Copy 8 of 8 copies

SECRET

I/L 3601-A
Report F-46
12 June 1944

SUBJECT: Fish Notes
TO : CO, SSA, War Dept.

1. SSA 161 left me in some doubt as to just what was wanted and also as to the urgency of the request. I assumed that a cabled reply was not expected because of the quantity of data requested but suggest that in the future requests of this nature indicate whether a reply is wanted by cable or through a report sent via regular channels.

2. I also assume that the inquiry was directed only at patterns used on links employing the P_5 limitation and that you are not interested, for example, in Stickleback patterns. The scope of your questions was puzzling not only to me but also to Maj. Tester and Mr. Newman because the number of crosses in X, Ψ and ΔX patterns is always the nearest integer (one side or the other) to half the period of the wheel (necessarily an even integer in the case of ΔX). It is conceivable that there may have been an isolated instance or two in which this rule has been violated but no-one here remembers such a case and the patterns have not been checked from this viewpoint. It is difficult to see how an occasional variation in this respect could be of cryptanalytic significance.

3. Even as to the $\Delta\Psi$ patterns the reason for the inquiry is not clear because the average answer is dependent on the number of dots in μ37 and I have been sending data on this regularly in my reports. Presumably, however, you are interested in deviations from the average so this information is set forth below. The people here are curious as to whether you have in mind any techniques based on these deviations.

	BREAM						JELLYFISH				
	Jan.	Feb.	Mar.	Apr.	May	June	Mar.	Apr.	May	2nd May	June
μ37(dots)	24	19	24	17	20	27	22	24	20	22	20
$\Delta\Psi$1(crosses)	34	28	32	28	30	34	28	32	30	30	30
$\Delta\Psi$2 "	36	32	36	30	32	38	34	34	32	34	32
$\Delta\Psi$3 "	40	34	38	32	34	40	36	38	34	36	34
$\Delta\Psi$4 "	40	36	40	34	36	42	38	40	36	38	36
$\Delta\Psi$5 "	46	40	44	38	40	46	42	44	40	42	40

4. Obtaining the correct answers to the "Questions on Tunny Conventions" which Capt. Seaman sent me as Annex G was rather difficult. There appears to be considerable misapprehension on the subject in the sections themselves and this accounts for the fact that the illustration given in Report #F 5 (IL 3365) is not wholly correct. I believe, however, that the following description is completely accurate. The Xs, Ψs and motor all start at the first position. The motor is not written above the column it is about to affect, as suggested in Annex G, but rather, as stated in Report # F 5 (IL 3365), the motor elements determine whether the Ψs in the column

MY

SECRET

DECLASSIFIED
Authority NND 963016
By SB NARA, Date

immediately below shall change or not, viz.- M_T x..x..x.xx.x. with ψ .x.x..x is written M_T x..x..x.xx.x. and then the pattern is extended so that the
ψ .x . x .. x
final picture is M_T x..x..x.xx.x. . The first position of M_T is always
ψ .xxx...xx...xx
x and its second position is always the same as M_2. Therefore the sum of $\bar{X}_2 + \bar{P}_5$ in the second position must be cross. In the first position its sum is immaterial but by adopting the convention that the sum must be dot we compel a cross in M_T regardless of the sign of M_2. Thus it becomes unnecessary to have any other special rules for the beginning of the message. The $\bar{X}_2$ sign is determined from the pattern, that is, if the X_2 pattern starts at 5, the $\bar{X}_2$ pattern starts at 4. If then elements 4 and 5 of X_2 are say x. the first two elements of P_5 must be (xx) The easiest way to think of it is that $\bar{X}_2$ is known and that the first two columns must be given arbitrary P_5 symbols which will combine with the X_2 to produce dot in the first column and cross in the second. The third column of $\bar{P}_5$ depends on P_5 of the first letter of the actual message. With these conventions nothing need be discarded and none of the initial letters of the deciphered text will be incorrect.

5. June Bream was broken very early in the month and is very favorable with 27 dots in 437. The Jellyfish wheels also changed on 1 June but efforts to break them by the usual statistical methods were unsuccessful. However, they have just been solved by the cribbing method hereinafter described. Bream went off the air when Rome fell but reappeared on 10 June. The new station is near Florence.

