5
Consequences of Order

One of the natural questions that accompanies any inequality is the possibility that it admits a converse of one sort or another. When we pose this question for Cauchy's inequality, we find a challenge problem that is definitely worth our attention. It not only leads to results that are useful in their own right, but it also puts us on the path of one of the most fundamental notions in the theory of inequalities — the systematic exploitation of order relationships.

Problem 5.1 (The Hunt for a Cauchy Converse)

Determine the circumstance which suffice for nonnegative real numbers a_k, b_k, $k = 1, 2, ..., n$ to satisfy an inequality of the type

$$\left(\sum_{k=1}^{n} a_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} b_k^2\right)^{\frac{1}{2}} \leq \rho \sum_{k=1}^{n} a_k b_k \tag{5.1}$$

for a given constant ρ.

Orientation

Part of the challenge here is that the problem is not fully framed — there are circumstances and conditions that remain to be determined. Nevertheless, uncertainty is an inevitable part of research, and practice with modestly ambiguous problems can be particularly valuable.

In such situations, one almost always begins with some experimentation, and, since the case $n = 1$ is trivial, the simplest case worth study is given by taking the vectors $(1, a)$ and $(1, b)$ with $a > 0$ and $b > 0$. In this case, the two sides of the conjectured Cauchy converse (5.1) relate the quantities

$$(1 + a^2)^{\frac{1}{2}}(1 + b^2)^{\frac{1}{2}} \quad \text{and} \quad 1 + ab,$$

and this calculation already suggests a useful inference. If a and b are chosen so that the product ab is held constant while $a \to \infty$, then one finds that the right-hand expression is bounded, but the left-hand expression is unbounded. This observation shows in essence that for a given fixed value of $\rho \geq 1$ the conjecture (5.1) cannot hold unless the ratios a_k/b_k are required to be bounded from above and below.

Thus, we come to a more refined point of view, and we see that it is natural to conjecture that a bound of the type (5.1) will hold provided that the summands satisfy the ratio constraint

$$m \leq \frac{a_k}{b_k} \leq M \quad \text{for all } k = 1, 2, ...n, \tag{5.2}$$

for some constants $0 < m \leq M < \infty$. In this new interpretation of the conjecture (5.1), one naturally permits ρ to depend on the values of m and M, though we would hope to show that ρ can be chosen so that it does not have any further dependence on the individual summands a_k and b_k. Now, the puzzle is to find a way to exploit the betweenness bounds (5.2).

Exploitation of Betweeness

When we look at our *unknown* (the conjectured inequality) and then look at the *given* (the betweenness bounds), we may have the lucky idea of hunting for clues in our earlier proofs of Cauchy's inequality. In particular, if we recall the proof that took $(a-b)^2 \geq 0$ as its departure point, we might start to suspect that an analogous idea could help here. Is there some way to obtain a useful quadratic bound from the betweenness relation (5.2)?

Once the question is put so bluntly, one does not need long to notice that the two-sided bound (5.2) gives us a cheap quadratic bound

$$\left(M - \frac{a_k}{b_k}\right)\left(\frac{a_k}{b_k} - m\right) \geq 0. \tag{5.3}$$

Although one cannot tell immediately if this observation will help, the analogy with the earlier success of the trivial bound $(a-b)^2 \geq 0$ provides ground for optimism.

At a minimum, we should have the confidence needed to unwrap the bound (5.3) to find the equivalent inequality

$$a_k^2 + (mM)\, b_k^2 \leq (m+M)\, a_k b_k \quad \text{for all } k = 1, 2, ..., n. \tag{5.4}$$

Now we seem to be in luck; we have found a bound on a sum of squares by a product, and this is precisely what a converse to Cauchy's inequality requires. The eventual role to be played by M and m is still unclear, but, the scent of progress is in the air.

The bounds (5.4) call out to be summed over $1 \le k \le n$, and, upon summing, the factors mM and $m+M$ come out neatly to give us

$$\sum_{k=1}^{n} a_k^2 + (mM) \sum_{k=1}^{n} b_k^2 \le (m+M) \sum_{k=1}^{n} a_k b_k, \tag{5.5}$$

which is a fine additive bound. Thus, we face a problem of a kind we have met before — we need to convert an additive bound to one that is multiplicative.

