Catastrophic Cancellation on the High Seas

Captain Bob is the captain of a large cruise ship which is currently leaving the port of Baltimore, Maryland. The first destination on this Mediterranean Cruise is Athens, Greece. The crew will be at sea for seven consecutive days until its first stop. Captain Bob, a recent Weight Watchers member, is in a predicament. Each week he must attend a Weight Watchers meeting, updating the group on his diet and weight loss progress. Weight Watchers has agreed to allow Captain Bob to conduct this week's meeting over the phone. Captain Bob's problem lies in the fact that his cruise ship has no personal scales and none are available at the port in Athens. The ship's engineer has persuaded Captain Bob to weigh himself using the large scales which all major ports have to weigh the ship and its cargo. The engineer explains that when they reach Athens, and after the passengers have disembarked, they could simply weigh the ship with the captain aboard. The captain could step off the ship, and the ship alone could then be weighed. A simple subtraction of the weight of the ship and the captain together minus the weight of the ship alone would give the captain's weight. The scales are known to be accurate to six digits. By the engineer's reasoning that should be plenty enough to accurately capture the captain's weight.

On the seventh day, after all the passengers had disembarked from the ship, the captain went through with the procedure. The engineer reported the weight of the ship plus the weight of the captain to be $1.000004 \cdot 10^8$ pounds and the weight of the ship to be $1.000001 \cdot 10^8$ pounds. After the simple subtraction, the captain must then weigh $.000003 \cdot 10^8 = 300$ pounds.

The captain was shocked, "300 lbs! Just last week I weighed 200 lbs!"

"But this is a cruise ship and the cook did serve filet mignon and cheesecake all week," countered the engineer.

Still in disbelief, the captain postponed his call to the Weight Watchers club until he got to the bottom of this mystery. Since the ship had docked for the day in Athens, the captain decided to take the engineer's latest piece of advice and visit the numerical analysis professor of a nearby university. By lunch, Captain Bob had caught up with one of the world's most renowned numerical analysts, Dr. Socrates.

"Ahh, I see. Such a common problem–the problem of catastrophic cancellation," Dr. Socrates remarked.

"Great! So you can explain this to me," said the captain excitedly.

"Sure, but it may take some calculations and fiddling." And so begins the professor's explanation of catastrophic cancellation.

"I'll explain exactly what catastrophic cancellation refers to in just a bit, but first, here's an intuitive explanation for it. It happens when we are subtracting two numbers of the same sign and the two numbers are in error. Specifically, when the two operands are in error and the result of the subtraction is much[1] smaller than the two operands, we can encounter this problem because then the result of the subtraction is of the same magnitude as the error. Scales, just like computers, do not always represent measurements exactly. In your situation, Captain, you used a scale that was accurate to 6 digits to represent your approximate weight and you subtracted the weight of the ship from the weight of the ship plus yourself, two numbers of equal magnitude. Therefore, catastrophic cancellation, the loss of accuracy in a subtraction, has occurred. Your weight as recorded on the scale may be completely wrong.

"Let's be more specific. A scale gives an approximation to your actual weight. Naturally, with any approximation, we are interested in how close the approximation is to the actual value. We call $w(x)$ the approximation of the actual value, x. In your case, Captain Bob, we will let

$w(b) =$ approximate weight of both the ship and the captain $= 1.000004 \cdot 10^8$,

$w(s) =$ approximate weight of the ship $= 1.000001 \cdot 10^8$.

Then $w(b) - w(s) =$ approximate weight of the captain
$$= 1.000004 \cdot 10^8 - 1.000001 \cdot 10^8 = .000003 \cdot 10^8.$$

By the engineer's calculation, you weigh 300 pounds. The engineer did the subtraction properly; there is no error in the subtraction operation. Yet you claim the resulting answer from the subtraction is still in error. You are right. The problem lies in the last digit of the subtraction. While it may be true that $4 - 1 = 3$, we are not sure that the 4 and the 1 are accurate. The scale you used only guarantees 6 digits of accuracy. The leading 5 zeros in the mantissa of the result are accurate. However, the 3 may not be accurate. Therefore, we have no idea whether the result of 300 pounds is accurate.

"Consider the weight of the ship plus Captain Bob. The seventh digit, 4, may be inaccurate. Suppose the next time we weigh the ship plus the captain the scale reads $1.000003 \cdot 10^8$. Then $w(b) - w(s) = .2000000 \cdot 10^3 = 200$ pounds. You, Captain, would be much happier with this answer. Yet every digit is still in error. Now suppose we do the procedure one more time. The ship plus the captain weighs $1.000007 \cdot 10^8$ this time. Then you would weigh 600 pounds. Preposterous? Yes. But this teaches you to compare the magnitude of the error in the operands with the magnitude of the result."

"So I should just call Weight Watchers and tell them that I don't know my weight this week. I weigh so much less than the ship that catastrophic cancellation has occurred," said Captain Bob.

"Yes, and if your next stop is Rome, I know they have personal scales there."

"Thanks for all your help, Dr. Socrates," said the captain with relief.

"Sure!"

The captain departed for Rome thoroughly satisfied with Dr. Socrates' explanation of his phony weight, yet Dr. Socrates continued to ponder the precise nature of catastrophic cancellation. In fact, later that day Dr. Socrates sat at his desk and revised the notes he planned to present to his class that evening. Some excerpts from his lecture on "Catastrophic Cancellation and the Captain Bob Story" follow.

