the times for the 5202 set-up are considered below. We usually run message films in groups of three in order to save time in the developing process as it takes no longer to develop 3 films on the same strip than to develop 1 film. The first process in England was to transfer the data from the 5-hole tape to the 6-hole tapes used on the Generator. This, presumably, would not be necessary in America. This was done before we started operations on the machine and so need not be counted in the time taken to run a message.

We will now give an estimate of the times taken to run a single film through the three parts of the process.

(i) The Generator.

The time taken for the generator to run through three message films of 5000 letters each seemed to average 40 minutes. As the actual speed obtained on the generator was 550 letters per minute, this gave a total running time of 27 minutes leaving 13 minutes for changing tapes, loading the camera, and resetting the dials after each message. This time seems reasonable.

In the case of a conditional de-chi extra time is needed to set the chi

wheels to the predetermined settings. This on an average seems to take about 5 minutes putting the time for three films up to 55 minutes. This, of course, assumes no mistakes by the operator or machine faults such as lamps burning out.

(ii) Developing

The times obtained here were not a fair sample due to the fact that the temperature control on our air conditioning system was not adequate. This meant that a considerable time had to be spent in bringing the various baths to the required temperature of 70°F ± 1°F. This usually took as long as 10 minutes. The actual developing time was 15 minutes. The other stage which seemed to take a long time was drying the film. It seemed to be necessary to keep the film in the drier for 30 minutes in order to dry it adequately.

(iii) Comparator

The time taken to run a 3-wheel film through the comparator, to set it and to count the score, seems to be about 10 minutes. This assumes good films and clear identification on the master film. One of our chief troubles with the comparator was the fact that the identification was often very nearly unreadable, with the result that settings could often only be determined by finding the nearest clear reading and calculating the settings, a rather tedious process. To sum up, the times that it seems reasonable to expect are:-

Generator	(i)	Break-in run 13-15 minutes per message
	(ii)	Conditional de-chi 18-20 minutes per message.
Developing	(i)	Actual developing 15 minutes
	(ii)	Drying 30 minutes. This should be reducible.
Comparator		Ten minutes per run.

Since each message will require at least one conditional de-chi in addition to the break-in run, it follows that at least 30 minutes of generator time, 30 minutes of developing time, 1 hour of drying time and 20 minutes of comparator time are required to completely set a message. Hence the number of messages that could be dealt with completely by the unit cannot exceed 2 an hour, and this will cause an accumulation in the drier. The time taken to set a message is 2 hours and 20 minutes as a minimum. This, as a whole, does not compare particularly favourably with Colossus as an operational method. But considered from another angle, the 5202 compares very favourably. This is the very large number of positions which can be examined by the comparator in a short time. The comparator can examine 2000 $\Delta\chi$ settings in one second. Colossus can only examine, with multiple testing, 5 $\Delta\chi$ settings in one second.

(c) Routines employed in practice.

Hence it seemed to us that the best use of the comparator was on 3-wheel or 4-wheel runs and we decided to start doing 3-wheel runs. The materials used were 2 months Squid traffic namely Squid of January and February, 1944. Both months were on χ_4 limitation. The January Squid had a motor with 26 dots and the February, 22 dots. Due to considerable trouble with the machinery we were not able to try as many messages of these months as we had hoped and in fact only 34 messages of January and 10 of February were attempted.

We decided on the following routine for these messages. The following films were made.

(a) "Standard" 1=2=4
(b) "Standard" 1=2=5
(c) Straight 345 film

The Standard films mentioned above have not yet been described, and it seems convenient to do so now. The standard message film was designed to do runs on 4 impulses for any set of letters which contained with any letter, the

complete opposite of that letter.

Hence the message film plugging was of the form

$T_1\ T_{16} \rightarrow C_1\ C_{16}$ $\quad$ $T_5\ T_{12} \rightarrow C_5\ C_{12}$

$T_2\ T_{15} \rightarrow C_2\ C_{15}$ $\quad$ $T_6\ T_{11} \rightarrow C_6\ C_{11}$

$T_3\ T_{14} \rightarrow C_3\ C_{14}$ $\quad$ $T_7\ T_{10} \rightarrow C_7\ C_{10}$

$T_4\ T_{13} \rightarrow C_4\ C_{13}$ $\quad$ $T_8\ T_9 \rightarrow C_8\ C_9$

The plugging for the 1=2=4, and 1=2=5 chi films was exactly the same but in the first case the fourth pair of lines are both made active and in the second case the third pair of lines are both made active. In these cases the first pair of lines carried $\Delta\chi_1$, the second $\Delta\chi_2$, the third $\Delta\chi_4$ and the fourth $\Delta\chi_5$.

The other standard film of the ordinary type which was made, was made to run 1+2 = . 3+4 = x This was run with a standard message film using impulses 1,2,3,4 instead of 1,2,4,5 as above. The plugging for the master was

$T_1\ T_{16} \rightarrow C_{15}\ C_{16}\ C_{13}\ C_{14}$ $\quad$ $T_5\ T_{12} \rightarrow C_1\ C_4\ C_9\ C_{12}$

$T_2\ T_{15} \rightarrow C_6\ C_7\ C_{10}\ C_{15}$ $\quad$ $T_6\ T_{11} \rightarrow C_2\ C_3\ C_{10}\ C_{11}$

$T_3\ T_{14} \rightarrow C_6\ C_7\ C_{10}\ C_{15}$ $\quad$ $T_7\ T_{10} \rightarrow C_2\ C_3\ C_{10}\ C_{11}$

$T_4\ T_{13} \rightarrow C_5\ C_8\ C_{13}\ C_{16}$ $\quad$ $T_8\ T_9 \rightarrow C_1\ C_4\ C_{11}\ C_{12}$

It was hoped that by making a single message film and running 1=2=4, 1=2=5 we could set with certainty $\Delta\chi_1$ $\Delta\chi_2$ and either or both of $\Delta\chi_4$ $\Delta\chi_5$. To do this it was essential that the counting of the scores obtained should be reliable, since it was necessary that we should have a numerical check of the significance of the score obtained. Unfortunately the counter was never completely reliable, chiefly due, I suspect, to faults in the film, such as stretching, imperfectly exposed spots on the film, and fortuitous spots on the film. Hence the only check on the runs which we would have had would have been agreement between the $\Delta\chi_1$ $\Delta\chi_2$ settings. In fact, we could always, if necessary, check the settings obtained as these messages were already set on Colossus and the Colossus dossiers available. In practice we found that the $\Delta\chi_1$ $\Delta\chi_2$ settings obtained were usually correct but that the $\Delta\chi_4$ $\Delta\chi_5$ settings were not always reliable. The next step was to run a conditional de-chi using the $\Delta\chi_1\Delta\chi_2$ settings obtained. The letters chosen for these de-chis were /5UF, one of the first two being usually the best letter in a 32-letter count on Squid and the other two being reliable letters. (It should be mentioned that Squid messages seem to be of two main types, one with high /'s and the other with high 5's. On both types of message U and F score fairly high.) This run was usually successful, when the film was made correctly. However difficulty in reading the identification film occasionally led to wrong settings being used for $\Delta\chi_1$ $\Delta\chi_2$ if the settings had not been checked against Colossus first. Of the 34 January Squid message tried, 24 came out without much trouble on the runs mentioned, 2 failed to set on χ_3 although the other chis set correctly, 6 were abandoned after unsuccessful initial runs and two were retried successfully on 4-wheel runs.

It had been thought that 4-wheel runs would be impossible since the length of the film for a 1245 run with the usual tenfold chi film would be 380 feet. Portunately, a way was found round this difficulty as follows. A 1245 chi film (with standard plugging) was made with every setting occuring somewhere on the film. The standard message film was then made ten times over starting at the first, second, third etc. letter in the message. Then the setting for the beginning of one of the films will occur at the head of a column of the $\Delta\chi$ film. As a result of this method of making films it took ten times as long to set a message on the comparator since we had to record at least one reading on each of the ten message films. The length of the $\Delta\chi$ film, however, was reduced to 38 feet, and the time taken to make the $\Delta\chi$ film worked out, in practice, at $2\frac{1}{2}$ hours instead of (theoretically) 21 hours.

The February Squid messages were not particularly notable and gave a certain amount of trouble frpm the machine point of view. However 5 of them

were set by the original method of running 1=2=4, 1=2=5 and 345/12.

One other class of message in which we were particularly interested was messages which set with certainty on two impulses, usually 1 and 2, but failed to set on any other impulses by the normal Colossus technique of one or two impulse runs, using the two known impulses. We obtained 4 of these messages on July, 1943 Squid, using the 1 and 2 settings given (which were certain according to our accepted conventions) and tried 3-wheel runs. Two such runs were tried, namely the run mentioned before, for /5UP and a run for letters which are strong on German language counts, namely 3JGP. These were not very successful. The second best score on one turned out to be correct but had been missed by Colossus. It seems that not much is gained by running for the last 3 wheels together but the size of the sample is very small and it seems hardly fair to damn the method on such faint hearing.

As the rest of the Fish section broke up 10 days before the end of the experiment, we decided to run a few unset messages of Squid of May, 1943, using Colossus to do letter counts to check the settings afterwards . We had 4 of these messages, all 4 being set on a 4-wheel run followed by a 3-wheel run using the 1 and 2 settings. The 4-wheel run used for these messages was 1=2=4=5.

Lastly a few Dace messages of December, 1943 were tried on 4-wheel runs using the run which seemed best on the Dace type of message namely 3+4 = x, 1+2 = . Only one was successful, this being a message which set with difficulty on Colossus, but set with ease on the 4-wheel run. In addition a run was done to set a crib. This is the subject of 91D.

To finish, here is a complete summary of the time taken to make various films.

