INEQUALITIES FOR THE DERIVATIVES OF POLYNOMIALS

R. P. BOAS, JR., Northwestern University

Some years after the chemist Mendeleev invented the periodic table of the elements he made a study of the specific gravity of a solution as a function of the percentage of the dissolved substance [16]. This function is of some practical importance: for example, it is used in testing beer and wine for alcoholic content, and in testing the cooling system of an automobile for concentration of anti-freeze; but present-day physical chemists do not seem to find it as interesting as Mendeleev did. Nevertheless, Mendeleev's study led to mathematical problems of great interest, some of which are still inspiring research today.

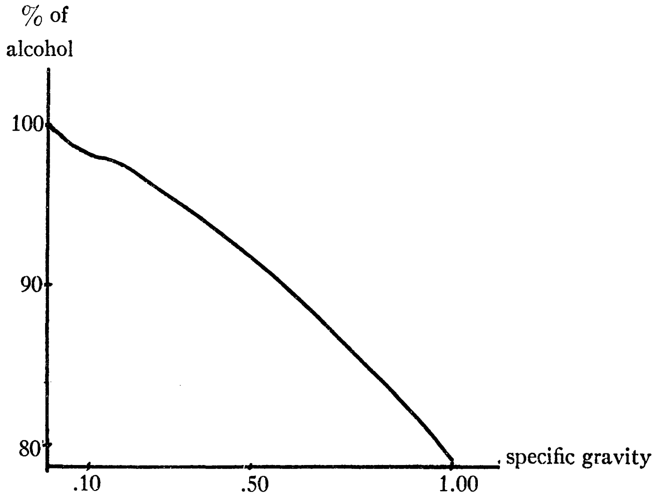

FIG. 1

An example of the kind of curve that Mendeleev obtained is shown in Figure 1 (alcohol in water, percentage by weight). He noticed that the curves could be closely approximated by successions of quadratic arcs, and he wanted to know whether the corners where the arcs joined were really there, or just caused by errors of measurement. In mathematical terms, this amounts to considering a quadratic polynomial $P(x)=px^2+qx+r$ on an interval $[a, b]$, with $\max P(x)-\min P(x)=L$, and asking how large $P'(x)$ can be on $[a, b]$. For, if the slope of one arc exceeds the largest possible slope for an adjacent arc, it follows that these arcs must come from different quadratic functions. We can reduce the problem to a simpler one by changing the horizontal scale and shifting the coordinate axes until the interval $[a, b]$ becomes $[-1, 1]$, and then changing the vertical scale and shifting the axes until we have $|P(x)| \leqq 1$. Our question then becomes, if $P(x)$ is a quadratic function and $|P(x)| \leqq 1$ on $[-1, 1]$, how large can $|P'(x)|$ be on $[-1, 1]$? The answer that Mendeleev found is that $|P'(x)| \leqq 4$; and this is the most that can be said, since when $P(x)=1-2x^2$ we have $|P(x)| \leqq 1$ and $|P'(\pm 1)|=4$. By using this result, Mendeleev convinced himself that the corners in his curves were genuine; and he was presumably right, since his measurements were quite accurate (they agree with modern tables to three or more significant figures).

One can readily imagine that a chemist who discovers such a pretty mathematical result would tell a mathematician about it; and in fact Mendeleev told it to A. A. Markov, who naturally investigated the corresponding problem for polynomials of degree n [13]. In particular, he proved what has come to be known as Markov's Theorem:

If $P(x)$ is a real polynomial of degree n, and $|P(x)| \leqq 1$ on $[-1, 1]$ then $|P'(x)| \leqq n^2$ on $[-1, 1]$, with equality attainable only at ± 1 and only when $P(x) = \pm T_n(x)$, where $T_n(x)$ (the so-called Chebyshev polynomial) is $\cos n \cos^{-1} x$ *(which actually is a polynomial, since* $\cos n\theta$ *is a polynomial in* $\cos \theta$).

