z lying in the first quadrant and $|z|$ large, we will certainly have

$$|e^{\gamma z}\varphi(z)| \leqslant e^{O(|z|)}e^{O(|z|\log|z|)} \leqslant e^{|z|^{3/2}},$$

say.

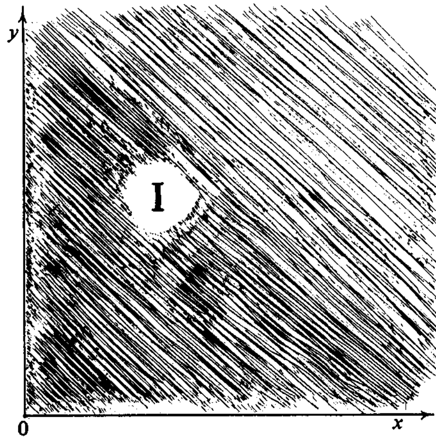


Figure 15

Therefore, because the opening of the quadrant, 90°, is *less than* $\frac{2}{3}\pi$ *radians,* the function $e^{\gamma z}\varphi(z)$, bounded on the *sides* of that quadrant, is *bounded in its interior,* and we see that

$$|\varphi(z)| \leqslant \text{const.}e^{|\gamma||z|}$$

in I. A similar argument works in each of the remaining three quadrants, and $\varphi(z)$ *is* therefore of exponential type.

Knowing φ to be of exponential type, we refer once more to the property $|\varphi(x)| \leqslant |f(x)|$, $x \in \mathbb{R}$, in order to obtain the condition

$$\int_{-\infty}^{\infty} \frac{\log^+|\varphi(x)|}{1+x^2}\,dx \;<\; \infty,$$

from the similar one assumed for f in the hypothesis. *We are in a position to apply the theorem of the previous article with our function* φ.

Write

$$B = \limsup_{y\to\infty} \frac{\log|\varphi(iy)|}{y}, \quad B' = \limsup_{y\to-\infty} \frac{\log|\varphi(iy)|}{|y|},$$

and call $\nu(t)$ the number of points λ'_n (counting multiplicities) in $[0,t]$ if $t \geqslant 0$. For $t < 0$, let $\nu(t)$ be *minus* that number of points λ'_n in $[t, 0)$. The theorem of the previous article is directly applicable to the function

$$e^{i(B-B')z/2}\varphi(z),$$

and tells us that

$$\frac{\nu(t)}{t} \longrightarrow \frac{B+B'}{2\pi} \quad \text{for} \quad t \to \pm\infty.$$

It is now claimed that $(B+B')/2 = A$, the common value of $\limsup_{y\to\infty} \log|f(iy)|/y$ and $\limsup_{y\to-\infty} \log|f(iy)|/|y|$.

First of all, for real y,

$$\begin{aligned}\left|1-\frac{iy}{\lambda_n}\right|^2 &= 1-2y\frac{\Im\lambda_n}{|\lambda_n|^2}+\frac{y^2}{|\lambda_n|^2}\\ &= \left(1+\frac{y^2}{|\lambda_n|^2}\right)\left(1-\frac{2y\Im\lambda_n}{|\lambda_n|^2+y^2}\right).\end{aligned}$$

Noting that

$$\frac{1}{|\lambda_n|^2} = \frac{1}{|\lambda_n'|^2}+\left(\frac{\Im\lambda_n}{|\lambda_n|^2}\right)^2,$$

we see that the last expression is in turn

$$\leqslant \left|1-\frac{iy}{\lambda_n'}\right|^2\left(1+\left(\frac{y\Im\lambda_n}{|\lambda_n|^2}\right)^2\right)\left(1-\frac{2y\Im\lambda_n}{|\lambda_n|^2+y^2}\right).$$

Since $(1-s)e^s \leqslant 1$ for real s we also have

$$1-\frac{2y\Im\lambda_n}{|\lambda_n|^2+y^2} \leqslant \exp\left(-\frac{2y\Im\lambda_n}{|\lambda_n|^2+y^2}\right).$$

Comparing the product representations for $f(iy)$ and $\varphi(iy)$, we see from the inequalities just written that

$$|f(iy)| \leqslant |\varphi(iy)|\prod_n\left(1+\left(\frac{y\Im\lambda_n}{|\lambda_n|^2}\right)^2\right)^{1/2}\cdot\exp\left(-\sum_n\frac{y\Im\lambda_n}{|\lambda_n|^2+y^2}\right).$$

Because $\sum_n(|\Im\lambda_n|/|\lambda_n|^2) < \infty$,

$$\left|\sum_n\frac{y\Im\lambda_n}{|\lambda_n|^2+y^2}\right| = o(|y|) \quad \text{for} \quad y\to\pm\infty.$$

The same is true for the logarithm of the *product* on the right side of the previous relation. To verify that, denote by $N(t)$ the *number* of the quantities $|\lambda_n|^2/|\Im\lambda_n|$ lying between 0 and t (counting repetitions in the usual way), and rewrite

$$\sum_n \log\left(1+\left(\frac{y\Im\lambda_n}{|\lambda_n|^2}\right)^2\right)^{1/2}$$

as

$$\frac{1}{2}\int_0^\infty \log\left(1+\frac{y^2}{t^2}\right)\mathrm{d}N(t).$$

Since $\sum_n|\Im\lambda_n|/|\lambda_n|^2<\infty$, we have $N(t)=\mathrm{o}(t)$ for $t\to\infty$, from which the integral just written is easily seen to be $\mathrm{o}(|y|)$ for $y\to\pm\infty$ after an integration by parts.

In view of these facts, the above relation between $|f(\mathrm{i}y)|$ and $|\varphi(\mathrm{i}y)|$ shows that

$$\log|f(\mathrm{i}y)|\leqslant\log|\varphi(\mathrm{i}y)|+\mathrm{o}(|y|)$$

for $y\to\pm\infty$. Therefore

$$A=\limsup_{y\to\infty}\frac{\log|f(\mathrm{i}y)|}{y}\leqslant\limsup_{y\to\infty}\frac{\log|\varphi(\mathrm{i}y)|}{y}=B,$$

and, in like manner, $A\leqslant B'$. We have thus proved that $A\leqslant(B+B')/2$.

We wish now to prove the reverse inequality. Take any $\delta>0$. We showed at the very beginning of this demonstration that, *for large r, all but $\mathrm{o}(r)$ of the original zeros λ_n of $f(z)$ with modulus $\leqslant r$ lie in one of the two sectors*

$$|\arg z|<\delta,\quad |\arg z-\pi|<\delta.$$

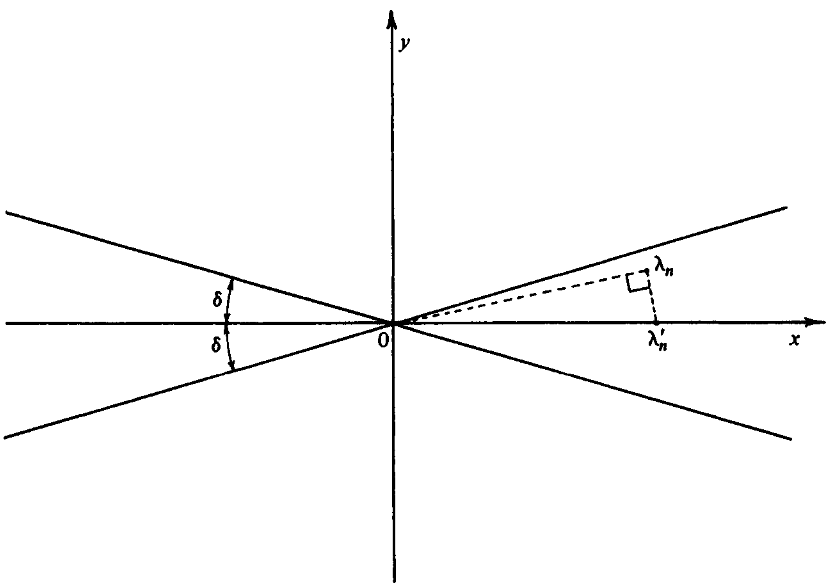


Figure 16

For λ_n in *either* of those two sectors,

$$|\lambda_n| \leqslant |\lambda_n'| = |\lambda_n|^2/|\Re\lambda_n| \leqslant |\lambda_n|\sec\delta,$$

so, for $r>0$ and large,

$$\begin{aligned} \nu(r) &\leqslant n_+(r) \leqslant \nu(r\sec\delta)+\mathrm{o}(r), \\ -\nu(-r) &\leqslant n_-(r) \leqslant -\nu(-r\sec\delta)+\mathrm{o}(r). \end{aligned}$$

(Recall that $\nu(t)$ is by its definition *negative* for $t<0$!) Since $\delta>0$ is arbitrary, these relations and the known asymptotic behaviour of $\nu(t)$ for $t\to\pm\infty$ imply that

$$\frac{n_+(r)}{r} \longrightarrow \frac{B+B'}{2\pi}$$

and

$$\frac{n_-(r)}{r} \longrightarrow \frac{B+B'}{2\pi}$$

for $r\to\infty$.

According to the theorem of §E (the first and simplest one of its kind!),

$$\log|f(z)| \leqslant A\Im z + \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\Im z}{|z-t|^2}\log^+|f(t)|\,\mathrm{d}t$$

for $\Im z>0$, and similarly, referring to the *lower* half plane,

$$\log|f(z)| \leqslant A|\Im z| + \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{|\Im z|}{|z-t|^2}\log^+|f(t)|\,\mathrm{d}t$$

for $\Im z<0$. Taking any $\delta>0$, we see from these two formulas that

$$\log|f(Re^{i\vartheta})| \leqslant AR|\sin\vartheta| + \mathrm{o}(R)$$

holds *uniformly* in each of the two sectors

$$\delta<\vartheta<\pi-\delta, \quad \pi+\delta<\vartheta<2\pi-\delta$$

when $R\to\infty$ on account of the finiteness of $\int_{-\infty}^{\infty}(\log^+|f(t)|/(1+t^2))\mathrm{d}t$.* In the remaining sectors $|\vartheta|\leqslant\delta$, $|\vartheta-\pi|\leqslant\delta$ we surely have

$$\log|f(Re^{i\vartheta})| \leqslant KR.$$

Therefore, for large R,

$$\frac{1}{2\pi}\int_0^{2\pi}\log|f(Re^{i\vartheta})|\,\mathrm{d}\vartheta \leqslant \frac{2A}{\pi}R + \mathrm{o}(R) + \frac{2\delta}{\pi}KR,$$

* In the integrals figuring in the two preceding relations the denominator of the integrand, $|z-t|^2$, is, for $z=Re^{i\vartheta}$, equal to $R^2+t^2-2Rt\cos\vartheta$. When $\delta<\vartheta<\pi-\delta$ or $\pi+\delta<\vartheta<2\pi-\delta$, this is $\geqslant(1-\cos\delta)(t^2+R^2)$.

or, since $\delta > 0$ is arbitrary,

$$\frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\vartheta})|\,d\vartheta \;\leqslant\; \frac{2A}{\pi}R + o(R)$$

for $R \to \infty$.

Now apply Jensen's formula, recalling that $n(r) = n_+(r) + n_-(r)$. By what has already been proved, we know that

$$\frac{n(r)}{r} \longrightarrow \frac{B+B'}{\pi}$$

for $r \to \infty$. Therefore

$$\int_0^R \frac{n(r)}{r}\,dr = \frac{B+B'}{\pi}R + o(R)$$

for large R. The left-hand side is, however, equal to

$$\frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\vartheta})|\,d\vartheta$$

which, as we have just seen, is $\leqslant (2A/\pi)R + o(R)$ for large R. Therefore $(B+B')/2 \leqslant A$.

The reverse inequality has already been shown. Therefore, $(B+B')/2 = A$, which means that

$$\frac{n_+(r)}{r} \longrightarrow \frac{A}{\pi} \quad \text{and} \quad \frac{n_-(r)}{r} \longrightarrow \frac{A}{\pi}$$

as $r \to \infty$. The first (principal) affirmation of our theorem is thus established, and we are done.

Remark. The above proof of the Cartwright version of Levinson's theorem depends on the elementary material of §§A and C, Kolmogorov's result, the formulas in §§E and G.1, and the first (easy) theorem in §G.3. The more delicate results in §§G.2 and G.3 (the one involving Blaschke products, in particular) are not used, nor are Lindelöf's theorems.

Problem 5

Let $f(z)$ be entire and of exponential type, with

$$\int_{-\infty}^{\infty} \frac{\log^+|f(x)|}{1+x^2}\,dx < \infty$$

and suppose that

$$\limsup_{y\to\infty} \frac{\log|f(iy)|}{y} = A.$$

(a) Show that

$$\int_0^{\pi} \left| \frac{\log|f(Re^{i\vartheta})|}{R} - A\sin\vartheta \right| d\vartheta \longrightarrow 0$$

as $R \to \infty$.

(b) Let $\{\lambda_n\}$ be the sequence of zeros of $f(z)$ in $\{\Im z > 0\}$ (repetitions according to multiplicities), and put

$$B(z) = \prod_n \left(\frac{1 - z/\lambda_n}{1 - z/\bar{\lambda}_n} \right)$$

for $\Im z > 0$. Show that

$$\int_0^{\pi} \frac{1}{R} \log|B(Re^{i\vartheta})| \, d\vartheta \longrightarrow 0$$

for $R \to \infty$.

IV

Quasianalyticity

One of the first applications of the relation

$$\int_{-\infty}^{\infty} \frac{\log^{-}|f(x)|}{1+x^2}\,dx \;<\; \infty$$

established in §G.2 of the previous chapter was to the study of quasianalyticity. This subject may now be treated by purely real-variable methods (thanks, in particular, to the work of Bang); the older function-theoretic approach of Carleman and Ostrowski is still, however, an excellent illustration of the power of complex-variable technique when applied in the investigation of real-variable phenomena, and it will be outlined here.

The material in the present chapter is due to Denjoy, Carleman, Ostrowski, Mandelbrojt and H. Cartan. We are only presenting an introduction to the subject of quasianalyticity; there are, for instance, other notions of that concept besides the one adopted here. One such, due to Beurling, will be taken up in Chapter VII, but for others, the reader should consult the books of Mandelbrojt and, regarding more recent work, Kahane's thesis published in the *Annales de l'Institut Fourier* in the early 1950s. Mandelbrojt's books on quasianalyticity also contain, of course, more elaborate treatments of the material given below. His 1952 book is the most complete, but his two earlier ones are easier to read.

A. Quasianalyticity. Sufficiency of Carleman's criterion

1. Definition of the classes $\mathscr{C}_I(\{M_n\})$

Suppose that $f(x)$ is infinitely differentiable on $\mathbb{R}$. The familiar example

$$f(x) = \begin{cases} \exp(-1/x^2), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

shows that, if, at some x_0, f and all of its derivatives *vanish*, we *cannot* conclude that $f(x) \equiv 0$. Under certain restrictions on f and its derivatives, however, such a conclusion may become legitimate. Consider, for example, functions f subject to the inequalities

$$|f^{(n)}(x)| \leqslant K^n n!, \quad x \in \mathbb{R},$$

on their successive derivatives. By looking, for instance, at Taylor's formula with Lagrange's form of the remainder, we see that

$$f(x) = f(x_0) + \sum_{n=1}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

for $|x - x_0| < 1/K$, x_0 being any point of $\mathbb{R}$, and this means that f is in fact the restriction to $\mathbb{R}$ of a function *analytic* in $|\Im z| < K$. Such a function cannot vanish together with all its derivatives at any point of $\mathbb{R}$ without being identically zero.

Are there perhaps some *other* systems of inequalities which, imposed on the successive derivatives of f, will *imply* the uniqueness property in question *without*, however, forcing f to be actually *analytic*? This question (which, like so many others in analysis, comes from mathematical physics) was raised at the beginning of the present century. The answer turns out to be *yes*, and the classes of functions thus obtained which, without necessarily being *themselves* analytic, share with the latter ones the property of being uniquely determined by their values and those of their successive derivatives at any point, are called *quasianalytic*. (Note: There are also *pseudoanalytic functions* in analysis. Those have *nothing to do* with the present discussion.)

Definition. Given any interval $I \subseteq \mathbb{R}$ and a sequence of numbers $M_n > 0$, we say that a function f, infinitely differentiable on I, belongs to the class

$$\mathscr{C}_I(\{M_n\})$$

if there are two numbers c and ρ, depending on f, such that

$$|f^{(n)}(x)| \leqslant c\rho^n M_n \quad \text{for} \quad x \in I$$

and $n = 0, 1, 2, 3, \ldots$.

Remarks. The number c is introduced *mainly for convenience*, because we want $\mathscr{C}_I(\{M_n\})$ to be a *vector space*. The number ρ is introduced because, in the case $I = \mathbb{R}$, we want $f(\rho x)$ to belong to $\mathscr{C}_{\mathbb{R}}(\{M_n\})$ when f belongs to that class.

Scholium. Suppose I has a finite endpoint, say a, but that we are *not* taking a to be in I. Then, by requiring

$$|f^{(n)}(x)| \leqslant c\rho^n M_n$$

for $x \in I$, we obtain the existence of

$$\lim_{\substack{x \to a \\ x \in I}} f^{(n)}(x)$$

for $n = 0, 1, 2, \ldots$, so that f *and all its derivatives* may be *defined by continuity* at a. We will *then still have*

$$|f^{(n)}(a)| \leqslant c\rho^n M_n$$

for $n = 0, 1, 2, \ldots$. This means that for the classes $\mathscr{C}_I(\{M_n\})$ as we have defined them, we may *always assume that the intervals I are closed.*

Definition. A class $\mathscr{C}_I(\{M_n\})$ is called *quasianalytic* if, given *any* $x_0 \in I$, *the only* $f \in \mathscr{C}_I(\{M_n\})$ such that

$$f^{(n)}(x_0) = 0, \quad n = 0, 1, 2, \ldots,$$

has $f(x) \equiv 0$, $x \in I$.

Now we have the problem: which classes $\mathscr{C}_I(\{M_n\})$ are quasianalytic and which are not?

2. The function $T(r)$. Carleman's criterion

The quasianalyticity of the class $\mathscr{C}_I(\{M_n\})$ turns out to be governed by the function

$$\boxed{T(r) = \sup_{n \geqslant 0} \frac{r^n}{M_n}}$$

defined for $r > 0$, whose use is due to Ostrowski.

