1. For this problem, a shorthand notation is used in which ciphertext is capitalized (e.g. RSZWO) and plaintext is lowercase (e.g. ¡plaintextexample¿).

   A frequency analysis program written in C tells us the frequency distribution is as follows:

B	–	2
C	–	18
D	–	7
G	–	13
H	–	11
J	–	3
K	–	6
L	–	1
N	–	13
O	–	18
P	–	14
Q	–	2
R	–	11
S	–	16
T	–	3
V	–	7
W	–	11
X	–	5
Y	–	4
Z	–	19

   A cursory look of our ciphertext shows Z, C, and O to be the most common letters. We can make use of our knowledge of bigrams - as well as the fact that 'he' is very common, but 'eh' is not - to determine that O likely maps to e, and W to h. Next, we see that Z occurs 19 times and is doubled - though this could be across a word break, it is likely that it is t - most common by far of all commonly doubled letters, and it makes sense with bigrams too - th is the most common. A trip through the dictionary searching for a commonly occurring pattern in this ciphertext reveals that what appears to be the word "the" is likely a part of the word "mathematics"; following this hunch confirms many of the conclusions I had previously come to but was unable to confirm regarding the substitutions, and other words and patterns start showing up as I complete the substitutions to fill in just the first word - which is "mathemeticians," not "mathematics." Continuing along this path we quickly find the solution: "MATHEMATICIANS AND MUSICIANS MAY SPEND MOST OF THEIR TIME IN THE MATHEMATICAL WORLD OF HYPOTHESIS AND REASON BUT THE INNER LIFE OF THEIR ARTS IS IN THE WORLD OF FORMS IN THE PROCESS OF DIALECTIC AND ITS ARGUMENT BY METAPHOR."

2. (a) Using the formula

   $$IC = \frac{\sum_{i=1}^{c} n_i(n_i - 1)}{N(N-1)}$$

   and substituting appropriate values for this ciphertext, we arrive with an IC of 0.04285 (in reality, this was found via an online IC calculator tested to match the results of IC calculations in class).

   (b) Estimated IC for a key length m can be calculated by

   $$\frac{1}{m}\kappa_S + \frac{m-1}{m}\kappa_R$$

   Using this, we will test key lengths 3 through 10 (assuming a longer key is excessive and a shorter key makes decryption too easy).

   $$E_1(IC) = 0.06600$$

   $$E_2(IC) = 0.05200$$

   $$E_3(IC) = 0.04733$$

   $$E_4(IC) = 0.04500$$

   $$E_5(IC) = 0.04360$$

$$E_6(IC) = 0.04266$$

$$E_7(IC) = 0.04200$$

$$E_8(IC) = 0.04150$$

$$E_9(IC) = 0.04111$$

$$E_{10}(IC) = 0.04080$$

It looks from this method as though the key length is likely 6.

(c) Again, I wrote a C program to do Kasiki's method on this text - its output is a line for every key length tested (the range of which is supplied by the user), with the IC of every key position's subtext on that row as well as the key position itself. This is the output of that program using our sample text and the same range of key lengths as in the last problem.

```
Total number of chars: 337
Key length = 3, ics:    0.055626 0.047136 0.043758
Key length = 4, ics:    0.037255 0.043029 0.037866 0.047332
Key length = 5, ics:    0.035996 0.043898 0.047490 0.038896 0.042515
Key length = 6, ics:    0.048872 0.055195 0.037013 0.059740 0.044805 0.042857
Key length = 7, ics:    0.030612 0.045213 0.042553 0.045213 0.045213 0.040780 0.042553
Key length = 8, ics:    0.043189 0.048780 0.037166 0.041812 0.037166 0.046458 0.040650
>                       0.046458
Key length = 9, ics:    0.083926 0.048364 0.069701 0.071124 0.061562 0.063063 0.052553
>                       0.067568 0.054054
Key length = 10, ics:   0.030303 0.051693 0.032086 0.051693 0.044563 0.040998 0.033868
>                       0.053030 0.039773 0.037879
```

The means of these values are (3)0.48840, (4)0.041371, (5)0.041759, (6)0.048080, (7)0.041734, (8)0.42710, (9)0.063546, (10)0.041589. As 9 is by far closest to the expected IC for English plaintext, our key is almost certainly 9 characters long.

(d) Using this method and a nifty Vigènere Cipher tool at simonsingh.net/The_Black_Chamber as well as some freqency analysis (largely provided by the aforementioned Vigènere tool) we can find rather easily that the key phrase is "ANIELTKYW", and the plaintext:

"I LEARNED HOW TO CALCULATE THE AMOUNT OF PAPER NEEDED FOR A ROOM WHEN I WAS AT SCHOOL. YOU MULTIPLY THE SQUARE FOOTAGE OF THE WALLS BY THE CUBIC CONTENTS OF THE FLOOR AND CEILING COMBINED AND DOUBLE IT. YOU THEN ALLOW HALF THE TOTAL FOR OPENINGS SUCH AS WINDOWS AND DOORS. THEN YOU ALLOW THE OTHER HALF FOR MATCHING THE PATTERN THEN YOU DOUBLE THE WHOLE THING AGAIN TO GIVE A MARGIN OF ERROR AND THEN YOU ORDER THE PAPER."