Passage to a Product

If we cling to our earlier pattern, we might now be tempted to introduce normalized variables $\hat{a}_k$ and $\hat{b}_k$, but this time normalization runs into trouble. The problem is that the inequality (5.5) may be applied to $\hat{a}_k$ and $\hat{b}_k$ only if they satisfy the ratio bound $m \le \hat{a}_k/\hat{b}_k \le M$, and these constraints rule out the natural candidates for the normalizations $\hat{a}_k$ and $\hat{b}_k$. We need a new idea for passing to a product.

Conceivably, one might get stuck here, but help is close at hand provided that we pause to ask clearly what is needed — which is just a lower bound for a sum of two expressions by a product of their square roots. Once this is said, one can hardly fail to think of using the AM-GM inequality, and, when it is applied to the additive bound (5.5), one finds

$$\begin{aligned}\left(\sum_{k=1}^{n} a_k^2\right)^{\frac{1}{2}} \left(mM \sum_{k=1}^{n} b_k^2\right)^{\frac{1}{2}} &\le \frac{1}{2}\left\{\sum_{k=1}^{n} a_k^2 + (mM) \sum_{k=1}^{n} b_k^2\right\} \\ &\le \frac{1}{2}\left\{(m+M) \sum_{k=1}^{n} a_k b_k\right\}.\end{aligned}$$

Now, with just a little rearranging, we come to the inequality that completes our quest. Thus, if we set

$$A = (m+M)/2 \quad \text{and} \quad G = \sqrt{mM},$$

then, for all nonnegative a_k, b_k, $k = 1, 2, ..., n$ with

$$0 < m \le a_k/b_k \le M < \infty,$$

we find the we have established the bound

$$\left(\sum_{k=1}^{n} a_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} b_k^2\right)^{\frac{1}{2}} \le \frac{A}{G} \sum_{k=1}^{n} a_k b_k; \tag{5.6}$$

thus, in the end, one sees that there is indeed a natural converse to Cauchy's inequality.

On the Conversion of Information

When one looks back on the proof of the converse Cauchy inequality (5.6), one may be struck by how quickly progress followed once the two *order relationships*, $m \le a_k/b_k$ and $a_k/b_k \le M$, were put together to build the simple *quadratic inequality* $(M - a_k/b_k)(a_k/b_k - m) \ge 0$. In the context of a single example, this could just be a lucky accident, but something deeper is afoot.

In fact, the device of order-to-quadratic conversion is remarkably versatile tool with a wide range of applications. The next few challenge problems illustrate some of these that are of independent interest.

Monotonicity and Chebyshev's "Order Inequality"

One way to put a large collection of order relationships at your finger tips is to focus your attention on monotone sequences and monotone functions. This suggestion is so natural that it might not stir hight hopes, but in fact it does lead to an important result with many natural applications, especially in probability and statistics.

The result is due to Pafnuty Lvovich Chebyshev (1821–1894), who, incidentally, had his first exposure to probability theory from our earlier acquaintance Victor Vacovlevich Bunyakovsky. Probability theory was one of those hot new mathematical topics which Bunyakovsky brought back to St. Petersburg when he returned from his student days studying with Cauchy in Paris. Another topic was the theory of complex variables which we will engage a bit later.

Problem 5.2 (Chebyshev's Order Inequality)

Suppose that $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$ *are nondecreasing and suppose* $p_j \ge 0$, $j = 1, 2, ..., n$, *satisfy* $p_1 + p_2 + \cdots + p_n = 1$. *Show that for any nondecreasing sequence* $x_1 \le x_2 \le \cdots \le x_n$ *one has the inequality*

$$\left\{\sum_{k=1}^{n} f(x_k)p_k\right\}\left\{\sum_{k=1}^{n} g(x_k)p_k\right\} \le \sum_{k=1}^{n} f(x_k)g(x_k)p_k. \tag{5.7}$$

Connections to Probability and Statistics

The inequality (5.7) is easily understood without relying on it connection to probability theory, and it has many applications in other areas of mathematics. Nevertheless, the probabilistic interpretation of the bound

(5.7) is particularly compelling. In the language of probability, it says that if X is a random variable for which one has $P(X = x_k) = p_k$ for $k = 1, 2, ..., n$ then

$$E[f(X)]E[g(X)] \le E[f(X)g(X)], \tag{5.8}$$

where, as usual, P stands for probability and E stands for the mathematical expectation. In other words, if random variables Y and Z may be written as nondecreasing functions of a single random variable X, then Y and Z must be nonnegativity correlated. Without Chebyshev's inequality, the intuition that is commonly attached to the statistical notion of correlation would stand on shaky ground.