Clearly we can also assert that if $|P(x)| \leqq L$ on $[-1, 1]$ then $|P'(x)| \leqq Ln^2$. Having now found an upper bound for $|P'(x)|$, it would be natural to go on and ask for an upper bound for $|P^{(k)}(x)|$ (where $k \leqq n$). Iterating Markov's theorem yields $|P^{(k)}(x)| \leqq n^{2k}L$ if $|P(x)| \leqq L$. However, this inequality is not sharp; the best possible inequality was found by Markov's brother, V. A. Markov, who proved that $|P^{(k)}(x)| \leqq T_n^{(2k)}(1)$ when $|P(x)| \leqq 1$; here T_n is again the Chebyshev polynomial. Explicitly,

$$|P^{(k)}(x)| \leqq \frac{n^2(n^2 - 1^2)(n^2 - 2^2) \cdots (n^2 - (k-1)^2)}{1 \cdot 3 \cdot 5 \cdots (2k-1)}.$$

Later on we shall give a fairly simple proof of the inequality for $P'(x)$, but the inequality for $P^{(k)}(x)$ is considerably harder, except for $k = n$ (see, for example, [6], [21]; for $k = n$, Lemma 6, below).

The next similar question about polynomials was not asked for about 20 years, when S. Bernstein wanted, for applications in the theory of approximation of functions by polynomials, the analogue of Markov's theorem for the unit disk in the complex plane instead of for the interval $[-1, 1]$. He asked, if $P(z)$ is a polynomial of degree n and $|P(z)| \leqq 1$ for $|z| \leqq 1$, how large can $|P'(z)|$ be for $|z| \leqq 1$? The answer is that $|P'(z)| \leqq n$, with equality attained for $P(z) = z^n$. Bernstein's problem can be stated in a different way which suggests many interesting generalizations and has a number of applications. Since a polynomial $P(z)$ is an analytic function, it attains its maximum absolute value for $|z| \leqq 1$ on the circumference $|z| = 1$; so does its derivative. Hence if we want $\max|P'(z)|$ for $|z| \leqq 1$ given that $|P(z)| \leqq 1$ for $|z| \leqq 1$, it is enough to consider only values of z with $|z| = 1$, that is, $z = e^{i\theta}$ with $0 \leqq \theta < 2\pi$. Now $P(e^{i\theta})$ can be written as a linear combination of sines and cosines,

$$S(\theta) = \sum_{k=0}^{n} (a_k \cos k\theta + b_k \sin k\theta);$$

such an expression is called a trigonometric sum of degree n (or a trigonometric polynomial). Hence Bernstein's theorem can be restated as follows:

If $S(\theta)$ is a trigonometric sum of degree n (possibly with complex coefficients) and $|S(\theta)| \leqq 1$, then $|S'(\theta)| \leqq n$, with equality attained when $S(\theta) = \sin n(\theta - \theta_0)$.

As Bernstein observed, if $P(x)$ is our original polynomial on $(-1, 1)$, $P(\cos\theta)$ is a trigonometric sum of degree n, and so $|P'(\cos\theta)\sin\theta| \leqq n$ by Bernstein's theorem, which is to say that $|P'(x)| \leqq n(1-x^2)^{-1/2}$ for $|x| \leqq 1$. This gives an estimate for $|P'(x)|$ that is much better than Markov's when x is not near ± 1, but it does not yield Markov's theorem directly since it tells us nothing about $|P'(x)|$ when x is near ± 1. It is rather remarkable that Markov's theorem can nevertheless be deduced from Bernstein's theorem.

There are many proofs of Bernstein's and Markov's theorems. Those given here are interesting because they demand very little machinery, and illustrate how unexpected results can sometimes be obtained from very simple considerations. We begin by stating some almost obvious results as lemmas.

LEMMA 1. *The polynomial $T_n(x)$ takes the values $+1$ and -1 a total of $n+1$ times in the interval $[-1, 1]$, with alternating signs.*

In fact, $T_n(x) = \pm 1$ whenever $n \cos^{-1}x = k\pi$, $0 \leqq k \leqq n$.

LEMMA 2. *If two polynomials of degree at most n have the same values at $n+1$ points, they are the same polynomial.*

LEMMA 3. *If two trigonometric sums of degree n have the same values at $2n+2$ points in $0 \leqq \theta \leqq 2\pi$ (counting both ends), they are the same.*

This follows from Lemma 2. A trigonometric sum $S(\theta)$ of degree n can be written

$$\sum_{k=-n}^{n} c_k e^{ik\theta} = e^{-in\theta} \sum_{k=0}^{2n} c_{k-n} e^{ik\theta},$$

i.e., as $e^{-in\theta}$ times a polynomial of degree $2n$ in $e^{i\theta}$, and accordingly has (if not identically 0) at most $2n+1$ zeros in $[0, 2\pi]$ (allowing for the fact that a zero at 0 is repeated at 2π).