6. The new Colossus mentioned in Report #P 22 (IL 3452) has been installed and is in operation. It is designated Colossus # 2. It had been promised for 1 June and was in partial operation on that day although the engineers and maintenance people had to work the last night until 0300 to carry out their commitment. Colossus # 3 is expected to be ready by 1 July. 9 more have been ordered. All subsequent to # 1 are of the same type and represent a substantial improvement over # 1. The speed of operation is actually 5,000 per second. # 1 could theoretically attain this speed but never achieved it. The new machines have an extensive switchboard in addition to a plugboard. They carry a great many more different patterns than the old and these can be switched in or out with triggers. Furthermore, it is unnecessary to plug up Δ patterns because the machine itself can difference if required.

7. In Report #P 36 (IL 3482) I mentioned the discovery of re-encodements in Fish traffic. A statistical method has been developed which is designed to locate these (and thus solve hitherto unbroken keys) by sliding a long crib through a message. About 800 letters of crib are ordinarily needed. When correctly set this yields a length of key more than sufficient for resolution into its components.

DECLASSIFIED
Authority NND 963076
By SB NARA, Date

The method is apparently not applicable to relays because the relaying operator makes up a new tape and this will vary in minor details, notably punctuation. But if Berlin sends the same message to both Rome and Paris the same tape will be used on automatic transmission and the two texts will be lett letter perfect. An occasional garble, even if on the 5th impulse, will not invalidate the procedure although an omitted letter of course will. The real risk is interruption of the tape with stretches of hand transmission or "go-backs", that is, repetitions of portions of what has already been transmitted. "Go-backs" in the crib can be found and corrected for but if the crib happens to be in a portion of the unknown cipher text which includes "hand" or "go-backs" it will not be possible to locate it. A possible aid to setting cribs is provided by the fact that where a transmission includes several messages the receipts, which are sent in clear, state the exact time at which each message was concluded. Correlation of these times with time of interception and with each other should make it possible to determine approximately the portion of the message which contains the crib. However, this information has not been used because the Germans make mistakes and it is therefore not too reliable and because once the necessary preliminary work has been done it is almost as easy to try all possible settings.

6. In the description which follows I will assume that the $X_2 + P_5$ limitation is employed. The method could not be used without any limitation but is fully applicable if only the X_2 limitation is used. To develop the theory of this procedure we use the usual notation - a = proporation of crosses in M_A; b = proportion of crosses in $\Delta\Psi$ and assume that $ab = \frac{1}{2}$. Let a' = proportion of crosses in M_B. Then, with $X_2 + P_5$ limitation, which is assumed to produce equal numbers of dots and crosses,

$$1 - a = \tfrac{1}{2}(1 - a') \quad \text{or} \quad a' = 2a - 1$$

If $P_5 + X_2 = x$, the chance that $M_B = .$ is the same as the chance that $M_B = .$, namely $1 - a'$. Therefore the chance that $\Delta\Psi\frac{1}{2} = .$ is

$$1 - a' + a'(1 - b) = a'b = 1 - 2ab + b = b$$

If $P_5 + X_2 = .$, M_A must be x and the chance that $\Delta\Psi\frac{1}{2} = x$ is b. Consequently the chance is b that

$$\Delta\Psi\tfrac{1}{2} + P_5 + X_2 = x$$

or that

$$\Delta\Psi\tfrac{1}{2} + \Delta X_2 + P_5 = X_2 + \Delta X_2 + x$$

or that

$$\Delta Z_2 + \Delta P_2 + P_5 = X_2 + \Delta X_2 + x$$

The left hand side of the last equation becomes known when plain and cipher texts are properly juxtaposed. $X_2 + \Delta X_2$ (which is called $\hat{X}_2$) is unknown but it must be periodic. Therefore, $\Delta Z_2 + \Delta P_2 + P_5$ combined with itself at an interval of 31 should yield a high proportion of coincidences (or dots) when the crib is correctly placed. The run is designated $(\Delta Z_2 + \Delta P_2 + P_5)\Delta 31$. The proportion of dots expected is $b^2 + (1 - b)^2$. For a crib of 300 letters with b = .65 this will give a score about $2\frac{1}{2}\sigma$ above random. This is not enough but we can also make independent runs at intervals 62, 93, etc.

9. It is not difficult to determine the minimum length of crib required. If we use as intervals all possible multiples of 31 the number of values available for comparison (that is, the aggregate of dots and crosses which can be counted and scored) with a crib of length N is

DECLASSIFIED
Authority NND 963016
By SB NARA, Date

(1) $31c(c-1)\frac{1}{2} + cd$ where N is expressed as $31c + d$ $(d < 31)$

The "bulge" is

(2) $b^2 + (1-b)^2 - \frac{1}{2}$

so that the expected excess of dots at the right placement is the product of (1) and (2). σ is one half the square root of (1) so that if we determine what score we will require in terms of σ it is easy to determine the necessary length for a given value of b (which is unknown and must be estimated from past experience). A fairly good approximation for (1) is $N^2/64$. If we want a score of 5 σ and assume that $b = 2/3$ we get $N = 360$ (using the approximation).