Standard Message Film	13 minutes
Conditional de-chi	18 minutes
Standard 1=2=4 chi film	1 hour and 5 minutes
1=2=5 chi film	55 minutes
Standard 1=2=4=5 chi film	2½ hours
Standard 1+2, 3+4x chi film	3 hours

91D CRIB RUN

The crib chosen was a Jellyfish-Gurnard combination of May, 1944 which had previously been set on Robinson. Standard three impulse running tapes were made on Mrs. Miles by the method independent of limitation giving

$$\left.\begin{matrix}\Delta_{713}\\ \Delta_{447}\\ \Delta_{595}\end{matrix}\right\}\begin{matrix}(Z_2 + Z_5)\\ (Z_3 + Z_5)\\ (Z_4 + Z_5)\end{matrix}\qquad\begin{matrix}\text{in the second impulse}\\ \text{in the third impulse}\\ \text{in the fourth impulse}\end{matrix}$$

for the cipher tape (Z^*) and an identical set up for the plain language (P^*). As boards were plugged to put five letters of text on the films per exposure, we got all possible starting positions by copying the plain language (3259 letters of text) five times and filming the resulting tape. Identification was put on this film by engaging the "master-daily" clutch.

Boards were plugged in such a way that a coincidence in the upper half of the film, when films were superimposed, meant . . . in $\Delta P^* + \Delta Z^*$ and, in the lower half, a coincidence meant one of the letters . . x, . x . and x . . We were thus able to make a simultaneous run for all the letters of $\Delta P^* + \Delta Z^*$ which are expected to score well in a crib run of this type.

There was a single big hit on the comparator and a count that compared very well with the Robinson run. The score for . . . on the high side was indeed only four away from the original. The identification of the position at which the plain and cipher set relative to each other was not

entirely satisfactory since we arrived at a calculated setting of 973 compared with 993, the correct position. The films, however, were faulty inasmuch as the identifying arrow could not be seen, so the reading of the identification was guess work to some extent.

91E CONCLUSIONS

The machine is clearly a very good method for dealing with cryptographic problems at very high speeds. It chief trouble, as far as Tunny traffic was concerned, was lack of flexibility. This was not important in the case of the normal settable message but unfortunately a large proportion of messagesare not normal. Thus the standard procedure which we adopted could set messages which conformed to the usual long supply reports with a considerable amount of punctuation,but would fail on a message which was say, an appreciation of the chances of an expected operation, in which case strong language differences would predominate and different types of run would be successful. On Colossus it is very easy to try both hypotheses, but at the moment to try both hypotheses on the 5202 requires the making of two films. It would be possible to use one film if all the data from the message tape could be recorded on one film, but this would require 32 levels. This could be reduced to 16 if we only considered runs for groups of letters defined by relations of the type i + j = dot or cross, but this is not always a good method of setting the last wheel or wheels (e.g. a run for /8 is not usually as good as a run for /).

The other place where lack of flexibility was apparent was in the G-circuits. When the whole film was being used, only 4 of these circuits were available, unless we were prepared to cut the effective text down to 2500.

It seems that if the 5202 had been available whilst Tunny traffic was still acti it would have been of immense value for setting, by 3- or 4-wheel runs, messages whic failed on Colossus, but the setting of the remaining wheels would have been completed on Colossus.

In χ-setting the advantages of Colossus over 5202 include :-

1. Flexibility
2. Shorter time of preparation
3. Maximum message length 30,000 instead of 5000
4. Letter counts
5. Spanning
6. Not 99

χ-setting is only one of the Tunny-breaking operations performed on Colossus: The following are impossible or impractible on 5202:

7. χ-breaking
8. Rectangling
9. Ψ-setting

The lack of these facilities is due in some cases to inherent characteristics of the 5202 principle; in others their provision would require a very precise technique. At the other extreme spanning could easily be improvised.

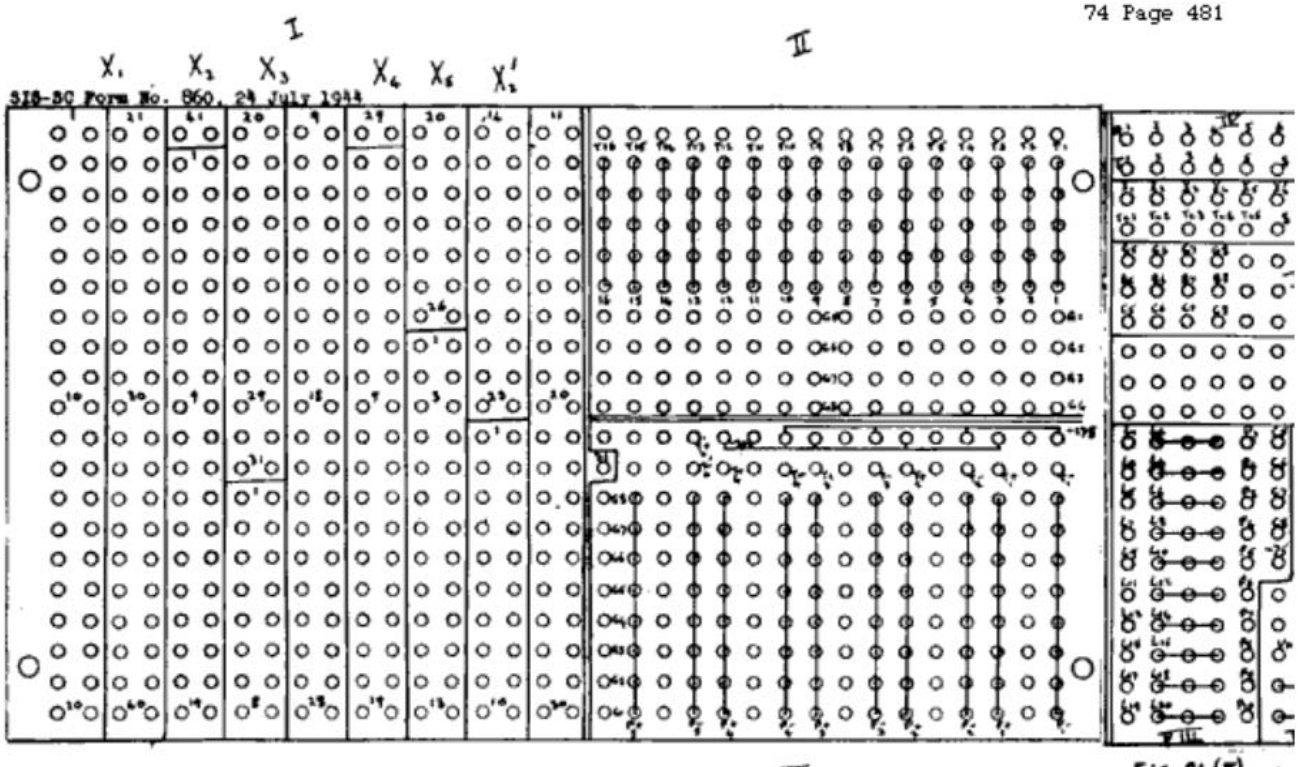


FIG. 91 (I) PLUGBOARDS I — III

FIG. 91 (II). PLUGBOARDS IV

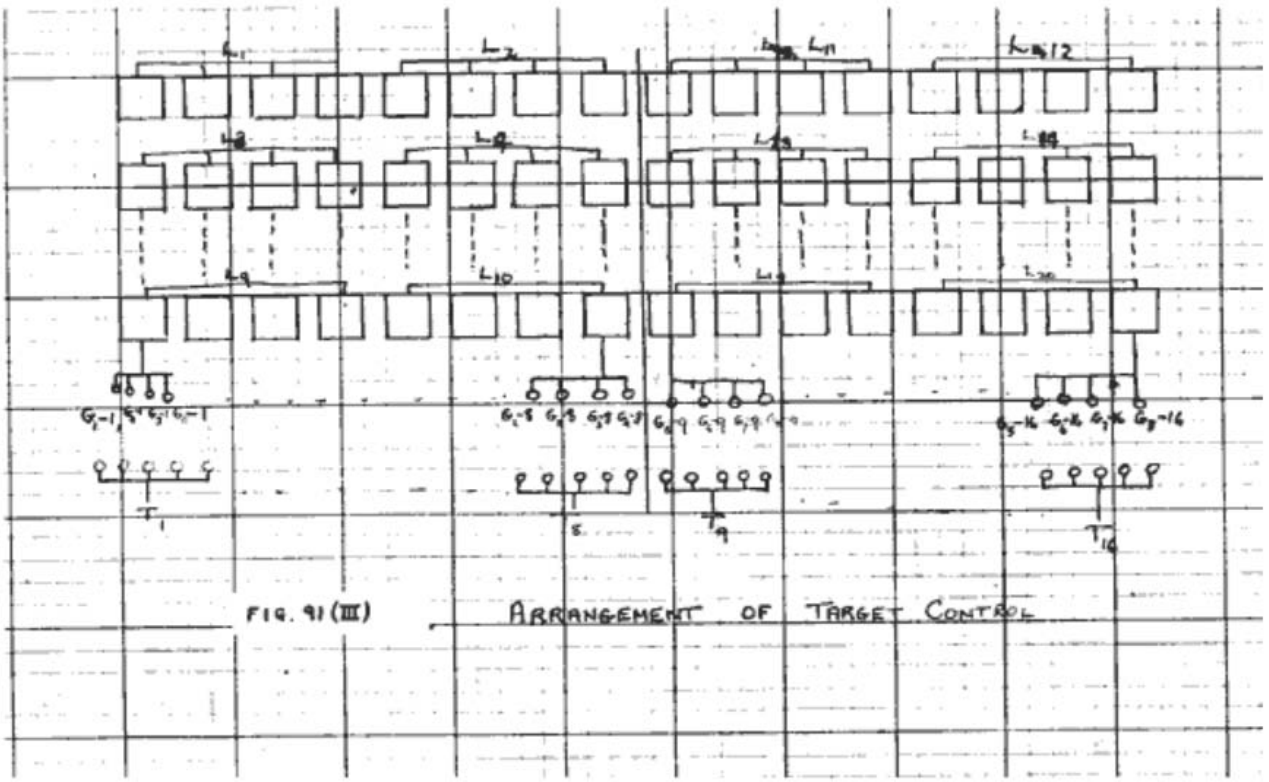


FIG. 91 (III) ARRANGEMENT OF TARGET CONTROL

92 RECOVERY OF MOTOR PATTERNS FROM DE-CHI

A. Introduction and outline

An account has already been given (28D) of the methods of recovering the patterns of μ_{37} and μ_{61} from a stretch of Ψ'. In that case a length of about 500 is needed, and eight lines of the motor rectangle are filled in; each square contains one of three entries - dot, cross or blank. The methods of recovering depend mainly on two ideas: first that consecutive columns of the complete motor rectangle match perfectly when compared level, if the earlier comes under a μ_{61} dot, and, when compared at constant slide (dependent only on the dottage of μ_{61}), if it comes under a μ_{61} cross; secondly that any column matches perfectly, if slid one down, with some column between 20 to the left and 10 to the right, and that the distance away of this repeat column is determined to within a range of about 4 by the dottage of μ_{61}.

