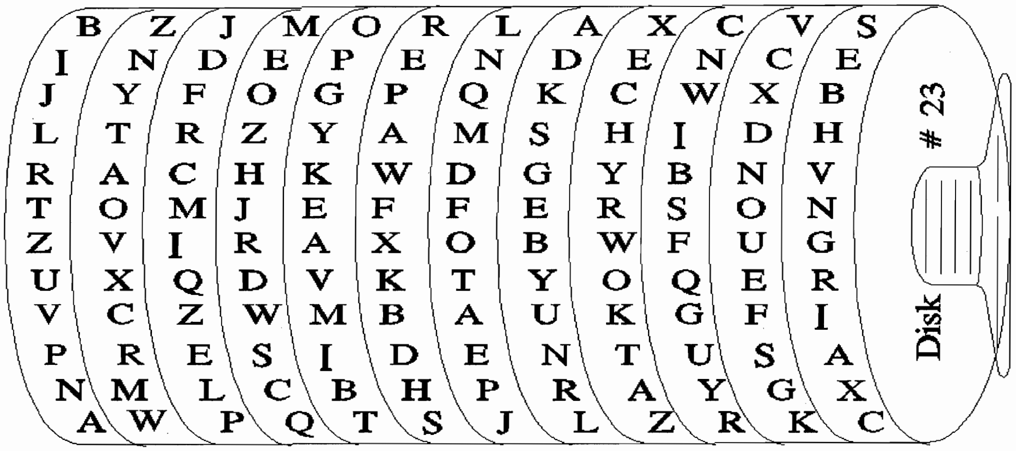
CIPHER WHEELS FOR CHEERFULLER HALLMARKING

CHRISTOPHER McMANUS
Silver Spring, Maryland

Cipher wheels are a well-known method of cryptography. Their embodiments range from the Jefferson wheel, invented by the American President Thomas Jefferson, to the Enigma machine, whose solution was critical to the course of World War II. Not surprisingly, cipher wheels are related to several basic topics in wordplay. These topics include non-crashing word sets, shiftwords, collinear words, halfway words and reflectors.

The original Jefferson wheel involved 34 disks, each with a hole in the middle and all the letters of the alphabet displayed along the edge. A simplified 12-disk version is shown below. Each Jefferson wheel disk was numbered and displayed the alphabet scrambled in a different way. To create a coded message, the sender strung the disks on a spindle, with the disk numbers in an order previously known to the recipient. The sender aligned the disks so that the message appeared on one row of the wheel. The other 25 rows would then be largely gibberish. The sender then simply sent any one of the nonsense rows to the recipient. That person would likewise align the disks in the agreed order, and turn the individual disks so that the communicated nonsense sequence appeared in a single row. The recipient would then rotate the entire set of disks until an intelligible message resulted on a single line.



In practice, the chance of a 34-disk arrangement leading to two conflicting solutions is extremely minimal. In a November 1970 article, Ross Eckler cited a finding in communications theory that even simple substitution ciphers of length 27 or greater have an extremely high probability of a unique solution. In wordplay, and where the number of disks is much smaller than 27, the chance of multiple solutions is related to the incidence of non-crashing n-letter words. If each disk contains the alphabet in certain non-random orders, number of multiple solutions is related to the incidence of collinear and halfway words. If the problem is further simplified, so that all disks arrange the letters in alphabetic order, then the number of multiple solutions is related to the incidence of shiftwords.

Non-crashing word sets

Jeff Grant reported in a May 1982 article that he found 20 non-crashing five-letter words (using only words in OED). This would be equivalent to the following arrangement of five disks, where each row (rather than column) represents a single disk. The columns of capitalized words are those found in OED, beginning with AFFIX and BWRCH. These twenty capitalized words represent conflicting solutions to that five-disk cipher wheel. The non-capitalized columns below do not form words, and could be re-assigned in different combinations.

```
Disk1: ABCDEFGHIjKLMNOPqRSTUVwxyz
Disk2: FWHISJLEGzNAUDXOvYCRTPmkqb
Disk3: FRALCEYXHqIKSUBPzGOWTMjnvd
Disk5: XHMOYDNLScFAEUWIgTBPRKzvjq
Disk4: ICUDRLNYTvJKSGOPwHBMEAxqzf
```

These same disks could be rotated so that HAIRY appears in one column; the PROSE will appear in another column. Other rotations of the above disks will produce the pairs CUPID-FOODS and LLAMA-MELEE. No rotation produces nearly as rich a trove of five letter words as the alignment above.