Theorem. *If* $\int_0^\infty (\log T(r)/(1 + r^2))\mathrm{d}r = \infty$, *the class* $\mathscr{C}_I(\{M_n\})$ *is quasianalytic.*

Proof. Suppose that $x_0 \in I$ and that $f \in \mathscr{C}_I(\{M_n\})$ and $f^{(n)}(x_0) = 0$ for $n = 0, 1, 2, \ldots$. To prove that $f(x) \equiv 0$ on I it is enough to show that $f(x) \equiv 0$ *on any interval* $J \subseteq I$ *having* x_0 *as an endpoint.*

Without loss of generality, take $J = [0, 1]$ and suppose that x_0 is 1, i.e., that $f^{(n)}(1) = 0$, $n = 0, 1, 2, \ldots$. (Having an interval J of length $\neq 1$ only means that the parameter ρ in the bounds on $\sup_{x\in J}|f^{(n)}(x)|$ gets changed, while the M_n remain *unchanged*.) We have

$$f^{(n)}(x) \leqslant c\rho^n M_n, \quad n = 0, 1, 2, \ldots \quad \text{and} \quad 0 \leqslant x \leqslant 1.$$

For $\Re\sigma \geqslant 0$, put

$$\varphi(\sigma) = \int_0^1 t^\sigma f(t)\,\mathrm{d}t.$$

$\varphi(\sigma)$ is clearly *analytic* for $\Re\sigma > 0$ and *continuous* for $\Re\sigma \geqslant 0$, and

$$|\varphi(\sigma)| \leqslant cM_0, \quad \Re\sigma \geqslant 0.$$

We are going to show that $\varphi(\sigma) \equiv 0$. By the theorem of §G.2 in the preceding chapter (applied, of course to the *right* half plane instead of the *upper* half plane), this will certainly follow from the relation

$$\int_{-\infty}^{\infty} \frac{\log|\varphi(\mathrm{i}\tau)|}{1+\tau^2}\,\mathrm{d}\tau = -\infty,$$

which we now set out to establish.

Since $f(1) = 0$, we have, when $\Re\sigma \geqslant 0$,

$$\begin{aligned}\varphi(\sigma) &= f(t)\frac{t^{\sigma+1}}{\sigma+1}\bigg|_0^1 - \frac{1}{\sigma+1}\int_0^1 t^{\sigma+1} f'(t)\,\mathrm{d}t \\ &= -\frac{1}{\sigma+1}\int_0^1 t^{\sigma+1} f'(t)\,\mathrm{d}t.\end{aligned}$$

Again, $f'(1) = 0$, so a similar integration by parts gives us

$$\varphi(\sigma) = \frac{1}{(\sigma+1)(\sigma+2)}\int_0^1 t^{\sigma+2} f''(t)\,\mathrm{d}t.$$

Repeating this process yields

$$\varphi(\sigma) = \frac{(-1)^n}{(\sigma+1)(\sigma+2)\ldots(\sigma+n)}\int_0^1 t^{\sigma+n} f^{(n)}(t)\,\mathrm{d}t.$$

Therefore, for $\Re\sigma \geqslant 0$,

$$|\varphi(\sigma)| \;\leqslant\; \frac{1}{|\sigma+1||\sigma+2|\cdots|\sigma+n|}\cdot c\rho^n M_n, \quad n = 1, 2, 3, \ldots.$$

We have already seen that an analogous inequality holds for $n = 0$.

Putting $\sigma = \mathrm{i}\tau$ with τ real, we get

$$|\varphi(\mathrm{i}\tau)| \;\leqslant\; \frac{c\rho^n M_n}{|\tau|^n}, \qquad n = 0, 1, 2, \ldots,$$

that is,

$$\frac{1}{M_n}\left(\frac{|\tau|}{\rho}\right)^n |\varphi(\mathrm{i}\tau)| \;\leqslant\; c, \quad n = 0, 1, 2, \ldots .$$

Since by definition

$$T\left(\frac{|\tau|}{\rho}\right) \;=\; \sup_{n\geqslant 0} \frac{1}{M_n}\left(\frac{|\tau|}{\rho}\right)^n,$$

we see that

$$|\varphi(\mathrm{i}\tau)|\, T\left(\frac{|\tau|}{\rho}\right) \;\leqslant\; c,$$

so that

$$\log|\varphi(\mathrm{i}\tau)| \;\leqslant\; \log c - \log T\left(\frac{|\tau|}{\rho}\right).$$

Since $T(r) \geqslant 1/M_0$ is *bounded below* (wlog $M_0 > 0$, for otherwise *surely* $f \equiv 0$), the relation

$$\int_0^\infty \frac{\log T(r)}{1+r^2}\,\mathrm{d}r = \infty$$

implies that

$$\int_0^\infty \frac{\log T(r)}{1+\rho^2 r^2}\,\mathrm{d}r = \infty.$$

Therefore

$$\int_{-\infty}^\infty \frac{\log|\varphi(\mathrm{i}\tau)|}{1+\tau^2}\,\mathrm{d}\tau \;\leqslant\; \pi\log c - \int_{-\infty}^\infty \frac{\log T(|\tau|/\rho)}{1+\tau^2}\,\mathrm{d}\tau$$

$$= \; \pi \log c - 2\rho\int_0^\infty \frac{\log T(r)}{1+\rho^2 r^2}\,\mathrm{d}r \;=\; -\infty,$$

as claimed above.

For this reason, $\varphi(\sigma) \equiv 0$ in $\Re\sigma \geqslant 0$ and in particular $\varphi(0) = \varphi(1) = \varphi(2) = \cdots = 0$. In other words,

$$\int_0^1 t^k f(t)\,\mathrm{d}t = 0, \quad k = 0, 1, 2, \ldots .$$

By Weierstrass' theorem on polynomial approximation, this makes the (continuous!) function $f(t)$ vanish identically on $[0, 1]$. That's what we had to prove. We're done.

B. Convex logarithmic regularization of $\{M_n\}$ and the necessity of Carleman's criterion

We are going to see that the *converse* of the theorem at the end of the previous § is true. This requires us to make a preliminary study of the geometrical relationship between the sequence $\{M_n\}$ and the function $T(r)$.

1. Definition of the sequence $\{\underline{M}_n\}$. Its relation to $\{M_n\}$ and $T(r)$

By the definition of $T(r)$ given at the beginning of §A.2, we have

$$\log T(r) = \sup_{n\geqslant 0}(n\log r - \log M_n);$$

moreover (unless $\mathscr{C}_I(\{M_n\})$ consists *only* of the function *identically zero* on I, which situation we henceforth *exclude* from consideration),

$$\log T(r) \geqslant -\log M_0 > -\infty.$$

The function $\log T(r)$ clearly *increases* with $\log r$. This description of $\log T(r)$ is conveniently shown by the following diagram:

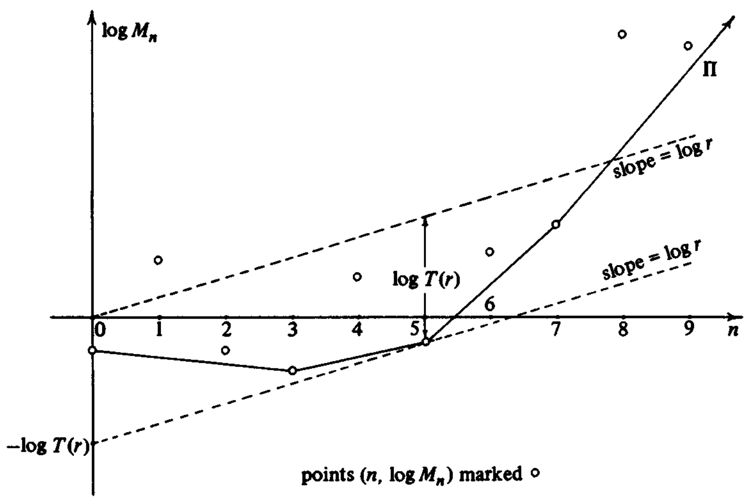


Figure 17

We see that $-\log T(r)$ is the y-intercept of the *highest* straight line of slope $\log r$ that lies *under all the points* $(n, \log M_n)$. It is convenient at this time to introduce the *highest convex curve*, Π, *lying under all the points* $(n, \log M_n)$. Π is the so-called *Newton polygon* of that collection of points (first applied by Isaac Newton in the computation of power series expansions of algebraic functions!). The straight line of ordinate

$n \log r - \log T(r)$ on the above diagram is, for each fixed $r > 0$, the *supporting line to* Π *having slope* $\log r$.

At any abscissa n, the *ordinate* of Π is simply the supremum of the ordinates of all of its supporting lines, since Π is convex. Therefore, the *ordinate* of Π at n is

$$\sup_{r>0} (n \log r - \log T(r)).$$

This quantity is henceforth denoted by $\log \underline{M}_n$; *and* $\{\underline{M}_n\}$ *is called the convex logarithmic regularization of the sequence* $\{M_n\}$; $\log \underline{M}_n$ is clearly a *convex function of* n. The following diagram shows the relation between $\log M_n$ and $\log \underline{M}_n$:

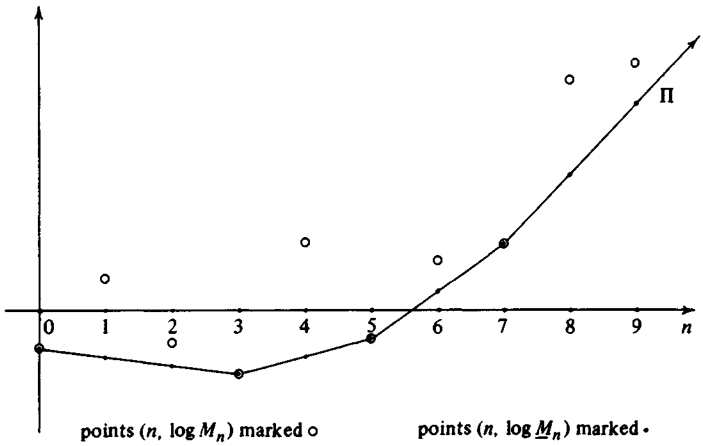


Figure 18

It is evident that $\underline{M}_n \leqslant M_n$ for each n.

The Newton polygon Π is *also* the highest convex curve lying under all the points $(n, \log M_n)$, as the figures show. Therefore, by the above geometric characterization of $\log T(r)$ in terms of supporting lines, we also have

$$\log T(r) = \sup_{n\geqslant 0} [n \log r - \log \underline{M}_n],$$

i.e.,

$$T(r) = \sup_{n\geqslant 0} \frac{r^n}{\underline{M}_n}.$$

We see that the function $T(r)$ *cannot distinguish between the sequence* $\{M_n\}$ *and its convex logarithmic regularization* $\{\underline{M}_n\}$.

It will be convenient to consider the ordinates of the Newton polygon

Π at *non-integer* abscissae v. We denote this ordinate simply by $\Pi(v)$; $\Pi(v)$ is a convex, piecewise linear function.

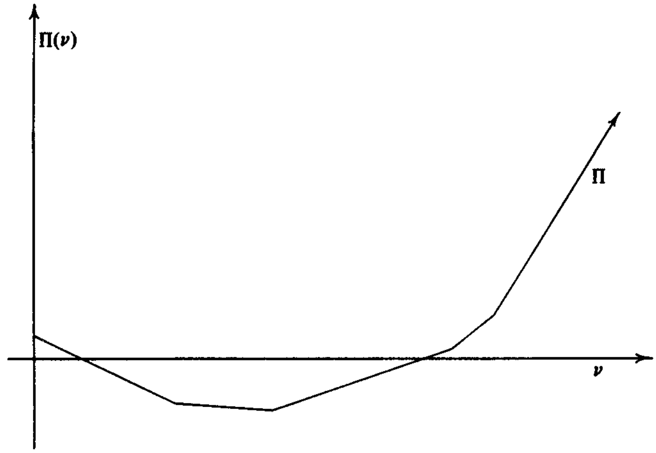


Figure 19

Lemma. *If $T(r)$ is finite for all finite r, then $\underline{M}_n^{1/n}$ is an increasing function of n when $n \geqslant$ some n_0.*

Proof. Under the hypothesis, the slope $\Pi'(v)$ of Π must tend to ∞ as $v \to \infty$. Otherwise, for some $r_0 < \infty$, $\Pi'(v)$ would remain $\leqslant \log r_0$ for all v, and *certainly no straight line of slope* $\log r$, *with* $r > r_0$, *could lie below* Π.

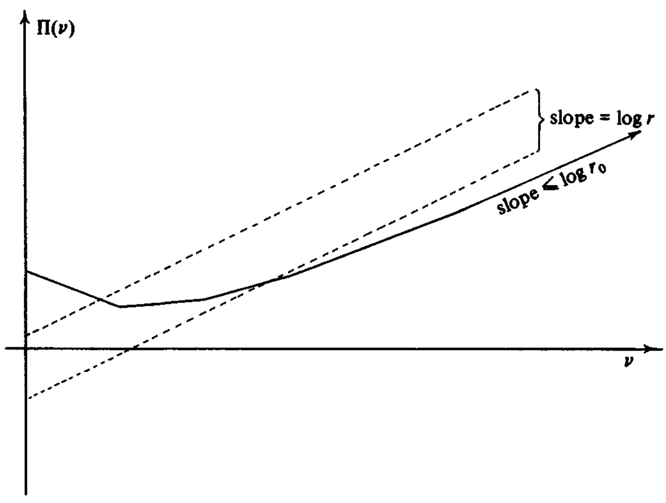


Figure 20

This means that $T(r) = \infty$ for $r > r_0$, contrary to hypothesis.

Because $\Pi'(\nu) \to \infty$ as $\nu \to \infty$ we certainly have $\Pi(\nu) \to \infty$. Pick any ν_0 with $\Pi(\nu_0)/\nu_0 > 0$, and draw a line $\mathscr{L}$ through the origin with slope *bigger than* $\Pi(\nu_0)/\nu_0$. Then $\mathscr{L}$ certainly passes *above* the point $(\nu_0, \Pi(\nu_0))$, lying on the convex curve Π.

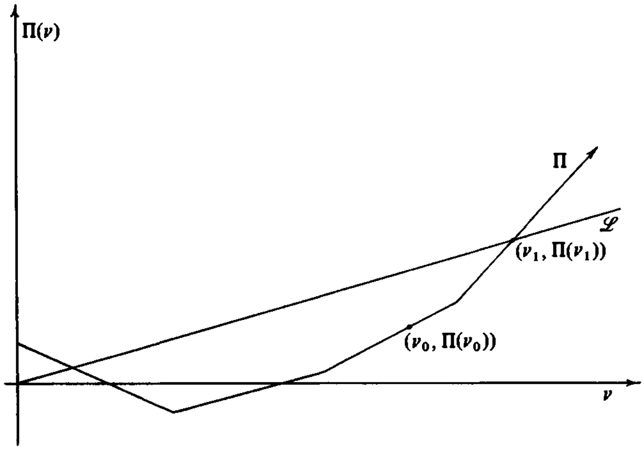


Figure 21

However, since $\Pi'(\nu) \to \infty$ as $\nu \to \infty$, $\mathscr{L}$ cannot lie above Π *forever*. Let ν_1 be the *last abscissa to the right* where Π cuts $\mathscr{L}$ – there *is* such a last abscissa because Π is convex. The figure shows that, at ν_1, Π *cuts* $\mathscr{L}$ *from below*. This means that

$$\frac{\Pi(\nu_1)}{\nu_1} < \Pi'(\nu_1).^*$$

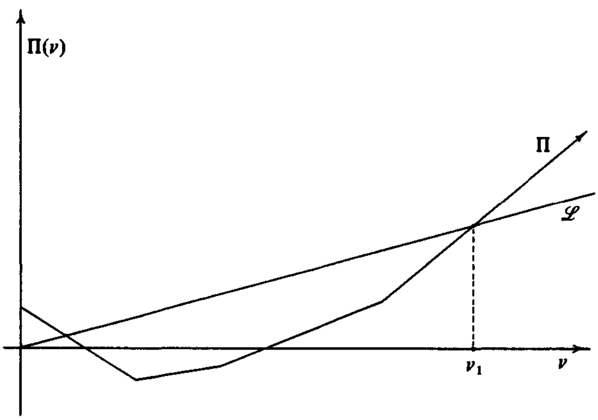


Figure 22

* If necessary, turn $\mathscr{L}$ slightly about O to ensure that $(\nu_1, \Pi(\nu_1))$ is not a vertex of Π.

It is claimed that $\Pi(\nu)/\nu$ *increases for* $\nu \geqslant \nu_1$.

To see this, observe that

$$\frac{\mathrm{d}}{\mathrm{d}\nu}\left(\frac{\Pi(\nu)}{\nu}\right) = \frac{\nu\Pi'(\nu) - \Pi(\nu)}{\nu^2};$$

so it is enough to verify that

$$\nu\Pi'(\nu) - \Pi(\nu) > 0 \quad \text{for} \quad \nu \geqslant \nu_1.$$

However, since $\Pi(\nu)$ is *convex*, $\mathrm{d}(\nu\Pi'(\nu) - \Pi(\nu)) = \nu \mathrm{d}\Pi'(\nu)$ is certainly $\geqslant 0$, so $\nu\Pi'(\nu) - \Pi(\nu)$ *increases*. And $\nu_1\Pi'(\nu_1) - \Pi(\nu_1)$ *is* > 0 by our construction. So $\nu\Pi'(\nu) - \Pi(\nu)$ remains > 0 for $\nu \geqslant \nu_1$, and therefore $\Pi(\nu)/\nu$ increases for such ν.

Let n_0 be any *integer* $\geqslant \nu_1$. Then, if $n \geqslant n_0$,

$$\frac{\log \underline{M}_n}{n} = \frac{\Pi(n)}{n} \leqslant \frac{\Pi(n+1)}{n+1} = \frac{\log \underline{M}_{n+1}}{n+1}.$$

This proves the lemma.

Lemma. *Suppose that* $\underline{M}_n^{1/n} \to \infty$ *for* $n \to \infty$. *Then, for sufficiently large* n, $\underline{M}_n^{1/n}$ *is increasing as a function of* n, *and*

$$\underline{M}_n^{1/n} \leqslant \frac{\underline{M}_{n+1}}{\underline{M}_n}.$$

Proof. Since $(\log \underline{M}_n)/n = \Pi(n)/n$ tends to ∞ with n, the *slope* of the convex curve Π *cannot remain bounded* as $\nu \to \infty$, and $\Pi'(\nu) \to \infty$, $\nu \to \infty$. We are therefore back in the situation of the previous lemma, and the argument used there shows that $\Pi(\nu)/\nu$ *increases* for large ν, and in particular $(\log \underline{M}_n)/n$ increases for large n.

The reasoning used in the above proof also showed that $\Pi'(\nu) > \Pi(\nu)/\nu$ if ν is large. Therefore, for large n,

$$\Pi'(n) > \frac{\Pi(n)}{n}.$$

Since, for $\nu \geqslant n$, $\Pi'(\nu) \geqslant \Pi'(n)$, we thus have, by the mean value theorem,*

$$\Pi(n+1) - \Pi(n) > \frac{\Pi(n)}{n},$$

i.e.,

$$\log \underline{M}_{n+1} - \log \underline{M}_n > \frac{\log \underline{M}_n}{n}$$

for large n. We're done.