The following account shows how motors can be recovered from ΔD. Probability methods must be used, and the dots, crosses and blanks in the $\Delta\Psi'$ motor rectangle replaced by deciban scores in favour of dot or cross, (with the prior decibanage not added in). The sort of length required appears to be about 8,000; additional lengths will nearly always be worth including. The count of ΔD is examined and an estimate made of the ΔP count underlying it: from this estimate it is possible to work out, for any given letter of ΔD, the probability that it comes against a basic motor dot, or rather the factor that the occurrence contributes to the odds on that character being a dot. These 32 factors are expressed as decibanages.

If we have a length of 10,000, any given character of the motor rectangle (which has 2257 squares) will occur 4 or 5 times in the de-chi and so there will be 4 or 5 letters of ΔD contributing factors to the prior odds on that character being a dot. All this evidence, in deciban form, is put together and entered into a 61 x 37 rectangle.

The basic operation, from now on, is to see whether two columns match well at a given slide: in the $\Delta\Psi'$ motor rectangle the match is either admissible or inadmissible; in the present case the answer is a decibanage in favour of the match (except, again, that the prior decibanage is not usually included). The scoring is done from a table and will be described in detail later. Such small experience as there is available, indicates the method of finding repeat columns to be less powerful than that of finding the slide at which adjacent columns are to be matched. The reason seems to be that this slide is uniquely determined by the μ_{61} dottage, but the distance between repeat columns is only determined to lie in a range of about four; since the decibanages in favour of slides are small compared with the prior decibanage, it is important to be able to collect a large number of scores for a given hypothesis.

Once the dottage of μ_{61} is established, it is a simple hand process to construct an approximate μ_{61} and an approximate μ_{37}. The hand work should be checked on Colossus: indeed Colossus gives the final scores for μ_{37} in a manner not very liable to error, and the possible variation of μ_{61}, can be fairly conveniently assessed. If certainty is reached for both patterns, well and good; if not, there should be little trouble in setting the probable patterns on another de-chi and using this message to settle the doubts.

When there is χ_2 limitation or no limitation at all the total motor is now known and the Ψ's can easily be recovered on Colossus by normal wheel-breaking methods. If the limitation is more complex, the natural way to recover the Ψ's would be to have recourse to the usual method of Ψ breaking from de-chi (28C). this is made considerably easier by the recovery of the basic motor.

Before work is started, some assessment can be made of the probability of success. Setting aside limitations on the number of dots in μ_{37} and μ_{41}, there are 2^{78} different pairs of patterns; so the prior decibanage of a given pair is 295. The decibanages in favour of a hypothesis in virtue of a score of 3σ above the average is approximately $2{\cdot}17s^2$; consequently a pair of patterns scoring 12σ is rather better than evens. Now it is a routine operation to estimate the expected σ-age of a motor run. Let this be done for all letters. The sums of the squares of the σ-ages on the separate letters gives the square of the σ-age for the correct patterns.

Sometimes a single letter in ΔD can be sufficient, as in the motor broken in August, 1943 (MO 11, 14 etc). This simplifies the work a lot.

8. Decibanage of ΔD letters

The first job is to determine the decibanage contributed by each of the 32 letters of ΔD to the hypothesis that the basic motor underlying it is a dot. The problem is slightly different with $\overline{\chi}_2$ limitation and not $\overline{\chi}_2$ limitation.

(a) With $\overline{\chi}_2$ limitation

Places where $\overline{\chi}_2 = \cdot$ provide no evidence about the BM and we need only consider places where $\overline{\chi}_2 = x$

Let $37D$ = Number of dots in μ_{37}

C = Number of crosses in χ_2

$N_\alpha^x, N_\alpha^\cdot$ = Number of occurrences of α at $\overline{\chi}_2$ = cross, dot.

Then, by regarding the $\overline{\chi}_2$ dot positions as a sample of what happens at motor crosses (e.g. 23), the factor in favour of a BM dot can be seen to be

$$\frac{1}{D}\,\frac{N_\alpha^x}{N_\alpha^\cdot}\quad\frac{31-C}{C}\quad\frac{1-D}{D}$$

We can estimate D from the relations

$$\frac{P.B.(\Delta D_{ij} = \cdot \mid \overline{\chi}_2 = x)}{P.B.(\Delta D_{ij} = \cdot \mid \overline{\chi}_2 = \cdot)} = \frac{2-\beta}{\beta}$$

Both formulae are liable to error due to random variation and prior knowledge must be taken into account.

(b) With other limitation, or no limitation

Let n_d = number of occurrences of d in ΔD

T = Text length

$\phi_d = P(\Delta D = d \mid T.M. = x)$

Then factor in favour of BM dot is

$$\frac{n_d}{T D \phi_d} - \frac{1-D}{D}$$

$\phi_d = \frac{1}{32}$, with an adjustment that depends on how good are the

complements and near complements of α (with respect to δ). The adjustment is made by judgement - there is in any case no theoretical solution without using prior knowled e (see e.g.R3, 48).

C. Construction of Motor Rectangle

The most thorough method of transferring the available information into the rectangle would be something of this kind: - Print out the delta D in widths of 2257: if there is χ_2 limitation, strike out those characters occurring against a χ_2 dot; using the decibanages already worked out for each letter, total up the decibanages for each of the 2257 columns; enter these in a 61 x 37 rectangle writing them in horizontally from the top left hand corner in 37 rows of 61. When this has been done what appears in a given square is the decibanage in favour of that character being a dot (except that the prior decibanage has not been added in.)

There is a short cut to this thorough method, making use of the plugging facilities on Garbo. The decibanages for ΔD can be taken to the nearest deciban or half-deciban; and instead of printing the letter itself, the plugging can cause the machine to print the score in the unit chosen; the ordinary numbers can be used for negative scores and A,B,C... for ①, ②, ③ .. This makes the totalling in the 2257 columns a great deal easier and sacrifices little in accuracy. If there is χ_2 limitation, it will still be necessary first to strike out the scores opposite χ_2 dots. Of course the machine does not print in widths of 2257 but in widths of 61 so that there are 37 rows. The technique of operating Garbo for this job has been described in the section describing the making of Garbo rectangles (24).

D. The Scoring of Columns against each other.

It has already been stated that the elementary operation that underlies all attempts to recover μ_{37} and μ_{61} is that of comparing two columns against each other at some slide, and of determining how likely it is that the true columns have a perfect match at that slide. Any such comparison consists of making 37 separate comparisons of the factors in the first column against the corresponding factors in the second. f_1 and f_2 may be regarded as a typical plan of corresponding factors. Consequently the first thing to do is to find what factor is contributed to the odds on the slide being correct by the pair f_1 and f_2.

Let the probability of observing a factor f_i in a square that is really a cross be ϕ_i; then the probability of seeing it in a square that is a dot is $f_i \phi_i$.

Then $$P(f_1, f_2 / \text{slide correct}) = D f_1 \phi_1 f_2 \phi_2 + (1-D)\phi_1 \phi_2$$

and $$P(f_1, f_2 / \text{slide incorrect}) = D^2 f_1 \phi_1 f_2 \phi_2 + D(1-D) f_1 \phi_1 \phi_2 + D(1-D)\phi_1 f_2 \phi_2 + (1-D)^2 \phi_1 \phi_2$$

$\therefore$ factor in favour of slide being correct given f_1, f_2

$$= \frac{D f_1 f_2 + (1-D)}{[D f_1 + (1-D)][D f_2 + (1-D)]}$$

Let $d = 10 \log_{10}(f)$ or $f = 10^{d/10}$

Let $10 \log_{10}\left[D\, 10^{x/10} + (1-D)\right] \equiv \theta(x)$

Decibanage in favour $\equiv \phi(d_1, d_2)$

$= \theta(d_1 + d_2) - \theta(d_1) - \theta(d_2)$

(See RO, 63) $\phi(d_1, d_2)$ can now be tabulated.

E. The Recovery of patterns (A). Finding the dottage of μ_{61}.

In our present state of inexperience, it would be absurd to dogmatise on the subject of techniques. One point that seems clear is that the first thing is to determine the dottage of μ_{61}, and then the finding of the approximate patterns is easy. In this section some account will be given of four methods of finding this dottage.

(a) The method of determining the slide between adjacent columns.

This is the simplest of the four methods and probably the most reliable. The general idea is that every pair of columns has a correct match; for a pair under a μ_{61} dot this match is level; for a pair under a μ_{61} cross it is at a slide determined uniquely by the dottage of μ_{61}.

Let us adopt a convention about slides; a slide between a column n and a column $(n+a)$ in which the $(S+1)$st character of n is opposite the first of $(n+a)$ is called a slide of S. In this paragraph 'a' is always 1 because we are comparing consecutive columns; but later we shall want to compare a column with another several places to the right.

It can be proved without difficulty that, if C = number of crosses in μ_{61},

$$SC \equiv 1 \pmod{37}.$$

This method consists of comparing a large block of pairs of consecutive columns at all possible slides and entering the scores in a 61 x 37 rectangle. The slide of S between columns n and $(n+1)$ is scored and the score entered in the sth row of the nth column. It will be most worth while comparing columns which score badly as a level match, because for these the chance of a μ_{61} cross is high.

In this scheme of entering, the last row of the rectangle contains all the scores for level comparisons. To estimate the relative probabilities of two slides, we could add up the scores for each slide and see by how much one exceeded the other. This would be accurate only if μ_{6} had no dots in the places considered, and better methods can be devised.