LEMMA 4. *Suppose that ϕ and f are real-valued continuous functions on $-1 \leqq x \leqq 1$, that ϕ takes the values ± 1 at $k+1$ points x_j with alternating signs, and that $|f(x_j)| < 1$ for all the points x_i. Then there are k points where $f(x) = \phi(x)$.*

In geometrical terms, if the graph of ϕ has k arcs connecting the line $y=1$ with $y=-1$, and the graph of f is between these lines at the points where the graph of ϕ meets them, the graphs of f and ϕ have k intersections. Indeed, if (for example) $\phi(x_j) = -1$ and $\phi(x_{j+1}) = +1$, we have $\phi(x) - f(x)$ negative at x_j and positive at x_{j+1}, and so zero somewhere in between.

LEMMA 5. *Under the hypotheses of Lemma 4, if the graph of f crosses the graph of ϕ from below to above on an arc that rises from -1 to $+1$, then the graphs of f and ϕ cross at at least $k+2$ points.*

The situation is illustrated in Figure 2. Let the arc specified in the lemma connect $(a, -1)$ with $(b, 1)$, $b > a$, and let the crossing occur at $x = c$. Since the graph of f is above that of ϕ at $x = a$, below it somewhere between $x = a$ and $x = c$, above it between c and b, and below it again at $x = b$, there are at least 3

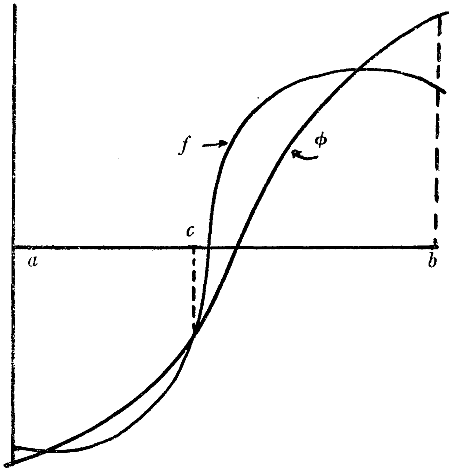

FIG. 2

crossings on this arc, and $k-1$ on the remaining arcs of the graph of ϕ, so $k+2$ in all.

We begin by using Lemma 5 to prove Bernstein's theorem (for a similar proof, cf. [**11**]). We begin with a real trigonometric sum (i.e., one with real coefficients); the extension to sums with complex coefficients depends on a trick, which we give at the end of the proof .

We are, then, given a real trigonometric sum $S(\theta)$ with $|S(\theta)| \leqq 1$; we shall suppose that $S'(\theta_0) > n$ for some θ_0, and obtain a contradiction. Since we can, if necessary, replace $S(\theta)$ by $\lambda S(\theta)$ with λ slightly less than 1, and still have $\lambda S'(\theta_0) > n$, we may assume $|S(\theta)| < 1$. Take the graph of sin $n\theta$ and slide it horizontally until one of its arcs with positive slope meets the graph of $S(\theta)$ at θ_0. At this point $S(\theta)$ has a larger slope than the shifted sine curve (whose slope is at most n). By Lemma 5, the shifted sine curve and the graph of $S(\theta)$ meet at $2n+2$ points between 0 and 2π; by Lemma 3, $S(\theta)$ is itself a shifted sine curve, a contradiction since $|S(\theta)| < 1$.

Now we consider trigonometric sums with complex coefficients. It is easy to show that for each θ_0 there is a trigonometric sum $S_0(\theta)$ of degree n for which $|S_0(\theta)| \leqq 1$ and $|S_0'(\theta)|$ has the largest possible value—this depends on the fact that $S(\theta)$ has only a finite number of coefficients, whose possible values are in a bounded set since $|S(\theta)| \leqq 1$; we then appeal to the principle that a real-valued continuous function (namely $|S_0'(\theta_0)|$) defined over a compact set (the $(2n+1)$-dimensional set of possible coefficients) attains its maximum. Since $S(\theta-\theta_0)$ is a trigonometric sum if $S(\theta)$ is, there is no loss of generality in assuming that $\theta_0=0$. Take, then, an $S_0(\theta)$ such that $|S_0'(0)|$ is as large as possible; we have to show that $|S_0'(0)| \leqq n$. We can choose a real λ such that $e^{i\lambda}S_0'(0) > 0$; having done this, consider the real trigonometric sum $\mathrm{Re}(e^{i\lambda}S_0(\theta))$, whose absolute value does not exceed 1. Its derivative has absolute value not exceeding n, by what we already know; in particular, this is true at $\theta=0$, so that $0 < e^{i\lambda}S_0'(0) \leqq n$. Hence $|S_0'(0)| \leqq n$, as asserted.