Shiftwords

If each disk is not an alphabetic jumble, but simply the letters in alphabetic order, then the incidence of conflicting solutions is related to the number of shiftword pairs and higher-order shiftword sets. In a February 1990 article, Len Gordon reported on the incidence of letter-shift pairs for words of length four through seven and triples for words of length four. Using the OSPD, as he did, his results can be extended to three-letter words, and the following table can be constructed.

Two-letter words: 7 shiftword pairs; 5 triplets; 2 quadruplets; 3 quintuplets; 4 sextuplets; 1 septuplet
Three-letter words: 191 shiftword pairs; 76 triplets; 19 quadruplets; 1 quintuplet
Four-letter words: 278 pairs; 11 triplets
Five-letter words: 85 pairs; no triplets
Six-letter words: 9 pairs; no triplets
Seven-letter words: 2 pairs; no triplets

The above table illustrates the exponentially increasing unlikelihood of having two conflicting solutions as the number of disks increases beyond five, even if each disk is arranged in a strictly alphabetic order.

In the February 1982 Kickshaws column, Charles Bostick extended the shiftword paradigm. In traditional shiftwords, each corresponding position in a shiftword pair is offset a given number of spaces from the letter value of the same position in the first word, just as each position in TIGER has the value of the position in PECAN plus four, or in his notation W(2,i) = W(1,i) +4 (mod 26). Bostick suggested extending the offset operation to multiplication, so that W(2,i) = n x W(1,i) (mod 26). Three restrictions are in order: n must not be 2, 13 or a multiple of either because those numbers produce only a partial alphabet; mod 26 reduces numbers to the range 1-26 by subtracting 26 from the answer as often as needed; n need only range between 1 and 25 - larger values of n are equivalent to one of those smaller values mod 26.

Bostick's paradigm produces the following patterns, which I call Bostick sequences. Considered as each forming a circle with last letter touching the first, the patterns will be called Bostick n-disks.

n	(n) letter value (mod 26)
1	ABCDEFGHIJKLMNOPQRSTUVWXYZ
3	CFILORUXADGJMPSVYBEHKNQTWZ
5	EJOTYDINSXCHMRWBGLQVAFKPUZ
7	GNUBIPWDKRYFMTAHOVCJQXELSZ
9	IRAJSBKTCLUDMVENWFOXGPYHQZ
11	KVGRCNYJUFQBMXITEPALWHSDOZ
15	ODSHWLAPETIXMBQFUJYNCRGVKZ
17	QHYPGXOFWNEVMDULCTKBSJARIZ
19	SLEXQJCVOHATMFYRKDWPIBUNGZ
21	UPKFAVQLGBWRMHCXSNIDYTOJEZ
23	WTQNKHEBYVSPMJGDAXUROLIFCZ
25	YXWVUTSRQPONMLKJIHGFEDCBAZ

Two features of the above table are immediately obvious: each ordering preserves the positions of M and Z (because n x 13 = 13 mod 26 and n x 26 = 0 mod 26 for all odd n); the last row is simply the alphabet reversed, rotated one position left.

We may also reiterate that there are no full-alphabet sequences for n divisible by 2 or 13. The Bostick 2-sequence is "BDFHJLNPRTVXZBDFHJLNPRTVXZ," which omits the odd-numbered letters (including all vowels!). I would suggest adding a complementary sequence "ACEGIKMOQSUWYACEGIKMOQSUWY," which I call the Bostick 2a-sequence. Similar complementary sequences can be constructed for the remaining even n: each sequence begins with A; successive letter values increase by n. The cases of n=13 and n=26 are best unused. The Bostick 13-sequence is simply "MZ" repeated 13 times and the Bostick 26-sequence is simply "ZZZZZZ…Z."

These Bostick n-disks (complemented by n-a disks when n is even) will be repeatedly referenced throughout this article. Keep in mind that there are only 36 different Bostick disks. They form a minuscule subset of all possible Jefferson disks, which number some 25 factorial (that is, 15 trillion trillion). Surprisingly, this tiny set of 36 Bostick disks is enough to reproduce many letter substitution quests.