Incidentally, there is another inequality due to Chebyshev that is even more important in probability theory; it tells us that for any random variable X with a finite mean $\mu = E(X)$ one has the bound

$$P(|X-\mu| \ge \lambda) \le \frac{1}{\lambda^2} E\big(|X-\mu|^2\big). \tag{5.9}$$

The proof of this bound is almost trivial, especially with the hint offered in Exercise 5.9, but it is such a day-to-day workhorse in probability theory that Chebyshev's *order* (5.8) inequality is often jokingly called Chebyshev's *other* inequality.

A Proof from Our Pocket

Chebyshev's inequality (5.7) is quadratic, and the hypotheses provide order information, so, even if one were to meet Chebyshev's inequality (5.7) in a dark alley, the *order-to-quadratic conversion* is likely to come to mind. Here the monotonicity of f and g give us the quadratic bound,

$$0 \le \big\{f(x_k) - f(x_j)\big\}\big\{g(x_k) - g(x_j)\big\},$$

and this may be expanded in turn to give

$$f(x_k)g(x_j) + f(x_j)g(x_k) \le f(x_j)g(x_j) + f(x_k)g(x_k). \tag{5.10}$$

From this point, we only need to bring the p_j's into the picture and meekly agree to take whatever arithmetic gives us.

Thus, when we multiply the bound (5.10) by $p_j p_k$ and sum over $1 \le j \le n$ and $1 \le k \le n$, we find that the left-hand sum gives us

$$\sum_{j,k=1}^{n} \big\{f(x_k)g(x_j) + f(x_j)g(x_k)\big\} p_j p_k = 2\bigg\{\sum_{k=1}^{n} f(x_k)p_k\bigg\}\bigg\{\sum_{k=1}^{n} g(x_k)p_k\bigg\},$$

while the right-hand sum gives us

$$\sum_{j,k=1}^{n} \{f(x_j)g(x_j) + f(x_k)g(x_k)\}p_j p_k = 2\left\{\sum_{k=1}^{n} f(x_k)g(x_k)p_k\right\}.$$

Thus, the bound between the summands (5.10) does indeed yield the proof of Chebyshev's inequality.

Order, Facility, and Subtlety

The proof of Chebyshev's inequality leads us to a couple of observations. First, there are occasions when the application of the order-to-quadratic conversion is an automatic, straightforward affair. Even so, the conversion has led to some remarkable results, including the versatile *rearrangement inequality* which is developed in our next challenge problem. The rearrangement inequality is not much harder to prove than Chebyshev's inequality, but some of its consequences are simply stunning. Here, and subsequently, we let $[n]$ denote the set $\{1, 2, ..., n\}$, and we recall that a permutation of $[n]$ is just a one-to-one mapping from $[n]$ into $[n]$.

Problem 5.3 (The Rearrangement Inequality)

Show that for each pair of ordered real sequences

$$-\infty < a_1 \le a_2 \le \cdots \le a_n < \infty \quad \textit{and} \quad -\infty < b_1 \le b_2 \le \cdots \le b_n < \infty$$

and for each permutation $\sigma : [n] \to [n]$, *one has*

$$\sum_{k=1}^{n} a_k b_{n-k+1} \le \sum_{k=1}^{n} a_k b_{\sigma(k)} \le \sum_{k=1}^{n} a_k b_k. \tag{5.11}$$

Automatic — But Still Effective

This problem offers us a hypothesis that provides order relations and asks us for a conclusion that is quadratic. This familiar combination may tempt one to just to dive in, but sometimes it pays to be patient. After all, the statement of the rearrangement inequality is a bit involved, and one probably does well to first consider the simplest case $n = 2$.