Observe that although Markov's theorem cannot be iterated, as we saw, Bernstein's theorem can; in other words, $|S''(\theta)| \leqq n^2$, and so on for higher derivatives. The bounds so obtained are best possible, as is shown by $S(\theta) = \sin n\theta$.

We now turn to the proof of Markov's theorem. It is conceivable that it could be proved directly that $|P'(x)|$ attains its maximum at $x = \pm 1$. If we could do this, Markov's theorem would follow at once from Bernstein's, since $P(\cos\theta)$ is a trigonometric sum $S(\theta)$ and we have

$$|S''(\theta)| = |P''(\cos\theta)\sin^2\theta - P'(\cos\theta)\cos\theta| \leqq n^2;$$

putting $\theta = 0$ and $\theta = \pi$, we would obtain $|P'(\pm 1)| \leqq n^2$.

Unfortunately we do not yet know that $|P'(x)|$ attains its maximum at ± 1. However, since $P(\cos\theta)$ is a trigonometric sum, we do have $|P'(x)| \leqq n(1-x^2)^{-1/2}$, and if $|x| \leqq \cos(\frac{1}{2}\pi/n)$ this gives us

$$|P'(x)| \leqq n\{1 - \cos^2(\tfrac{1}{2}\pi/n)\}^{-1/2} = n \csc(\tfrac{1}{2}\pi/n) \leqq n^2$$

because $|\sin n\theta| \leqq n|\sin\theta|$ for all real θ (here $\theta = \frac{1}{2}\pi/n$). Note that if $|P(x)| < 1$ for $-1 \leqq x \leqq 1$ we have strict inequality here. Thus Markov's theorem is established except for $\cos(\frac{1}{2}\pi/n) < x < 1$ (and the symmetric interval).

To handle the excluded intervals we need another auxiliary result, which is of independent interest.

LEMMA 6 (Chebyshev's theorem). *Let $P(x)$ be a real polynomial of degree n, such that $|P(x)| \leqq 1$ on $[-1, 1]$; then the leading coefficient of $P(x)$ has absolute value at most 2^{n-1} (which is the leading coefficient of $T_n(x) = \cos n \cos^{-1} x$).*

Suppose, in fact, that $P(x)$ has a leading coefficient larger than 2^{n-1}; this is still true for $\lambda P(x)$ with some $\lambda < 1$, so that we may assume $|P(x)| < 1$ on $[-1, 1]$. The polynomial $T_n(x)$ takes the values ± 1 with alternating signs at $n+1$ points, so Lemma 4 applies with $\phi = T_n$, $f = P$: we have $P(x) = T_n(x)$ at n distinct values of x on $[-1, 1]$. In addition $P(1) < T_n(1) = 1$ but $P(x) > T_n(x)$ for large positive x, since $P(x)$ has a larger leading coefficient than $T_n(x)$. Hence there is an $(n+1)$th point where $P(x) = T_n(x)$; consequently $P(x) \equiv T_n(x)$, a contradiction. (For another elementary proof, see [19].)

We have actually proved a stronger result: we only need to assume $|P(x)| \leqq 1$ at the points where $|T_n(x)| = 1$, and we can add to the conclusion the statement that $|P(x)| \leqq |T_n(x)|$ for $|x| > 1$.

We now return to the problem of establishing Markov's theorem for $\cos(\frac{1}{2}\pi/n) < x < 1$. We shall actually prove somewhat more, namely that $|P'(x)| \leqq T_n'(x)$ for $x > \cos(\frac{1}{2}\pi/n)$ (including $x > 1$). Suppose, in fact, that $P'(x_0) > T_n'(x_0)$ for some $x_0 > \cos(\frac{1}{2}\pi/n)$; we seek to obtain a contradiction. As before, we may assume that $|P(x)| < 1$ for $|x| \leqq 1$. We may also suppose $n > 1$, since the theorem is trivial when $n = 1$. There are $n-1$ arcs of the graph of $T_n(x)$, each connecting $y = -1$ with $y = 1$, to the left of the point $\cos(\frac{1}{2}\pi/n)$, and x_0 is on the nth one or on its continuation to $x > 1$. The graph of $P(x)$ crosses each of the $n-1$ arcs, so that $P(x) = T_n(x)$ at (at least) $n-1$ points to the left of $\cos(\frac{1}{2}\pi/n)$, and therefore $P'(x) = T_n'(x)$ at $n-2$ points to the left of $\cos(\frac{1}{2}\pi/n)$. We have already seen that $|P'(\cos(\frac{1}{2}\pi/n))| < T_n'(\cos(\frac{1}{2}\pi/n))$; we have assumed $P'(x_0) > T_n'(x_0)$; and finally, by Lemma 6, we have $P'(x) < T'(x)$ for very large x. Therefore $P' = T_n'$ at (at least) two points to the right of $\cos(\frac{1}{2}\pi/n)$, making n points in all. But $P' - T_n'$ is of degree $n-1$ at most, and so vanishes identically. Hence P and T_n differ at most by a constant, which must be zero since they coincide at $n-1$ points. Since $|P(x)| < 1$, we have a contradiction. (Cf. [**12**].)

