PHYSICAL PROOF OF THE CAUCHY-SCHWARZ INEQUALITY

DAN CRYTSER

Abstract. We present a physical proof of the Cauchy-Schwarz inequality shown by Tadashi Tokieda during his special undergrad talk at Dartmouth on May 1, 2014.

1. Physics

Suppose that you have water separated into n layers, stacked vertically. Thus the first layer touches the second layer, the second layer touches the first and third, and so on. The first layer contains m_1 kilograms of water, the second layer contains m_2 kilograms, and so on, so that the kth layer contains m_k kilograms. Each layer begins with a velocity and we will for simplicity's sake assume that the motion of all the layers is parallel; that is, every layer moves in the same direction as every other layer (with possibly different speeds). The first layer has velocity v_1 (meters per second), the second has velocity v_2, and so on, so that the kth layer has velocity v_k. Let's consider the initial momentum of the system. The momentum of the system is the sum of the momenta of the parts, so

$$p_i = m_1 v_1 + m_2 v_2 + \ldots + m_n v_n.$$

Similarly we can consider the initial kinetic energy of the system. The kinetic energy of the system adds over the parts, so

$$KE_i = \frac{1}{2} m_1 (v_1)^2 + \frac{1}{2} m_2 (v_2)^2 + \ldots + \frac{1}{2} m_n (v_n)^2.$$

We allow the water to move without outside interference, so that no external force acts upon the system. Here a few physical facts come into play.

(1) After a while all the layers move with common final velocity v.
(2) The momentum p of the system is conserved, so the initial momentum equals the final momentum.
(3) The kinetic energy KE does not increase. Also, if two adjacent layers do not have equal velocity, then the kinetic energy decreases because those layers will "grind" against one another. (This energy changes into heat, I think.)

These are all the physical facts we need. The final momentum of the system is the sum of the momenta of the layers, and each layer comes to move with velocity v. Thus

$$p_f = m_1 v + m_2 v + \ldots + m_n v = (m_1 + m_2 + \ldots + m_n) v.$$

Conservation of momentum requires that $p_i = p_f$, so that we can solve for the final velocity as

$$v = \frac{m_1 v_1 + m_2 v_2 + \ldots + m_n v_n}{m_1 + m_2 + \ldots + m_n}.$$

The final kinetic energy of the system is the sum of the kinetic energies of the individual layers. Thus

$$KE_f = \frac{1}{2}m_1v^2 + \frac{1}{2}m_2v^2 + \ldots + \frac{1}{2}m_nv^2.$$

We can rewrite this as

$$KE_f = \frac{1}{2}(m_1 + m_2 + \ldots + m_n)^{-1}(m_1v_1 + \ldots + m_nv_n)^2.$$

The diminution of kinetic energy tells us that $KE_f \leq KE_i$. Mathematically this translates to

$$(m_1 + m_2 + \ldots + m_n)^{-1}(m_1v_1 + \ldots + m_nv_n)^2 \leq m_1(v_1)^2 + m_2(v_2)^2 + \ldots + m_n(v_n)^2,$$

where we have canceled out the $\frac{1}{2}$. We multiply over the total mass and obtain the inequality

$$(m_1v_1 + \ldots + m_nv_n)^2 \leq (m_1 + \ldots + m_n)(m_1v_1^2 + \ldots + m_nv_n^2),$$

from which we can extract square roots to obtain

$$|m_1v + 1 + \ldots + m_nv_n| \leq (m_1 + \ldots + m_n)^{1/2}(m_1v_1^2 + \ldots + m_nv_n^2)^{1/2}.$$

This shows that for any choice of positive numbers $m_1, \ldots, m_n$ and $v_1, \ldots, v_n$, we have the previous inequality. We can even relax the restriction on positive numbers to include negative values. If a layer has a "negative" mass we let it run in the opposite direction, if it has a negative velocity it also runs in the opposite direction, and if both the mass and velocity are negative then we treat it the same.

Moreover, we have that

$$|m_1v_1 + \ldots + m_nv_n| < (m_1 + \ldots + m_n)^{1/2}(m_1v_1^2 + \ldots + m_nv_n^2)^{1/2}$$

whenever two adjacent layers don't have the same velocity. This is the same as requiring that there be a constant k such that $(v_1, v_2, \ldots, v_n) = (k, k, \ldots, k)$.

2. The Cauchy-Schwarz Inequality

Theorem 1. *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. *Then*

$$|\mathbf{x} \cdot \mathbf{y}| \leq ||\mathbf{x}||||\mathbf{y}||$$

and $|\mathbf{x} \cdot \mathbf{y}| = ||\mathbf{x}||||\mathbf{y}||$ *exactly when* $\{\mathbf{x}, \mathbf{y}\}$ *is linearly dependent.*

Proof. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and let $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. We can set $m_i = x_i^2$ and v_i such that $y_i = \sqrt{m_i}v_i$. This gives us a two lists of numbers $m_1, m_2, \ldots, m_n$ and $v_1, v_2, \ldots, v_n$ just as in the previous section. Thus we have

$$|m_1v_1 + \ldots + m_nv_n| \leq (m_1 + \ldots + m_n)^{1/2}(m_1v_1^2 + \ldots + m_nv_n^2)^{1/2}.$$

But $m_iv_i = x_iy_i$, $m_i = x_i^2$ and $m_iv_i^2 = y_i^2$. So we can rewrite this inequality as

$$|x_iy_i + \ldots + x_ny_n| \leq (x_1^2 + \ldots + x_n^2)^{1/2}(y_1^2 + \ldots + y_n^2)^{1/2}.$$

This is the inequality $|\mathbf{x} \cdot \mathbf{y}| \leq ||\mathbf{x}||||\mathbf{y}||$. The inequality is an equality exactly when there is some constant k such that $(v_1, v_2, \ldots, v_n) = (k, k, \ldots, k)$. But then $kx_i = y_i$ for all i, and $\{\mathbf{x}, \mathbf{y}\}$ is linearly dependent. So the inequality is an equality only if $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent. □