* Note that all the *vertices* of Π have *integer* abscissae.

Corollary. *If* $\sum_{n=1}^{\infty} \underline{M}_n^{-1/n} < \infty$, *then*

$$\sum_{n=1}^{\infty} \frac{\underline{M}_{n-1}}{\underline{M}_n} < \infty.$$

Proof. Clear.

Theorem. *If* $\int_0^{\infty} (\log T(r)/(1+r^2))\mathrm{d}r < \infty$, *then*

$$\sum_{n=1}^{\infty} \underline{M}_n^{-1/n} < \infty.$$

Proof. (Rudin). Since $T(r)$ is increasing, convergence of $\int_0^{\infty}(\log T(r)/(1+r^2))\mathrm{d}r$ certainly implies that $T(r) < \infty$ for all finite r, so, by the first of the above lemmas, $\underline{M}_n^{1/n}$ is *increasing* for $n \geqslant n_0$, say.

As we saw during the discussion preceding the above two lemmas, $T(r) = \sup_{k \geqslant 0}(r^k/\underline{M}_k)$, and this is $\geqslant \mathrm{e}^n$ when $r \geqslant \mathrm{e}\underline{M}_n^{1/n}$. If $n \geqslant n_0$ so that $\underline{M}_n^{1/n} \leqslant \underline{M}_{n+1}^{1/(n+1)}$, we therefore get

$$\int_{\mathrm{e}\underline{M}_n^{1/n}}^{\mathrm{e}\underline{M}_{n+1}^{1/(n+1)}} \frac{\log T(r)}{r^2}\mathrm{d}r \geqslant \frac{n}{\mathrm{e}}\left[\frac{1}{\underline{M}_n^{1/n}} - \frac{1}{\underline{M}_{n+1}^{1/(n+1)}}\right].$$

Similarly,

$$\int_{\mathrm{e}\underline{M}_n^{1/n}}^{\infty} \frac{\log T(r)}{r^2}\mathrm{d}r \geqslant \frac{n}{\mathrm{e}}\underline{M}_n^{-1/n}.$$

Using these inequalities and taking an arbitrary $m > n_0$, we find that

$$\begin{aligned}\int_{\mathrm{e}\underline{M}_{n_0}^{1/n_0}}^{\infty} \frac{\log T(r)}{r^2}\mathrm{d}r &= \sum_{n=n_0}^{m-1} \int_{\mathrm{e}\underline{M}_n^{1/n}}^{\mathrm{e}\underline{M}_{n+1}^{1/(n+1)}} \frac{\log T(r)}{r^2}\mathrm{d}r \\ &\quad + \int_{\mathrm{e}\underline{M}_m^{1/m}}^{\infty} \frac{\log T(r)}{r^2}\mathrm{d}r \\ &\geqslant \frac{1}{\mathrm{e}}\sum_{n=n_0}^{m-1} n(\underline{M}_n^{-1/n} - \underline{M}_{n+1}^{-1/(n+1)}) + \frac{m}{\mathrm{e}}\underline{M}_m^{-1/m} \\ &= \frac{n_0}{\mathrm{e}}\underline{M}_{n_0}^{-1/n_0} + \sum_{n=n_0+1}^{m} \frac{1}{\mathrm{e}\underline{M}_n^{1/n}}.\end{aligned}$$

We see that

$$\sum_{n=n_0+1}^{m} \underline{M}_n^{-1/n} \leqslant \mathrm{e}\int_{\mathrm{e}\underline{M}_{n_0}^{1/n_0}}^{\infty} \frac{\log T(r)}{r^2}\mathrm{d}r.$$

Since $\log T(r)$ is bounded below, the hypothesis makes the integral on the right finite. Therefore, making $m \to \infty$, we get

$$\sum_{n_0+1}^{\infty} \underline{M}_n^{-1/n} < \infty,$$

Q.E.D.

Corollary. *If* $\int_0^\infty (\log T(r)/(1+r^2))\mathrm{d}r < \infty$,

$$\sum_{n=1}^{\infty} \frac{\underline{M}_{n-1}}{\underline{M}_n} < \infty.$$

Proof. By the theorem and the corollary just before it.

2. Necessity of Carleman's criterion and the characterization of quasianalytic classes

Using the work of the preceding article, we can now establish the

Theorem. *If* $\int_0^\infty (\log T(r)/(1+r^2))\mathrm{d}r < \infty$, $\mathscr{C}_I(\{M_n\})$ *is not quasianalytic for any interval* I *of positive length.*

Proof. Take the convex logarithmic regularization $\{\underline{M}_n\}$ of $\{M_n\}$; then, by the corollary at the end of the last article, we have

$$\sum_1^{\infty} \frac{\underline{M}_{n-1}}{\underline{M}_n} < \infty.$$

The following construction works with the ratios $\mu_n = \underline{M}_{n-1}/\underline{M}_n$. Picking any $\varepsilon > 0$, we fix an $n_0 > 1$ such that

$$\sum_{n_0}^{\infty} \mu_n < \varepsilon.$$

For each n, $\sin \mu_n z/\mu_n z$ is entire, of exponential type μ_n, and bounded by 1 in absolute value on the real axis. Hence, by Phragmén–Lindelöf (§C of Chapter III),

$$\left|\frac{\sin \mu_n z}{\mu_n z}\right| \leqslant e^{\mu_n |\Im z|}.$$

The product $\prod_{n \geqslant n_0}(\sin \mu_n z/\mu_n z)$ will therefore equal a non-zero entire function $\varphi(z)$ with $|\varphi(z)| \leqslant e^{\varepsilon|\Im z|}$, if only it is convergent. However, if $|\mu_n z|$ is *small*,

$$\frac{\sin \mu_n z}{\mu_n z} = 1 - \frac{1}{3!}\mu_n^2 z^2 + O(\mu_n^4 |z|^4).$$

Since $\sum_n \mu_n < \infty$, we also have $\sum_n \mu_n^2 < \infty$; the product in question is therefore u.c.c. convergent in the complex plane.

Put

$$f(z) = \underline{M}_{n_0-1}\left(\frac{\sin(\varepsilon/n_0)z}{(\varepsilon/n_0)z}\right)^{2n_0} \prod_{n\geqslant n_0} \frac{\sin\mu_n z}{\mu_n z};$$

$f(z)$ is a non-zero entire function with

$$|f(z)| \leqslant \underline{M}_{n_0-1}e^{3\varepsilon|\Im z|}$$

and

$$\int_{-\infty}^{\infty} |f(x)|\,dx < \infty.$$

This last relation and the boundedness of f on the real axis certainly make $\int_{-\infty}^{\infty}|f(x)|^2\,dx < \infty$, and $f(z)$ is obviously of exponential type $\leqslant 3\varepsilon$. We therefore conclude by the Paley–Wiener theorem (Chapter II, §D) that

$$(*) \qquad F(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda x} f(x)\,dx$$

vanishes identically for $\lambda < -3\varepsilon$ and for $\lambda > 3\varepsilon$. (Vanishes identically there and not just a.e., because here, $f(x)$ being in $L_1(\mathbb{R})$, $F(\lambda)$ is *continuous*.) Because $f(z) \not\equiv 0$, $F(\lambda)$ *cannot* be everywhere zero.

As we just remarked, $F(\lambda)$ is continuous on $\mathbb{R}$. It is even infinitely differentiable there. Indeed, if $k < n_0$,

$$|x^k f(x)| \leqslant \underline{M}_{n_0-1}\left(\frac{n_0}{\varepsilon}\right)^k \left|\frac{\sin(\varepsilon/n_0)x}{(\varepsilon/n_0)x}\right|^{2n_0-k}$$

certainly belongs to $L_1(\mathbb{R})$, so we can differentiate $(*)$ k times with respect to λ under the integral sign. In this way we see that

$$F^{(k)}(\lambda) = \int_{-\infty}^{\infty} (ix)^k e^{i\lambda x} f(x)\,dx$$

is in absolute value

$$\leqslant \left(\frac{n_0}{\varepsilon}\right)^k \underline{M}_{n_0-1} \int_{-\infty}^{\infty} \left|\frac{\sin(\varepsilon/n_0)x}{(\varepsilon/n_0)x}\right|^{n_0+1} dx,$$

a *finite constant*, for $k < n_0$. (Remember, we took $n_0 > 1$.) When $k \geqslant n_0$, we can start to use products of the factors $(\sin\mu_n x)/\mu_n x$, $n \geqslant n_0$, to absorb powers of x:

$$|x^k f(x)| \leqslant \left(\frac{n_0}{\varepsilon}\right)^{n_0-1} \left|\frac{\sin(\varepsilon/n_0)x}{(\varepsilon/n_0)}\right|^{n_0+1} \cdot \underline{M}_{n_0-1} \cdot \frac{1}{\mu_{n_0}} \cdot \cdot \cdot \frac{1}{\mu_k}$$

$$= \left(\frac{n_0}{\varepsilon}\right)^{n_0-1}\left|\frac{\sin(\varepsilon/n_0)x}{(\varepsilon/n_0)x}\right|^{n_0+1}\cdot \underline{M}_{n_0-1}\cdot\frac{\underline{M}_{n_0}}{\underline{M}_{n_0-1}}\cdots\frac{\underline{M}_k}{\underline{M}_{k-1}}$$

$$= \underline{M}_k\left(\frac{n_0}{\varepsilon}\right)^{n_0-1}\left|\frac{\sin(\varepsilon/n_0)x}{(\varepsilon/n_0)x}\right|^{n_0+1}$$

Therefore $|F^{(k)}(\lambda)| \leqslant \underline{M}_k(n_0/\varepsilon)^{n_0-1}\int_{-\infty}^{\infty}|\sin(\varepsilon/n_0)x/(\varepsilon/n_0)x|^{n_0+1}\,dx$ for $k \geqslant n_0$. We see that we can choose c in such a way that

$$|F^{(k)}(\lambda)| \leqslant c\underline{M}_k \quad \text{on} \quad \mathbb{R}$$

for *all* k (including the finite number of values from 0 to $n_0 - 1$).

We know, however, that $\underline{M}_k \leqslant M_k$ for each k. Therefore our function F belongs to $\mathscr{C}_{\mathbb{R}}(\{M_n\})$, does not vanish identically, but *is* identically zero outside the interval $[-3\varepsilon, 3\varepsilon]$. This means that F and all of its derivatives must *vanish* at both points 3ε and -3ε. Since $F(\lambda) \not\equiv 0$ on $[-3\varepsilon, 3\varepsilon]$, the class $\mathscr{C}_I(\{M_n\})$ cannot be quasianalytic when $I = [-3\varepsilon, 3\varepsilon]$. By translating F, we see that the same is true when I is any interval of length 6ε. Here, $\varepsilon > 0$ is arbitrary, so we're done.

The work just done can now be combined with that in §A.2 to give a *complete characterization* of the quasianalytic clases $\mathscr{C}_I(\{M_n\})$.

Theorem. *Given any interval I of positive length, the class $\mathscr{C}_I(\{M_n\})$ is quasianalytic iff any one of the following equivalent relations holds:*

(a) $\int_0^\infty (\log T(r)/(1 + r^2))dr = \infty$
(b) $\sum_n \underline{M}_n^{-1/n} = \infty$
(c) $\sum_n (\underline{M}_{n-1}/\underline{M}_n) = \infty$.

Proof. By the preceding material and logic-chopping.

We know that (a) implies quasianalyticity of $\mathscr{C}_I(\{M_n\})$ for any I by the theorem of §A.2. Also, *not*-(a) implies that $\sum_n \underline{M}_n^{-1/n} < \infty$ by the theorem of the preceding subsection, and this last relation *by itself* implies that

$$(\dagger) \qquad \sum_n (\underline{M}_{n-1}/\underline{M}_n) < \infty$$

according to the corollary immediately preceding that theorem.

Now, the *proof* of the preceding theorem is entirely based on the condition (†). Therefore, (†) *by itself* implies that $\mathscr{C}_I(\{M_n\})$ is *not* quasianalytic (for any I). So *quasianalyticity* of $\mathscr{C}_I(\{M_n\})$ *implies* (c) (the negation of (†)) which implies (b) which implies (a), which, however, *itself* implies the quasianalyticity of $\mathscr{C}_I(\{M_n\})$ as we have already seen. So complete equivalence of the latter property with any of the conditions (a), (b) and (c) is fully established. Q.E.D.

C. Scholium, Direct establishment of the equivalence between the three conditions $\int_0^\infty (\log T(r)/(1+r^2))\,dr < \infty$, $\sum_n \underline{M}_n^{-1/n} < \infty$ and $\sum_n \underline{M}_{n-1}/\underline{M}_n < \infty$.

These three conditions are of course equivalent according to the theorem at the end of the preceding §. In §B.1 we gave direct arguments, based upon the geometric properties of the Newton polygon Π of the set of points $(n, \log M_n)$, to show that the *first* condition implied the *second*, and the *second* the *third*. The above establishment of the *reverse* implications depended, however, on the theorem of §A.2 as well as on the construction used to prove the first result of §B.2; in other words, on lots of complex variable theory. Since the relationship between $T(r)$ and the convex logarithmic regularization $\{\underline{M}_n\}$ is *strictly graphical*, i.e., *geometric*, we should also give a *direct proof* of the reverse implications. Let's do that. The reasoning used involves a peculiar change of variable and a summation by parts.

Theorem. *$\sum_{n=1}^\infty (\underline{M}_{n-1}/\underline{M}_n) < \infty$ implies that*

$$\int_0^\infty \frac{\log T(r)}{1+r^2}\,dr \;<\; \infty.$$

Proof. Look at the Newton polygon Π, denoting, as in §B.1, its ordinate corresponding to the (real) abscissa ν by $\Pi(\nu)$:

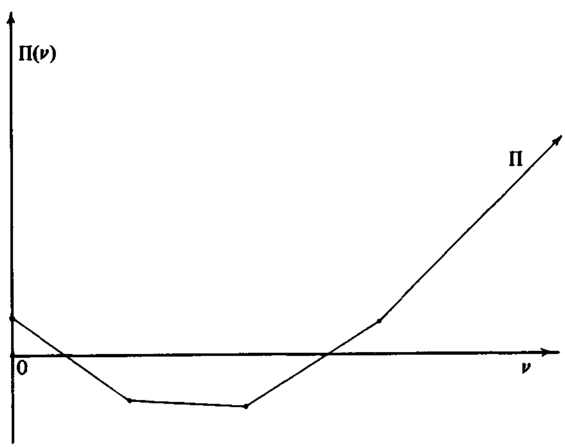


Figure 23

Π consists of certain straight segments, and each vertex of Π has a certain *integer* abscissa. Therefore, on each segment $[n-1, n]$, Π is a *straight line* *with the constant slope*

$$\Pi'(\nu) = \Pi(n) - \Pi(n-1) = \log\frac{\underline{M}_n}{\underline{M}_{n-1}}.$$

If, then,

$$\sum_n \frac{\underline{M}_{n-1}}{\underline{M}_n} < \infty,$$

we must have $\underline{M}_n/\underline{M}_{n-1} \xrightarrow[n]{} \infty$, and so the slope $\Pi'(\nu)$ of Π must tend to ∞ as $\nu \to \infty$. Either, then, Π has an *infinite* number of vertices, or else, if it has only finitely many, its *last* side must be *vertical* and have *infinite slope.* We examine only the *first* of these situations; treatment of the *second* is similar (and easier).

We are dealing, then, with a Newton polygon Π having an infinite number of vertices and thus an infinite number of straight sides whose slopes increase without limit. Denote by ν_1 the abscissa of the *first vertex* of Π *where two sides of positive slope meet*, and by ν_2, ν_3, etc. those of the successive vertices lying to *the right* of $(\nu_1, \Pi(\nu_1))$.

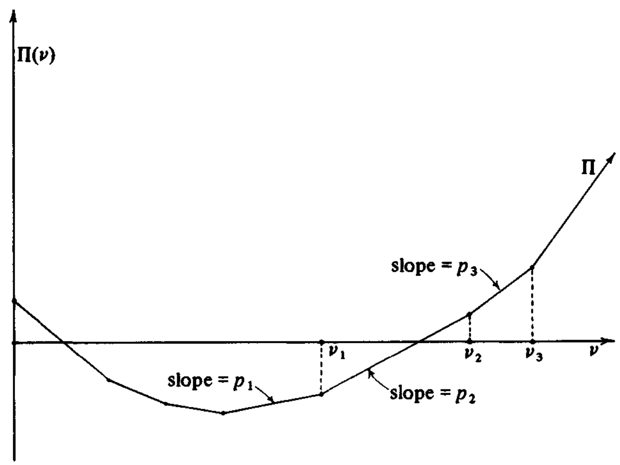


Figure 24

We call p_k the slope of the side of Π meeting the vertex $(\nu_k, \Pi(\nu_k))$ *from the left*; thus, $\Pi'(\nu) = p_k$ for $\nu_{k-1} < \nu < \nu_k$, and therefore $\underline{M}_{n-1}/\underline{M}_n = e^{-p_k}$ for $\nu_{k-1} < n \leqslant \nu_n$. (Keep in mind that the vertex abscissae ν_k are *integers*.) Since $p_k \xrightarrow[k]{} \infty$ and $p_{k+1} > p_k$ (convexity of Π), we can break up

$$\int_{\exp p_1}^{\infty} \frac{\log T(r)}{r^2}\,dr$$

into a sum of integrals of the form

$$\int_{\exp p_k}^{\exp p_{k+1}} \frac{\log T(r)}{r^2}\,\mathrm{d}r.$$

Make the change of variable $\log r = p$ and put $\log T(r) = \tau(p)$. Integrating by parts between $r_k = \mathrm{e}^{p_k}$ and $r_{k+1} = \mathrm{e}^{p_{k+1}}$, we have

$$\int_{r_k}^{r_{k+1}} \frac{\log T(r)}{r^2}\,\mathrm{d}r = \frac{\log T(r_k)}{r_k} - \frac{\log T(r_{k+1})}{r_{k+1}} + \int_{r_k}^{r_{k+1}} \frac{1}{r}\,\mathrm{d}T(r)$$

$$= \mathrm{e}^{-p_k}\tau(p_k) - \mathrm{e}^{-p_{k+1}}\tau(p_{k+1}) + \int_{p_k}^{p_{k+1}} \mathrm{e}^{-p}\,\mathrm{d}\tau(p).$$

However, for $p_k \leqslant p \leqslant p_{k+1}$, $\tau(p) = \nu_k p - \Pi(\nu_k)$:

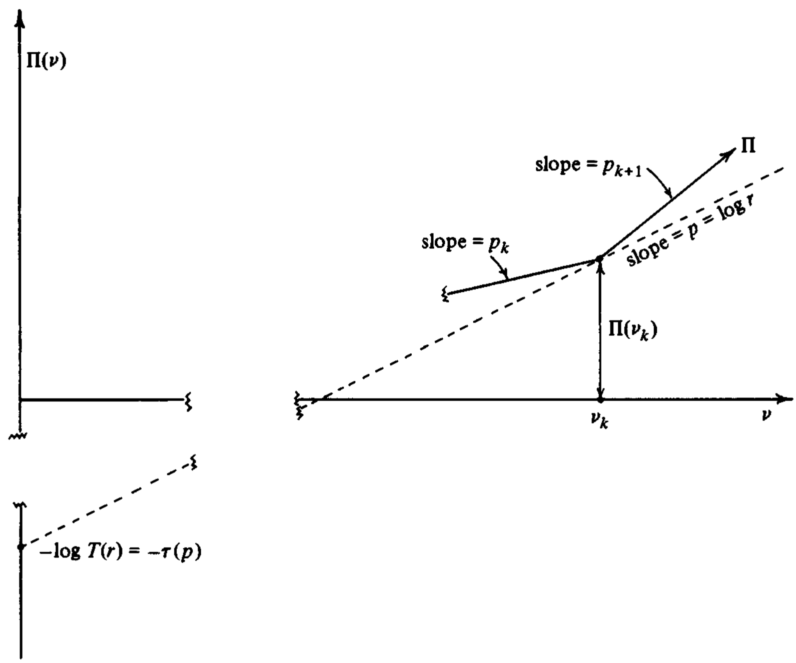


Figure 25

Therefore

$$\int_{p_k}^{p_{k+1}} \mathrm{e}^{-p}\,\mathrm{d}\tau(p) = \nu_k \int_{p_k}^{p_{k+1}} \mathrm{e}^{-p}\,\mathrm{d}p = \nu_k(\mathrm{e}^{-p_k} - \mathrm{e}^{-p_{k+1}}).$$

Adding, we thus get

$$\int_{r_1}^{r_{m+1}} \frac{\log T(r)}{r^2}\mathrm{d}r = \sum_{k=1}^{m} (\mathrm{e}^{-p_k}\tau(p_k) - \mathrm{e}^{-p_{k+1}}\tau(p_{k+1})) + \sum_{k=1}^{m} \nu_k(\mathrm{e}^{-p_k} - \mathrm{e}^{-p_{k+1}}).$$

The first sum on the right telescopes, and to the second we apply summation by parts. In this way, we see that the right side of the previous relation equals

$$\mathrm{e}^{-p_1}\tau(p_1) - \mathrm{e}^{-p_{m+1}}\tau(p_{m+1}) + \nu_1\mathrm{e}^{-p_1} + \sum_{k=2}^{m} (\nu_k - \nu_{k-1})\mathrm{e}^{-p_k} - \nu_m\mathrm{e}^{-p_{m+1}}.$$

Recall that $\log T(r)$ is increasing, so, if $\tau(p_{m+1}) = \log T(r_{m+1})$ remains <0 for $m \to \infty$,

$$\int_0^\infty \frac{\log T(r)}{1+r^2}\mathrm{d}r$$

will *certainly be* $<\infty$. There is thus no loss of generality in assuming $\tau(p_{m+1}) \geqslant 0$ for large m. If that is the case, we may drop the two terms prefixed by – signs from the previous expression and make $m \to \infty$, getting finally

$$\int_{r_1}^\infty \frac{\log T(r)}{r^2}\mathrm{d}r \leqslant \frac{\log T(r_1)}{r_1} + \nu_1\mathrm{e}^{-p_1} + \sum_{k=2}^{\infty} (\nu_k - \nu_{k-1})\mathrm{e}^{-p_k},$$

since, as we know, $r_{m+1} = \mathrm{e}^{p_{m+1}} \to \infty$ for $m \to \infty$.

As we saw at the beginning of this discussion, we have $\underline{M}_{n-1}/\underline{M}_n = \mathrm{e}^{-p_k}$ for $\nu_{k-1} < n \leqslant \nu_k$; there are clearly $\nu_k - \nu_{k-1}$ such values of n. Therefore the preceding inequality becomes

$$\int_{r_1}^\infty \frac{\log T(r)}{r^2}\mathrm{d}r \leqslant \frac{\log T(r_1)}{r_1} + \nu_1\mathrm{e}^{-p_1} + \sum_{n>\nu_1} \frac{\underline{M}_{n-1}}{\underline{M}_n}.$$

Hence $\int_{r_1}^\infty (\log T(r)/r^2)\,\mathrm{d}r < \infty$ if $\sum_n \underline{M}_{n-1}/\underline{M}_n$ converges, and thence $\int_0^\infty (\log T(r)/(1+r^2))\mathrm{d}r < \infty$, $T(r)$ being increasing.

Q.E.D.

This theorem, combined with that of §B.1, establishes equivalence of the two conditions

$$\int_0^\infty \frac{\log T(r)}{1+r^2}\mathrm{d}r < \infty \quad \text{and} \quad \sum_n \frac{\underline{M}_{n-1}}{\underline{M}_n} < \infty$$

since, as we saw in §B.1, the latter condition *is implied* by the inequality $\sum_n \underline{M}_n^{-1/n} < \infty$. In order to obtain full equivalence of our three conditions, it is still necessary to show that $\sum_n (\underline{M}_{n-1}/\underline{M}_n) < \infty$ *also implies* the relation $\sum_n \underline{M}_n^{-1/n} < \infty$. It was in order to establish such an implication that Carleman proved his celebrated inequality which says that

$$\sum_{n=1}^{\infty} (a_1 a_2 \cdots a_n)^{1/n} \leqslant c \sum_{n=1}^{\infty} a_n$$

with an absolute constant c for any sequence of numbers $a_k \geqslant 0$. Given this fact, we need only observe that

$$\frac{\underline{M}_0}{\underline{M}_1} \cdot \frac{\underline{M}_1}{\underline{M}_2} \cdots \frac{\underline{M}_{n-1}}{\underline{M}_n} = \frac{\underline{M}_0}{\underline{M}_n}.$$

Since $\underline{M}_0^{1/n} \xrightarrow[n]{} 1$, the desired implication follows directly from the inequality.

Problem 6

Prove Carleman's inequality using Lagrange's method of undetermined multipliers. (Hint: Take $x_1, x_2, \ldots, x_N \geqslant 0$; the problem is to find the *maximum* of $\sum_{n=1}^{N} (x_1 x_2 \cdots x_n)^{1/n}$ subject to the condition that $\sum_{k=1}^{N} x_k = 1$. Show first that the maximum is attained at a place where *all* the x_k are *strictly positive* for $1 \leqslant k \leqslant M$, say, $M \leqslant N$, and the x_k with $M < k \leqslant N$ (if there are any) *all vanish*. The effect of this is to merely lower N, so we may always take the maximum to be an *internal* one, obtained for $x_k > 0$, $1 \leqslant k \leqslant N$.

Now apply Lagrange's method with the undetermined multiplier λ, and show that at the presumed maximum, $\sum_1^N (x_1 x_2 \cdots x_n)^{1/n} = \lambda$ by *adding equations*. The whole problem reduces to *getting a bound* on λ. To this end, write *each* of the N equations involving λ, and in them, make the substitutions

$$x_r = \frac{\xi_r}{r}, \quad \xi_r > 0.$$

Pick out the equation obtained by doing $\partial/\partial x_k$ in the Lagrange procedure, with k so chosen that ξ_k (at the sought maximum) is $\geqslant$ *all the other* ξ_r. This will give you the estimate

$$\lambda \leqslant \sum_{n \geqslant k} \frac{k}{n(n!)^{1/n}},$$

which yields a bound on λ independent of k.)

D. The Paley–Wiener construction of entire functions of small exponential type decreasing fairly rapidly along the real axis

Suppose we are given an *increasing* function $S(r)$, defined for $r \geqslant 0$ and (say) $\geqslant 1$ there. Do there exist any non-zero entire functions $\varphi(z)$ of exponential type such that

$$|\varphi(x)| \leqslant \frac{1}{S(|x|)}, \quad x \in \mathbb{R}?$$

According to the *first* theorem of Chapter III, §G.2, there are *no* such φ if

$$\int_0^\infty \frac{\log S(r)}{1+r^2} \mathrm{d}r = \infty.$$

If, however, the integral on the left is *convergent, we can use the construction in* §B.2 (applied in proving the first theorem of that article) to *obtain* such φ with, indeed, *arbitrarily small* exponential type. This application requires us to go a little further with the graphical work of §§B.1 and C.

As far as the problem taken up in this § is concerned, there is no loss of generality in assuming that $S(r) \equiv 1$ for $0 \leqslant r \leqslant 1$, say.

Lemma. *Let $S(r)$ be increasing on $[0, \infty)$, with $S(r) \equiv 1$ for $0 \leqslant r \leqslant 1$, and suppose that*

$$\int_0^\infty \frac{\log S(r)}{r^2} \mathrm{d}r < \infty.$$

Then there is an increasing function $T(r) \geqslant S(r)$ with $\log T(r)$ a convex function of $\log r$ and also

$$\int_0^\infty \frac{\log T(r)}{r^2} \mathrm{d}r < \infty.$$

Proof. Just put

$$\log T(r) = \int_0^{er} \frac{\log S(\rho)}{\rho} \mathrm{d}\rho \ !$$

Then, since $S(\rho)$ is increasing and $\geqslant 1$,

$$\log T(r) \geqslant \log S(r) \int_r^{er} \frac{\mathrm{d}\rho}{\rho} = \log S(r).$$

Again,

$$\frac{\mathrm{d}\log T(r)}{\mathrm{d}\log r} = r\frac{\mathrm{d}\log T(r)}{\mathrm{d}r} = \log S(er),$$

an increasing function of r, so $\log T(r)$ is convex in $\log r$. Finally,

$$\int_0^\infty \frac{\log T(r)}{r^2}\mathrm{d}r = \int_0^\infty \int_0^{er} \frac{\log S(\rho)}{\rho r^2}\mathrm{d}\rho\,\mathrm{d}r$$
$$= \int_0^\infty \int_{\rho/e}^\infty \frac{\log S(\rho)}{\rho}\frac{\mathrm{d}r}{r^2}\mathrm{d}\rho = \int_0^\infty \frac{e\log S(\rho)}{\rho^2}\mathrm{d}\rho < \infty.$$

Suppose now that we are given a function $S(r)$ satisfying the hypothesis of the lemma. We may, if we like, first obtain the function $T(r)$ and then search for entire functions φ satisfying the inequality

$$(*)\qquad |\varphi(x)| \leqslant \frac{1}{T(|x|)}.$$

Any such φ will also satisfy

$$|\varphi(x)| \leqslant \frac{1}{S(|x|)}$$

and hence solve our original problem. Our task thus reduces to the construction of entire functions φ of exponential type satisfying $(*)$, given that $T(r) \geqslant 1$, *that* $\log T(r)$ *is a convex function of* $\log r$, and that

$$\int_0^\infty \frac{\log T(r)}{r^2}\mathrm{d}r < \infty.$$

Starting with such a function $T(r)$, put

$$M_n = \sup_{r>0}\frac{r^n}{T(r)} \quad \text{for} \quad n = 0, 1, 2, \ldots.$$

Since $\log M_n = \sup_{r>0}\{n\log r - \log T(r)\}$, *the sequence* $\{\log M_n\}$ *is already convex in n.* Now take

$$\tilde{T}(r) = \sup_{n\geqslant 0}\frac{r^n}{M_n}.$$

Lemma. *For* $r \geqslant 1$,

$$\log \tilde{T}(r) \leqslant \log T(r) \leqslant \log \tilde{T}(r) + \log r.$$

Proof. Uses graphs *dual* to the ones employed up to now to study the sequence $\{M_n\}$. Because $\log T(r)$ is a convex function of $\log r$, we have the following picture:

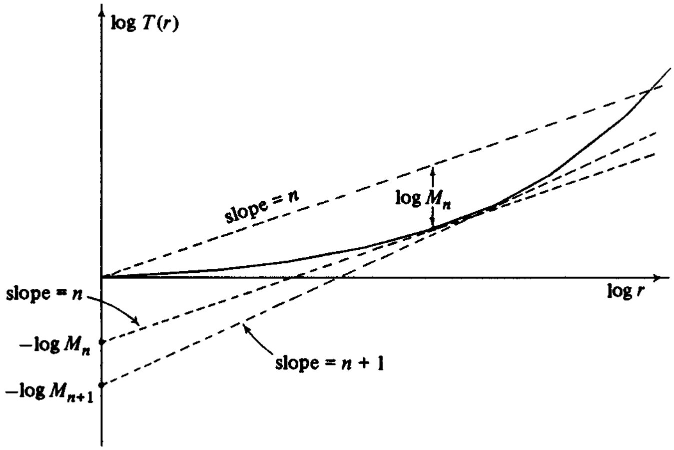


Figure 26

The supporting line to the graph of $\log T(r)$ vs $\log r$ *with integral slope* n has ordinate $n(\log r) - \log M_n$ at the abscissa $\log r$. It is therefore clear that $\log \tilde{T}(r)$, the *largest ordinate* of *those supporting lines with integral slope*, must lie *below* $\log T(r)$. This proves *one* of our desired inequalities.

To show the *other* one, take any $r > 1$, and look at any supporting line through the point $(\log r, \log T(r))$ of our graph. Since $\log T(r)$ is increasing, the *slope*, ν, of that supporting line must be $\geqslant 0$.

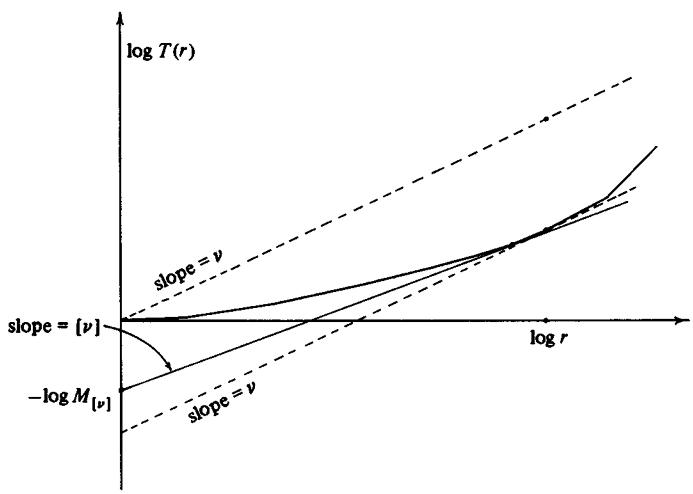


Figure 27

If $[\nu]$ denotes the largest integer $\leqslant \nu$, it is clear from the figure that

$$\nu \log r - \log M_{[\nu]} \geqslant \log T(r).$$

Therefore,

$$\begin{aligned}\log T(r) &\leqslant ([\nu]+1)\log r - \log M_{[\nu]}\\ &= [\nu]\log r - \log M_{[\nu]} + \log r \leqslant \log \tilde{T}(r) + \log r\end{aligned}$$

by definition of the function $\tilde{T}(r)$. We are done.

Theorem. *If* $T(r) \geqslant 1$ *is increasing, with* $T(r) \equiv 1$ *for* $0 \leqslant r \leqslant 1$ *and* $\log T(r)$ *a convex function of* $\log r$, *and if*

$$\int_0^\infty \frac{\log T(r)}{r^2}\,dr < \infty,$$

then, given any $\eta > 0$ *there is a non-zero entire function* $\varphi(z)$ *of exponential type* $< 2\eta$ *with*

$$|\varphi(x)| \leqslant \frac{1}{T(|x|)}, \quad x \in \mathbb{R}.$$

Proof. Form the sequence $\{M_n\}$ and then the function $\tilde{T}(r)$ in the manner described above. According to the preceding lemma, it is enough to find an entire function $\varphi \not\equiv 0$ of exponential type $< 2\eta$ with $|\varphi(x)| \leqslant 1$ on $\mathbb{R}$ and

$$|\varphi(x)| \leqslant \frac{1}{|x|\tilde{T}(|x|)} \quad \text{for} \quad |x| \geqslant 1, x \in \mathbb{R}.$$

The lemma and the hypothesis taken together tell us that

$$(\dagger) \qquad \int_1^\infty \frac{\log \tilde{T}(r)}{r^2}\,dr < \infty.$$

Also, since $\log M_n$ is here a convex function of n, the sequence $\{M_n\}$ is *identical* with its convex logarithmic regularization $\{\underline{M}_n\}$. Therefore (†) implies that

$$\sum_{n=1}^{\infty} \frac{M_{n-1}}{M_n} < \infty$$

by the corollary at the end of §B.1.

Write $\mu_n = M_{n-1}/M_n$ and take N so large that

$$\sum_{n=N}^{\infty} \mu_n < \eta.$$

Then put

$$f(z) = M_{N-1}\left(\frac{\sin(\eta/N)z}{(\eta/N)z}\right)^N \prod_{n=N}^{\infty} \frac{\sin \mu_n z}{\mu_n z}.$$

We see as in the proof of the first theorem of §B.2 that $f(z)$ is entire, not identically zero, and of exponential type $< \eta + \eta = 2\eta$. Arguing as in §B.2, we see also that $|x^k f(x)|$ is bounded on $\mathbb{R}$ for each positive integer k and moreover, when $k \geqslant N$, that

$$|x^k f(x)| \;=\; \frac{|x^{k+1}f(x)|}{|x|} \;\leqslant\; \left(\frac{N}{\eta}\right)^N \frac{M_{N-1}}{|x|\mu_N\mu_{N+1}\cdots\mu_k} \;=\; \left(\frac{N}{\eta}\right)^N \frac{M_k}{|x|},$$

whence

$$\frac{|x|^k}{M_k}|f(x)| \;\leqslant\; \frac{1}{|x|}\left(\frac{N}{\eta}\right)^N, \quad x \in \mathbb{R}.$$

Since similar inequalities hold also for $k = 0, 1, \ldots, N-1$, we have

$$\frac{|x|^k}{M_k}|f(x)| \;\leqslant\; \frac{C}{|x|}$$

for $x \in \mathbb{R}$ and $k = 0, 1, 2, \ldots$, C being a certain constant. Given $x \in \mathbb{R}$, take the supremum of the left-hand side for $k = 0, 1, 2, \ldots$. Referring to the definition of $\tilde{T}(r)$, we get $\tilde{T}(|x|)|f(x)| \leqslant C/|x|$, i.e.

$$|f(x)| \;\leqslant\; \frac{C}{|x|\tilde{T}(|x|)}, \quad x \in \mathbb{R}.$$

It is now evident that we can take φ as a suitable constant multiple of f, and φ will satisfy the required conditions. The theorem is proved.

Now we may refer to the lemma at the beginning of this §, and to the discussion given there. In that manner, we deduce from the result just established the following.