When $x > 1$, we get, in particular, $|P'(x)| < n^2(x + (x^2-1)^{1/2})^{n-1}$, which was known to Mendeleev when $n=2$.

Markov's and Bernstein's theorems have been generalized in a great variety of ways. Unfortunately, most of the generalizations cannot be proved by the elementary methods we have just used, so we shall only be able to state some typical results without proof. Each of the three theorems (Markov's, Bernstein's, and Chebyshev's) is a special case of a general problem: we have a class of functions (polynomials of degree at most n, or trigonometric sums of degree at most n) each of absolute value not exceeding 1, and to each function we attach a number (the maximum of the derivative in the Bernstein and Markov cases, the leading coefficient in Chebyshev's case); in technical language, we have a functional defined over our class of functions. We then want to maximize the functional over the given class.

Recall that Bernstein's theorem, when applied to polynomials P of absolute value at most 1 on $[-1, 1]$, gave the result that $|P(x)| \leqq n(1-x^2)^{-1/2}$. We can put this in our current framework by picking a value of x and taking our functional to be $|P'(x)|$. In other words, we ask, how large can $|P'(x)|$ be, for a given x, when $|P(x)| \leqq 1$ on $[-1, 1]$? Call this maximum $M_n(x)$. Then Markov's theorem says that $M_n(x) \leqq n^2$ for $-1 \leqq x \leqq 1$, and Bernstein's theorem says that $M_n(x) \leqq n(1-x^2)^{-1/2}$ for $-1 < x < 1$. The problem of finding $M_n(x)$ exactly was attacked by Markov himself, and solved explicitly for $n=2$ and $n=3$. Since it is easy to see that $M_n(-x) = M_n(x)$, it is enough to find $M_n(x)$ for $x \geqq 0$. Although we already know that $M_n(1) = T_n'(1)$, it is clear that $M_n(x)$ cannot always be $|T_n'(x)|$, since $T_n'(x)$ is sometimes zero; we know, however, from our proof of Markov's theorem that $M_n(x) = T_n'(x)$ for $x > \cos(\frac{1}{2}\pi/n)$, and in particular for $x > 1$. Markov found that $M_2(x) = T_2'(x) = 4x$ for $(1/2) < x < 1$, but $M_2(x) = 1/(1-x)$ for $0 \leqq x < (1/2)$. It follows that Bernstein's estimate for $M_2(x)$ is exact for just one x in $[0, 1]$, namely $x = 2^{-1/2}$ (see Figure 3). The function $M_3(x)$ is much more complicated (see the appendix to this paper, and Figure 4). Calculation of $M_n(x)$ for larger values of n had to wait until quite recently, when, as a culmination of some 30 years of work, E. V. Voronovskaja [**25**] produced a technique that not only lets one calculate $M_n(x)$ for any n (graphs for $n=4$ and 5 are given in [**26**] and [**25**]), but makes it possible to solve many other problems. For example, Gusev [**8**] finds the function corresponding to $M_n(x)$ for the functional $P^{(k)}(x)$ $(1 < k < n)$. If $P(x)$ is a polynomial with real coefficients and $|P(x)| \leqq 1$ on $[-1, 1]$ then Voronovskaja and Zinger determined $\max|\operatorname{Re} P(z)|$ and $\max|\operatorname{Im} P(z)|$ for a given complex z [**27**], and Zinger [**28**], [**29**], determined the corresponding maxima for the derivatives of P.

The maximizing functions for a given functional are usually not Chebyshev polynomials; and when they are not, elementary methods do not usually work.

