# Mathematical Apologies

*Professor John Butcher, The University of Auckland*

"Apology" is used in the sense of justification or defence of an opinion firmly held. From time to time mathematicians are called upon to explain why they are what they are and they have to be apologists for what they regard as important, just as much as holders of an unfashionable belief or ideology have to be prepared to explain themselves to sceptics. For some reason it is difficult to make anything reasonable out of what we do, whereas other behaviour that is no more obviously sensible seems to justify itself. The famous mathematician G. H. Hardy felt compelled to call his autobiography "A Mathematician's Apology" and my choice of the name for this column is partly as a tribute to him. I hope that these apologies will become a regular feature of New Zealand Mathematics Magazine.

In these articles I will explore some small mathematical ideas and see what can be learnt from them. The starting point will usually be something quite elementary that can be explained to a high-school pupil, if he or she is willing to work at it. It will not, however, be something that can easily be dismissed as either true or false and therefore of no further interest, any more than a musical phrase is dismissed by a composer who still sees some potential new development or variation in it.

The general theme of the first few Apologies will be the computation of $\pi$. Today we will go back to one of the oldest known methods for the evaluation of this constant. In fact it goes back to Archimedes (circa 287BC - 212BC) and starts from the inequality $n\sin(\pi/n) < \pi < n\tan(\pi/n)$, which can be interpreted as stating that the circumference of a circle lies between the lengths of the perimeters of an inscribed and of a circumscribed regular polygon with $n$ sides. For $n$ a low integer it is sometimes possible to evaluate the upper and lower bound exactly and to then improve the result by obtaining the values of $2n\sin(\pi/2n)$ and $2n\tan(\pi/2n)$ in terms of the known bounds. Write $P_n = n\tan(\pi/n)$ and $p_n = n\sin(\pi/n)$. By elementary trigonometry, we find that

$$P_{2n} = \frac{2p_nP_n}{p_n+P_n}, \qquad p_{2n} = \sqrt{P_{2n}p_n}.$$

Using the known values of $p_6 = 3$, $P_6 = 2\sqrt{3}$, the following table gives the first few approximations

| $n$ | $p_n$ | $P_n$ |
|---:|---|---|
| 6 | 3.0000000000 | 3.4641016151 |
| 12 | 3.1058285412 | 3.2153903092 |
| 24 | 3.1326286133 | 3.1596599421 |
| 48 | 3.1393502030 | 3.1460862151 |
| 96 | 3.1410319509 | 3.1427145996 |
| 192 | 3.1414524723 | 3.1418730500 |
| 384 | 3.1415576079 | 3.1416627471 |
| 768 | 3.1415838921 | 3.1416101766 |

¿From the $n = 192$ entries we derive the famous bounds, due to Archimedes: $3\frac{10}{71} < \pi < 3\frac{1}{7}$. This takes quite a lot of work, especially if we do not have the benefit of modern computing aids. In fact, apart from the basic data and some multiplications, divisions and additions, the heavy part of the work is an extra square root calculation for each iteration.

Can we squeeze more information out of this data? Can we avoid so many square-root calculations and still obtain good accuracy? As a first step in investigating these questions, let us see if we can find any pattern in the way the errors behave. The following table gives not $p_n$ and $P_n$ but the values of $\pi - p_n$ and $P_n - \pi$.

| $n$ | $\pi - p_n$ | $P_n - \pi$ |
|---|---|---|
| 6 | 0.1415926536 | 0.3225089615 |
| 12 | 0.0357641124 | 0.0737976556 |
| 24 | 0.0089640403 | 0.0180672885 |
| 48 | 0.0022424505 | 0.0044935615 |
| 96 | 0.0005607027 | 0.0011219461 |
| 192 | 0.0001401813 | 0.0002803964 |
| 384 | 0.0000350457 | 0.0000700935 |
| 768 | 0.0000087614 | 0.0000175230 |

It looks as though, as we move down the list, the value of an entry divided by the entry above it becomes closer and closer to $\frac{1}{4}$. It also looks as though the values of $P_n - \pi$ become closer and closer to twice the corresponding values of $\pi - p_n$. Both of these facts can be verified by using the series for the sine and tangent functions

$$\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \cdots, \qquad \tan(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \cdots,$$

from which it follows that

$$p_n = n\sin\left(\frac{\pi}{n}\right) = \pi - \frac{\pi^3}{6n^2} + \frac{\pi^5}{120n^4} - \cdots, \qquad P_n = n\tan\left(\frac{\pi}{n}\right) = \pi + \frac{\pi^3}{3n^2} + \frac{2\pi^5}{15n^4} + \cdots.$$