In this case, the order-to-quadratic conversion reminds us that

$$a_1 \le a_2 \quad \text{and} \quad b_1 \le b_2 \quad \text{imply} \quad 0 \le (a_2 - a_1)(b_2 - b_1),$$

and, when this is unwrapped, we find

$$a_1 b_2 + a_2 b_1 \le a_1 b_1 + a_2 b_2,$$

which is precisely the rearrangement inequality (5.11) for $n = 2$. Nothing could be easier than this warm-up case; the issue now is to see if a similar idea can be used to deal with the more general sums

$$S(\sigma) = \sum_{k=1}^{n} a_k b_{\sigma(k)}.$$

INVERSIONS AND THEIR REMOVAL

If σ is not the identity permutation, then there must exist some pair $j < k$ such that $\sigma(k) < \sigma(j)$. Such a pair is called an *inversion*, and the observation that one draws from the case $n = 2$ is that if we switch the values of $\sigma(k)$ and $\sigma(j)$, then the value of the associated sum will increase — or, at least not decrease. To make this idea formal, we first introduce a new permutation τ by the recipe

$$\tau(i) = \begin{cases} \sigma(i) & \text{if } i \neq j \text{ and } i \neq k \\ \sigma(j) & \text{if } i = k, \\ \sigma(k) & \text{if } i = j. \end{cases} \tag{5.12}$$

By the definition of τ and by factorization, we then find

$$\begin{aligned} S(\tau) - S(\sigma) &= a_j b_{\tau(j)} + a_k b_{\tau(k)} - a_j b_{\sigma(j)} - a_k b_{\sigma(k)} \\ &= a_j b_{\tau(j)} + a_k b_{\tau(k)} - a_j b_{\tau(k)} - a_k b_{\tau(j)} \\ &= (a_k - a_j)(b_{\tau(k)} - b_{\tau(j)}) \geq 0. \end{aligned}$$

Thus, the transformation $\sigma \mapsto \tau$ achieves two goals; first, it increases S, so, $S(\sigma) \leq S(\tau)$, and, second, the number of inversions of τ is forced to be strictly few than the number of inversions of the permutation σ.

REPEATING THE PROCESS — CLOSING THE LOOP

A permutation has at most $n(n-1)/2$ inversions and only the identity permutation has no inversions, so there exists a finite sequence of inversion removing transformations that move in sequence from σ to the identity. If we denote these by $\sigma = \sigma_0, \sigma_1, ..., \sigma_m$ where σ_m is the identity and $m \leq n(n-1)/2$, then, by applying the bound $S(\sigma_{j-1}) \leq S(\sigma_j)$ for $j = 1, 2, ..., m$, we find

$$S(\sigma) \leq \sum_{k=1}^{n} a_k b_k.$$

This completes the proof of the upper half of the rearrangement inequality (5.11).

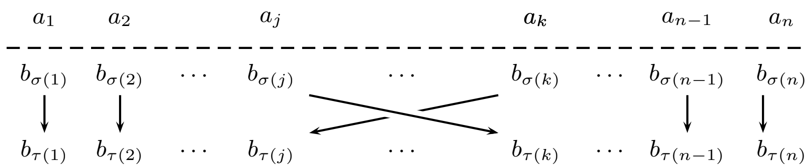


Fig. 5.1. An interchange operation converts the permutation σ to a permutation τ. By design, the new permutation τ has fewer inversions than σ; by calculation, one also finds that $S(\sigma) \leq S(\tau)$.

The easy way to get the lower half is then to notice that it is an immediate consequence of the upper half. Thus, if we consider $b'_1 = -b_n, b'_2 = -b_{n-1}, ..., b'_n = -b_1$ we see that

$$b'_1 \leq b'_2 \leq \cdots \leq b'_n,$$

and, by the upper half of the rearrangement inequality (5.11) applied to the sequence $b'_1, b'_2, ..., b'_n$ we get the lower half of the inequality (5.11) for the sequence $b_1, b_2, ..., b_n$.

Looking Back — Testing New Probes

The statement of the rearrangement inequality is exceptionally natural, and it does not provide us with any obvious loose ends. We might look back on it many times and never think of any useful variations of either its statement or its proof. Nevertheless, such variations can always be found; one just needs to use the right probes.

Obviously, no single probe, or even any set of probes, can lead with certainty to a useful variation of a given result, but there are a few generic questions that are almost always worth our time. One of the best of these asks: "Is there a *non-linear* version of this result?"

Here, to make sense of this question, we first need to notice that the rearrangement inequality is a statement about sums of linear functions of the ordered n-tuples

$$\{b_{n-k+1}\}_{1\leq k\leq n}, \quad \{b_{\sigma(k)}\}_{1\leq k\leq n} \quad \text{and} \quad \{b_k\}_{1\leq k\leq n},$$

where the "linear functions" are simply the n mappings given by

$$x \mapsto a_k x \qquad k = 1, 2, ..., n.$$

Such simple linear maps are usually not worth naming, but here we have a higher purpose in mind. In particular, with this identification behind us, we may not need long to think of some ways that the monotonicity condition $a_k \leq a_{k+1}$ might be re-expressed.