Corollary. *Let $S(r) \geqslant 1$ be increasing. A necessary and sufficient condition for there to exist entire functions $\varphi \not\equiv 0$ of exponential type with*

$$|\varphi(x)| \;\leqslant\; \frac{1}{S(|x|)} \quad \textit{on} \quad \mathbb{R}$$

is that

$$\int_0^\infty \frac{\log S(r)}{1+r^2}\,dr \;<\; \infty.$$

If that condition is met, there are entire $\varphi \not\equiv 0$ of arbitrarily small exponential type satisfying the inequality in question.

This result, which is due to Paley and Wiener, has found extensive use. Generalizations of it will be take up in Chapters X and XI.

E. Theorem of Cartan and Gorny on equality of $\mathscr{C}_{\mathbb{R}}(\{M_n\})$ and $\mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\})$. $\mathscr{C}_{\mathbb{R}}(\{M_n\})$ an algebra

The criteria for quasianalyticity of the class $\mathscr{C}_I(\{M_n\})$ given in §B.2 all depend on the *convex logarithmic regularization* $\{\underline{M}_n\}$ of $\{M_n\}$ rather than on the latter sequence itself. This makes it seem plausible that our initial consideration of classes $\mathscr{C}_I(\{M_n\})$ with *completely general* sequences $\{M_n\}$ was in fact *superfluous*, and that any such class is *in reality equal* to one of the form $\mathscr{C}_I(\{M'_n\})$ with a sequence $\{M'_n\}$ having fairly regular behaviour.

We are going to verify this hunch for the case where $I = \mathbb{R}$ by proving that $\mathscr{C}_{\mathbb{R}}(\{M_n\})$ always equals $\mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\})$. A similar result holds when the interval I is *not* the whole real line, but then the regularized sequence $\{M'_n\}$ such that $\mathscr{C}_I(\{M_n\}) = \mathscr{C}_I(\{M'_n\})$ is no longer necessarily $\{\underline{M}_n\}$. In that circumstance, which we do not treat here, the regularization process used to pass from $\{M_n\}$ to $\{M'_n\}$ is more complicated than the one yielding $\{\underline{M}_n\}$. Interested readers may find a discussion of this and other related matters in Mandelbrojt's 1952 book.

Lemma (S. Bernstein). *Let $P(z)$ be a polynomial of degree n, and suppose that $|P(x)| \leqslant M$ for $-R \leqslant x \leqslant R$. If $a > 1$, we have $|P(z)| \leqslant Ma^n$ for all z of the form $\frac{1}{2}R(w + w^{-1})$ with $1 \leqslant |w| \leqslant a$.*

Remark. For fixed a, the set of z in question fills out a certain *ellipse* with foci at $\pm R$.

Proof of lemma Under the conformal mapping $w \longrightarrow z = \frac{1}{2}R(w + w^{-1})$, the region $|w| > 1$ goes onto the complement of the segment $[-R, R]$, and the unit circumference is taken onto that segment.

With z related to w in the manner described, put, for $|w| > 1$,

$$f(w) = \frac{P(z)}{w^n};$$

$f(w)$ is then certainly analytic outside the unit circle, and continuous up to it.

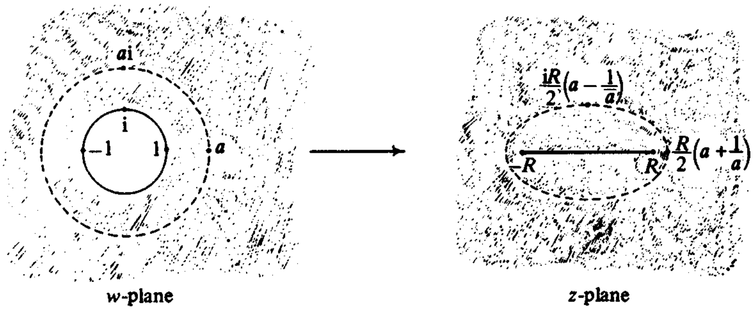


Figure 28

For large $|w|$, $z \sim Rw/2$. Since $P(z)$ is of degree n, we see that $f(w)$ is *bounded* in $\{|w|>1\}$. Therefore, by the extended principle of maximum (Chapter III, §C!), we have, in that region,

$$|f(w)| \leqslant \sup_{|\omega|=1} |f(\omega)|.$$

However, for $|\omega|=1$, $z=(R/2)(\omega+1/\omega)$ is a *real number x on the segment* $[-R,R]$, so $|f(\omega)|=|P(x)|$ is $\leqslant M$. Hence $|f(w)|\leqslant M$ for $|w|\geqslant 1$, i.e. $|P(z)|=|f(w)|\,|w|^n \leqslant M|w|^n$. For $1\leqslant |w| \leqslant a$, the right side is $\leqslant Ma^n$. We are done.

Problem 7

(a) Let P be a polynomial of degree $n-1$ with $|P(x)|\leqslant M$ for $-R\leqslant x\leqslant R$. Show that

$$|P'(0)| \leqslant \frac{en}{R} M.$$

(Hint: Apply Cauchy's inequality, using a circle of suitably chosen radius with its center at 0.)

(b) Let $f(x)$ be infinitely differentiable on $\mathbb{R}$, and *bounded* thereon. Suppose that each of f's derivatives is also bounded on $\mathbb{R}$, and write

$$B_k = \sup_x |f^{(k)}(x)|.$$

Show that there is a constant C independent of n such that

$$B_1 \leqslant C B_0^{(n-1)/n} B_n^{1/n}$$

for $n = 1, 2, 3, \ldots$.
(Hint: To show that

$$|f'(0)| \leqslant C B_0^{(n-1)/n} B_n^{1/n},$$

take

$$P(x) = f(0) + x f'(0) + \cdots + \frac{x^{n-1}}{(n-1)!} f^{(n-1)}(0),$$

and apply (a) to $P(x)$ with a suitably chosen R, using Lagrange's formula for the remainder to estimate $\sup_{-R\leqslant x\leqslant R}|P(x)|$.)

(c) By iterating the result found in (b), show that

$$B_k \leqslant C^k B_0^{(n-k)/n} B_n^{k/n} \quad \text{for } 1 \leqslant k \leqslant n-1.$$

(Hint. $f''(x)$ is $\mathrm{d}f'(x)/\mathrm{d}x$, and so forth.)

***Remarks on problem** 7.* In the result of (a), the factor e is not necessary. The inequality without e requires a more sophisticated proof. The result of (c) was first established (independently) by Gorny and by Cartan. In it, the factor C^k may be replaced by 2. This improvement, due to Kolmogorov, is quite a bit deeper. A discussion of it is found in Mandelbrojt's 1952 book. Another treatment is in the complements near the end of Akhiezer's book on the theory of approximation.

The final result of the last problem is used to establish the following

Theorem (Cartan, Gorny). *Let $M_n > 0$, and let $\{\underline{M}_n\}$ be the convex logarithmic regularization of $\{M_n\}$. Then*

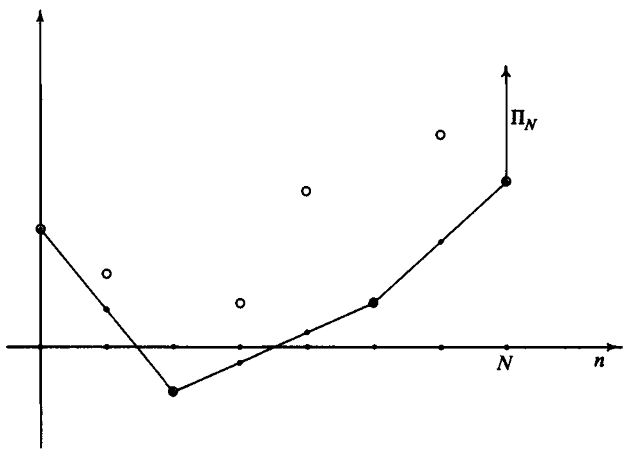
$$\mathscr{C}_{\mathbb{R}}(\{M_n\}) = \mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\}).$$

Proof. Since $\underline{M}_n \leqslant M_n$, it is manifest that $\mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\}) \subseteq \mathscr{C}_{\mathbb{R}}(\{M_n\})$, so our real task is to prove the opposite inclusion.

For each $N = 2, 3, 4, \ldots$, put

$$M_n(N) = \begin{cases} M_n & 0 \leqslant n \leqslant N, \\ \infty, & n > N, \end{cases}$$

and form the convex logarithmic regularization $\{\underline{M}_n(N)\}$ of $\{M_n(N)\}$ in the usual fashion:



points $(n, \log M_n(N))$ marked ○ points $(n, \log \underline{M}_n(N))$ marked

Figure 29

For $0 \leqslant n \leqslant N$, $\log \underline{M}_n(N)$ is the ordinate at n of Π_N, the *highest convex polygon* lying under the first $N+1$ points $(n, \log M_n)$, $n = 0, 1, \ldots, N$. As $N \to \infty$, the polygons Π_N *go down* towards Π, the Newton polygon of the set of *all* the points $(n, \log M_n)$, $n = 0, 1, 2, \ldots$. This means that $\underline{M}_n(N) \to \underline{M}_n$ for each n as $N \to \infty$.

Let $f \in \mathscr{C}_{\mathbb{R}}(\{M_n\})$. In order to show that $f \in \mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\})$ (which will complete the proof of our theorem) it is enough, according to the observation just made, to show that there are constants a and σ independent of N with

$$|f^{(n)}(x)| \leqslant a\sigma^n \underline{M}_n(N)$$

for every x and $n = 0, 1, \ldots, N$.

Pick any N, which we fix for the moment, and denote by n_k, $0 = n_0 < n_1 < n_2 < \cdots < n_p = N$, the *abscissae of the vertices* of Π_N. Since $f \in \mathscr{C}_{\mathbb{R}}(\{M_n\})$, we have $|f^{(k)}(x)| \leqslant b\rho^k M_k$ with some constants b and ρ for all $x \in \mathbb{R}$ and each $k = 0, 1, 2, 3, \ldots$. Therefore, *if n is one of the n_j, we already have*

$$|f^{(n)}(x)| \leqslant b\rho^n \underline{M}_n(N), \quad x \in \mathbb{R},$$

since in that case $\underline{M}_n(N) = \underline{M}_{n_j}(N) = M_{n_j}(N) = M_n$. Suppose, then, that $n_j < n < n_{j+1}$. We at least have

$$(*) \qquad |f^{(n_j)}(x)| \leqslant b\rho^{n_j} M_{n_j}, \quad x \in \mathbb{R},$$

and

$$(\ast\!\ast) \qquad |f^{(n_{j+1})}(x)| \leqslant b\rho^{n_{j+1}} M_{n_{j+1}}, \quad x \in \mathbb{R}.$$

Now apply result (c) of problem 7 to the function

$$g(x) = f^{(n_j)}(x).$$

With $(*)$ and $(\ast\!\ast)$, that result yields

$$|f^{(n)}(x)| = |g^{(n-n_j)}(x)| \leqslant C^{n-n_j} \{b\rho^{n_j} M_{n_j}\}^{(n_{j+1}-n)/(n_{j+1}-n_j)} \times \{b\rho^{n_{j+1}} M_{n_{j+1}}\}^{(n-n_j)/(n_{j+1}-n_j)}$$

for $x \in \mathbb{R}$. Here, C is an *absolute constant which we can wlog take* $\geqslant 1$. However,

$$M_{n_j}^{(n_{j+1}-n)/(n_{j+1}-n_j)} M_{n_{j+1}}^{(n-n_j)/(n_{j+1}-n_j)} = \underline{M}_n(N):$$

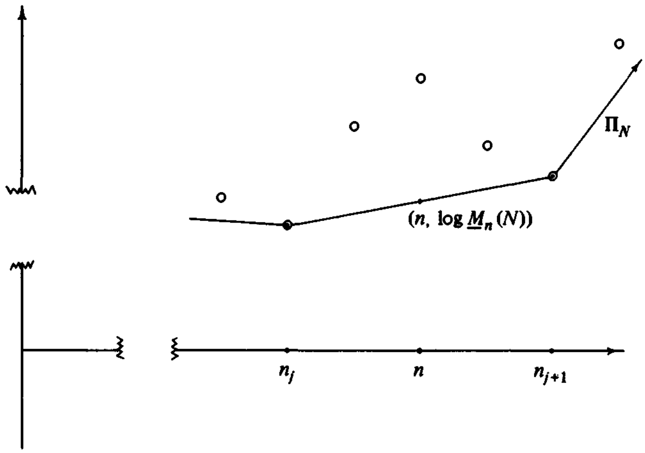


Figure 30

So the preceding relation becomes

$$|f^{(n)}(x)| \leqslant C^{n-n_j} b\rho^n \underline{M}_n(N), \quad x \in \mathbb{R},$$

or, since $C \geqslant 1$,

$$|f^{(n)}(x)| \leqslant b(C\rho)^n \underline{M}_n(N), \quad x \in \mathbb{R}.$$

We have thus established the desired inequality for $n = 0, 1, 2, \ldots, N$ with $a = b$ and $\sigma = C\rho$. As remarked above, this is enough to prove the theorem, and we are done.

The result just established has an important theoretical consequence.

Theorem. *$\mathscr{C}_{\mathbb{R}}(\{M_n\})$ is an algebra, i.e., if f and g belong to that class, so does $f \cdot g$.*

Proof. Let f and g belong to $\mathscr{C}_{\mathbb{R}}(\{M_n\})$; since $\underline{M}_n \leqslant M_n$ it is enough to prove that $f \cdot g \in \mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\})$.

By the above theorem, we certainly have f and g in $\mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\})$, so, for $n \geqslant 0$,

$$|f^{(n)}(x)| \leqslant a\rho^n \underline{M}_n, \quad x \in \mathbb{R},$$

and

$$|g^{(n)}(x)| \leqslant a\sigma^n \underline{M}_n, \quad x \in \mathbb{R},$$

if the constant a is chosen sufficiently large. According to Leibniz' formula,

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^n (f(x)g(x)) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x),$$

and, by the inequalities just given, the sum on the right is in modulus

$$\leqslant \sum_{k=0}^{n} \binom{n}{k} a\rho^k \underline{M}_k \cdot a\sigma^{n-k} \underline{M}_{n-k}.$$

However, since $\log \underline{M}_n$ is a *convex function* of m,

$$\log \underline{M}_n - \log \underline{M}_{n-k} \geqslant \log \underline{M}_k - \log \underline{M}_0,$$

so $\underline{M}_k \underline{M}_{n-k} \leqslant \underline{M}_0 \underline{M}_n$.

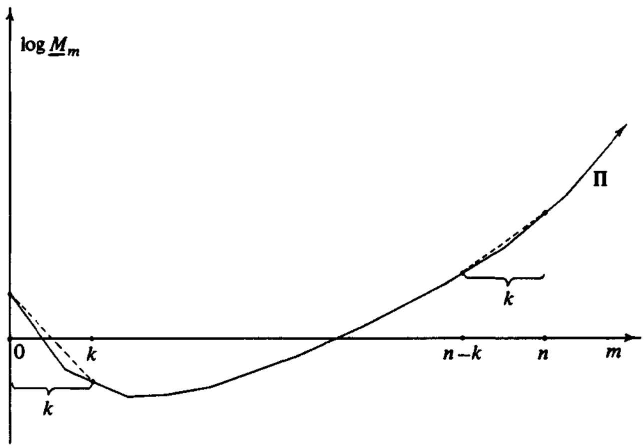


Figure 31

The preceding sum is therefore

$$\leqslant a^2 \sum_{k=0}^{n} \binom{n}{k} \rho^k \sigma^{n-k} \underline{M}_0 \underline{M}_n = a^2 \underline{M}_0 (\rho + \sigma)^n \underline{M}_n,$$

in other words,

$$\left| \left(\frac{\mathrm{d}}{\mathrm{d}x} \right)^n \Big(f(x) g(x) \Big) \right| \leqslant a^2 \underline{M}_0 (\rho + \sigma)^n \underline{M}_n, \quad x \in \mathbb{R},$$

when $n = 1, 2, 3, \ldots$.

We also clearly have $|f(x)g(x)| \leqslant a^2 M_0$ on $\mathbb{R}$. Therefore $f \cdot g \in \mathscr{C}_{\mathbb{R}}(\{\underline{M}_n\})$, and the theorem is proved.

Here is a good place to end our elementary discussion of quasianalyticity. Several ideas introduced in this chapter find applications in other parts of analysis, and will be met with again in this book. The Paley–Wiener construction in §D has various uses, and is the starting point of some important further investigations, which we will take up in Chapters X and XI. The whole notion of convex regularization, which played such a big rôle in this chapter, turns out to be especially important. A similar kind of regularization is used in some work of Beurling, and in the proof of Volberg's theorem. Those matters will be studied in Chapter VII.

V

The moment problem on the real line

The moment problem on $\mathbb{R}$, known also as the Hamburger moment problem, consists of two questions:

1. Given a numerical sequence S_0, S_1, $S_2, \ldots$, when is there a *positive* Radon measure μ on $\mathbb{R}$ with all the integrals $\int_{-\infty}^{\infty} |x|^k \mathrm{d}\mu(x)$, $k \geqslant 0$, convergent, such that

$$S_k = \int_{-\infty}^{\infty} x^k \, \mathrm{d}\mu(x), \quad k = 0, 1, 2, 3, \ldots ?$$

2. If the answer to 1 is *yes*, is there *only one positive measure* μ on $\mathbb{R}$ with

$$S_k = \int_{-\infty}^{\infty} x^k \, \mathrm{d}\mu(x), \quad k = 0, 1, 2, 3, \ldots ?$$

When the answer to 1 is *yes*, $\{S_k\}$ is called a *moment sequence* and the numbers S_k are called *the moments* of the measure μ. If, for a moment sequence $\{S_k\}$, the answer to 2 is *yes*, we say that $\{S_k\}$ is *determinate*. If the answer to 2 is *no*, the moment sequence $\{S_k\}$ is said to be *indeterminate*.

The study of various kinds of moment problems goes back to the second half of the last century, when Tchebyshev and Stieltjes investigated the moment problem on the *half-line* $[0, \infty)$. Stieltjes' research thereon led him to invent the integral bearing his name. A lot of familiar ideas and notions in analysis did in fact originate in work on the moment problem, and the subject as we now know it has many of its roots in such work. The following discussion will perhaps give the reader some perceptions of this relationship.

It is really only question 2 (the one involving *uniqueness* of μ) that has to do with the subject of this book, mainly through its connection with the material in the previous and next chapters. It would not, however, make much sense to discuss 2 without at first dealing with 1 (on the *existence* of μ). We give what is essentially M. Riesz' treatment of 1 in §A. Most of the rest of the material in this chapter is also based on M. Riesz' work.

The reader should not believe that our discussion of the moment problem reflects its real scope. We do not even touch on some very important approaches to it. There is, for instance, a vast formal structure involved with the recurrence relations of orthogonal polynomials and the algebra of continued fractions which is relevant to other parts of analysis such as Sturm–Liouville theory and the problem of interpolation by bounded analytic functions, as well as to the moment problem. Our subject is also connected in various ways with the theory of operators in Hilbert space, Krein's work being especially important in this regard. It would require a whole book to deal with all of these matters.

