Example 5 (Worksheet) — Cracking a Shift Cipher by Frequency

Problem. We intercepted the ciphertext `ZNK KGXRE HOXJ MKZY ZNK CUXS` produced by a shift cipher. What was the original plaintext?

Why this works

In English text, some letters appear more often (E, T, A, O, I, N). A shift cipher preserves *relative* frequencies, just moves them around the alphabet. If a letter occurs most often in the ciphertext, it likely corresponds to one of the most common plaintext letters. Hypothesize a mapping, compute the shift k, and test by decrypting.

Step 1 — Count letter frequencies

Ignore spaces/punctuation and count:

K	Z	X	N	G	R	E	H	O	J	M	Y	C	U S
4	3	3	2	1	1	1	1	1	1	1	1	1	1

The most frequent letter is `K`.

Step 2 — Form a hypothesis

In normal English, `E` is often the most frequent letter. Hypothesize that E (which is 4 with A=0) was shifted to K (which is 10). Then the encryption used

$$k \equiv 10 - 4 \equiv 6 \pmod{26}.$$

So decryption should be $p \equiv c - 6 \pmod{26}$.

Step 3 — Test by decrypting

Try a few letters to check the hypothesis:

$$\texttt{Z}\ (25) \mapsto 25 - 6 = 19 \Rightarrow \texttt{T}, \qquad \texttt{N}\ (13) \mapsto 7 \Rightarrow \texttt{H}, \qquad \texttt{K}\ (10) \mapsto 4 \Rightarrow \texttt{E}.$$

The first three letters become `THE`, which is promising. Decrypt the whole string with $k = 6$.

Step 4 — Conclusion

Full decryption yields:

`THE EARLY BIRD GETS THE WORM`

Because this makes excellent English, our hypothesis $k = 6$ is accepted.

Tips, tricks, and pitfalls

- **A=0** convention: $E = 4,\ K = 10$. Off-by-one mistakes derail the shift quickly.
- **Test, then trust.** A frequency guess is just a hypothesis; always decrypt a chunk to confirm.
- **One-letter words** in ciphertext often map to `A` or `I`; common bigrams like `TH`, `HE`, `TO` are great anchors.
- **Decrypt rule:** $p \equiv c - k \pmod{26}$. Negative values? Add 26.

Practice — Your Turn

Problem A (easier). Decrypt the ciphertext `URYYB JBEYQ` given it was made with a shift $k = 13$.

Hint: subtract 13 from each letter mod 26.

Problem B (similar). The ciphertext below was made with an *unknown* shift:

`ZHOFRPH WR FODVV`

Find k and the plaintext.

Hints: the block `WR` might be `TO`, or `FODVV` looks like `CLASS`.

Problem C (harder). The ciphertext was produced by a shift cipher with *unknown* k:

`YMJ VZNHP GWTBS KTC OZRUX TAJW YMJ QFED ITL`

Determine k using a smart guess (look for a repeated common word), then decrypt the whole message.

Reflection. Briefly explain why frequency analysis defeats a shift cipher but not a one-time pad.