10. However, there are both theoretical and practical limitations. The theoretical limitation arises from the fact that if, at a particular position we get a high random proportion of dots, say $\frac{1}{2} + \lambda$, on a run at interval 31, then at other intervals we will not get random scores but instead proportions of $(\frac{1}{2} + \lambda)^2 + (\frac{1}{2} - \lambda)^2 = \frac{1}{2} + 2\lambda^2$. This factor makes it necessary to require slightly higher relative scores for runs made at several intervals than for runs at interval 31 only. However, it is almost negligible. The practical limitation which is much more important arises from the diminishing utility of the runs at higher multiples of 31 and the labor of making tapes.

11. For the ideal run we would make one tape containing $(\Delta P_2 + \overline{P_5})\,\Delta 31$, then $(\Delta P_2 + \overline{P_5})\,\Delta 62$, etc, up to the maximum available intervals. A space would be left after the $\Delta 31$ portion equal to the difference in length between the cipher text and the crib and the same amount of space after the $\Delta 62$, $\Delta 93$, and succeeding portions. One additional space is left after the final block. The other tape contains $(\Delta Z_2)\Delta 31$, $(\Delta Z_2)\Delta 62$, etc. with no spaces at all between the blocks. When these two tapes are run against each other single scores are produced using all available data. The extra space at the end of the crib tape causes the tapes to step against each other and thus tries all juxtapositions. Symbols must be placed on the crib tape, at the beginning and end of each blank stretch, to stop the machine from counting and to start it again. Another symbol, after the final stretch, causes the scores to be recorded. It is apparent, however, that this method requires tremendously long tapes so that the first practical simplification is to use only a few multiples of 31. However a much easier method which avoids the expenditure of a lot of preliminary time in making tapes is to use the double Robinson machine with $2Z_2$ tapes and $2\Delta P_2 + \overline{P_5}$ tapes. Z_2 can be used instead of ΔZ_2 because the machine can do the differencing. Separate runs are then made for intervals 31, 62, etc. and the high scores compared. This saves preparation time but takes a little longer to check and evaluate results.

12. It would seem that location of cribs by this method could very easily be accomplished with IC equipment. If equipment is used which will not record scores the method described at the beginning of paragraph 11

DECLASSIFIED
Authority NND 963016
By SG NARA, Date

is probably the best. Plates present the usual drawbacks but with the film machine the testing would be much more rapid than on the Robinson. The preparatory work is, of course, substantial.

13. The crib on which Jellyfish was broken by this method was from Bream and was about 1600 letters long. Both messages were from Berlin. The crib turned out to have been fairly correct for about two-thirds of its length but the balance was all wrong. The message through which it was dragged was about 2200 letters long so that approximately 600 positions had to be tried.

14. Enclosed is description, prepared by Maj. Tester's section, covering the method (mentioned in Report #F 43) used to recover wheel patterns from Stickleback key. Apparently the method is not as new as I had been informed.

15. Some Sturgeon type traffic of July 1943 has finally been read and as a result the machine, as then used, has been reconstructed. The link read was Halibut. It consisted of practice messages and is no longer passing traffic. It is believed that the current permuted Fish traffic is sent on the same type of machine but with an auto-key element added. The Halibut machine has 10 wheels with periods ranging from 47 to 73. The patterns are fixed and are the same as wheels previously solved. 5 of the wheels act in a combining capacity (the British call these subtractor wheels and designate the key they produce by Σ) and the other 5 act as permuters. The permutation key is designated π. The enciphering equation is written $Z = \pi (P + \Sigma)$ which indicates that combination takes place before permutation. In deciphering the permutation is the first step and the equation is $P = \pi^{-1} Z + \Sigma$. Each wheel can act in any capacity so there are 10! wheel orders.

16. The permutation consists of a series of exchanges between adjacent impulses (considered cyclically so that 1 and 5 are adjacent). A dot in the permutation key means that the permuting factor is operative, a cross that it is inoperative. The first position of the permutation key represents an interchange between impulses 1 and 5, the second between 5 and 4, the third between 4 and 3, the fourth between 3 and 2 and the fifth between 2 and 1. The order in which the permutations take place is significant and in enciphering they proceed in the left to right order of the key, i.e. .xxx. would cause 12345 to become 52341 after the first permutation and 25341 after the second. In deciphering the order is reversed and this is the reason for using the symbol π^{-1}.