There is a very good book, namely, the one by Akhiezer (*The Classical Moment Problem*). That book has been translated into English; unfortunately, both the Russian and the English versions are now out of print and very hard to find – at present there is *one* copy that I know of in the city of Montreal! The older work of Tamarkin and Shohat is somewhat more accessible. The reader may also be interested in looking at the original papers by M. Riesz; they are in the *Arkiv för Matematik, Astronomi och Fysik*, and appeared around 1922–3. I am indebted to Professor R. Vermes for showing me those papers.

A. Characterization of moment sequences. Method based on extension of positive linear functionals

Theorem. *There is a positive measure μ on $\mathbb{R}$ with*

$$S_k = \int_{-\infty}^{\infty} x^k \, d\mu(x), \quad k = 0, 1, 2, 3, \ldots,$$

if and only if

$$(*) \qquad \sum_{j=0}^{N} \sum_{i=0}^{N} S_{i+j}\xi_i\xi_j \geqslant 0$$

for any N and any choice of the real numbers $\xi_0, \xi_1, \ldots, \xi_N$.

Proof. (M. Riesz) The condition $(*)$ is certainly *necessary*, for if

$$P(x) = \sum_{i=0}^{N} \xi_i x^i$$

is any real polynomial of degree N, we certainly have

$$\int_{-\infty}^{\infty} (P(x))^2 \, \mathrm{d}\mu(x) \geqslant 0;$$

the integral is, however, clearly equal to the left side of (*). The real work is to prove (*) sufficient.

Denote by $\mathscr{P}$ the set of *real polynomials*, and for $P(x) \in \mathscr{P}$ put

$$L(P) = \sum_{k=0}^{N} S_k a_k,$$

where $\sum_{k=0}^{N} a_k x^k = P(x)$. L is then a *real linear form* on the *vector space* $\mathscr{P}$; it is claimed that L is *positive* on $\mathscr{P}$, i.e., that if $P \in \mathscr{P}$ and $P(x) \geqslant 0$ on $\mathbb{R}$, we have $L(P) \geqslant 0$. *Take* a real polynomial $P(x)$ which is non-negative on $\mathbb{R}$. By Schwarz' reflection principle, $\overline{P(\bar{z})} = P(z)$, so if $\alpha \notin \mathbb{R}$ is a root of P, so is $\bar{\alpha}$, and $\bar{\alpha}$ has the same multiplicity as α. Again, every *real* root of P must have *even* multiplicity. Factoring $P(x)$ completely, we see that $P(x)$ must be of the form $|g(x)|^2$ (for real x), where $g(x)$ is a certain polynomial with *complex coefficients*. We can write

$$g(x) = R(x) + \mathrm{i}S(x)$$

where R and S are *polynomials* with *real coefficients*, and then we will have $P(x) = (R(x))^2 + (S(x))^2$, so that

$$L(P) = L(R^2) + L(S^2).$$

However, if, for example,

$$R(x) = \sum_i \xi_i x^i,$$

the coefficient of x^k in $(R(x))^2$ is $\sum_{i+j=k} \xi_i \xi_j$, whence

$$L(R^2) = \sum_k S_k \left(\sum_{i+j=k} \xi_i \xi_j \right) = \sum_i \sum_j S_{i+j} \xi_i \xi_j$$

which is $\geqslant 0$ by (*). In the same way, we see that $L(S^2) \geqslant 0$, so finally $L(P) \geqslant 0$ as asserted.

In order to prove the sufficiency of (*), we have to obtain a positive measure μ on $\mathbb{R}$ with $S_k = \int_{-\infty}^{\infty} x^k \, \mathrm{d}\mu(x)$ for $k \geqslant 0$; for this purpose, M. Riesz brought into play the rather peculiar space $\mathscr{E} = \mathscr{P} \dot{+} \mathscr{C}_0$ consisting of *all sums* $P(x) + \varphi(x)$, where $P \in \mathscr{P}$ and $\varphi \in \mathscr{C}_0$, the set of *real continuous functions* on $\mathbb{R}$ *tending to zero as* $x \to \pm\infty$. Our linear form L is *defined and positive* on the vector subspace $\mathscr{P}$ of $\mathscr{E}$, and the idea is to *extend* L *to all of* $\mathscr{E}$, in such fashion that *it remain positive on this larger space.* The extension is carried out inductively.

Let us take any fixed countable subset $\{\varphi_n: n=1,2,3,\ldots\}$ of $\mathscr{C}_0$, *dense therein with respect to the usual sup-norm* $\|\ \|_\infty$, and, for $n=1,2,\ldots$, call $\mathscr{E}_n$ the *vector subspace* of $\mathscr{E}$ generated by $\mathscr{P}$ and $\varphi_1,\ldots,\varphi_n$. In order to have a uniform notation, we write $\mathscr{E}_0=\mathscr{P}$; then $\mathscr{E}_0\subset\mathscr{E}_1\subset\mathscr{E}_2\subset\cdots$, so the union $\mathscr{E}_\infty=\bigcup_{n=0}^{\infty}\mathscr{E}_n$ is also a vector subspace of $\mathscr{E}$. We first extend L from $\mathscr{E}_0=\mathscr{P}$ to $\mathscr{E}_\infty$ so as to keep it positive on $\mathscr{E}_\infty$.

Suppose, for $n\geqslant 0$, that L has already been extended to $\mathscr{E}_n$ and is positive thereon. We show how to extend it to $\mathscr{E}_{n+1}$ so that it stays positive. In case $\varphi_{n+1}\in\mathscr{E}_n$, we have $\mathscr{E}_{n+1}=\mathscr{E}_n$, and then nothing need be done. We must examine the situation where $\varphi_{n+1}\notin\mathscr{E}_n$. Because L is already defined on $\mathscr{E}_n$, the two quantities

$$A=\sup\{Lf\colon\ f\in\mathscr{E}_n \text{ and } f\leqslant\varphi_{n+1}\}$$

and

$$B=\inf\{Lg\colon\ g\in\mathscr{E}_n \text{ and } g\geqslant\varphi_{n+1}\}$$

are available. (If peradventure there were *no* $f\in\mathscr{E}_n$ with $f\leqslant\varphi_{n+1}$ we would put $A=-\infty$. And if there were no $g\in\mathscr{E}_n$ with $g\geqslant\varphi_{n+1}$, we'd take $B=\infty$. Neither of these possibilities can, however, occur in the present situation as we shall see immediately.) It is important to verify that

$$-\infty<A\leqslant B<\infty.$$

We do this by using the function 1, which, *as a polynomial* (!), *belongs to* $\mathscr{P}=\mathscr{E}_0$, *hence to* $\mathscr{E}_n$. Because $\varphi_{n+1}\in\mathscr{C}_0$, we have the evident inequality

$$-\|\varphi_{n+1}\|_\infty\cdot 1\ \leqslant\ \varphi_{n+1}(x)\ \leqslant\ \|\varphi_{n+1}\|_\infty\cdot 1,$$

so $A\geqslant-\|\varphi_{n+1}\|_\infty L(1)$ and $B\leqslant\|\varphi_{n+1}\|_\infty L(1)$. Moreover, if f and $g\in\mathscr{E}_n$ are such that $f\leqslant\varphi_{n+1}\leqslant g$ (*note that we have just seen that there* ***are*** *such functions* f *and* g!), we have $g-f\in\mathscr{E}_n$ and $g-f\geqslant 0$, so $L(g-f)\geqslant 0$ *by positivity of* L *on* $\mathscr{E}_n$, i.e.,

$$L(f)\leqslant L(g).$$

This shows that $A\leqslant B$.

Take any number c with

$$A\leqslant c\leqslant B,$$

and put

$$L(\varphi_{n+1})=c,$$

which, taking L as linear, gives an extension of L from $\mathscr{E}_n$ to $\mathscr{E}_{n+1}$. It is claimed that L as thus extended is still positive on $\mathscr{E}_{n+1}$.

Any element of $\mathscr{E}_{n+1}$ can be expressed as $h + a\varphi_{n+1}$ with $h \in \mathscr{E}_n$ and $a \in \mathbb{R}$. If $h + a\varphi_{n+1} \geqslant 0$ and $a = 0$, then we already know $L(h + a\varphi_{n+1}) = L(h)$ is $\geqslant 0$. If $h + a\varphi_{n+1} \geqslant 0$ and $a > 0$, $(1/a)h + \varphi_{n+1} \geqslant 0$, i.e., $-(1/a)h \leqslant \varphi_{n+1}$. Since the left side belongs to $\mathscr{E}_n$, we have by the definition of A that

$$L\left(-\frac{1}{a}h\right) \leqslant A \leqslant c = L(\varphi_{n+1}),$$

whence $0 \leqslant L((1/a)h) + L(\varphi_{n+1})$ and finally $L(h + a\varphi_{n+1}) \geqslant 0$. There remains the case where $h + a\varphi_{n+1} \geqslant 0$ and $a < 0$. Here $(1/a)h + \varphi_{n+1} \leqslant 0$ and $\varphi_{n+1} \leqslant -(1/a)h \in \mathscr{E}_n$. Therefore $L(-(1/a)h) \geqslant B \geqslant c = L(\varphi_{n+1})$, $L(-(1/a)h) - L(\varphi_{n+1}) \geqslant 0$, and finally $L(h + a\varphi_{n+1}) \geqslant 0$, since $-a > 0$. Taking $L(\varphi_{n+1}) = c$ thus produces a positive extension of L from $\mathscr{E}_n$ to $\mathscr{E}_{n+1}$, as asserted.

Indefinite continuation of this process yields a positive extension of L from $\mathscr{E}_0 = \mathscr{P}$ to $\mathscr{E}_\infty$. Once this has been done, however, we may extend L from $\mathscr{E}_\infty$ to all of $\mathscr{E}$ *by continuity*.

Take any $f \in \mathscr{E}$; we have $f = P + \varphi$ where $P \in \mathscr{P}$ and $\varphi \in \mathscr{C}_0$. Since the functions φ_n used in forming the $\mathscr{E}_n$ are *dense* in $\mathscr{C}_0$, there is some subsequence of them, $\{\varphi_{n_k}\}$, with

$$\|\varphi - \varphi_{n_k}\|_\infty \longrightarrow 0 \quad \text{for} \quad k \to \infty.$$

Then $\lim_{k\to\infty} L(\varphi_{n_k})$ exists. Indeed, since L is positive on $\mathscr{E}_\infty$, which includes all the φ_n, we have

$$|L(\varphi_{n_k}) - L(\varphi_{n_l})| \leqslant \|\varphi_{n_k} - \varphi_{n_l}\|_\infty \cdot L(1),$$

because

$$-\|\varphi_{n_k} - \varphi_{n_l}\|_\infty \leqslant \varphi_{n_k}(x) - \varphi_{n_l}(x) \leqslant \|\varphi_{n_k} - \varphi_{n_l}\|_\infty.$$

Here, $\|\varphi_{n_k} - \varphi_{n_l}\|_\infty \xrightarrow[k,l]{} 0$, so the limit in question does exist. If $\{\psi_k\}$ is *any other* sequence of functions in $\mathscr{E}_\infty$ with $\|\varphi - \psi_k\|_\infty \xrightarrow[k]{} 0$, we have $\|\varphi_{n_k} - \psi_k\|_\infty \xrightarrow[k]{} 0$, so, by the argument just given,

$$L(\psi_k) - L(\varphi_{n_k}) \xrightarrow[k]{} 0,$$

and $\lim_{k\to\infty} L(\psi_k)$ exists, equalling $\lim_{k\to\infty} L(\varphi_{n_k})$. We see in this way that the latter limit is independent of the choice of the particular sequence, $\{\varphi_{n_k}\}$, of φ_n used to approximate φ in norm $\|\ \ \|_\infty$, so it makes sense to define

$$L(\varphi) = \lim_{k\to\infty} L(\varphi_{n_k}).$$

We can then put $L(f) = L(P) + L(\varphi)$, and L is in this way extended so as to be a linear form on $\mathscr{E}$.

L as thus extended is *positive* on $\mathscr{E}$. Suppose that $f \in \mathscr{E}$ is non-negative on $\mathbb{R}$. Writing as above $f = P + \varphi$ with $P \in \mathscr{P}$ and $\varphi \in \mathscr{C}_0$, we can take a sequence $\{\varphi_{n_k}\}$ of the φ_n with $\|\varphi - \varphi_{n_k}\|_\infty \xrightarrow[k]{} 0$, and we'll have

$$L(f) = L(P) + \lim_{k \to \infty} L(\varphi_{n_k}).$$

Since $P + \varphi \geqslant 0$, we see that

$$P + \|\varphi - \varphi_{n_k}\|_\infty + \varphi_{n_k} \geqslant 0;$$

this function, however, belongs to $\mathscr{E}_\infty$ (the sum of the first two terms is a polynomial!). Therefore, L being positive on $\mathscr{E}_\infty$,

$$L(P) + \|\varphi - \varphi_{n_k}\|_\infty \cdot L(1) + L(\varphi_{n_k}) \geqslant 0,$$

and, making $k \to \infty$, we get $L(f) \geqslant 0$, as claimed.

The linear form L *is in particular positive on* $\mathscr{C}_0$. Therefore, *by F. Riesz' representation theorem*, there is a *positive measure* μ on $\mathbb{R}$ with

$$L(\varphi) = \int_{-\infty}^{\infty} \varphi(x)\, \mathrm{d}\mu(x)$$

for $\varphi \in \mathscr{C}_0$. We would like now to show that in fact

$$L(f) = \int_{-\infty}^{\infty} f(x)\, \mathrm{d}\mu(x)$$

for all f in $\mathscr{E}$. This seems at first unlikely, because there is *so little connection* between the two vector spaces $\mathscr{P}$ and $\mathscr{C}_0$ used to make up $\mathscr{E}$ – there doesn't seem to be much hope of relating L's behaviour on $\mathscr{P}$ to that on $\mathscr{C}_0$. The formula in question turns out nevertheless to be correct.

In order to accomplish the passage from $\mathscr{C}_0$ to $\mathscr{P}$, M. Riesz *used a trick* (which was later codified by Choquet into the so-called 'method of adapted cones'). Let us start with an even power x^{2k} of x, and show that

$$\int_{-\infty}^{\infty} x^{2k}\, \mathrm{d}\mu(x)$$

is *finite and equal to* $L(x^{2k})$. For each large N, take the function $\varphi_N \in \mathscr{C}_0$ defined thus:

$$\varphi_N(x) = \begin{cases} x^{2k}, & |x| \leqslant N, \\ 0, & |x| \geqslant 2N, \\ \textit{a linear function} \text{ on } [N, 2N] \\ \textit{and on } [-2N, -N]. \end{cases}$$

There is clearly a quantity ε_N, *tending to zero as* $N \to \infty$, *such that*

$$(\dagger) \qquad x^{2k} \leqslant \varphi_N(x) + \varepsilon_N x^{2k+2}, \quad x \in \mathbb{R}.$$

(That's the *main idea* here!) We also have

$$\varphi_N(x) \leqslant x^{2k}, \quad x \in \mathbb{R},$$

with equality for $-N \leqslant x \leqslant N$, and of course $\varphi_N \geqslant 0$.

Since the measure μ is positive,

$$\begin{aligned} \int_{-N}^{N} x^{2k}\,\mathrm{d}\mu(x) &= \int_{-N}^{N} \varphi_N(x)\,\mathrm{d}\mu(x) \\ &\leqslant \int_{-\infty}^{\infty} \varphi_N(x)\,\mathrm{d}\mu(x) = L(\varphi_N) \leqslant L(x^{2k}), \end{aligned}$$

the last inequality holding because L is positive on $\mathscr{E}$. At the same time, by $(\dagger)$ and the positivity of L,

$$L(x^{2k}) \leqslant L(\varphi_N) + \varepsilon_N L(x^{2k+2}),$$

whilst

$$L(\varphi_N) = \int_{-2N}^{2N} \varphi_N(x)\,\mathrm{d}\mu(x) \leqslant \int_{-2N}^{2N} x^{2k}\,\mathrm{d}\mu(x).$$

Combining these two relations with the preceding one, we get

$$\int_{-N}^{N} x^{2k}\,\mathrm{d}\mu(x) \leqslant L(x^{2k}) \leqslant \int_{-2N}^{2N} x^{2k}\,\mathrm{d}\mu(x) + \varepsilon_N L(x^{2k+2}).$$

Making $N \to \infty$, we see that

$$L(x^{2k}) = \int_{-\infty}^{\infty} x^{2k}\,\mathrm{d}\mu(x),$$

since $\varepsilon_N \to 0$. Because $|x|^k \leqslant 1 + x^{2k}$ and L is finite (!), this reasoning also shows that all the integrals

$$\int_{-\infty}^{\infty} |x|^k\,\mathrm{d}\mu(x)$$

are *convergent*.

We must still treat the odd powers of x. This can be done by going through an argument like the one just made, working with $x^{2k} + x^{2k+1} + x^{2k+2}$ (a *non-negative* function of x!) instead of with x^{2k}. In that way, we can conclude that

$$L(x^{2k} + x^{2k+1} + x^{2k+2}) = \int_{-\infty}^{\infty} (x^{2k} + x^{2k+1} + x^{2k+2})\,\mathrm{d}\mu(x),$$

whence, using what we already know for the even powers of x,

$$L(x^{2k+1}) = \int_{-\infty}^{\infty} x^{2k+1}\,\mathrm{d}\mu(x).$$

The relation $L(x^n) = \int_{-\infty}^{\infty} x^n\,\mathrm{d}\mu(x)$ is now established for $n = 0, 1, 2, 3 \ldots$. *However*, $L(x^n) = S_n$ *according to our original definition of the linear form* L! So

$$S_n = \int_{-\infty}^{\infty} x^n\,\mathrm{d}\mu(x),$$

and $\{S_n\}$ *is* a moment sequence. We are done.

Remark. The argument at the beginning of the above proof can be followed so as to establish a general theorem about the extension of positive linear functionals on real linear spaces $\mathscr{E}$ with positive cones. (A *positive cone* in $\mathscr{E}$ is a cone $\mathscr{K}$ with vertex at 0 such that $\mathscr{E} = \mathscr{K} - \mathscr{K}$, and a functional on $\mathscr{E}$ is called *positive* if it takes non-negative values on $\mathscr{K}$.) The reader is invited to formulate and prove such a theorem.

Remark. The reader's attention is directed to the similarity of the inductive extension procedure used in the above proof and the inductive step in the proof of the Hahn–Banach theorem. What is the relation of the general theorem alluded to in the previous remark and Hahn–Banach? Can either one be obtained from the other?