Several variations of the rearrangement inequality may come to mind, and our next challenge problems explores one of the simplest of these. It was first studied by A. Vince, and it has several informative consequences.

Problem 5.4 (An Nonlinear Rearrangement Inequality)
Let f_1, f_2, ..., f_n be functions from the interval I into $\mathbb{R}$ such that

$$f_{k+1}(x) - f_k(x) \quad \textit{is nondecreasing for all} \quad 1 \le k \le n. \tag{5.13}$$

Let $b_1 \le b_2 \le \cdots \le b_n$ be an ordered sequence of elements of I, and show that for each permutation $\sigma : [n] \to [n]$, one has the bound

$$\sum_{k=1}^{n} f_k(b_{n-k+1}) \le \sum_{k=1}^{n} f_k(b_{\sigma(k)}) \le \sum_{k=1}^{n} f_k(b_k). \tag{5.14}$$

Testing the Waters

This problem is intended to generalize the rearrangement inequality, and, we see immediately that it does when we identify $f_k(x)$ with the map $x \mapsto a_k x$. To be sure, there are far more interesting non-linear examples which one can find after even a little experimentation.

For instance, one might take $a_1 \le a_2 \le \cdots \le a_n$ and consider the functions $x \mapsto \log(a_k + x)$. Here one finds

$$\log(a_{k+1} + x) - \log(a_k + x) = \log\left(\frac{(a_{k+1} + x)}{(a_k + x)}\right),$$

and, if we set $r(x) = (a_{k+1} + x)/(a_k + x)$, then direct calculation gives

$$r'(x) = \frac{a_k - a_{k+1}}{(a_k + x)^2} \le 0,$$

so, if we take

$$f_k(x) = -\log(a_k + x) \quad \text{for } k = 1, 2, ..., n,$$

then condition (5.13) is satisfied. Thus, by Vince's inequality and exponentiation one finds that for each permutation $\sigma : [n] \to [n]$ that

$$\prod_{k=1}^{n} (a_k + b_k) \le \prod_{k=1}^{n} (a_k + b_{\sigma(k)}) \le \prod_{k=1}^{n} (a_k + b_{n-k+1}). \tag{5.15}$$

This interesting product bound (5.15) shows that there is power in Vince's inequality, though in this particular case the bound was known earlier. Still, we see that a proof of Vince's inequality will be worth our time — even if only because of the corollary (5.15).

Recycling an Algorithmic Proof

If we generalize our earlier sums and write

$$S(\sigma) = \sum_{k=1}^{n} f_k(b_{\sigma(k)}),$$

then we already know from the definition (5.12) and discussion of the inversion decreasing transformation $\sigma \mapsto \tau$ that we only need to show

$$S(\sigma) \leq S(\tau).$$

Now, almost as before, we calculate the difference

$$\begin{aligned} S(\tau) - S(\sigma) &= f_j(b_{\tau(j)}) + f_k(b_{\tau(k)}) - f_j(b_{\sigma(j)}) - f_k(b_{\sigma(k)}) \\ &= f_j(b_{\tau(j)}) + f_k(b_{\tau(k)}) - f_j(b_{\tau(k)}) - f_k(b_{\tau(j)}) \\ &= \{f_k(b_{\tau(k)}) - f_j(b_{\tau(k)})\} - \{f_k(b_{\tau(j)}) - f_j(b_{\tau(j)})\} \geq 0, \end{aligned}$$

and this time the last inequality comes from $b_{\tau(j)} \leq b_{\tau(k)}$ and our hypothesis that $f_k(x) - f_j(x)$ is a nondecreasing function of $x \in I$. From this relation, one then sees that no further change is needed in our earlier arguments, and the proof of the non-linear version of the rearrangement inequality is complete.