17. The most interesting feature of the machine is the motor action. Each wheel is driven by two of the others. In order to avoid the possibility of the machine getting into a rut both dots and crosses act as motor impulses. Wheel A, for example, might be driven by dots in B and crosses in C. Wheel Either a dot in B or a cross in C will cause it to move and it will fail to advance only if there is a cross in B and a dot in C. The motor action is controlled by a portion of the pattern different than the portion operative

DECLASSIFIED
Authority NND 963016
By SB NARA, Date

for encipherment, but there is a fixed relation. The operative motor element is behind the enciphering element an interval which is the nearest integer under 2/3 of the wheel period (or ahead by an interval which is the nearest integer above 1/3 of the period). In the case of the 69 wheel the figure is exactly 46 (or 23). If, for example, the 47 wheel is at position 8 for enciphering, its 24th position will control its effect as a motor wheel.

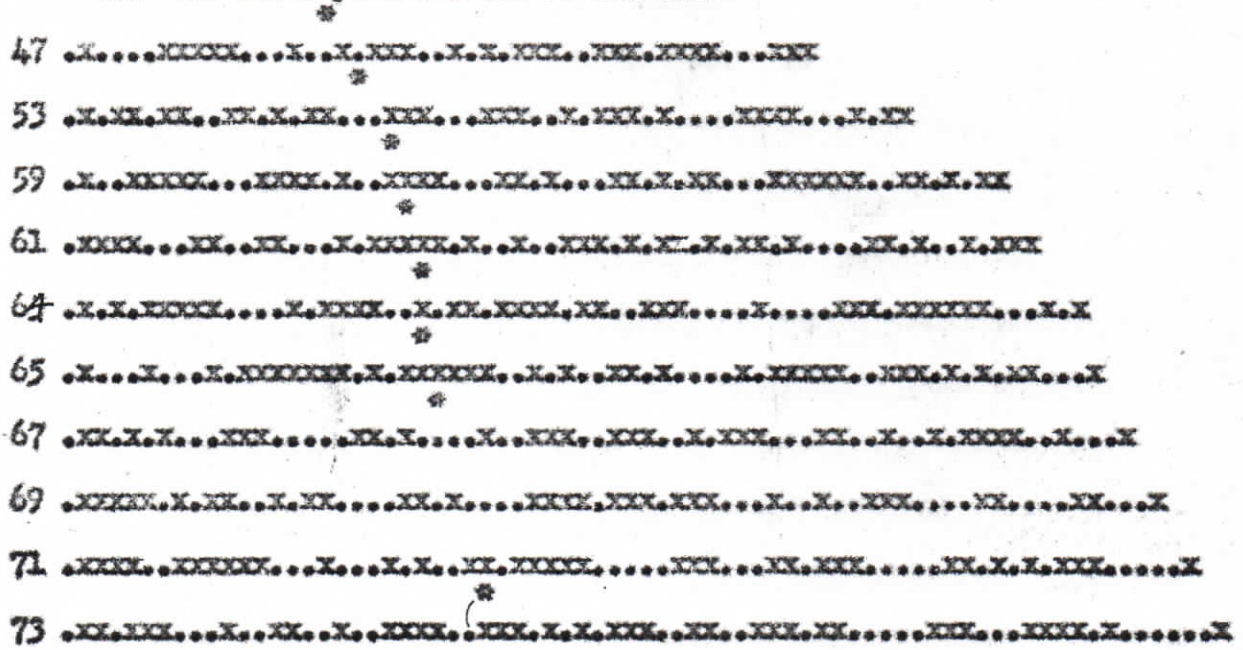
18. The motor relations, which seem to be invariant, are as follows:-

Wheel	Driven by
47	65 crosses, 67 dots
53	47 crosses, 73 dots
59	47 dots, 53 crosses
61	53 dots, 59 dots
64	59 crosses, 61 crosses
65	61 dots, 64 crosses
67	64 dots, 65 dots
69	" "
71	" "
73	" "

Apparently the reason for having the 4 largest wheels driven by the same motor is to insure a minimum key length of the product of their cycles. The machine must contain a switch with which the motor action can be thrown on or off because instances have been observed where the motor is not used. In such cases each wheel steps forward one position for each letter enciphered.

19. The daily variant seems to be the wheel order and the message variant is the initial setting of the 10 wheels. These are given by QEP numbers. It seems, however, that the settings corresponding to a given QEP do not change daily, and possible not even monthly. Messages of 9 July and 3 August with the same QEP's have the same initial settings but since they were practice messages this may not be too significant.