B. Scholium. Determinantal criterion for $\{S_n\}$ to be a moment sequence

The necessary and sufficient condition for $\{S_k\}$ to be a moment sequence furnished by the theorem of the preceding §, namely, that the forms

$$\sum_{i=0}^{N} \sum_{j=0}^{N} S_{i+j}\xi_i\xi_j$$

be *positive*, is equivalent to another one involving the principal determinants of the infinite matrix

$$\begin{bmatrix} S_0, & S_1, & S_2, & S_3, & \cdots \\ S_1, & S_2, & S_3, & S_4, & \cdots \\ S_2, & S_3, & S_4, & S_5, & \cdots \\ S_3, & S_4, & & & \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}.$$

This determinantal condition played an important rôle in the older investigations on the moment problem, and we give it here for the sake of completeness. Matrices like the one just written are called *Hankel matrices* and were extensively studied towards the end of the last century, in particular, by Frobenius.

We need two lemmas from linear algebra.

Lemma. *Given the symmetric matrix*

$$S = \begin{bmatrix} s_{0,0} & s_{0,1} & \cdots & s_{0,N+1} \\ s_{1,0} & s_{1,1} & \cdots & s_{1,N+1} \\ \vdots & & & \\ s_{N+1,0} & s_{N+1,1} & \cdots & s_{N+1,N+1} \end{bmatrix}$$

where $N \geqslant -1$ (*sic!*), *the form*

$$\sum_{i=0}^{N+1} \sum_{j=0}^{N+1} s_{i,j}\xi_i\xi_j$$

is strictly (*sic!*) *positive definite if and only if all the principal determinants*

$$\det \begin{bmatrix} s_{0,0} & s_{0,1} & \cdots & s_{0,M} \\ s_{1,0} & s_{1,1} & \cdots & s_{1,M} \\ \vdots & & & \\ s_{M,0} & s_{M,1} & \cdots & s_{M,M} \end{bmatrix}$$

are *strictly positive* for $M = 0, 1, \ldots, N + 1$.

► ***Remark and warning.*** If we replace '*strictly* positive definite' by 'positive definite' and merely require the principal determinants to be $\geqslant 0$, the corresponding statement is *false*. Example:

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

The danger of this pitfall (in which I myself landed during one of my lectures!) was pointed out to me by Professor G. Schmidt.

Proof of lemma. If the quadratic form in question *is* strictly positive definite, then so is each of the forms

$$\sum_{i=0}^{M} \sum_{j=0}^{M} s_{i,j}\xi_i\xi_j$$

for $M = 0, 1, \ldots, N + 1$. This means that the *characteristic values* of the *matrix* of any such form are all *strictly positive*. But the *product* of the characteristic values of the form just written is equal to the *determinant* figuring in the lemma's statement. So, in *one* direction, the lemma is *clear*.

To go in the *opposite* direction, we argue by *induction* on N. For $N = -1$ we have the quadratic form $s_{0,0}\xi_0^2$ whose determinant is just $s_{0,0}$. In this case, the desired result is manifest.

Let us therefore *assume* that the lemma is *true* with N standing in place of $N+1$, and then *prove* that it is *also true* as stated, with $N+1$. We are given that the determinants in question are all >0 for $M = 0, 1, \ldots, N+1$. In particular, then, we have $s_{0,0} > 0$ (*this is the place in the proof where strict positivity of the principal determinants is used*!) so we may wlog take $s_{0,0} = 1$, since multiplication of the quadratic form by a constant >0 does not affect its strict positivity. With this normalization, our $(N+2) \times (N+2)$ matrix S take the form

$$S = \left[\begin{array}{c|cccc} 1 & \sigma_1 & \sigma_2 & \cdots & \sigma_{N+1} \\ \hline \sigma_1 & & & & \\ \sigma_2 & & & & \\ \vdots & & & S' & \\ \sigma_{N+1} & & & & \end{array}\right]$$

where S' is a certain $(N+1) \times (N+1)$ symmetric matrix.

To show that S is strictly positive definite, it is enough to show that the matrix T congruent to it equal to

$$(\ast)\qquad \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\sigma_1 & 1 & 0 & \cdots & 0 \\ -\sigma_2 & 0 & 1 & 0\cdots & 0 \\ \vdots & & & & \\ -\sigma_{N+1} & 0 & 0 & \cdots & 1 \end{bmatrix} \times S \times \begin{bmatrix} 1 & -\sigma_1 & -\sigma_2 & \cdots & -\sigma_{N+1} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}$$

is strictly positive definite. Observe first of all that

$$T = \left[\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & \\ 0 & & T' & & \\ \vdots & & & & \\ 0 & & & & \end{array}\right]$$

with a certain $(N+1) \times (N+1)$ symmetric matrix T'. It is therefore clear that T will be strictly positive definite if T' is. *On account of the particular triangular forms* of the matrices standing on each side of S in $(\ast)$, we have, however, for any *principal minor*

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & t_{1,1} & t_{1,2} & \cdots & t_{1,R} \\ 0 & t_{2,1} & t_{2,2} & \cdots & t_{2,R} \\ \vdots & & & & \\ 0 & t_{R,1} & t_{R,2} & \cdots & t_{R,R} \end{bmatrix}$$

of T $(1 \leqslant R \leqslant N+1)$:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & t_{1,1} & \cdots & t_{1,R} \\ \vdots & & & \\ 0 & t_{R,1} & \cdots & t_{R,R} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\sigma_1 & 1 & \cdots & 0 \\ \vdots & & & \\ -\sigma_R & 0 & \cdots & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} s_{0,0} & \cdots & s_{0,R} \\ \vdots & & \\ s_{R,0} & \cdots & s_{R,R} \end{bmatrix} \begin{bmatrix} 1 & -\sigma_1 & \cdots & -\sigma_R \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

The determinant of the matrix on the left is therefore equal to

$$\det \begin{bmatrix} s_{0,0} & \cdots & s_{0,R} \\ s_{R,0} & \cdots & s_{R,R} \end{bmatrix}$$

which by hypothesis is > 0 for $1 \leqslant R \leqslant N+1$. The determinant of the left-hand matrix is, however, just

$$\det \begin{bmatrix} t_{1,1} & \cdots & t_{1,R} \\ \vdots & & \\ t_{R,1} & \cdots & t_{R,R} \end{bmatrix},$$

i.e., *the determinant of the Rth principal minor* of the $(N+1) \times (N+1)$ *symmetric matrix* T'. Those determinants are therefore all > 0, so, since T' has one row and one column *less* than T, it is strictly positive definite by our induction hypothesis. So, therefore, is T, and hence S, as we wished to show. The lemma is proved.

Kronecker's lemma. *Let a sequence* $s_0, s_1, s_2, \dots$ *be given, and denote, for* $n \geqslant 0$, *the matrix*

$$\begin{bmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & & & \\ s_n & s_{n+1} & \cdots & s_{2n} \end{bmatrix}$$

by Δ_n. *Suppose there is a number* $m \geqslant 1$ *such that* $\det \Delta_n \neq 0$ *for* $n = 0, 1, \dots, m-1$ *while* $\det \Delta_n = 0$ *for all* $n \geqslant m$. *Then there are numbers* $\alpha_0, \alpha_1, \dots, \alpha_{m-1}$ *such that, for all* $p \geqslant 0$,

$$s_{m+p} + \alpha_{m-1} s_{m-1+p} + \cdots + \alpha_0 s_p = 0.$$

Proof. Since $\det \Delta_m = 0$, there is a non-trivial relation of linear dependence

$$\alpha_0 \begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_m \end{bmatrix} + \alpha_1 \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{m+1} \end{bmatrix} + \cdots + \alpha_m \begin{bmatrix} s_m \\ s_{m+1} \\ \vdots \\ s_{2m} \end{bmatrix} = 0$$

between the columns of Δ_m. Since $\det \Delta_{m-1} \neq 0$, we *cannot have* $\alpha_m = 0$, and may as well take $\alpha_m = 1$. Then the desired relation clearly holds for $p = 0, 1, 2, \ldots, m$, and we want to show that it holds for $p > m$. This we do by induction.

Write, for $p \geqslant 0$,

$$\Sigma_{m+p} = s_{m+p} + \alpha_{m-1} s_{m-1+p} + \cdots + \alpha_0 s_p,$$

and *assume* that $\Sigma_{m+p} = 0$ for $p = 0, 1, \ldots, r-1$, with $r - 1 \geqslant m$, i.e. $r \geqslant m + 1$. Let us then *prove* that $\Sigma_{m+r} = 0$.

We have $\det \Delta_r = 0$. Since $r \geqslant m + 1$, we can write

$$\Delta_r = \left[\begin{array}{ccc|cccc} & & & s_m & s_{m+1} & \cdots & s_r \\ & \Delta_{m-1} & & s_{m+1} & \cdots & & \\ & & & \vdots & & & \vdots \\ & & & s_{2m-1} & \cdots & & s_{m+r-1} \\ \hline s_m & \cdots & \cdots & s_{2m} & \cdots & & s_{m+r} \\ \vdots & & & \vdots & & & \vdots \\ s_r & & & s_{m+r} & \cdots & & s_{2r} \end{array}\right]$$

Denote by $\boldsymbol{\sigma}_k$ the kth *column* of this matrix, whose *initial* column is called the *zeroth* one. The *determinant* of the matrix is then *unchanged* if, for each $k \geqslant m$ we *add* to $\boldsymbol{\sigma}_k$ the linear combination

$$\alpha_{m-1}\boldsymbol{\sigma}_{k-1} + \alpha_{m-2}\boldsymbol{\sigma}_{k-2} + \cdots + \alpha_0\boldsymbol{\sigma}_{k-m}$$

of the m columns preceding it. These column operations convert Δ_r to the matrix

$$\left[\begin{array}{ccc|ccc} & & & \Sigma_m & \cdots & \Sigma_r \\ & \Delta_{m-1} & & \vdots & & \\ & & & \Sigma_{2m-1} & \cdots & \Sigma_{m+r-1} \\ \hline s_m & \cdots & s_{2m-1} & \Sigma_{2m} & \cdots & \Sigma_{m+r} \\ \vdots & & & \vdots & & \\ s_r & \cdots & s_{m+r-1} & \Sigma_{m+r} & \cdots & \Sigma_{2r} \end{array}\right]$$

which, by our induction hypothesis, equals

$$\left[\begin{array}{ccc|cccc} & & & 0 & 0 & \cdots & 0 \\ & & & 0 & 0 & \cdots & 0 \\ & \Delta_{m-1} & & \vdots & \vdots & & \vdots \\ & & & 0 & 0 & \cdots & 0 \\ \hline s_m & \cdots & s_{2m-1} & 0 & \cdots & 0 & \Sigma_{m+r} \\ \vdots & & \vdots & \vdots & & & \vdots \\ & & & 0 & \Sigma_{m+r} & & \\ s_r & \cdots & s_{m+r-1} & \Sigma_{m+r} & \cdots & & \Sigma_{2r} \end{array}\right]$$

The determinant of this latter matrix is, however, just

$$\det \Delta_{m-1} \cdot \det \begin{bmatrix} 0 & 0 & \cdots & \Sigma_{m+r} \\ \vdots & & & \\ 0 & \Sigma_{m+r} & & \\ \Sigma_{m+r} & \Sigma_{m+r+1} & \cdots & \Sigma_{2r} \end{bmatrix}$$

$= \pm \det \Delta_{m-1}(\Sigma_{m+r})^{r-m+1}$. This quantity, then, is equal to $\det \Delta_r$ which we know must be *zero* since $r \geqslant m+1$. But, according to the hypothesis of the lemma, $\det \Delta_{m-1} \neq 0$. Hence $\Sigma_{m+r} = 0$, which is what we wanted to prove. The lemma is established.

Now we are able to prove the main result of this §.

Theorem. *Given a sequence of numbers* $s_0, s_1, s_2, \ldots,$ *form the matrices*

$$\Delta_n = \begin{bmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & \\ \vdots & & & \\ s_n & s_{n+1} & \cdots & s_{2n} \end{bmatrix}.$$

A necessary and sufficient condition for the s_k *to be the moments of a non-zero positive measure* μ *is that either*

(i) *all the quantities* $\det \Delta_n$ *are* > 0 (*sic*!) *for* $n = 0, 1, 2, \ldots,$

or else

(ii) *for some* $m \geqslant 1$, $\det \Delta_n > 0$ *for* $n = 0, 1, \ldots, m-1$, *while* $\det \Delta_n = 0$ *for all* $n \geqslant m$.

Remarks. The condition that $\det \Delta_n \geqslant 0$ for $n \geqslant 0$ is *necessary, but not sufficient* for $\{s_k\}$ to be a moment sequence. Case (ii) of the theorem is *degenerate* and, as we shall see, happens iff the s_k are the moments of a positive measure *supported on a finite set of points.*

Proof of theorem. Suppose, in the first place, that we *have* a positive non-zero measure μ on $\mathbb{R}$ with

$$s_k = \int_{-\infty}^{\infty} x^k \, d\mu(x), \quad k = 0, 1, 2, \ldots.$$

Then, as we observed at the beginning of the proof of the theorem in §A,

$$\sum_{i=0}^{n} \sum_{j=0}^{n} s_{i+j}\xi_i\xi_j = \int_{-\infty}^{\infty} \left(\sum_{k=0}^{n} \xi_k x^k \right)^2 d\mu(x).$$

If μ is *not* supported on a finite set of points, the integral on the right can only vanish when

$$\xi_0 = \xi_1 = \xi_2 = \cdots = \xi_n = 0,$$

so in this case all the forms $\sum_{i=0}^n \sum_{j=0}^n s_{i+j}\xi_i\xi_j$ are *strictly positive definite.* Here, $\det \Delta_n > 0$ for all $n \geqslant 0$ by the *first* of the above two lemmas.

Suppose now that our positive measure μ is supported on m points, call them $x_1, x_2, \ldots, x_m$. If $n < m$, the polynomial

$$\sum_{k=0}^{n} \xi_k x^k$$

vanishes at *each* of those points only when $\xi_0 = \xi_1 = \cdots = \xi_n = 0$, so, if $\mu(\{x_p\}) > 0$ for $1 \leqslant p \leqslant m$, the form

$$\sum_{i=0}^{n} \sum_{j=0}^{n} s_{i+j}\xi_i\xi_j$$

is *strictly positive definite* when $0 \leqslant n < m$. By the first lemma, then, $\det \Delta_n > 0$, $0 \leqslant n < m$. Consider now a value of n which is $\geqslant m$. We can then take the polynomial

$$x^{n-m}(x - x_1)(x - x_2)\ldots(x - x_m)$$

which vanishes on the support of μ. Rewriting that polynomial as

$$\sum_{k=0}^{n} \xi_k x^k$$

we must therefore have

$$\sum_{i=0}^{n} \sum_{j=0}^{n} s_{i+j}\xi_i\xi_j = 0,$$

although here $\xi_n = 1 \neq 0$. For such n, our quadratic form, although *positive definite*, is *not strictly so*, and hence at least *one* characteristic value of the matrix Δ_n must be *zero*. This makes $\det \Delta_n = 0$ whenever $n \geqslant m$.

Our theorem is proved in one direction.

Going the other way, suppose, first of all, that we are in case (i). Then, by the first lemma, *all* the quadratic forms

$$\sum_{i=0}^{n} \sum_{j=0}^{n} s_{i+j}\xi_i\xi_j$$

are (strictly) positive definite, so $\{s_k\}$ *is* a moment sequence by the theorem of §A. The argument just given shows that, *here*, *no* positive measure of which the s_k are the moments *can be supported* on a *finite set of points.*

It remains for us to treat case (ii). By the theorem of §A, we will be *through* when we show that *all the forms*

$$\sum_{i=0}^{n} \sum_{j=0}^{n} s_{i+j}\xi_i\xi_j$$

are positive definite. For $0 \leqslant n < m$ we do have $\det \Delta_n > 0$, so we can by the *first* lemma conclude that those forms *are* positive definite for such n.

To handle the forms with $n \geqslant m$ we must apply *Kronecker's lemma.* According to that result, we have some quantities $\alpha_0, \alpha_1, \ldots, \alpha_{m-1}$ such that

$$\alpha_0 s_p + \alpha_1 s_{p+1} + \cdots + \alpha_{m-1} s_{m-1+p} + s_{m+p} = 0$$

for $p \geqslant 0$. For $n \geqslant m$, our matrix Δ_n takes the form

$$\left[\begin{array}{ccc|ccc} & \Delta_{m-1} & & s_m & \cdots & s_n \\ & & & s_{2m-1} & \cdots & s_{n+m-1} \\ \hline s_m & \cdots & \cdots & s_{2m} & \cdots & s_{n+m} \\ \vdots & & & \vdots & & \\ s_n & \cdots & \cdots & s_{n+m} & \cdots & s_{2n} \end{array}\right]$$

The $(n+1) \times (n+1)$ matrix

$$\left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \hline \alpha_0 & \alpha_1 & \cdots & \alpha_{m-1} & 1 & 0 & \cdots & 0 \\ 0 & \alpha_0 & \cdots & \alpha_{m-2} & \alpha_{m-1} & 1 & & \vdots \\ \vdots & & & & & & & \\ 0 & \cdots & \alpha_0 & & & \cdots & \alpha_{m-1} & 1 \end{array}\right]$$

is non-singular. Therefore, positive definiteness of Δ_n is *implied* by that of the product

$$
\left[\begin{array}{cccc|cccc}
1 & 0 & \cdots & 0 & 0 & & \cdots & 0 \\
0 & 1 & \cdots & 0 & & & & \\
\vdots & & & & \vdots & & & \vdots \\
0 & & \cdots & 1 & 0 & & \cdots & 0 \\
\hline
\alpha_0 & & \cdots & \alpha_{m-1} & 1 & & & \\
\vdots & & & & & & & \\
0 & \cdots & \alpha_0 & & & \cdots & \alpha_{m-1} & 1
\end{array}\right]
\left[\begin{array}{ccccc}
\multicolumn{3}{c|}{\Delta_{m-1}} & s_m \;\cdots & s_n \\
\cline{1-3}
 & & & \vdots & \vdots \\
s_m & \cdots & s_{2m} & \cdots & s_{m+n} \\
 & & \vdots & & \vdots \\
s_n & \cdots & s_{n+m} & \cdots & s_{2n}
\end{array}\right]
$$

$$
\times\left[\begin{array}{cccc|cccc}
1 & 0 & \cdots & 0 & \alpha_0 & & \cdots & 0 \\
0 & 1 & & & \vdots & & & \vdots \\
 & & & & & & & \alpha_0 \\
0 & & \cdots & 1 & \alpha_{m-1} & & & \vdots \\
\hline
0 & & \cdots & 0 & 1 & & & \\
 & & & & & & & \vdots \\
 & & & & & & & \alpha_{m-1} \\
0 & & \cdots & 0 & & \cdots & & 1
\end{array}\right].
$$

Using the relation furnished by Kronecker's lemma, we see that the product is just

$$
\left[\begin{array}{ccc|ccc}
 & & & 0 & \cdots & 0 \\
 & \Delta_{m-1} & & \vdots & & \vdots \\
 & & & 0 & \cdots & 0 \\
\hline
0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & & \vdots & & & \\
0 & \cdots & 0 & 0 & \cdots & 0
\end{array}\right].
$$

This matrix is *certainly* positive definite (although, of course, not *strictly* so!), because Δ_{m-1} is, as we already know. So Δ_n is positive definite (*not* strictly) also for $n \geqslant m$, and the proof is finished for case (ii). We are all done.