---

[1]The mathematical meaning of "much" will be clarified in the next section.

An Analysis of Catastrophic Cancellation

The absolute error of a measurement is the difference between the measured value and the actual value. The absolute error in the weight of the ship is $|w(s) - s|$. In two different scenarios an absolute error of .1 units might have contrasting meanings. For example, in measuring the distance from a point in Athens to a point in Rome an absolute error of .1 inches would be laudable, while in a heart surgery procedure, which requires a .2 inch incision, an absolute error of .1 inches would be unacceptable[2].

To remedy this problem, we need to consider relative error, that is, we talk about the error relative to the magnitude of one of the values involved. There are two ways to make an absolute error relative. One type of relative error starts with the absolute error and makes it relative by dividing by the actual value, resulting in $|\frac{w(s)-s}{s}|$. The other type of relative error divides the absolute error by the approximated value, giving $|\frac{w(s)-s}{w(s)}|$. Both relative errors normalize the absolute error. Therefore, the .1 inch absolute error in the measurement of the distance from Athens to Rome (646 miles) gives a relative error of $\frac{.1}{40930560} = .0000000244$, a miniscule value and thus an excellent result. On the other hand, the .1 inch absolute error in the heart surgery gives a relative error of $\frac{.1}{.2} = .5$. The error in the size of the incision is half the size of the actual incision! Looking at relative error as opposed to absolute error can often tell us whether an error is acceptable in a particular context.

Let's take this notion of relative error and get back to Captain Bob. Recall that $w(s)$ represents the approximate weight of the ship and s represents the actual, yet unknown, weight of the ship. Let's use the first type of relative error. Let δ_s represent this relative error, i.e., $\delta_s = |\frac{w(s)-s}{s}|$. Similarly, δ_b represents the relative error associated with the ship plus the captain. Thus, $\delta_b = |\frac{w(b)-b}{b}|$. With any measuring device, we would like to know just how "good" its approximations are, compared to the actual values. Let's assume that the scale is calibrated and its relative errors do not exceed U. Therefore, $|\delta_b| \leq U$ and $|\delta_s| \leq U$. In Captain Bob's case, $U = .5 \cdot 10^{-6}$. Hence the scale guarantees 6 digits of accuracy.

Now we have all the machinery in place to analyze the relative quality of an approximation such as the captain's weight. Specifically, how well does $w(b) - w(s)$ approximate $b - s$? Consider the relative error associated with the captain's weight:

$$
\begin{aligned}
\frac{|w(b) - w(s) - (b - s)|}{|b - s|} &= \frac{|(w(b) - b) - (w(s) - s)|}{|b - s|} \\
&\leq \frac{|b| \cdot \delta_b + |s| \cdot \delta_s}{|b - s|} \\
&\leq \frac{|b| + |s|}{|b - s|} \cdot U = A \cdot U,
\end{aligned}
$$

where

$$
A = \frac{|b| + |s|}{|b - s|}.
$$

A is called the amplification factor or the condition number of the subtraction. A condition number for a mathematical problem indicates how much the input error is amplified in the final result. Here the input error is U, the error in the operands. The error in the captain's weight can be as large as $A \cdot U$. From the formula for A, we observe that the condition number for subtraction is large when the result

---

[2] The idea for this example is due to Carl Meyer.

of the subtraction, $b - s$, has much smaller magnitude than the individual operands, b and s. If this is the case, the subtraction is ill-conditioned. This is what happens in the captain's example. In that example, $|b - s|$ is on the order of 10^2 and $|b| + |s|$ is on the order of 10^8, hence $|b - s| \ll |b| + |s|$, since the captain's weight is so small in comparison to the weight of the ship. Therefore, A is large, approximately 10^6, and the relative error associated with the captain's weight could be large. This means the error that occurs when weighing the ship can be amplified by as many as six orders of magnitude in the subtraction that approximates the weight of the captain. This explains the loss of digits of accuracy in the resulting answer of 300 pounds.

Catastrophic cancellation is not limited to Captain Bob's scale. The same considerations apply when we subtract two numbers, x and y, on a computer, such as a calculator or PC. The mere process of entering the numbers can cause errors. For instance, a computer might represent $\frac{1}{3}$ as .3333. To account for this rounding, let $fl(x)$ denote the floating point representation of a number x. For example, $fl(\frac{1}{3}) = .3333$. Since only 4 digits are used in the decimal representation of $\frac{1}{3}$, $U = .5 \cdot 10^{-4}$, where U represents the unit roundoff of the computer. In a floating point subtraction, $fl(x) - fl(y)$, the relative error associated with this difference is

$$\frac{|fl(x) - fl(y) - (x - y)|}{|x - y|} \leq A \cdot U,$$

where

$$A = \frac{|x| + |y|}{|x - y|}$$

is the amplification factor.

Now it should be clear why subtraction of floating point numbers of almost equal magnitude should be avoided. The result of the subtraction has a magnitude similar to the error in the operands. In this case, the subtraction is ill-conditioned and the relative error can be large. The computed difference can be completely wrong. Unfortunately, catastrophic cancellation cannot be blamed the next time the bathroom scale registers an unsightly number, unless, of course, the Captain Bob procedure is used.