Bostick's original column ended with the request: "tell me, please, for which n you find the longest [shiftword] pair." The answer will be given in the five paragraphs following this one. We note first that individual multiplicative-shift alphabets do not extend the shiftword search greatly. All 13 odd-numbered choices of multiplier n return the same two seven-letter pairs that the simple alphabetic offset does, namely ABJURER-NOWHERE and PRIMERO-SULPHUR. To see behind the fact, note that for each n-sequence, there are numbers j and k for which the sequence and the sequence when rotated k positions differ by j places in the alphabet everywhere in the sequence. For example, if the columns below are read in order, one see GHIJ..., NOPQ..., UVWX..., BCDE...

```
n offset  sequence
7  0 GNUBIPWDKRYFMTAHOVCJQXELSZ
7 15 HOVCJQXELSZGNUBIPWDKRYFMTA
7  4 IPWDKRYFMTAHOVCJQXELSZGNUB
7 19 JQXELSZGNUBIPWDKRYFMTAHOVC
              ...
```

In the previous paragraph, I mentioned that individual multiplicative-shift alphabets do not greatly extend our search for shiftwords. If we use multiplicative shifts which repeat in pairs, we increase our success rate. All combinations of alternating Bostick n disks return ABJURER-NOWHERE as a shiftword pair. Each odd-n disk can be alternated with one different n-disk to produce each of the following shiftword pairs: AERIALS-CATECHU (5-3); CHEAPLY-RATTEEN (3-9); CLASHER-DUBBINS (9-3); DIBASIC-TARSIAS (5-17); EYEFULS-SASHING (9-5); GROUSED-SHAKEUP (5-11); NIFTIES-TALLOWY (7-21); RETIRED-UNWRUNG (3-9). The numbers in parentheses show just one pair of alternating n-disks which will produce the shiftword pair.

For maximal efficiency, we should allow a choice of Bostick n-disk for each letter position, just as the Jefferson wheel used different disks at each position. In that case, any non-crashing word-pair whose letter values differ at each position by a number other than 13 can be trivially re-created as adjoining columns, when the rows of n-disks reflect those position differences. For example, among 11-letter words in OSPD I found 72 word pairs where the letter difference at each position was an odd number other than 13. These pairs include CHEERFULLER-HALLMARKING. Align the Bostick n-disks, where the n's are letter differences, so that the first column reads CHEERFULLER. Then HALLMARKING emerges as the second column.

```
Disk01 ( 5) CHMRWBGLQVAFKPUZEJOTYDINSX
Disk02 (19) HATMFYRKDWPIBUNGZSLEXQJCVO
Disk03 ( 7) ELSZGNUBIPWDKRYFMTAHOVCJQX
Disk04 ( 7) ELSZGNUBIPWDKRYFMTAHOVCJQX
Disk05 (21) RMHCXSNIDYTOJEZUPKFAVQLGBW
Disk06 (21) FAVQLGBWRMHCXSNIDYTOJEZUPK
Disk07 (23) UROLIFCZWTQNKHEBYVSPMJGDAX
Disk08 (25) YXWVUTSRQPONMLKJIHGFEDCBAZ
Disk09 (23) LIFCZWTQNKHEBYVSPMJGDAXURO
Disk10 ( 9) ENWFOXGPYHQZIRAJSBKTCLUDMV
Disk11 (15) RGVKZODSHWLAPETIXMBQFUJYNC
```

The above-mentioned algorithm will work as long as no letter differences are zero or 13. If any differences are multiples of 2, then the Bostick n-disks at those positions need to be n-a disks (rather than n-disks) wherever the first letter of the pair is odd-numbered. The OSPD has very few 12-letter words, and all non-crashing word pairs differed in at least one position by 13.

The OSPD has been used throughout this study as a consistent frame of reference. We can certainly improve our record by using a word list which features a richer mix of long words. One such source is Stedman's Medical Dictionary (26th Edition). In an electronic version of this dictionary, the longest non-crashing pair (and with no letter differences equal to 13) were each 31 letters long: CONJUNCTIVODACRYOCYSTOSTOMIZING - HYPERPREBETALIPOPROTEINEMICALLY By the same process employed with CHEERFULLER-HALLMARKING above, this pair can quickly be reproduced on Bostick n-disks and n-a disks as consecutive columns.