After about 15 of the columns of this slide rectangle have been completed (particularly if the columns with large negative scores for the level comparison are selected), it should be possible to narrow down the preference to three or four slides. Not all of these need correspond to admissible dottages. (It would be possible, of course, not to make the comparisons for inadmissible dottages, but the job is a mass-produced affair and probably the time saved by ommitting some of the rows would be lost in the mistakes arising). The best admissible slides can now be scored in some more of the

columns - if necessary in all, - and in this way there seems to be an excellent chance of finding with great confidence the correct slide (and hence the correct dottage of μ_{61}).

(b) The Method of Repeat Columns.

The general idea will be familiar from the account of ordinary motor-breaking : it is that of a search for the approximate distance away of the column which is a match at a slide of -1. This distance is only roughly determined by the dottage of μ_{61} : the formula is :-

Dottage of μ_{61} given a repeat column at a distance x to the right, is approximately

$$\frac{61(x+24)}{x+61}$$

This will not be wrong by more than 2, (c.f. R0,69).

The scheme of work recommended is to compare every column at a slide of -1 with every admissible column, i.e. from 20 to the left to 10 to the right. These scores should be entered in a square 61 by 61 : the score between column n and column n + x at a slide of -1 is entered in the nth column and the (n + x)th row. In this way there will be at least one 'correct' score in each column and in each row : the presence of μ_{61} dots can increase the number from one even up to three or four. The 'correct' scores will form a connected pattern (allowing diagonal connection) and will lie in a band going down diagonally to the right : the width of the band including them will be 4 or 5 squares.

The plan would be to guess what that band approximately is, by noticing the occurrence of very high scores : when the band has been settled, the dottage of μ_{61} should be clear with a possible error of ±1. The candidates can then be tried out by the basic method of (a), and (it is hoped) the correct one established.

To put this method in perspective, - it is hoped by doing rather less work than that entailed in (a), to get an approximate μ_{61} dottage; and, with that information, to use the method of (a), to distinguish between these candidates.

It should here be mentioned that in extremely favourable cases something more could be done. When the 61 x 61 rectangle has been filled with all the relevant scores, the attempt can be made to trace the pattern of 'correct' scores. This pattern is determined uniquely by the μ_{61} pattern (and is quite uninfluenced by the μ_{37} pattern); each character of μ_{61} influences the pattern of 'correct' scores in two places, and judicious use of this fact might enable one to construct a nearly complete μ_{61}. However, even at that stage the dottage of μ_{61} may conceivably not be uniquely settled.

(c) The method of looking for good slides.

Select a column that has a lot of large decibanages, or a pair of consecutive columns which score so well level that a μ_{μ} dot can be assumed. Try this column (or the composite column got by adding the scores of the pair together) against all other columns at all slides. The idea of this method is that it should be done unsystematically; the good portions are selected by eye and scored and recorded. For any given slide at a given distance, it is possible to narrow down the dottages of μ_{μ} that could reasonably have given rise to it. This is done for all the good positions recorded, and a dottage of μ_{μ} is credited with any of the good scores that it could reasonably have picked up. In this way a total score is given for each

possible μ_{61} dottage and preferences between the dottages now exist. The final choice can be made by the method (a).

This method has had a success (RO, 11, 14).

(d) The smooth μ_{61} method.

For all conceivable μ_{61} dottages make up a μ_{61} pattern or a partial μ_{61} pattern, with the correct numbers of dots evenly spaced. Assuming that pattern correct, add together all the columns and observe the sum of the moduli of the scores. The hope is that the smooth μ_{61} constructed for the correct μ_{61} dottage will sufficiently resemble the true μ_{61} for the μ_{37} scores to be significantly high. If this hope is fulfilled the μ_{61} dottage immediately emerges.

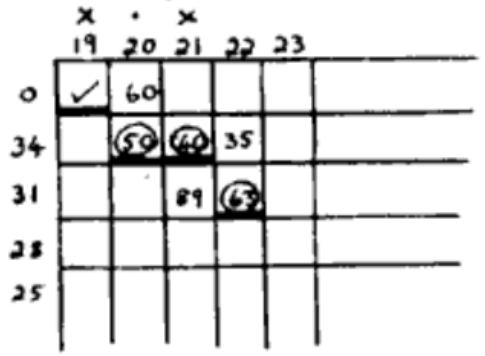
This method has the advantage that it can be conveniently done on Colossus. For every assumed μ_{61} dottage a smooth μ_{61} is plugged up and wheel-breaking run(s) are done for μ_{37}. The correct assumption is picked out by the significance of the wheel-breaking runs. This is identically the same as the hand process of adding together all the columns, described above. When the dottage of μ_{61} is selected, a μ_{37} can be put up and with the use of that, μ_{61} can be improved; and so on backwards and forwards between the two wheels.

Several smooth μ_{61}'s can be tried in the time taken for any hand process.

F. The Recovery of Patterns (B). The approximate μ_{37} and μ_{61}

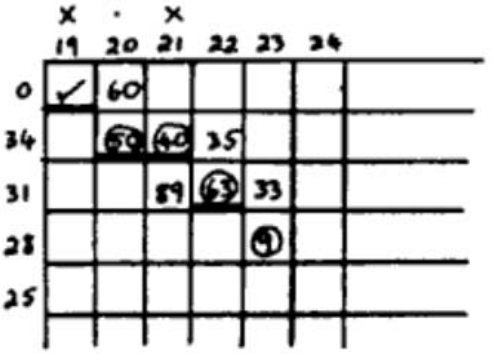
It is now assumed that the dottage of μ_{61} has been established. The μ_{37} and μ_{61} can now be worked out approximately by a 'snaking' process, similar to that sometimes used in anagramming a de-chi after the Ψ's have been set. Select a few columns where the μ_{61} pattern can be written in with confidence : such a patch will almost certainly exist if method (a) or (b) has been used. The slide S between consecutive columns is now known; so that the left hand column of the patch can be used as a start and the scores of the other columns of the patch can be added to it at the slides determined by S and the partial μ_{61} pattern.

Now make a rectangle 61 x 37; label across the top with the column numbers starting with the selected column, and down with the numbers 0, 5, 25.....365, (all reduced modulo 37) which stand for the slides of the selected first column against the other different columns). Suppose there are 12 dots in μ_{61}, this means that $5 = 34$; and that it can be assumed that the 19th character of μ_{61} is a cross, the 20th a dot and the 21st a cross; let the evidence for these three characters be scores of (50), (40), (63) as opposed to 60, 89, 35, for the contrary hypotheses. These facts can be entered in the rectangle as shown :-



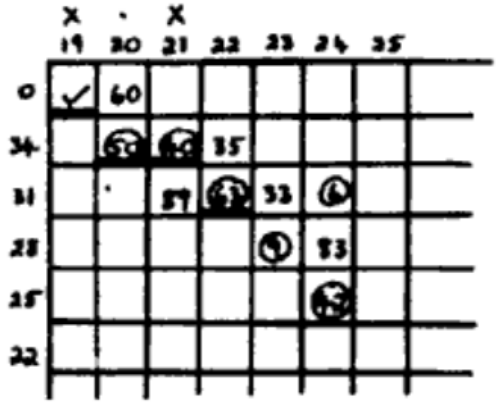
We now have a composite column made from the scores of columns 19,20,21 and 22 at their correct slides (of 0,34,34 and 31); this is our best approximation so far to the true pattern in a column. There are only two possible slides at which this should be compared with column 23; if the 22nd character of μ_{61} is a dot, the slide is 31, if it is a cross the slide is 28. Make these two comparisons and enter the scores in the rectangle.

Suppose we now have:-



This gives an advantage of 42 centibans to the hypothesis of this μ_{61} character being a cross. It already had the advantage of 60 centibans (because μ_{61} has 12 dots and 49 crosses) so that it would be reasonable to accept it. But if that is felt to be a risk, then the column already used can be compared at slides 31,28 and 25 against column 24.

This might give:-



Now there are two characters of μ_{61} in question, the 22nd and 23rd. The most probable are clearly xx which gain (51) centibans from these scores; the next best is . . which loses 27. This advantage of (78) centibans combined with the prior advantage of (120) centibans is overwhelming. When this has been decided the μ_{61} characters can be filled in and the accepted scores underlined in red; and, the most important, the two new columns can be added into the composite column at the slides now determined. There is one point of apparent discontinuity in the path of 'correct' scores; this occurs between column 61 and column 1. It is advisable to draw a heavy line down between these columns to avoid having this fact forgotten. If the correct slide for the 61st column is K, then the slide against the first column is $(K-1)$ if the 61st character of μ_{61} is a dot, and is $(K-1+5)$ if that character is a cross.

Doubts as to the exact cause of 'correct squares' can be left until the snake is completed and μ_{37} is nearly certain; indeed it may not be possible to settle such a doubt on the message itself. It is a check that the right number of dots in μ_{61} have been taken that the snake leaves the rectangle eithe through its bottom right hand or its top right hand corner. The μ_{61} pattern will have been constructed during the operation and the μ_{37} can be written in from the decibanages collected from the cumulative total of all the columns. It is reasonable to hope that there will be only two or three doubts in either

wheel. The pattern deduced from the composite column will not be the μ_{37} itself but will have to be disentangled. The details of the disentangling depend on the value of S, or on the dottage of μ_{61}. If there are C crosses in μ_{61} then the first character in the column labelled 1 is the first character of μ_{37}; the second character in this column is the $(C+1)$st character (of course reduced modulo 37); the third is the $(2C+1)$st and so on.

6. Finishing off the μ's

Once the μ_{61} pattern has been approximately established the rest of the work can be done on Colossus. Set up the μ_{61} pattern as near as may be: it will have to be remembered that the hand work did not start from the first column of the motor rectangle; also set up the μ_{37} pattern, about which a necessary warning has been given at the end of the last paragraph. Do a count against basic motor dots (also against $\overline{\chi}_2 \cdot \chi$ if there is the χ_2 limitation) and against basic motor crosses (or against $\overline{\chi}_2 \cdot$ dot if there is χ_2 limitation). From these counts the letters can be re-decibanned with a considerable improvement in accuracy over the original estimates. From this new decibanning the letters can be grouped for further wheelbreaking runs for μ_{37}, - in just the usual manner of grouping letters in wheelbreaking. Do wheelbreaking runs for μ_{37} in the way to be described, decibanning the runs and totalling the scores, and so get a more accurate approximation to the wheel than the hand methods gave.