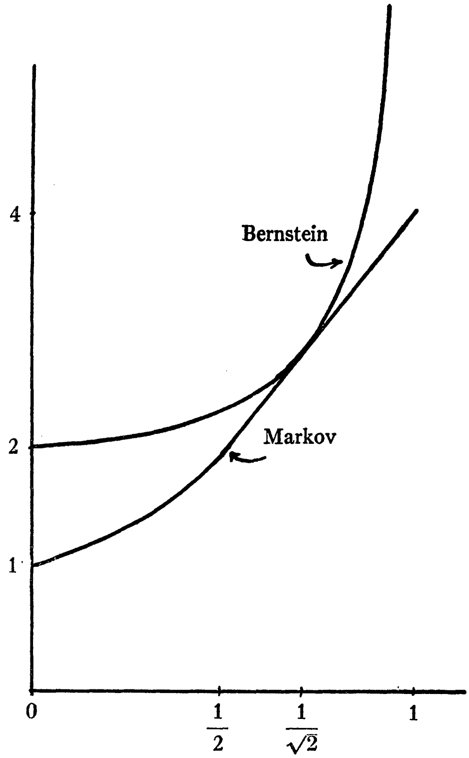

FIG. 3

FIG. 4

In fact, every polynomial is the maximizing polynomial for some functional [25], [18].

The situation for functionals over trigonometric sums is similar.

However, merely considering different functionals does not exhaust the possibilities of generalization. In the first place one can try to maximize a functional under some side-condition, for example that all the functions taken into consideration vanish at a given point. The simplest such problem is perhaps Schur's problem of maximizing $P_n(x)$ (for a given x) when P_n is a polynomial of degree n such that $|P_n(x)| \leqq 1$ on $[-1, 1]$ and $P_n(0) = 0$. This is again an elementary problem, and the answer is that $|P_n(x)| \leqq |\sin m \sin^{-1} x|$ for $|x| < \sin \frac{1}{2}\pi/m$, where $m = n$ or $n-1$ according as n is odd or even; otherwise no more than $|P_n(x)| \leqq 1$ can be asserted. Hyltén-Cavallius [10] solved the more general problem when the value of $P_n(z_0)$ is assigned for an arbitrary (real or complex) z_0. The presence of side-conditions does not complicate the problems appreciably, but neither does it simplify them.

In the Markov theorems we found a bound for a functional given $\max|P_n(x)|$ on $[-1, 1]$; in Bernstein's theorem, the maximum was taken over the unit disk. We can ask the same questions given a bound for $|P_n(z)|$ on any specified subset of the plane, or for $|S_n(\theta)|$ on a subset of a period (or indeed on a subset of the plane). (Cf. [3].)

An essential difference, which is reflected in a difference in method, arises if one restricts the class of functions in a different way. For example, consider polynomials of degree n all of whose roots are real and outside $[-1, 1]$; then if $|P(x)| \leqq 1$ on $[-1, 1]$ it follows that $|P'(x)| < \frac{1}{2}en$ (and the constant is best possible) [7].

There are also many new problems when the polynomials or trigonometric sums are restricted to be nonnegative. For example, if $S_n(\theta) = \sum_{k=-n}^{n} c_k e^{ik\theta}$ is a trigonometric sum of degree n then $|c_k| \leqq \frac{1}{2}\max|S_n(\theta)|$ for $k > \frac{1}{2}n$ (the inequality with 1 instead of $\frac{1}{2}$ is trivial); many similar but more complicated inequalities are known. A related condition to impose on a trigonometric polynomial is that it is a partial sum of the Fourier series of a nonnegative function (this does not force it to be nonnegative itself). For example, if the function is even as well as nonnegative, we have $|c_k|^2 \leqq \frac{1}{2}(1+c_{2k})$, an inequality that seems to have been discovered by crystallographers before it was noticed by mathematicians (cf. [9]).

On the other hand, one can also generalize the problem by enlarging the class of functions under consideration. In general, the maximum of a given functional over a larger class of functions can of course be expected to be larger than over the smaller class. In some cases, however, it turns out to be the same. A trigonometric sum $\sum_{k=-n}^{n} c_k e^{ikx}$ is a special case of a finite Fourier-Stieltjes transform $\int_{-n}^{n} e^{ixt} d\alpha(t)$; another special case is an integral of the form $\int_{-n}^{n} e^{ixt} g(t)dt$. This occurs in communication theory under the name of a band-limited signal (see, e.g., [17]), and it also occurs in the theory of optical instruments, antennas, and other kinds of electromagnetic apparatus [4]. It turns out that Bernstein's theorem on trigonometric sums extends to functions of this form (and even to a larger class of functions) without change (but naturally with a quite different proof), and has physical significance. There are also inequalities for nonnegative finite Fourier transforms corresponding to those for nonnegative trigonometric sums.