So the answer to Bostick's request is that, for the OSPD, no single choice of n produces a breakthrough, but that by allowing n to vary for each position, almost any non-crashing word pair up to length 11 become shiftwords. For a larger reservoir of long words, a combination of Bostick n-disks and n-a disks can produce word pairs up to at least length 31.

Collinear words

In a November 1984 Kickshaws column, Bostick examined the concept of collinear words. Len Gordon explicated the use of modular alphabetic sequences and derived some remarkable unbroken collinear word chains in May and August 1991 Kickshaws columns. The name collinear refers to the fact that the letter values of a collinear word set will all lie on a single line in n-space.

Collinear words are best explained by an example. If we begin with the word GYP, and replace the first letter by the next letter alphabetically, the second letter by the letter 4 places earlier than it in the alphabet, and the third letter by the letter three places earlier than it in the alphabet, we get the sequence HUM. Continuing the process we successively get IQJ, JMG, KID and LEA. Three letters earlier than A is X, so the sequence continues MAX, NWU, OSR, POO, and beyond. IQJ, JMG, NWU and OSR are not words, but the other seven sequences are words, and form a collinear word set together with YEN and COB.

The process of creating collinear words is substantially duplicated by a wheel of appropriate disks. Consider the word GYP. The first instruction, " replace the first letter by the next letter alphabetically," corresponds to use of a Bostick-1 disk rotated six places left ("GHIJKLMNOPQRSTUVWXYZABCDEF"). The second instruction requires use of the Bostick 22a-sequence, rotated seven places right (YUQMIEAWSOKGCYUQMIEAWSOKGC). The third instruction nearly corresponds to the Bostick 23-sequence, when rotated eleven places left (PMJGDAXUROLIFCZWTQNKHEBYVS). Then the disks align as below:

```
Disk 1 (  1): GHIJKLMNOPQRSTUVWXYZABCDEF
Disk 2 (22a): YUQMIEAWSOKGCYUQMIEAWSOKGC
Disk 3 ( 23): PMJGDAXUROLIFCZWTQNKHEBYVS
```

The first column spells GYP, the second HUM, the fifth KID, the sixth LEA. Continuing along, we also read MAX, POO, YEN, and COB.

Halfway and reflected words

Halfway words were introduced by myself in a February 1992 article. Halfway words are a trio, such as VAGUE-TEMPO-RISKY, where each letter of the middle word lies halfway between the corresponding letters of the first and last word. In the example, the first letters successively decrease in value by four, the second letters increase by four, the third letters increase by six, the fourth decrease by five, and the fifth increase by ten. By modularity, a decrease of four is equivalent to an increase of 22. We can therefore choose as our five disks the Bostick 24a-sequence, the 4a-sequence, the 6a-sequence, the 21 sequence, and the 10a-sequence. If we rotate each disk so that the first column reads VAGUE, then the disks line up:

```
Disk 1 (24a): VTRPNLJHFDBZXVTRPNLJHFDBZX
Disk 2 ( 4a): AEIMQUYCGKOSWAEIMQUYCGKOSW
Disk 3 ( 6a): GMSYEKQWCIOUAGMSYEKQWCIOUA
Disk 4 ( 21): UPKFAVQLGBWRMHCXSNIDYTOJEZ
Disk 5 (10a): EOYISCMWGQAKUEOYISCMWGQAKU
```

Voila! The first three columns produce our halfway-word set. Any halfway word set can be reproduced as consecutive columns in a disk arrangement, when suitable Bostick n-disks are chosen and rotated.

My definition of halfway words specifically excluded circular alphabetic sequences such as ABC…XYZABC…XYZ. For my definition the only letter halfway between X and B was M (13 is halfway between 24 and 2). In the extended alphabetic sequence, we see …XYZABC…. In this sequence, Z is halfway between X and B. In his November 1987 Kickshaws column, Dave Morice had coined the terms 'reflector' and 'reflected words.' This category, as I now notice, is equivalent to halfway words when modular alphabetic sequences are allowed. The use of n-disks to find halfway word sequences will also find reflected word sequences as a bonus.

In summary, cipher wheels are a physical form which embodies a variety of wordplay topics. Their mechanism throws light on the incidence of non-crashing words, and on construction of shiftwords, collinear words, halfway words and reflectors. Thank you for your idea, Mr. Jefferson.