In doing a wheelbreaking run for μ_{37} first count the score with a trigger consisting only of dots - giving R. As usual put a cross in the last position and do a run. The 37 scores must all be subtracted from R to give the scores for the various positions; and, as a check, the total of the 37 answers must be R.

In selecting the groups of letters it is important to include all letters except those scoring almost nothing: unless a correct balance is preserved between positive and negative letters there may be confusion about the number of dots in μ_{37}. For instance, if no negative letters are run for, all scores will be positive and the distinction between dots and crosses will only be that the scores against the dots are much larger than the cross scores.

μ_{61} can also be confirmed in a slightly less convenient way. Use the new μ_{37} and select the five groups of letters that give the best decibanages; count them on the five different counters. Now vary the position of the dots in μ_{61}. The correct μ_{61} must be reachable from the one set up by a comparatively small number of moves, each of which consist of moving a dot one place to the right or to the left. If this series of moves is made in a sensible manner, after each operation the basic motor pattern resembles more closely the true pattern. Consequently, a test of the μ_{61} is afforded by moving all dots in either direction and seeing whether the score improves. So, first count the five groups with the supposed μ_{61} pattern; list all possible dot moves (there will be twice as many as there are groups of dots in the original μ_{61}); perform these moves in order, counting the five scores, with each move and remembering, between moves, to return the dot to its original position. Comparison of the sets of the five scores, with the use of decibanning of the groups, leads to a decibanage in favour of each of the possible moves. If some dot move seems clearly to be an improvement, then it is necessary to continue that move, or (more generally) to try all dot moves newly made available. This does no more to μ_{61} than can be done by hand, but it easier to do it accurately.

It is quite likely that the patterns will remain in doubt at the end of these operations; it would be a mistake to spend too long on refinements on one message before trying to set the other messages on the presumed patterns. Once a new message had been set, the additional decibanages can be added in.

(H) Recovery of the Ψ patterns

This is a mere appendix but it gives completeness. It is well, before attacking the Ψ's to think whether any χ was at all doubtful; in particular a wrong χ_2 makes the Ψ's unobtainable by statistical methods. Any χ thought to be worth confirming can be run for against basic motor dots.

If there is some limitation other than χ_2, no workable proposals have yet been made for statistical attack on the Ψ's. For instance, if there is $\chi_2+\Psi_1'$ limitation, it is necessary to know the complete Ψ_1 before the motorisation is known. In cases of this kind the best plan is to print out the de-chi with basic motor above it and hope that the Ψ's can be broken by the usual method of guessing the plain language. The basic motor will certainly be of great assistance.

If there is no limitation at all, or χ_2 limitation, then the total motor is known, and the Ψ's can be broken on Colossus without difficulty. The ordinary short runs for Ψ's (P_1 = ., P_2 = ., P_3 = x , P_4 = . , P_5 = .) can be tried as wheelbreaking runs; it will be very surprising if none of these is significant. If that should happen, and if careful checking disclosed no new doubts or mistakes it would still be possible to converge the P 1 +2 rectangle for Ψ_1 and Ψ_2 from a random start. The only difficulty is the technical one, that there is no facility on Colossus for doubting a Ψ wheel. However P 1 + 2 is so powerful that the convergence should not be much held up by the necessity to take a complete wheel each time.

(I) Example of method (b)

For the interest of comparison the work of method (b) was done (later). The method was not found strong enough, on its own, to determine a substantially right μ_{61}. The best use of the method would have been to stop after about the 20th column, by which time the μ_{61} dottage had been narrowed down to 13, 14, or 15. These could have been tested out by method (a).

(K) Experiment in recovery by method of the smooth μ_{61}

There are here, as with other methods, two distinct phases of the work. The first phase is devoted to determining the dottage of μ_{61} and the second to the recovery of the patterns. The distinctive feature of this method is that it is done entirely on Colossus although it would be possible after the successful completion of phase one, to return to hand methods for the patterns themselves. The results and times seem to encourage the belief that machine methods are the most promising.

An experiment was done on a message KOA 4400 of length 7005. It was known to have no limitation and 11 dots in μ_{37}. It was reasonably certain that the μ_{61} dottage lay between 11 and 19. These facts were known because the message was of a date when the monthly keys were used, and before the introduction of limitation. An earlier experiment had been done with 31 dots in μ_{61}. This made the earlier experiment more difficult, since the smooth μ_{61} is a more powerful method when the μ_{61} dottage is not close to $\frac{1}{2} \times 61$.

The message TOA 4400 had an expected 17 sigma for the motor run on ///, so only /// was used at first.

To estimate the significance of μ_{37} runs, a crude approximation was used. If the number of occurrences of the letters used for a run is R, these R occurrences will be spread over the 37 places; so R/37 is used as a norm and subtracted from each of the 37 scores. The resulting numbers are treated just as the numbers in a wheel-breaking run, i.e. their absolute values are summed. This sum is considered as having an expected score of $\cdot 8\sqrt{37R} = 4{\cdot}9\sqrt{R}$ and a standard deviation of $\cdot 6\sqrt{R}$. If it exceeds $6{\cdot}1\sqrt{R}$ it is considered to be significant. This ignores the fact that the distribution is Poissonian rather than normal, and that the sum of the scores is restricted, - (in fact they must add up to R.) It is thought to be a reasonable approximation becoming unreliable as R/37 decreases. The accurate test would be to sum squares of the scores and apply the χ^2 test, but that is in practice laborious. (RO, 65).

For each μ_{61} dottage from 11 to 19 a smooth μ_{61} was chosen. With this pattern a wheel-breaking run was done on /// for μ_{37}. In order to try shorter patches of smooth motor, the doubting trigger was set up with crosses in the last 40 places; runs were done on the remaining 1/3 of the wheel set at 01, 21 and 41 by plugging the special μ_{61} pattern to a dot. So 4 runs were done for each dottage; for each run the scores (with R/37 subtracted) were entered and the sums taken to test for significance; $x/\sqrt{R}$ should reach 6.1 before any attention need be paid. The values of $x/\sqrt{R}$ for the four runs and the different dottages are given:-

μ_{61} Dottage	11	12	13	14	15	16	17	18	19
$x/\sqrt{R}$ on Full Wheel	5·9	5·1	5·2	4·4	4·3	6·8	6·0	5·4	5·2
" " " " at 01	5·1	3·8	4·9	5·0	4·6	7·0	6·4	4·1	5·3
" " " " at 21	5·3	5·7	4·6	4·9	4·4	5·4	6·7	5·0	5·2
" " " " at 41	6·0	4·9	5·0	4·7	4·9	6·6	6·1	5·0	4·2

This gave a clear preference for 16 dots, with 17 dots as a rival to consider and 11 dots as a poor third. The smooth μ_{61}, with 16 dots was then used for further runs and from now on four groups of letters were used: -

///	at 7
P, U, O, 3	at 2
N, D, X	at 6
V, 8, W, B, E,	at $3\frac{1}{2}$

Runs for the other three groups were done, both on the full wheel and on the partial wheel set at 01 and 41; the 4 runs were added together at their own decibanages and the totals were tested for significance. If runs with R's of R_i and decibanages of k_i are added in this way the usual test can be used on the totals if $\sum_i k_i R_i$ is used as the R for the composite run.

The additional evidence was decisive. It gave $x/\sqrt{R}$ for the full wheel as 8.3 and for the parts as 7.3 and 6.5 At this stage phase one is over; the μ_{41} dottage is certain. The time for this work was of the order of one shift not counting the time for preliminary calculations. As already mentioned, phase two can be carried out by hand methods without much difficulty. But it can be done much more quickly on Colossus, provided that it is always held in mind that the first approximation to μ_{37} is liable to be the result of adding together better approximations at a slide

Therefore it is best to see what are the contributions to μ_{37} due to a number of different patches of μ_{61} and to see if the most significant contributions to μ_{37} can be set well at a short slide. Only after this is the next approximation to μ_{61} attempted. This can be done by moving a 'doubting cross by hand. If instead this is done as an ordinary wheelbreaking run it is necessary to allow for the fact that the smooth μ_{61} is also going round. By this method phase two was completed in one shift, but not until after another method had been used which took 3 or 4 shifts. The patterns were proved correct by the setting of the Ψ's (which were already known).

References to the subject of breaking the motor by statistical means may be found in the following places:-

R0, 1, 7, 8, 11, 12, 13, 14, 16, 45, 47, 56, 63, 68, 69, 70. R3, 50. 58.

The account given here of motor breaking from a de-chi is a paraphrase of work that has been filed away.

93 THRASHER

(a) General Description

Thrasher was a fish link whose only manifest abnormality was its QEP system. Between consecutive transmissions the QEP number increased not by one, but by an amount roughly proportional to the length of the earlier transmission, approx. 1 per 120 letters. After some 30,000 - 40,000 letters there was a change of 'Rolle'. The obvious interpretation was that the Rolle was an expendable key tape of which each terminal had a copy, containin 30,000 - 40,000 letters, marked with numbers at intervals of about 120 letter At the start of a new transmission the tape was presumably moved forward to the next mark.

Whether or not this interpretation was correct, it remained possible that the key was Tunny key: this is plausible only because Tunny machine would be an easy source of key tape, and the Germans were confident that it was unbreakable. Rectangles had already failed but it was thought not unlikely that the Tunny settings would be changed every 2,000 letters or so, which would defeat rectangling.

(b) The statistical method

This appendix describes an attempt to discover statistically whether Thrasher was Tunny. The interpretation of log evidence is dealt with in a Sixta report.