`Early Reader Note:  Later, I plan to add a brief discussion here that puts a ribbon around both the "order-to-quadratic conversion" and the notion of a nonlinear generalization.`

Exercises

Exercise 5.1 (Baseball and Cauchy's Third Inequality)

In the remarkable *Note II* of 1821 where Cauchy proved both his namesake inequality and the fundamental AM-GM bound, one finds a third inequality which is not as notable or as deep but which is still handy from time to time. The inequality asserts that for any positive real numbers $h_1, h_2, ..., h_n$ and $b_1, b_2, ..., b_n$ one has the ratio bounds

$$m = \min_{1 \leq j \leq n} \frac{h_j}{b_j} \leq \frac{h_1 + h_2 + \cdots + h_n}{b_1 + b_2 + \cdots + b_n} \leq \max_{1 \leq j \leq n} \frac{h_j}{b_j} = M. \tag{5.16}$$

Sports enthusiasts may imagine, as Cauchy never would, that b_j denotes the number of times a baseball player j goes to bat, and h_j denotes the number of times he gets a hit. The inequality confirms the intuitive fact that the batting average of a team is never worse that of its worst hitter and never better than that of its best hitter.

Prove the inequality (5.16) and put it to honest mathematical use by proving that for any polynomial $P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ with positive coefficients one has the monotonicity relation

$$0 < x \leq y \quad \Longrightarrow \quad \frac{x}{y} \leq \frac{P(x)}{P(y)} \leq 1.$$

Exercise 5.2 (Betweeness and an Inductive Proof of AM-GM)

One can build an inductive proof of the basic AM-GM inequality (2.3) by exploiting the conversion of an order relation to a quadratic bound. To get started, first consider $0 < a_1 \leq a_2 \leq \cdots \leq a_n$, set $A = (a_1 + a_2 + \cdots + a_n)/n$, and then show that one has

$$a_1 a_n / A \leq a_1 + a_n - A.$$

Now, complete the induction step of the AM-GM proof by considering the $n-1$ element set $S = \{a_2, a_3, ..., a_{n-1}\} \cup \{a_1 + a_n - A\}$.

Exercise 5.3 (Cauchy-Schwarz and the Cross-Term Defect)

If u and v are elements of the real inner product space V for which on has the upper bounds

$$\langle \mathbf{u}, \mathbf{u} \rangle \leq A^2 \quad \text{and} \quad \langle \mathbf{v}, \mathbf{v} \rangle \leq B^2,$$

then Cauchy's inequality tells us $\langle \mathbf{u}, \mathbf{v} \rangle \leq AB$. Show that one then also has a lower bound on the cross-term difference $AB - \langle \mathbf{u}, \mathbf{v} \rangle$, namely,

$$\left\{ A^2 - \langle \mathbf{u}, \mathbf{u} \rangle \right\}^{\frac{1}{2}} \left\{ B^2 - \langle \mathbf{v}, \mathbf{v} \rangle \right\}^{\frac{1}{2}} \leq AB - \langle \mathbf{u}, \mathbf{v} \rangle. \tag{5.17}$$

Exercise 5.4 (A Remarkable Inequality of I. Schur)

Show that for all values of $x, y, z \geq 0$, one has for all $\alpha \geq 0$ that

$$x^\alpha (x-y)(x-z) + y^\alpha (y-x)(y-z) + z^\alpha (z-x)(x-y) \geq 0. \tag{5.18}$$

Moreover, show that one has equality here if and only if one has *either* $x = y = x$ *or* two of the variables are equal and the third is zero.

Schur's inequality can sometimes saves the day in problems where the AM-GM inequality looks like the natural tool, yet it comes up short. Sometimes the two-pronged condition for equality also provides a clue that Schur's inequality may be of help.

```
 Early Reader Note:  The mod 3 sum problem by Jim Pitman looked
like it was going to provide a nice illustration of this possibility,
but Jim's problem has was "cooked" by Jim's own solution which
is simple and Schur-free.  I need to look again at Kedlaya's
nice examples to see if one of his can serve here, though I am
hesitant to poach.
```

Exercise 5.5 (The Pólya-Szegö Converse Restructured)

The converse Cauchy inequality (5.6) is expressed with the aid of bounds on the ratios a_k/b_k, but for many applications it is useful to know that one also has a natural converse under the more straightforward hypothesis that

$$0 < a \le a_k \le A \quad \text{and} \quad 0 < b \le b_k \le B \quad \text{for all} \quad k = 1, 2, ..., n.$$

Use the Cauchy converse (5.6) to prove that in this case one has

$$\left\{\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2\right\} \Big/ \left\{\sum_{k=1}^{n} a_k b_k\right\}^2 \le \frac{1}{4}\left\{\sqrt{\frac{AB}{ab}} + \sqrt{\frac{ab}{AB}}\right\}.$$

Exercise 5.6 (A Competition Perennial)

Show that if $a > 0$, $b > 0$, and $c > 0$ then one has the elegant symmetric bound

$$\frac{3}{2} \le \frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b}. \tag{5.19}$$

This is known as Nesbitt's inequality, and, in suitable variations and amplifications, it has served in numerous mathematical competitions, from Moscow 1962 to the Canadian Maritimes 2002.