Still another possibility is to recognize that $\max|S_n(\theta)|$, for example, is just the norm usually used for the space of continuous functions, of which the trigonometric sums form a subset. They are also a subset of the space of functions of integrable pth power, in which the norm is

$$\left\{\int_{-\pi}^{\pi} |f(\theta)|^p d\theta\right\}^{1/p}.$$

The exact analogue of Bernstein's theorem (due to Zygmund) holds here ([30], vol. 2, p. 11; see also [23]):

$$\left\{\int_{-\pi}^{\pi} |S_n'(\theta)|^p d\theta\right\}^{1/p} \leqq n\left\{\int_{-\pi}^{\pi} |S_n(\theta)|^p d\theta\right\}^{1/p}, \qquad p \geqq 1;$$

Bernstein's theorem is the limiting case $p \to \infty$.

Again, one can extend all the problems we have considered by asking for generalizations to higher dimensions. Here the difficulty is often not so much that of proving the theorems as of discovering what would be interesting to prove. The most interesting results are those that do not have an analogue in one dimension. One illustration is as follows. A polynomial $P(z)$ in the complex variable $z=x+iy$ can be written as $R(x, y)+iS(x, y)$, where $R(x, y)$ is a harmonic polynomial (i.e., a polynomial solution of Laplace's equation). Since

$$P'(z) = \frac{\partial P}{\partial r} = \frac{\partial R}{\partial r} + i\frac{\partial S}{\partial r} = \frac{\partial R}{\partial r} + \frac{i}{r}\frac{\partial R}{\partial \theta},$$

the real and imaginary parts of P' are the (polar) components of the vector grad R. Bernstein's theorem says that if $|P(z)| \leqq 1$ for $|z| \leqq 1$ then $|P'(z)| = |\text{grad } R| \leqq n$. Szegö (cf. [22], [24]) proved more: if we assume only that $|R| \leqq 1$ then $|\text{grad } R| \leqq n$; interpreting this in rectangular coordinates, we have

$$\left(\frac{\partial R}{\partial x}\right)^2 + \left(\frac{\partial R}{\partial y}\right)^2 \leqq n^2.$$

It would be interesting to have an elementary proof of this along the preceding lines. Szegö extended the theorem to three dimensions, where the bound for grad P is more complicated; in more than three dimensions the problem seems to be unsolved.

So much has been written on Bernstein's and Markov's theorems and their generalizations that it is hardly possible to give a complete bibliography. The following list of references includes, besides items specifically cited, a number of books and papers that contain additional results or interesting methods of proof.

Appendix

Here is the explicit form of Markov's function $M_3(x)$ for $0 \leqq x \leqq 1$; its graph is sketched in Figure 4.

x	$M_3(x)$
$0 \leqq x \leqq \frac{\sqrt{7}-2}{6} = 0.108$	$3(1-4x^2)$
$\frac{\sqrt{7}-2}{6} \leqq x \leqq \frac{2\sqrt{7}-1}{9} = 0.47$	$\frac{7\sqrt{7}+10}{9(x+1)}$
$\frac{2\sqrt{7}-1}{6} \leqq x \leqq \frac{1+2\sqrt{7}}{9} = 0.70$	$\frac{16x^3}{(1-9x^2)(1-x^2)}$
$\frac{1+2\sqrt{7}}{9} \leqq x \leqq \frac{\sqrt{7}+2}{6} = 0.774$	$\frac{7\sqrt{7}-10}{9(1-x)}$
$\frac{\sqrt{7}+2}{6} \leqq x \leqq 1$	$3(4x^2-1)$

References

1. N. I. Ahiezer, Lectures on the Theory of Approximation (Russian), 2d ed., Moscow, 1965. Trans. 1st ed., 1947. Theory of Approximation, Ungar, New York, 1956. Trans. 2nd ed. Vorlesungen über Approximationstheorie, Akademie-Verlag, Berlin, 1967.

2. ———, Extremal properties of entire functions of exponential type (Russian), Teor. Funkciĭ Funkcional. Anal. i Priložen. 1 (1965) 111–135. (Trans. to appear in Amer. Math. Soc. Transl.)