The basic method was the Δ_{1271} test (24E(c)) i.e. for Tunny cipher the proportional bulge,

$$P\beta\,(\Delta_{1271}\,\Delta\,Z_{,,} = \cdot) \doteqdot \delta^2$$

Since $\delta \doteqdot \frac{1}{10}$ the expected sigma-age was about $\frac{1}{100}\sqrt{N}$

Corruption would reduce this. If the settings were changed every 2,000 letters the score would be reduced in the ratio $\frac{2000-1271}{2000}$

The value of N required to reach definite conclusions implies the use of many messages.

Further evidence was obtained by differencing at intervals which are multiples of 1271, but this was kept separate lest the settings were changed frequently.

The test for each message was carried out by putting two identical cipher tapes on Robinson with the first character of A opposite the 1271 + 1st character of B.

Then the count $\Delta B_{12} + \Delta B_{12} = .$ is clearly equivalent to $\Delta_{1271} \Delta Z_{12} = \bullet$

The count should of course be made only on the overlap of the texts A and B : this was arranged by using A for start, B for stop.

A similar count was made with the tapes staggered by 2 x 1271, 3 x 1271 and so on, till the message was exhausted.

The aggregate score for many messages were

Δ_{1271}: effective text 381,701, bulge + 366 sigma-age 1.18
Δ_{2543}: effective text 268,573, bulge + 428 sigma-age 1.65

This was unconvincing, but the poor score might have been due to corrupt texts, therefore only messages which had scored well were used in similiar tests on other impulses; firstly $\Delta_{598} \Delta Z_{45}$, $\Delta_{1196} \Delta Z_{45}$.

Messages that still scored well were then tested by

$$\Delta_{784}\Delta Z_{34}, \Delta_{1601}\Delta Z_{35};$$
$$\Delta_{743}\Delta Z_{28}, \Delta_{1434}\Delta Z_{26}; \Delta_{804}\Delta Z_{29}, \Delta_{1612}\Delta Z_{25}$$

For each of the first three of these the aggregate score was negative.

It was concluded that Thrasher was almost certainly not Tunny.

(c) Note on precautions adopted

To provide checks and eliminate spurious effects in the tests just described, all the following were counted for each message and entered in appropriate columns:

Message number,
Text length,
Stagger,
Calculated effective text,
Measured effective text,
Average (half effective text),
9's in whole text (an excess indicates corruption),
ΔZ_{21} = . (A Tape) and its bulge.
ΔZ_{12} = . (B Tape) and its bulge.
Δ stagger ΔZ_{12} = . and its bulge.

It was found that the correction for 9's and consequent bulges on ΔZ_{12} = • was negligible.

(d) Mystery of alleged depths

Log evidence, which appeared to be unambiguous, suggested that Rolle 40,034 was used twice, once by each terminal. It seemed that in several instances two messages must be in depth though the exact settings were uncertain: an attempt was made to set them by running for ///'s in

$$Z^{(\psi)} + Z^{(\chi)} = P^{\psi} + K + P^{(\chi)} + K = P^{(\chi)} + P^{(\psi)}$$

The attempt was unsuccessful, which was equally unexpected whatever the nature of the key tape; it is inconceivable that the designer of a random tape machine would provide an autoclave.

It is however possible that the orginator of a message, not understanding the principle of random key, or merely mistrusting the new-fangled machine, might demand double encipherment, which would destroy the expected properties of $P^{(\chi)} + P^{(\omega)}$.

In fact at various times a machine fault caused clear text to be transmitted, and in one case what appeared to be Enigma, though it was not broken.

94 - RESEARCH INTO THE QEP SYSTEM

(a) The output of TUNNY decodes would obviously have been enormously greater if the indicating system in use in 1943 to 1945 had been broken. It was suspected, correctly, to be of the 'Book of Settings' type from the following known facts :-

(i) The only part of the preamble of messages which could possibly indicate the settings was the QEP number which was usually between 1 and 100, though very occasionally of three of four figures.

(ii) Messages on the same link and with the same QEP number were in depth only if the sequence of QEP numbers had not passed through a 100 between their transmission.

(iii) A book of settings was captured in the early days, which was thought to apply to similar traffic. The settings had four figure numbers. Settings were given for all twelve wheels.

Work on the QEP system before FISH traffic ceased in 1945 consisted of three pieces of analysis.

(i) Analysis of settings of messages set in the Sections.

(ii) Analysis of the captured QEP book.

(iii) Analysis of a partial QEP sheet transmitted in a Whiting message in November, 1944.

(iv) Analysis of allocation of QEP Books (not discussed below).

(b) Analysis of message settings.

When the number of messages set per month began to reach three figures the possibility of breaking the QEP book had to be considered. Obviously everything depended on the size of the book. It might be infinite in the sense that fresh books of settings might be issued as the old ones were used. On the other hand it might consist of a limited number of settings, say 10,000, used over and over again.

The first step was to record the settings of all messages. This was done in Room 12 when they received the Red Forms back from the decoding section. The settings were recorded on cards, one for each message. The obvious method of attack was to sort the settings for repeats. Unfortunately the settings we obtained were only slides of the German settings, since the starting points of wheels are chosen arbitrarily by the wheel-breakers. The slide was of course constant for a day's traffic on a given link. The method adopted was to difference the settings or pairs of messages on the same day and link and sort these different settings for repeats over several months. There was a further complication. The initial break of a message was often a hundred or more letters from the beginning. It was necessary to work back the settings to the beginning. The working back and differencing which had to be done modulo the lengths of the various wheels, was done by Mr. Freeborn's Hollerith Section (Block C) mechanically. Only the chi settings were used for this purpose. This was sufficiently powerful to eliminate almost all random repeats.

The information was sent to Block C in the following form :-

Msge. No.	Date.	QEP	No. of ltrs. before initial break	Chi settings	Motor Settings	Psi Settings
JP 2374	5.7.44.	40	46	23.14.21.16.03.	36.21.	16.43.21.50.26.

All messages decoded between July and November 1944 were sent.

The following processes were carried out mechanically

(i) The settings were punched on Hollerith cards and printed out in book form for reference, sorted by Link and serial number.

(ii) These books were gone through by hand and a reference bigram given to each message corresponding to the day's keys on which it was decoded.

(iii) New settings cards including the bigram were punched from the book.

(iv) The number of letters before the initial break were subtracted modulo 41,31 etc. from the chi settings.

(v) The correct settings of messages on the same day's keys were differenced, again modulo 41, 31 etc.

(vi) The differences were sorted by Chi 1 difference, then Chi 2 difference etc. and the results were printed.

The results were examined for repeats, which could be further tested by looking up the μ_4 settings to find if they also gave identical differences. The psi differences though not the same should be simple slides of each other.

Only one case of a genuine repeat was found. The differences of the settings of two Stickleback messages in September, 1944 were the same as those of two others retransmitted a week later.

The conclusion to be drawn, was that no large scale reuse of QEP sheets was taking place over the period worked on. The isolated instance was probably due to a temporary use of an old sheet. Evidence from the Whiting QEP sheets of November, 1944 shows that the Germans number their wheel positions in the reverse order to us(e.g. the German Chi 1 set at 2 was in our sense 1 back compared with the wheel set at 1), but this unexpected fact makes no difference to the validity of the method used to test for repeated settings.

(c) The Captured QEP book

A QEP book was captured during the Sicilian campaign. The settings were evidently Tunny machine settings since there were twelve wheels and the limitations on the numbers involved corresponded to the lengths of the Tunny wheels arranged in the order:

$$\psi_1\ \psi_2\ \psi_3\ \psi_4\ \psi_5 \qquad \mu_{37}\ \mu_{61} \qquad \chi_1\ \chi_2\ \chi_3\ \chi_4\ \chi_5$$

The following further discoveries were made:-

For each value of k all QEP's of the form $21n+k$ were associated with a group of letters.

Thus :

k	Ψ_1	Ψ_2	Ψ_3	Ψ_4	Ψ_5	$\Psi_{\alpha 1}$	$\Psi_{\alpha 2}$	χ_1	χ_2	χ_3	χ_4	χ_5
1	E	P	C	N	D	B	U	M	F	I	G	A
2	C	G	M	I	L	K	P	D	J	F	A	H
3	J	N	H	A	N	E	T	M	B	D	F	G
4	D	A	Q	K	R	I	F	N	C	J	B	E
5	H	N	D	E	A	K	C	L	I	G	F	B
6	N	L	J	C	Q	D	B	F	K	E	H	G
7	I	O	K	F	J	H	A	C	E	G	B	D
8	G	C	M	B	F	I	O	L	J	H	E	A
9	F	B	A	P	I	G	R	J	[illegible]	D	H	C
10	A	L	E	J	H	M	O	B	K	I	C	D
11	K	E	I	O	S	C	H	N	G	J	D	F
12	N	F	G	L	E	A	K	H	D	C	I	B
13	H	M	K	D	T	G	J	F	B	I	A	C
14	L	J	Q	G	C	F	I	A	H	B	D	E
15	F	D	P	H	B	L	Q	E	C	A	I	G
16	A	K	O	G	S	C	E	I	D	B	F	H
17	B	I	F	D	M	L	S	K	H	A	G	E
18	O	H	L	N	K	M	G	J	A	C	E	B
19	M	P	I	R	O	E	L	D	G	F	A	C
20	C	E	N	Q	G	I	M	K	F	H	D	A
21	J	H	B	M	P	A	D	G	I	E	C	F

Where A means one of the settings 1,2 or 3; B one of 4,5,6; C 7.8.9; D 10,11,12; E 13,1415; F 16,17,18; G 19,20,21; H 22,23; I 24,25,26; J 27,28,29; K 30,31; L 32,33,34; M 35,36,37; N 38,39,40,41; O 42,43; P 44,45,46,47; Q 48,49,50,51; R 52,53; S 54,55,56; T 57,58,59; U 60,61.

It will be seen that if a letter has been used for a setting of one wheel for a given value of k, it has been used for no other wheel for the same k. In other words a letter occurs not more than once in a given row.

Further investigation suggested that within these limits the book was compiled by hand. Depths on different QEP's are almost completely ruled out by the design of the book.

Probably this system is not a regular feature of QHP books on Tunny links.