If we have two consecutive members of a column and they are equal to

$$A = \pi + E, \quad \text{and} \quad B = \pi + \frac{1}{4}E,$$

approximately, then $\pi$ is approximately equal to $\frac{4}{3}B - \frac{1}{3}A$. This suggests that we can get improved sequences by using $\frac{4}{3}p_{2n} - \frac{1}{3}p_n$ in place of $p_{2n}$ and $\frac{4}{3}P_{2n} - \frac{1}{3}P_n$ in place of $P_{2n}$. These sequences are

| $2n$ | $\frac{4}{3}p_{2n} - \frac{1}{3}p_n$ | $\frac{4}{3}P_{2n} - \frac{1}{3}P_n$ |
|---|---|---|
| 12 | 3.1411047216 | 3.1324865405 |
| 24 | 3.1415619706 | 3.1410831531 |
| 48 | 3.1415907330 | 3.1415616395 |
| 96 | 3.1415925335 | 3.1415907278 |
| 192 | 3.1415926461 | 3.1415925334 |
| 384 | 3.1415926531 | 3.1415926461 |
| 768 | 3.1415926536 | 3.1415926531 |

Unfortunately, even though the two columns now seem to converge to the correct answer more rapidly, the property of bracketing $\pi$ between the two columns has disappeared. We can recover this by carrying out the same operations to the new table but using the factors $\frac{16}{15}$ and $\frac{1}{15}$ instead of $\frac{4}{3}$ and $\frac{1}{3}$ respectively. This gives the following table

| $4n$ | $\frac{64}{45}p_{4n} - \frac{20}{45}p_{2n} + \frac{1}{45}p_n$ | $\frac{64}{45}P_{4n} - \frac{20}{45}P_{2n} + \frac{1}{45}P_n$ |
|---|---|---|
| 24 | 3.1415924539 | 3.1416562606 |
| 48 | 3.1415926505 | 3.1415935386 |
| 96 | 3.1415926535 | 3.1415926670 |
| 192 | 3.1415926536 | 3.1415926538 |
| 384 | 3.1415926536 | 3.1415926536 |

We don't need to go so far this time because, to the accuracy to which we have presented the answers, we get perfect accuracy even in these few steps.

The other observation we have made, that the errors in $P_n$ are approximately $-2$ times those in $p_n$, suggests that we might get better approximations from the values of $\frac{2}{3}p_n + \frac{1}{3}P_n$. This gives the table

| $n$ | $\frac{2}{3}p_n + \frac{1}{3}P_n$ |
|---|---|
| 6 | 3.1547005384 |
| 12 | 3.1423491305 |
| 24 | 3.1416390562 |
| 48 | 3.1415955404 |
| 96 | 3.1415928338 |
| 192 | 3.1415926649 |
| 384 | 3.1415926543 |
| 768 | 3.1415926536 |

This gives a rapidly converging sequence, but the advantage of the original Archimedes scheme, in which the value of $\pi$ is bracketed between two sequences, one increasing and one decreasing, is lost.

Another means of squeezing out extra information is to use extrapolation. Amongst the many formulations for the value of a polynomial at a given point which interpolates an existing function on a given data set is due to Neville and the New Zealander Aitken. Let $\phi(t)$ denote a function for which the values of $\phi(t_1)$ and $\phi(t_2)$ are given. If a polynomial of degree 1 is fitted to this data, the value at an aribtrary value of $t$ of this polynomial, which can be used to approximate the value of $\phi(t)$, is given by

$$\phi(t) \approx \frac{t - t_2}{t_1 - t_2}\phi(t_1) + \frac{t - t_1}{t_2 - t_1}\phi(t_2).$$

Denote this expression by $\phi(t; t_1, t_2)$, because it is the interpolated value based on data at $t_1$ and $t_2$. If we have only a single data point, then the appropriate approximation is just $\phi(t; t_1) = \phi(t_1)$. The nice thing is that the interpolated value based on many points can be evaluated recursively according to the formula

$$\phi(t; t_1, t_2, \ldots, t_n) = \frac{t - t_n}{t_1 - t_n}\phi(t; t_1, t_2, \ldots, t_{n-1}) + \frac{t - t_1}{t_n - t_1}\phi(t; t_2, \ldots, t_n).$$

We can apply this formula to each of two functions associated with the polygons of Archimedes

$$f(t) = t^{-\frac{1}{2}}\sin(\pi\sqrt{t}), \qquad F(t) = t^{-\frac{1}{2}}\tan(\pi\sqrt{t})$$

By using the values $t = n^{-2}$ for various $n$, we can interpret the $p_n$ and $P_n$ data in terms of these functions. Furthermore, by choosing $t = 0$ as the point at which approximations are to be found, we are effectively extrapolating the data to the limiting values at $t = 0$ for which $f(0) = F(0) = \pi$.

It is customary to arrange the computations in a table like this:

$$\begin{array}{c|lllll}
t_1 & \phi(t;t_1) & & & \\
 & & \phi(t;t_1,t_2) & & \\
t_2 & \phi(t;t_2) & & \phi(t;t_1,t_2,t_3) & \\
 & & \phi(t;t_2,t_3) & & \phi(t;t_1,t_2,t_3,t_4) \\
t_3 & \phi(t;t_3) & & \phi(t;t_2,t_3,t_4) & \\
 & & \phi(t;t_3,t_4) & & \phi(t;t_2,t_3,t_4,t_5) \\
t_4 & \phi(t;t_4) & & \phi(t;t_3,t_4,t_5) & \\
 & & \phi(t;t_4,t_5) & & \\
t_5 & \phi(t;t_5) & & &
\end{array}$$

Note that each entry, after the $\phi(t;t_i)$ column, is found from the two closest entries to its left. Note also that we will use tables like this only with $t = 0$, as we have described above.

