A one line proof of the Cauchy-Schwarz inequality

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

October 8, 2015

In Courant and Hilbert's famous book "Methods of Mathematical Physics", Volume 1, Wiley Classics, 1989 at page 2, the authors state and prove the Cauchy-Schwarz inequality in essentially one line. They say that:

$$\Big(\sum_{i=1}^{n} a_i\, b_i\Big)^2 \le \Big(\sum_{i=1}^{n} a_i^2\Big)\Big(\sum_{i=1}^{n} b_i^2\Big) \tag{1}$$

where equality holds if and only if the a_i and b_i are proportional ie if a relation of the form $\lambda\mathbf{a} + \mu\mathbf{b} = \mathbf{0}$ with $\lambda^2 + \mu^2 \neq 0$ being satisfied. Here $\mathbf{a} = (a_i, a_2, \dots, a_n)$ and $\mathbf{b} = (b_i, b_2, \dots, b_n)$. They go on to say:

"The proof of this 'Schwarz Inequality' follows from the fact that the roots of the quadratic equation:

$$\sum_{i=1}^{n}(a_i\, x + b_i)^2 = x^2 \sum_{i=1}^{n} a_i^2 + 2x \sum_{i=1}^{n} a_i\, b_i + \sum_{i=1}^{n} b_i^2 \tag{2}$$

for the unknown x can never be real and distinct, but must be imaginary, unless the a_i and b_i are proportional. The Schwarz inequality is really an expression of this fact in terms of the discriminant of the equation." This last sentence is the "one line proof".

According to J Michael Steele, "the Cauchy-Schwarz Master Class: An introduction to the art of mathematical inequalities", Cambridge University Press, 2004, pages 10-11, Schwarz devised his inequality (in the form of a two dimensional integral analog) in the context of minimal surfaces. Thus he wanted to show that if $A = \iint_S f^2 dx\, dy$, $B = \iint_S fg\, dx\, dy$ and $C = \iint_S g^2 dx\, dy$ then $|B| \le \sqrt{A}\,\sqrt{C}$. His proof involved an examination of the polynomial $p(t) = \iint_S \Big(t\, f(x,y) + g(x,y)\Big)^2 dx\, dy = At^2 + 2Bt + C$.

Going back to the Courant-Hilbert proof, it is a legitimate proof in terms of concept but the details are compressed. If you can see the details then don't bother reading on. If not the following may assist.

Because equation (2) is a quadratic in x we can make observations about its behaviour by considering the discriminant:

$$\Delta = \Big(2\sum_{i=1}^{n} a_i\, b_i\Big)^2 - 4\Big(\sum_{i=1}^{n} a_i^2\Big)\,\Big(\sum_{i=1}^{n} b_i^2\Big) = 4\Big\{\Big(\sum_{i=1}^{n} a_i\, b_i\Big)^2 - \Big(\sum_{i=1}^{n} a_i^2\Big)\,\Big(\sum_{i=1}^{n} b_i^2\Big)\Big\} \quad (3)$$

Because we are interested in the sign of Δ we can ignore the factor 4 in what follows. We know that if $\Delta > 0$ ie $\Big(\sum_{i=1}^{n} a_i\, b_i\Big)^2 > \Big(\sum_{i=1}^{n} a_i^2\Big)\,\Big(\sum_{i=1}^{n} b_i^2\Big)$ we would get real and distinct roots. However, this is not possible. Why? If you look at (2) it asserts that the sum of n squared numbers is zero which means that $(a_i x + b_i)^2 = 0$ for all $1 \le i \le n$. Thus $x = \frac{-b_i}{a_i}$ for all i. This is a contradiction since the roots would be equal.

When $\Delta = 0$ then the following is true for $\mu \neq 0$ and $\lambda \neq 0$:

$$\Big(\sum_{i=1}^{n} \frac{-\mu b_i}{\lambda}\, b_i\Big)^2 = (\frac{\mu}{\lambda})^2\,\Big(\sum_{i=1}^{n} b_i^2\Big)^2 = \sum_{i=1}^{n}\Big(\frac{-\mu b_i}{\lambda}\Big)^2 \sum_{i=1}^{n} b_i^2 \quad (4)$$

so that with $a_i = \frac{-\mu b_i}{\lambda}$ you get the required structure. Since $x = \frac{-b_i}{a_i}$ = constant we can write $\frac{-b_i}{a_i} = \frac{\lambda}{\mu}$.

Going in the other direction (note the "if and only if" in the Courant-Hilbert statement) if we let $x = \frac{-b_i}{a_i} = \frac{\lambda}{\mu}$ then $\Big(\sum_{i=1}^{n} a_i\, b_i\Big)^2 = \Big(\sum_{i=1}^{n}\Big(\frac{-\mu b_i}{\lambda}\Big)\, b_i\Big)^2 = (\frac{\mu}{\lambda})^2\,\Big(\sum_{i=1}^{n} b_i^2\Big)^2 = \sum_{i=1}^{n}\Big(\frac{-\mu b_i}{\lambda}\Big)^2 \sum_{i=1}^{n} b_i^2$ so $\Delta = 0$.

So we are left with the fact that $\Delta < 0$ ie the roots are complex, in which case the Cauchy-Schwarz inequality follows in the strict inequality sense.

Thus there really is a one line proof as long as you can see the implications of (2). Schwarz's proof is ingenious in its simplicity.

1 A novel viscosity proof of the inequality

A novel proof of the Cauchy-Schwarz inequality has recently been constructed by inventive Cambridge mathematician Tadashi Tokieda. The title of his proof is "A Viscosity

Proof of the Cauchy-Schwarz Inequality" and can be found at [1] or perhaps on his website (when I checked it was not yet there) [https://www.dpmms.cam.ac.uk/~tokieda/Tokieda_publications.html] The proof is short but the gist of it is that is you assume you have layers of fluids of mass $m_1, m_2, \ldots$ moving in parallel at speeds $v_1, v_2, \ldots$. After some time the effect of viscosity equalizes the speeds among the layers but momentum is conserved (assuming all stresses are internal). If $\langle v \rangle$ is the resulting common speed then we have:

$$\sum_i m_i \, v_i = \sum_i m_i \, \langle v \rangle \tag{5}$$

But energy is dissipated so that the kinetic energy after the viscous equalization must be less than the energy before. Thus:

$$\sum_i \frac{1}{2} m_i \, v_i^2 \geq \sum_i \frac{1}{2} m_i \, \langle v \rangle^2 \tag{6}$$

But $\langle v \rangle = \frac{\sum_i m_i \, v_i}{\sum_i m_i}$ so we have:

$$\begin{aligned} \sum_i m_i \, v_i^2 &\geq \sum_i m_i \left(\frac{\sum_i m_i \, v_i}{\sum_i m_i} \right)^2 \\ \therefore \sum_i m_i \sum_i m_i \, v_i^2 &\geq \left(\sum_i m_i \, v_i \right)^2 \end{aligned} \tag{7}$$

Tokieda then replaces each m_i by m_i^2 and each v_i by $\frac{v_i}{m_i}$ which results in:

$$\sum_i m_i^2 \sum_i v_i^2 \geq \left(\sum_i m_i \, v_i \right)^2 \tag{8}$$

2 References

[1] Tadashi Tokieda, The American Mathematical Monthly, Vol. 122, No. 8 (October 2015), p. 781, http://dx.doi.org/10.4169/amer.math.monthly.122.8.781

3 History

Created 21/02/2013
08/10/2015 - added section on Tadashi Tokieda's viscosity proof