3. N. I. Ahiezer, and B. J. Levin, Generalization of S. N. Bernstein's Inequality for Derivatives of Entire Functions, (Russian), Issledovanija po sovremennym problemam teorii funkciĭ kompleksnogo peremennogo, Moscow, 1960, pp. 111–165.

4. J. Arsac, Fourier Transforms and the Theory of Distributions, Prentice-Hall, Englewood Cliffs, N. J., 1966.

5. S. Bernstein, Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle, Paris, 1926.

6. R. J. Duffin and A. C. Schaeffer, On some inequalities of S. Bernstein and W. Markoff for derivatives of polynomials, Bull. Amer. Math. Soc., 43 (1938) 289–297.

7. P. Erdös, On extremal properties of the derivatives of polynomials, Ann. of Math., 2, 41 (1940) 310–313.

8. V. A. Gusev, Functionals of derivatives of an algebraic polynomial and V. A. Markov's theorem (Russian), Izv. Akad. Nauk SSSR. Ser. Mat., 25 (1961) 367–384.

9. D. Harker and J. S. Kasper, Phases of Fourier coefficients directly from crystal diffraction data, Acta Cryst., 1 (1948) 70–75. Several other papers in the same field are reviewed in Math. Reviews 12 (1951) 496.

10. C. Hyltén-Cavallius, Some extremal problems for trigonometrical and complex polynomials, Math. Scand., 3 (1955) 5–20.

11. D. Jackson, The Theory of Approximation, Amer. Math. Soc., New York, 1930.

12. O. D. Kellogg, On bounded polynomials in several variables, Math. Z., 27 (1927) 55–64.

13. A. Markov, On a problem of D. I. Mendeleev (Russian), Zapiski Imp. Akad. Nauk, 62 (1889) 1–24.

14. N. N. Meĭman, Solution of the fundamental problems of the theory of polynomials and entire functions deviating least from zero (Russian), Trudy Moskov. Mat. Obšč, 9 (1960) 507–535; Amer. Math. Soc. Transl., 2, 32 (1963) 359–393.

15. G. Meinardus, Approximation von Funktionen und ihre numerische Behandlung, Springer-Verlag, Berlin-Göttingen-Heidelberg-New York, 1964.

16. D. Mendeleev, Investigation of Aqueous Solutions Based on Specific Gravity (Russian), St. Petersburg, 1887.

17. H. O. Pollak, Energy distribution of band-limited functions whose samples on a half line vanish, J. Math. Anal. Appl., 2 (1961) 299–332.

18. W. W. Rogosinski, Extremum problems for polynomials and trigonometrical polynomials, J. London Math. Soc., 29 (1954) 259–275.

19. ———, Some elementary inequalities for polynomials, Math. Gaz., 39 (1955) 7–12.

20. ———, and G. Szegö, Extremum problems for nonnegative sine polynomials, Acta Sci. Math., (Szeged) 12, Pars B (1950) 112–124.

21. A. C. Schaeffer, Inequalities of A. Markoff and S. Bernstein for polynomials and related functions, Bull. Amer. Math. Soc., 47 (1941) 565–579.

22. ———, and G. Szegö, Inequalities for harmonic polynomials in two and three dimensions, Trans. Amer. Math. Soc., 50 (1941) 187–225.

23. E. M. Stein, Functions of exponential type, Ann. of Math., 2, 65 (1957) 582–592.

24. G. Szegö, On the gradient of solid harmonic polynomials, Trans. Amer. Math. Soc., 47 (1940) 51–65.

25. E. V. Voronovskaja, The Method of Functionals and its Applications (Russian), Leningrad, 1963. Trans. to be published by the Amer. Math. Soc.

26. ———, The first derivative functional and a sharpening of A. A. Markov's theorem, Izv. Akad. Nauk SSSR Ser. Mat., 23 (1959) 951–962.

27. ———, and M. J. Zinger, An estimate for polynomials in the complex plane (Russian), Dokl. Akad. Nauk SSSR, 143 (1962) 1022–1025; Trans. Soviet Math. Dokl., 3, 516–520.

28. M. J. Zinger, Functionals of derivatives in the complex plane (Russian), Dokl. Akad. Nauk SSSR, 166 (1966) 775–778. Trans. Soviet Math. Dokl., 7, 158–161.

29. ———, An estimate of the derivatives of an algebraic polynomial in the complex plane (Russian), Sibirsk. Mat. Z., 8 (1967) 952–957. Trans. Siberian Math. J., to appear.

30. A. Zygmund, Trigonometric Series, Cambridge, New York, 1959.