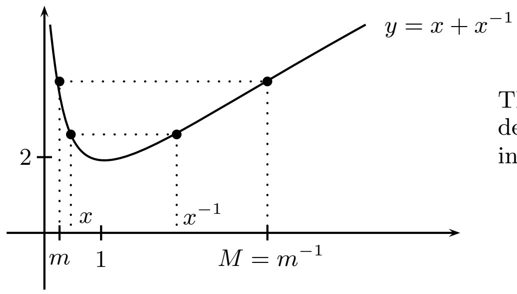
Exercise 5.7 (Rearrangement, Cyclic Shifts, and the AM-GM)

Skillful use of the rearrangement inequality often calls for one to exploit symmetry and to look for clever specializations of the resulting bounds. This problem outlines a proof of the AM-GM inequality that nicely illustrates these steps.

(a) Show that for positive a_k, $k = 1, 2, ..., n$ one has

$$n \le \frac{a_1}{a_n} + \frac{a_2}{a_1} + \frac{a_3}{a_2} + \cdots + \frac{a_n}{a_{n-1}}.$$

(b) Specialize the result of part (a) to show that one also has for all



The map $f(x) = x + x^{-1}$ decreases on (0,1] and increases on [1,∞)

Fig. 5.2. One key to the proof of Kantorovich's inequality is the geometry of the map $x \rightarrowtail x + x^{-1}$; another key is that a multiplicative inequality is sometimes proved most easily by first establishing an appropriate additive inequality. To say much more would risk giving away the game.

positive x_k, $k = 1, 2, ..., n$, that

$$n \leq \frac{x_1}{x_1 x_2 \cdots x_n} + x_2 + x_3 + \cdots + x_n.$$

(c) Specialize a third time to show that for $\rho > 0$ one also has

$$n \leq \frac{\rho x_1}{\rho^n x_1 x_2 \cdots x_n} + \rho x_2 + \rho x_3 + \cdots + \rho x_n,$$

and finally indicate how the right choice of ρ now yields the AM-GM inequality (2.3).

Exercise 5.8 (Kantorovich's Inequality for Reciprocals)

Show that if $0 < m = x_1 \leq x_2 \leq \cdots \leq x_n = M < \infty$ then for nonnegative weights with $p_1 + p_2 + \cdots + p_n = 1$ one has

$$\left\{ \sum_{j=1}^{n} p_j x_j \right\} \left\{ \sum_{j=1}^{n} p_j \frac{1}{x_j} \right\} \leq \frac{\mu^2}{\gamma^2} \tag{5.20}$$

where $\mu = (m + M)/2$ and $\gamma = \sqrt{mM}$. This bound provides a natural complement to the elementary inequality of Exercise 1.2, page 13, but it also has important applications in numerical analysis, where, for example, it has been used to estimated the rate of convergence of the method of steepest ascent. To get started with the proof, one might note that by homogeneity it suffices to consider the case when $\gamma = 1$; the geometry of Figure 5.2 then tells a powerful tale.

Exercise 5.9 (Chebyshev's Inequality for Tail Probabilities)

One of the most basic properties of the mathematical expectation $E(\cdot)$

that one meets in probability theory is that for any random variables X and Y with finite expectations the relationship $X \leq Y$ implies that $E(X) \leq E(Y)$. Use this fact to show that for any random variable Z with finite mean $\mu = E(Z)$ one has the inequality

$$P\big(|Z - \mu| \geq \lambda\big) \leq \frac{1}{\lambda^2} E\big(\,|Z - \mu|^2\big). \tag{5.21}$$

This bound provides one concrete expression of the notion that a random variable is not likely to be too far away from its mean, and it is surely the most used of the several inequalities that carry Chebyshev's name.