20. The wheel patterns are as follows:-



DECLASSIFIED
Authority NND 963016
By SB NARA, Date

The asterisks indicate the respective motor elements which correspond to the first enciphering elements. The statement about the 69 wheel in paragraph 17 is meaningless because it never acts as a motor wheel.

21. The machine was broken through reading a depth (5 for part of its length, 4 for the remainder). A depth of 4 seems about the absolute minimum. Catalogues were prepared to assist in reading. These show the possible alphabets which result from any assumption of a plain-cipher pair of letters. Another method tried involved masks and inverse probability calculations. The permuting mechanism can produce only 30 out of the 120 possible permutations. 32 seem possible but / and Z produce identical permutations and so do T and K. A substantial amount of research is being done on this problem at the moment. Solution through depths cannot be relied on for the future because of the apparent introduction of the auto-key element. One problem being considered is a statistical means of differentiating between this traffic and ordinary Fish traffic. Other special problems such as solution where wheel order is known but settings unknown, or where settings are known and wheel order unknown, are receiving consideration. Each of these can be subdivided depending on whether or not a depth is available. The possibilities of both statistical and cribbing methods are being studied.

22. Without a machine to assist in enciphering and deciphering, all the necessary operations are exceptionally laborious. Even the development of $\leq$ and π keys from a given wheel order and setting is very slow and tedious. No plans have yet been made to construct a machine. The actual machine is believed to have a zeroizing mechanism and is probably similar to a captured machine (apparently of earlier design) which Lt. Col. Rowlett and I saw on a recent trip to Knockholt. He is planning to take some photographs of this model.

23. Cryptanalysis of traffic enciphered with this machine seems to present extraordinary difficulties and I shall try to keep abreast of all developments here.

24. Paragraph 4 is not complete because it doesn't consider the situation where only the X_2 limitation is used. In such cases the first M_T position is still always cross but the second is dependent on X_2. If the X_2 pattern starts at 5, X_2 starts at 5 in the second column and a conventional dot is put in the first column to compel a cross in M_T. This was the meaning of the red dot in Report #F 5. Looked at from the viewpoint of the machine, rather than from the viewpoint of the conventions used here, we can generalize as to both types of limitation and say that M_T must be cross in the first position and that from there on M_T is M_3 as affected by the limitation - but the limitation does not operate until all elements comprising it have had an opportunity to become effective.

Walter J. Fried
Capt. Signal Corps

Encl. - 1 page

ARMY

DECLASSIFIED
Authority NND 963016
By SB NARA, Date

K5 METHOD

The success obtained from the method enunciated in the "History of November SQUID" led us to try the following method on January STICKLEBACK.

We wrote out DK 5 on a width of 31 and numbered the squares of our resulting rectangle thus:-

```
 1  2  3 . . . . .22 23  1  2 . . . . . .8
 9 10 11
17 18  .
 2  3  .
10 11  .
18 19  .
 3  4
11 12
19 20
```

We then argued in this way:- If there is a cross in $\bar{X}_2$ at the head of a column then there is a probability b that any particular sign in the column represents the sign on the corresponding character of ΔX_5, and similarly, if there is a dot in $\bar{X}_2$ at the head of a column, there is a probability b that it represents the reverse of the sign on the corresponding character of ΔX_5. Now suppose we compare the two columns of the rectangle which are numbered 1,9,17 etc., then each agreement in sign between corresponding numbers gives a factor $\frac{b^2 + (1-b)^2}{2b\,(1-b)}$ in favour of the two $\bar{X}_2$ characters being the same. (With 24 dots in $\mathcal{M}37$, this factor equals $\frac{1538}{962} = 1.6$). Thus if we get 7 agreements and 3 disagreements, it is 13 to 2 on the $\bar{X}_2$ characters being the same. We can thus select two columns which go together very well (i.e. which, when compared in the right position give a large number of agreements or disagreements) and combine the ΔX_5 signs which we deduce from them; for if our assumption about the $\bar{X}_2$ is correct, then it is 1369 to 169 on the ΔX_5 sign which occurs in both columns being right. We then compare this embryonic ΔX_5 with each column in turn and every time we make an assumption about a $\bar{X}_2$ sign, we incorporate in our ΔX_5 the signs deduceable from the assumption (of course, our start is quite arbitrary and we should bear in mind that we may be inside out on $\bar{X}_2$ and ΔX_5). In this way we cover the whole rectangle and deduce (a) $\bar{X}_2$ and (b) the hatted ΔX_5, from which, of course, X_5 may be inferred.

DECLASSIFIED
Authority NND 963016
By SB NARA, Date