(d) <u>The Whiting QEP Sheet</u>.

At 1105 on 13th November 1944 Berlin sent QEP 45 which was intercepted as W.B.3756, and decoded by us soon after the wheels had been broken on another message. QEP 45 consisted in part of a message which said : "To Heeresgruppe Nord. You will shortly receive the following QEP Blatt 1, from 001 to 035. Wheels 1,2,3 ..... 12.

001, 20 47 26 50 17 35 13 02 12 10 24 09
002, ... etc.

Though several messages on Whiting 12th November were read there was no information that Riga was without any cypher equipment.

At 1149 the first message on Blatt 1 (QEP 01) was sent. In QEP's 08-9 there was another transmission of wheel settings, this time Blatt 3 from 071 to 105 (WB 3761-2). It seems certain that Blatt 2 was never sent or used. QEP 30 was followed, at an interval, by QEP 77 and it certainly looked as if the transmitted Blatt 3 was used from QEP 77 to QEP 96. All these QEP's were on the same day's keys. It is improbable that any further QEP Blatt was transmitted.

All traffic intercepted on QEP Blatt 1 was decoded except for 6 short or corrupt messages. To get 'English Settings', the German settings were taken in the order Psi 1, Psi 2, Psi 3, Psi 4, Psi 5, M37, M61, Chi 1, Chi 2, Chi 3, Chi 4, Chi 5 and subtracted from
26 33 48 28 50 19 50 30 07 54 21 05. (Defined as rings).

On messages connected with QEP Blatt 3 such messages as had been set on Colossus had settings which bore no obvious relation to the expected settings deduced from the QEP number and the rings deduced from Blatt 1. Settings on Blatt 3 messages were not even consistent among themselves in giving a different set of rings.

It is possible that a slide may have been put on the QEP sheet or that some other transposition or substitution was applied to the printed settings. But it seems more probable, that in spite of the logs evidence, a courier with the missing QEP sheets had arrived at Riga before QEP 78 was sent. This would imply that Berlin transmitted for temporary use a set of QEP's different from those whose transit was delayed and which would otherwise have been used.

A fuller account (of the Whiting QEP sheet) is given in R4 pages 17,18, 36 to 38.

95 - MECHANICAL FLAGS

95A GENERAL DESCRIPTION

(a) Experiments to be described

This appendix deals with experiments in obtaining mechanically

(1) The complete ($\Delta\chi_2$) flag of an ordinary rectangle.
(2) The flag of a combined key rectangle.

The corresponding hand processes are described in 24D (b,c,d) and 26B (c).

(b) Results obtained

In neither case was mechanization operationally successful. This is probably due to a combination of circumstances, among them:

Both needed complicated processes on Miles D, which was unreliable.

Both used Super-Robinson before it was reliable.

The key-flag , in its final form, requires gadgets on Colossus and Robinsons wh.ch were not available till the end of the war in Europe was imminent.

The ordinary flag involved very long tapes.

Neither was even theoretically much faster than computation by hand.

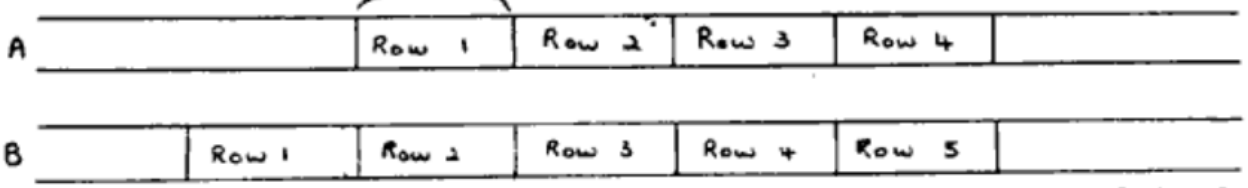
(c) Simplification of this account.

The account describes the processes in their final form, and indeed, for the ordinary flag, exceeds this by ignoring certain complications introduced to deal with a mistakenly alleged weakness of Super Robinson. Some ingenious improvisations are th.a omitted, but so are many tedious instances of failure to see the obvious solution.

(d) The common basis of mechanical flagging.

The basis of the run on Robinson is the same for both types of flag.

Two tapes, A,B, are used, each bearing the entries, in the cells of the rectangle arranged by rows, thus



To find the flag entry corresponding to rows 1 and 2 of the rectangle, it is arranged that row 2 of B is opposite row 1 of A, and row 1 of A is spanned. What is required is the sum

of products of corresponding scores on A and B.

When each score is ± 1 (as in the key flag) this is easy, for if ± 1 are represented by dot and cross, the correct result is obtained by counting +1 when A+B = .; -1 where A+B = x.

The complexity of the Mechanical Ordinary Flag is that of summing the products when the rectangle scores are not all ± 1.

The complexity of the Mechanical Combined Key Flag is that of combining the four rectangles on one tape.

NOTE. The flag of a single rectangle of depth 1 can be obtained on Super-Robinson directly from two message tapes, the selection of a row being made, not by spanning, but by a control tape (cf 54H).

95B. MECHANICAL ORDINARY FLAG.

(a) The Robinson run.

Suppose that the depth is 8 so that the score in a cell of the rectangle may be -8,-6,-4,-2,0,2,4,6,8, Because the common factor 2 is irrelevant these may be treated as -4,-3,-2,-1,0,1,2,3,4.

The contribution to the score of two rectangle entries opposite one another on tapes A,B is the product of these entries, which may be as great as 16. Super-Robinson can count only 1 at most for each sprocket hole, and thus each score of the rectangle must be represented on the tape by a symbol extending over 16 sprocket holes at least.

Further, Robinson cannot record negative scores, and accordingly positive and negative scores are counted separately, actually in the two halves of a split counter. The score is positive where A and B entries have like signs, negative when they have unlike signs : minus is represented by a cross in the fifth impulse, and the switching is $A_5 + B_5 = \cdot$ for positive scores, $A_5 + B_5 = x$ for negative scores.

The magnitude of each score, apart from sign, is represented in the first impulse :

A Tape (Tate)			B Tape (Lyle)		
0	by	.....	0	by	.....
1	by	x...x...x...x...	1	"	xxxx.....
2	"	xx..xx..xx..xx..	2	"	xxxxxxxx.....
3	"	xxx.xxx.xxx.xxx.	3	"	xxxxxxxxxxxx....
4	"	xxxxxxxxxxxxxxxx	4	"	xxxxxxxxxxxxxxxx

The switching $A_1 = x, B_1 = x$ gives the correct products, as will be seen on inspection.

For a $\Delta\chi_2$ flag, tape A comprises the 31 rows of the rectangle, each represented Tatewise by 41 x 16 characters, and further 41x16 blanks, with start and stop. Tape B comprises the 31 rows of the rectangle, each represented Lylewise by

41X16 characters, with a start (for checking relative positions).

Start tapes level. Span A 1-657. A will step relative to B, and the scores obtained will be those of the first row of the flag.

Then span A 657-1313, to obtain the second row, and so on.

(b) Tape-making.

This need two processes, on Colossus and Miles D respectively.

(c) Colossus rectangle tapes.

On Colossus with a punch [53M (h)] make an ordinary 1+2 rectangle, and in addition plug scores to punch thus 0 to 1, 2 to E, 4 to 4, 6 to 9, 8 to 3.

The machine adds a cross in the fifth impulse for negative scores so that on the resulting tape, the representation of scores is

	8	6	4	2	0
+	...x.	..x..	.x...	x....	x
−	...xx	..x.x	.x..x	x...x	

(d) Making Tate tapes on Miles D.

In the first transmitter a loop E///E///E///E//9 with the 1st E in the window.

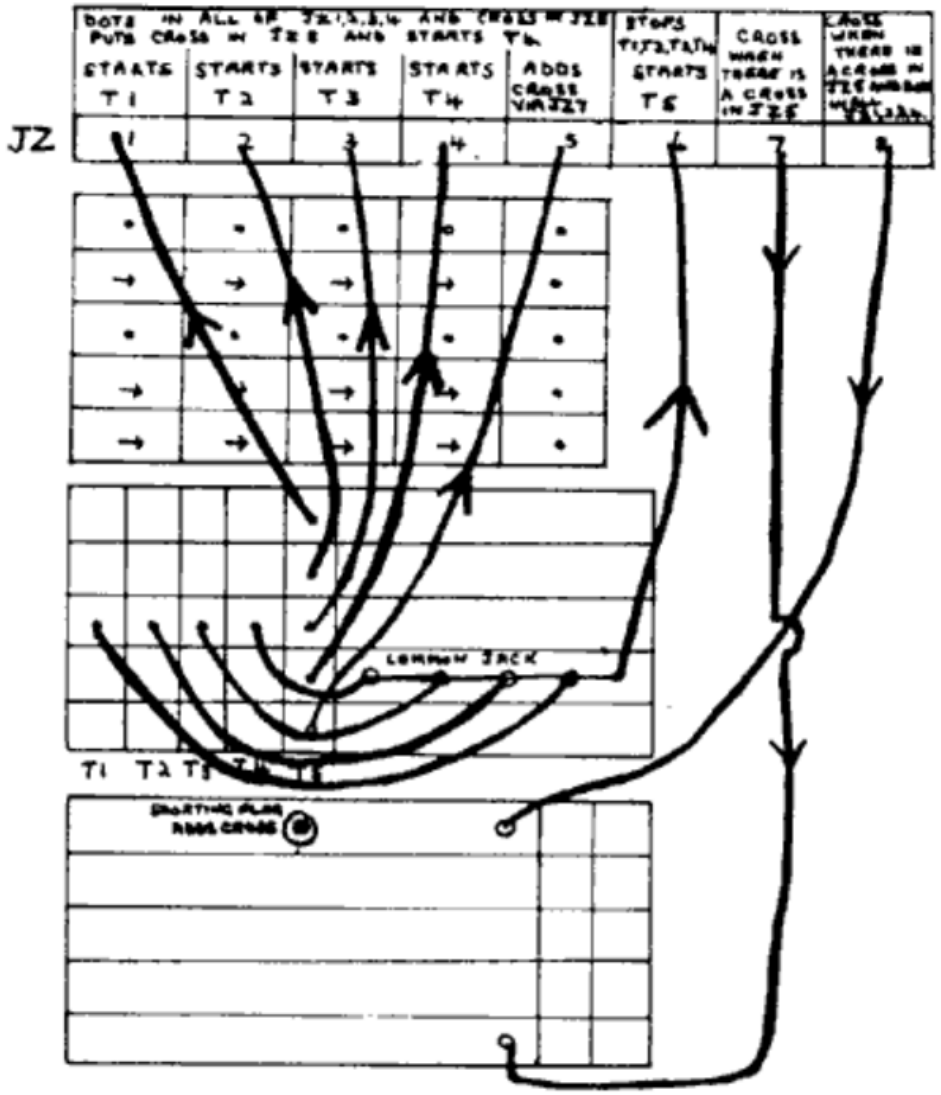
" " 2nd " " " EE//EE//EE//EE/9 "