To show how this works, the functions $f(t) = p_n$ and $F(t) = P_n$, are computed for values $t = \frac{1}{n^2}$ for $n = 3, 4, 5, 6, 8, 10, 12, 16, 20$ and the first four extrapolated columns are evaluated. Note that the first few of these can be evaluated using the table

| $n$ | $p_n$ | $P_n$ |
|---|---|---|
| 3 | $\frac{3}{2}\sqrt{3}$ | $3\sqrt{3}$ |
| 4 | $2\sqrt{2}$ | 4 |
| 5 | $\frac{5}{4}\sqrt{10-2\sqrt{5}}$ | $5\sqrt{5-2\sqrt{5}}$ |
| 6 | 3 | $2\sqrt{3}$ |
| 8 | $4\sqrt{2-\sqrt{2}}$ | $8(\sqrt{2}-1)$ |
| 10 | $\frac{5}{2}(\sqrt{5}-1)$ | $2\sqrt{25-10\sqrt{5}}$ |

Later members of the table can be found in terms of earlier members; it is interesting that this can be done with only a single square-root operation for each three successive values of $n$.

The extrapolated values for the function $f$ are in the following table

| $n$ | $f(n^{-2}) = p_n$ | | | | |
|---|---|---|---|---|---|
| 3 | 2.5980762114 | | | | |
| | | 3.1245925848 | | | |
| 4 | 2.8284271247 | | 3.1414310010 | | |
| | | 3.1353691712 | | 3.1415920333 | |
| 5 | 2.9389262615 | | 3.1415517753 | | 3.1415926527 |
| | | 3.1388039512 | | 3.1415925656 | |
| 6 | 3.0000000000 | | 3.1415823680 | | 3.1415926535 |
| | | 3.1404970490 | | 3.1415926394 | |
| 8 | 3.0614674589 | | 3.1415900716 | | 3.1415926536 |
| | | 3.1411965834 | | 3.1415926511 | |
| 10 | 3.0901699437 | | 3.1415920062 | | 3.1415926536 |
| | | 3.1414162628 | | 3.1415926532 | |
| 12 | 3.1058285412 | | 3.1415924915 | | 3.1415926536 |
| | | 3.1415236522 | | 3.1415926535 | |
| 16 | 3.1214451523 | | 3.1415926130 | | |
| | | 3.1415677871 | | | |
| 20 | 3.1286893008 | | | | |

And the similar table for $F$.

| $n$ | $F(n^{-2}) = P_n$ | | | | |
|---|---|---|---|---|---|
| 3 | 5.1961524227 | | | | |
| | | 2.4620897422 | | | |
| 4 | 4.0000000000 | | 3.2709453534 | | |
| | | 2.9797573334 | | 3.1254244368 | |
| 5 | 3.6327126400 | | 3.1618046660 | | 3.1426705290 |
| | | 3.0808947404 | | 3.1402452973 | |
| 6 | 3.4641016151 | | 3.1456351395 | | 3.1416487930 |
| | | 3.1203459211 | | 3.1414242336 | |
| 8 | 3.3137084990 | | 3.1424769601 | | 3.1415974655 |
| | | 3.1345097861 | | 3.1415673906 | |
| 10 | 3.2491969623 | | 3.1417947829 | | 3.1415930546 |
| | | 3.1385570065 | | 3.1415894456 | |
| 12 | 3.2153903092 | | 3.1416407799 | | 3.1415926860 |
| | | 3.1404361809 | | 3.1415921675 | |
| 16 | 3.1825978781 | | 3.1416043206 | | |
| | | 3.1411837903 | | | |
| 20 | 3.1676888065 | | | | |

Note that we have already achieved 10 decimal places accuracy for the $f$ extrapolations but that the accuracies of the results based on $F$ lag far behind.

If this pattern is taken further, so as to include altogether 32 values of $n$ and 6 levels of extrapolation, then $\pi$ can be bracketed to about 40 decimal places. If 8 levels of extrapolation are used instead, then 26 values of $n$ are enough to achieve this sort of accuracy. Even restricting ourselves to 4 levels of extrapolation just 22 values of $n$ will yield 20 decimal places. This mightn't sound very much but it is sufficient to calculate the circumference of a circle with radius the distance between the sun and the earth, to an accuracy of about the size of an atom.

I hope that there is something of interest to some of the readers of the Magazine in this first Apology. Please do not hestitate to tell me if I have got something wrong or if you feel that further details would be worth presenting in a later number. I have made many assertions without the benefit of proof and I shouldn't be allowed to get away with too much of this. Please try, yourself, to justify some statements that are not obvious but let me know if you want me to give my own attempt at justification.

I also write a series of "Mathematical Miniatures" for the Newsletter of the New Zealand Mathematical Society and you might find some of these interesting, especially a recent potted history of the computation of $\pi$. These can now be read on internet at
`http://math.auckland.ac.nz/~butcher/miniature`

J. C. Butcher
Department of Mathematics
The University of Auckland