" " 3rd " " " EEE/EEE/EEE/EEE9 "

" " 4th " " " EEEEEEEEEEEEEES "

" " 5th " " " the rectangle tape just made,

and plug as in the diagram: (not to scale)


DOTS IN ALL OR JZ1,2,3,4 AND CROSS IN JZ5 PUTS CROSS IN JZ6 AND STARTS T4
STARTS T1
STARTS T2
STARTS T3
STARTS T4
ADDS CROSS VIA JZ7
STOPS T1,T2,T3,T4 STARTS T5
CROSS WHEN THERE IS A CROSS IN JZ5
CROSS WHEN THERE IS A CROSS IN JZ6 AND NONE IN JZ5
JZ
1
2
3
4
5
6
7
8
COMMON JACK
T1 T2 T3
SHORTING PLUG MAKES CROSS

Suppose there is a score of +4 on the rectangle tape i.e. .x... in the eye of T_5: this starts T_2 and produces EE//EE//EE//EE/9. which is the sum of the tape in T_2, the three E's in T_1, T_3, T_4 and the cross in impulse 1 of output. When the 9 is reached, the x in the third impulse stops T_2 (via JZ6), and steps T_5 one place.

If the score is -4, i.e. .x..x, the x in the fifth impulse will add a cross to the output via JZ7 , giving ZZTTZZTTZZTTZZTH.

If the score is 0 i.e. ....x, this will start T_4 (via JZ5) and add a cross to impulse 5 (via JZ7) and a cross to impulse 1 (via JZ8), giving TTTT TTTT TTTT TTTH.

(e) Lyle tape.

A "Lyle" tape is made in an essentially similar manner.

(f) Checks

Tate and Lyle tapes are elaborately checked, being marked at intervals of 656; the letters following each mark are checked against the printed rectangle.

(g) Practical modifications.

In fact the Robinson bedstead can hold only 16 rows, and the runs would have to be done piecemeal : nearly all the rectangles used experimentally were of depth 6. Most of the experiments were made on the old Robinsons, so that strings of dots and crosses had to be avoided. What is here called the first impulse was really the sum of the first two. The original Tate tape bore the word "Tate" thrice for each score +1. In the fifth impulse + was represented by .x.x.x and - by x.x.x. etc.

The Colossus punch is now wired to punch / instead of T for zero.

Robinson is plugged :

$$\begin{cases} A_2 = B_2 = x & \text{(excluding a zero on either tape)} \\ A_5 + B_5 = \cdot & \text{in one half counter for positive products.} \\ A_5 + B_5 = x & \text{in the other for negative products.} \end{cases}$$

which will obviously give the correct result.

A consists of the 23 rows of the combined rectangle, 2 row lengths of blanks, start and stop.

B consists of the 23 rows of the combined rectangle. 1 row length of blanks, start and stop.

Span the first row of A : owing to the difference in tape lengths it will step so as to be opposite each row of B in turn : the scores obtained will be those of the first row of the flag. They can be identified by the position counter. Span the other rows in turn.

(b) Some natural checks.

When the 1st row of A is opposite the (identical) first row of B, thus

+1 +1 0 0 −1 0 0 −1 +1 .....
+1 +1 0 0 −1 0 0 −1 +1 .....

the score for positive products is the number of non-zero scores in the first row of the combined rectangle, i.e. the total number in the first rows of all constituent rectangles, and the score for negative products is zero.

Likewise when the n^{th} row is opposite the n^{th} row.

Further, when the spanned row of A is opposite the blanks in B both scores are zero.

(c) Tape-making.

The tapes R_1 R_2 R_3 R_4 of the four rectangles are made separately on Colossus and combined into a single rectangle on Miles D.

(d) Making the four rectangle tapes.

In a key rectangle there are many doubtful characters (26%) for which no provision is made in an ordinary Colossus rectangle.

The ordinary not 99 circuit is useless. Colossus rectangling works by counting places where $\Delta z_{ij} \neq 0$ and then doubling and subtracting the depth. With not 99 in use, a doubt is not counted, i.e. it is treated exactly as though

$\Delta z_{ij} = x$, and will score -1 not 0.

To overcome the difficulty Colossus 5 has been fitted with 'rectangle not 99' directly controlling the score, making it zero: it overrides the score produced normally. It works only for a depth of 1.

The punch is plugged 0 to punch E
+1 " R
-1 " G

Carriage Return to punch $\overline{9}$/ i.e. to punch / and add a cross in the 3rd impulse of the preceding character.

The four rectangle tapes R_1, R_2, R_3, R_4 are, otherwise, made normally from the cipher tape as rectangles of depth 1. Owing to the spanning from O_4, the first four scores are lost: it may be worth while to preface a short text with four 9's.

The first few scores in each rectangle are checked by hand. Machine faults generally produce an obvious irregularity in the pattern of 0's and 1's.

(e) Combining rectangles on Miles D.

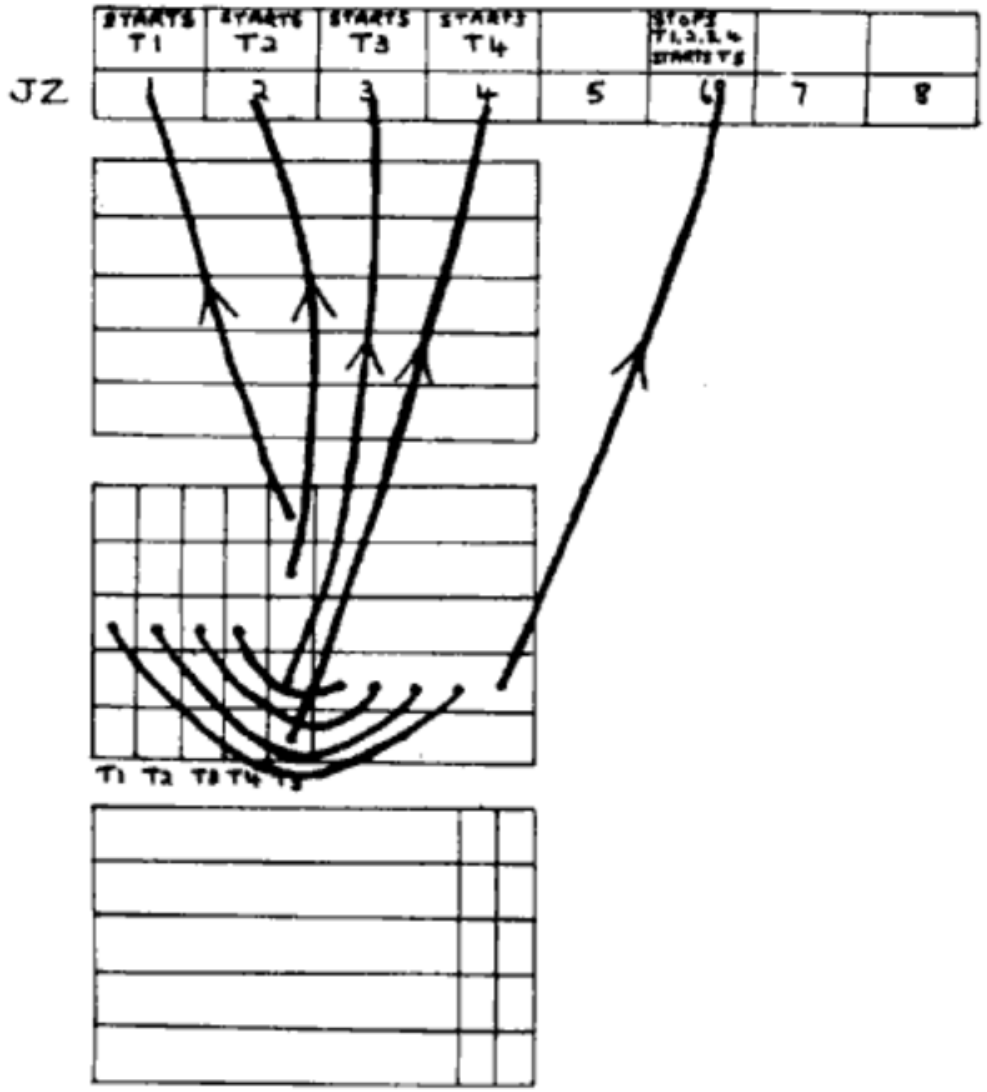
R_1, R_2, R_3, R_4 are placed in T_1, T_2, T_3, T_4 with the stroke before the text on the peckers.

In T_5 is a control tape consisting of E43T..... repeated 23 times. Plug as in the diagram below (not to scale).

The 'E' in T_5 starts R_1 (in T_1) which is reproduced till the cross in the 3rd impulse at the end of the first row stops it, leaving / on the peckers and steps T_5 to '4'. The '4' starts R_2 (in T_2) and so on.

The tape produced is

1st row of R_1,/, 1st row of R_2,/, 1st row of R_3,/, 1st row of R_4,/, 2nd row of R_1,/ and so on.


STARTS T1
STARTS T2
STARTS T3
STARTS T4
STOPS T1,2,3,4 STARTS T5
JZ
1
2
3
4
5
6
7
8
T1 T2 T3 T4 T5