Remark. Since large determinants are hard to compute, the theorem just proved may not seem to be of much use. It does, in any event, furnish the complete answer to a rather interesting question.

Suppose we *lift* the requirement that the measures considered be positive in our statement of the moment problem. Let us, in other words, ask *which real sequences $\{A_k\}$ can be represented in the form*

$$A_k = \int_{-\infty}^{\infty} x^k \, d\tau(x), \quad k = 0, 1, 2, \ldots,$$

with *real signed measures* τ such that $\int_{-\infty}^{\infty} |x|^k |d\tau(x)| < \infty$ for every $k \geqslant 0$. The rather surprising answer turns out to be that *every real sequence $\{A_k\}$ can be so represented.*

In order to establish this fact, it is enough to show that, *given any real* *sequence* $\{A_k\}$, *two moment sequences* $\{S_k\}$ and $\{S'_k\}$ can be found with

$$A_k = S'_k - S_k, \quad k = 0, 1, 2, \ldots.$$

We use an inductive procedure to do this.

Take first $S_0 > 0$, and sufficiently large so that $S'_0 = A_0 + S_0$ is *also* > 0. Put $S_1 = 0$ and $S'_1 = A_1$. It is clear that if $S_2 > 0$ is *large enough*, and $S'_2 = A_2 + S_2$, *both the determinants*

$$\det\begin{bmatrix} S_0 & S_1 \\ S_1 & S_2 \end{bmatrix}, \quad \det\begin{bmatrix} S'_0 & S'_1 \\ S'_1 & S'_2 \end{bmatrix}$$

will be strictly positive.

Now just keep going. We can take $S_3 = 0$ and $S'_3 = A_3$. *Because the above two determinants are* > 0, we can find $S_4 > 0$ *large enough* so that

$$\det\begin{bmatrix} S_0 & S_1 & S_2 \\ S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \end{bmatrix} \quad \text{and} \quad \det\begin{bmatrix} S'_0 & S'_1 & S'_2 \\ S'_1 & S'_2 & S'_3 \\ S'_2 & S'_3 & S'_4 \end{bmatrix}$$

are both > 0, where $S'_4 = A_4 + S_4$. There is clearly nothing to stop the continuation of this process. For each *odd* k we take $S_k = 0$ and $S'_k = A_k$. If $k = 2m + 2$, we can adjust $S_k > 0$ so as to make the corresponding $(m + 2) \times (m + 2)$ determinants involving the S_l and S'_l, $0 \leqslant l \leqslant k$, both > 0 (with, of course, $S'_k = A_k + S_k$) by *merely taking account* of the S_l and S'_l *already gotten* for $0 \leqslant l \leqslant k - 1$. This is because the *preceding step* has *already ensured* that

$$\det\begin{bmatrix} S_0 & S_1 & \cdots & S_m \\ S_1 & & & \\ \vdots & & & \\ S_m & \cdots & & S_{2m} \end{bmatrix} > 0.$$

and

$$\det\begin{bmatrix} S'_0 & S'_1 & \cdots & S'_m \\ S'_1 & & & \\ \vdots & & & \\ S'_m & \cdots & & S'_{2m} \end{bmatrix} > 0.$$

The sequences $\{S_k\}$ and $\{S'_k\}$ arrived at by following this procedure indefinitely are *moment sequences* according to the above theorem, and their construction is such that $A_k = S'_k - S_k$ for $k = 0, 1, 2, \ldots$. That is what we needed.

The result just found should have some applications. I do not know of any.

C. Determinacy. Two conditions, one sufficient and the other necessary

Having discussed the circumstances under which $\{S_k\}$ is a moment sequence, we come to the second question: if it *is*, when is the positive measure with moments S_k *unique*? In this §, we derive some simple partial answers to this question from earlier results.

1. Carleman's sufficient condition

Theorem (Carleman). *A moment sequence $\{S_k\}$ is determinate provided that*

$$\sum_{k=0}^{\infty} \frac{1}{S_{2k}^{1/2k}} = \infty.$$

Proof. Suppose we have *two* positive measures, μ and ν, with

$$S_k = \int_{-\infty}^{\infty} x^k \, d\mu(x) = \int_{-\infty}^{\infty} x^k \, d\nu(x), \quad k = 0, 1, 2, \ldots .$$

We have to show that $\mu = \nu$, and, as is well known, this will be the case if the Fourier–Stieltjes transform

$$f(\lambda) = \frac{1}{2} \int_{-\infty}^{\infty} e^{i\lambda x}(d\mu(x) - d\nu(x))$$

vanishes identically on $\mathbb{R}$.

It is now claimed that $f(\lambda)$ is infinitely differentiable on $\mathbb{R}$ and in fact belongs to a *quasianalytic class* thereon (see previous chapter). Observe that

$$\frac{1}{2} \int_{-\infty}^{\infty} x^{2k}(d\mu(x) + d\nu(x)) = S_{2k} < \infty;$$

therefore all the integrals

$$\frac{1}{2} \int_{-\infty}^{\infty} (ix)^k e^{i\lambda x}(d\mu(x) - d\nu(x))$$

are *absolutely convergent* (at least, first of all, for *even* $k \geqslant 0$ and hence for *all* $k \geqslant 0$), since the measures μ and ν are *positive* (*here* is where we *use* their positivity!). This means that $f(\lambda)$ is infinitely differentiable on $\mathbb{R}$, and that

$$f^{(k)}(\lambda) = \frac{1}{2} \int_{-\infty}^{\infty} (ix)^k e^{i\lambda x}(d\mu(x) - d\nu(x)).$$

For $\lambda \in \mathbb{R}$ we have

$$|f^{(k)}(\lambda)| \leqslant \frac{1}{2}\int_{-\infty}^{\infty} |x|^k (\mathrm{d}\mu(x) + \mathrm{d}\nu(x)),$$

a *finite quantity* independent of λ.

Denote $\sup_{\lambda\in\mathbb{R}} |f^{(n)}(\lambda)|$ by M_n. Then,

$$M_{2k} \leqslant \frac{1}{2}\int_{-\infty}^{\infty} x^{2k}(\mathrm{d}\mu(x) + \mathrm{d}\nu(x)) = S_{2k}.$$

Bringing in, as in §B.1 of the previous chapter, the *convex logarithmic regularization* $\{\underline{M}_n\}$ of the sequence $\{M_n\}$, we see that

$$\underline{M}_{2k} \leqslant S_{2k},$$

so

$$\sum_{n=0}^{\infty} \underline{M}_n^{-1/n} \geqslant \sum_{k=0}^{\infty} \underline{M}_{2k}^{-1/2k} \geqslant \sum_{k=0}^{\infty} S_{2k}^{-1/2k}.$$

The last sum on the right is, however, *infinite* by hypothesis. Therefore, by the *second* theorem of §B.2, Chapter IV, the class $\mathscr{C}_{\mathbb{R}}(\{M_n\})$ is *quasianalytic.*

However, $f(\lambda) \in \mathscr{C}_{\mathbb{R}}(\{M_n\})$ and

$$f^{(k)}(0) = \frac{1}{2}\int_{-\infty}^{\infty} (\mathrm{i}x)^k (\mathrm{d}\mu(x) - \mathrm{d}\nu(x)) = \mathrm{i}^k(S_k - S_k) = 0$$

for $k = 0, 1, 2, \ldots$ according to our initial supposition. Therefore $f(\lambda) \equiv 0$ on $\mathbb{R}$, as required, and we are done.

Scholium. If $\{S_k\}$ is a moment sequence, $\log S_{2k}$ is a *convex function of k.* This is an elementary consequence of Hölder's inequality. Taking, namely, a positive measure μ with

$$S_k = \int_{-\infty}^{\infty} x^k \mathrm{d}\mu(x),$$

we have, for $r, s \geqslant 0$ and $0 < \lambda < 1$,

$$\int_{-\infty}^{\infty} |x|^{\lambda r + (1-\lambda)s} \mathrm{d}\mu(x) \leqslant \left(\int_{-\infty}^{\infty} |x|^r \mathrm{d}\mu(x)\right)^{\lambda} \cdot \left(\int_{-\infty}^{\infty} |x|^s \mathrm{d}\mu(x)\right)^{1-\lambda},$$

so, if $r = 2k$, $s = 2l$, and $\lambda r + (1-\lambda)s$ is an *even integer*, $2m$ say, with (say) $2k < 2m < 2l$, we find that

$$S_{2m} \leqslant S_{2k}^{(2l-2m)/(2l-2k)} \cdot S_{2l}^{(2m-2k)/(2l-2k)}$$

the asserted convexity of $\log S_{2k}$.

An obvious adaptation of the work in Chapter IV, §§B.1 and C now shows that this convexity has the following consequence:

Theorem. *Let* $S(r) = \sup_{k\geq 0}(r^{2k}/S_{2k})$ *for* $r > 0$. *Then* $\sum_k S_{2k}^{-1/2k} = \infty$ *iff*

$$\int_0^\infty \frac{\log S(r)}{1+r^2}\,dr = \infty.$$

Corollary. *If* $\int_0^\infty (\log S(r)/(1+r^2))\,dr = \infty$ *with* $S(r)$ *as defined in the theorem, then the moment sequence* $\{S_k\}$ *is determinate.*

Proof. Combine the preceding theorem with that of Carleman.

2. A necessary condition

Theorem. *Let* $w(x) \geq 0$, *suppose that* $\int_{-\infty}^\infty |x|^k w(x)\,dx < \infty$ *for* $k = 0, 1, 2, \ldots$, *and put*

$$S_k = \int_{-\infty}^\infty x^k w(x)\,dx, \quad k = 0, 1, 2, \ldots.$$

If

$$\int_{-\infty}^\infty \frac{1}{1+x^2}\log\left(\frac{1}{w(x)}\right)dx < \infty,$$

the moment sequence $\{S_k\}$ *is indeterminate.*

Remark. Since $\int_{-\infty}^\infty (w(x)/(1+x^2))\,dx \leq \int_{-\infty}^\infty w(x)\,dx < \infty$, we have

$$\int_{-\infty}^\infty \frac{\log^+ w(x)}{1+x^2}\,dx < \infty$$

in any case, by the inequality between arithmetic and geometric means.

Proof of theorem. According to Problem 2 (at the end of §B, Chapter II!), if $\int_{-\infty}^\infty (\log w(x)/(1+x^2))\,dx > -\infty$, the infimum of $\int_{-\infty}^\infty |e^{-ix} - \sum_{\lambda\geq 0} A_\lambda e^{i\lambda x}|\,w(x)\,dx$, taken over *all finite sums* $\sum_{\lambda\geq 0} A_\lambda e^{i\lambda x}$, is *strictly positive.* By the Hahn–Banach theorem and the known form of linear functionals on $L_1(\mu)$ for σ-finite measures μ, we get a Borel function $\varphi(x)$, defined on

$$\{x:\ w(x) > 0\}$$

and *essentially bounded on that set*, with

$$\int_{-\infty}^\infty \varphi(x)\cdot e^{-ix} w(x)\,dx \neq 0$$

(hence φw is *not* almost everywhere zero!), whilst

$$(*)\qquad \int_{-\infty}^{\infty} \varphi(x)\cdot e^{i\lambda x} w(x)\,dx = 0, \quad \lambda \geqslant 0.$$

Under the conditions of this theorem, $w(x) > 0$ a.e., so $\varphi(x)$ is in fact defined a.e. on $\mathbb{R}$ and essentially bounded there, i.e., $\varphi \in L_\infty(\mathbb{R})$. Without loss of generality,

$$|\varphi(x)| \leqslant \tfrac{1}{2} \quad \text{a.e.,} \quad x \in \mathbb{R}.$$

Differentiating $(*)$ successively with respect to λ (which we can *do*, since the integrals $\int_{-\infty}^{\infty} |x|^k w(x)\,dx$ are all finite for $k \geqslant 0$) and looking at the resulting derivatives at $\lambda = 0$, we find that

$$\int_{-\infty}^{\infty} x^k \varphi(x) w(x)\,dx = 0, \quad k = 0, 1, 2, \ldots.$$

The functions $\Re\varphi(x)$ and $\Im\varphi(x)$ *can't both be zero* a.e.; say, wlog, that $\Re\varphi(x)$ isn't zero a.e. Then, from the preceding relation, we have

$$\int_{-\infty}^{\infty} x^k \Re\varphi(x) w(x)\,dx = 0, \quad k = 0, 1, 2, \ldots,$$

so that

$$S_k = \int_{-\infty}^{\infty} x^k (1 - \Re\varphi(x)) w(x)\,dx, \quad k = 0, 1, 2, \ldots,$$

as well as

$$S_k = \int_{-\infty}^{\infty} x^k w(x)\,dx, \quad k = 0, 1, 2, \ldots.$$

Here, $|\varphi(x)| \leqslant \frac{1}{2}$ a.e. but $\Re\varphi(x)$ is *not* a.e. equal to zero; therefore $(1 - \Re\varphi(x))w(x)\,dx$ is the *differential of a certain positive measure on* $\mathbb{R}$, *different from the positive measure with differential* $w(x)\,dx$, but having the *same moments*, S_k, as the latter. The moment sequence $\{S_k\}$ is thus indeterminate. Q.E.D.

Corollary. *Let* $T(r)$ *be* $\geqslant 1$ *for* $r \geqslant 0$, *and bounded near* 0. *Suppose that* $\log T(r)$ *is a convex function of* $\log r$, *and that*

$$\int_0^{\infty} \frac{r^k}{T(r)}\,dr \;<\; \infty \quad \textit{for} \quad k \geqslant 0.$$

The moment sequence

$$S_k \;=\; \int_{-\infty}^{\infty} \frac{x^k}{T(|x|)}\,dx, \quad k = 0, 1, 2, \ldots,$$

is determinate iff

$$\sum_{k=0}^{\infty} S_{2k}^{-1/2k} = \infty.$$

Remark. Here, of course, the S_k with *odd* k are all *zero*.

Proof of corollary. The *if* part follows by Carleman's theorem (preceding article).

To do *only if*, suppose that

$$\sum_k S_{2k}^{-1/2k} < \infty.$$

By the formula for the S_k, we have first of all

$$(\ast) \qquad S_{2k} = o(1) + 2\int_1^\infty \frac{x^{2k}}{(T(x)/x^2)} \cdot \frac{dx}{x^2}.$$

Put

$$M_n = \sup_{x \geqslant 0} \frac{x^n}{(T(x)/x^2)};$$

$\log M_n$ is then a *convex function* of n, and we proceed to apply to it and to $T(x)$ some of the work on convex logarithmic regularization from the preceding chapter.

From ($\ast$), we see that

$$S_{2k} \leqslant o(1) + 2M_{2k},$$

whence surely $\sum_k M_{2k}^{-1/2k} < \infty$. This certainly implies that $M_{2k}^{1/2k} \underset{k}{\longrightarrow} \infty$, so, since $\log M_n$ is a convex function of n, the proof of the *second* lemma in §B.1 of Chapter IV shows that the expression $M_n^{1/n}$ is *eventually increasing*. Therefore the convergence of $\sum_k M_{2k}^{-1/2k}$ implies that

$$\sum_n M_n^{-1/n} < \infty.$$

Taking, for $x > 0$,

$$P(x) = \sup_{n \geqslant 0} \frac{x^n}{M_n},$$

we will then get

$$(\dagger) \qquad \int_1^\infty \frac{\log P(x)}{x^2} dx < \infty$$

by the second lemma of §B.1, Chapter IV and the theorem of §C in that chapter.

Since $\log T(x)$ is a convex function of $\log x$, so is $\log(T(x)/x^2)$. The *second* lemma of §D, Chapter IV therefore shows that, for all sufficiently large x.

$$\log(T(x)/x^2) \leqslant \log P(x) + \log x,$$

in view of the relations between $T(x)/x^2$, M_n and $P(x)$. (The convex function $\log T(x)$ of $\log x$ must *eventually be increasing*, since all the integrals $\int_0^\infty (x^k/T(x))\,dx$, $k \geqslant 0$, converge!) Referring to (†), we see that

$$\int_1^\infty \frac{\log T(x)}{x^2}\,dx \;<\; \infty,$$

i.e.,

$$\int_{-\infty}^\infty \frac{1}{1+x^2}\log T(|x|)\,dx \;<\; \infty.$$

Indeterminateness of $\{S_k\}$ now follows by the above theorem.

Example. The sequence of moments

$$S_k(\alpha) = \int_{-\infty}^\infty x^k e^{-|x|^\alpha}\,dx$$

is *determinate* for $\alpha \geqslant 1$ and *indeterminate* if $0 < \alpha < 1$. (Note: In applying the above results it is better to work directly with $T(x) = e^{x^\alpha}$ for $\alpha \geqslant 1$ as well as for $0 < \alpha < 1$. Otherwise one should express $S_k(\alpha)$ in terms of the Γ-function and use Stirling's formula.)

Problem 8

The moment sequence

$$S_k = \int_1^\infty x^k e^{-x/\log x}\,dx$$

is *determinate*, but the Taylor series $\sum_0^\infty (S_n/n!)(i\lambda)^n$ of $\int_1^\infty e^{i\lambda x}e^{-x/\log x}\,dx$ *does not converge for any* $\lambda \neq 0$. (Hint: To see that the Taylor series can't converge for $\lambda \neq 0$, *estimate* S_n *from below* for large n. To do this, write

$$S_n = \int_1^\infty e^{-\varphi_n(x)}\,dx$$

with $\varphi_n(x) = x/\log x - n\log x$, and use *Laplace's method* to estimate the integral. To a first approximation, the *zero* x_0 of $\varphi_n'(x)$ has $x_0 \sim n\log n$, and this yields a good enough approximation to $\varphi_n''(x_0)$. To get a lower bound for $e^{-\varphi_n(x_0)}$, compute $\varphi_n(x)$ for $x = n\log n + n\log\log n$.)

D. M. Riesz' general criterion for indeterminacy

Let $\{S_k\}$ be a moment sequence. If we put

$$S(r) = \sup_{k\geqslant 0} \frac{r^{2k}}{S_{2k}} \quad \text{for} \quad r>0,$$

then, according to the corollary at the end of §C.1, $\{S_k\}$ is *determinate* when

$$\int_0^\infty \frac{\log S(r)}{1+r^2}\mathrm{d}r = \infty.$$

If, on the other hand, there is a density $w(x)\geqslant 0$ with

$$S_k = \int_{-\infty}^{\infty} x^k w(x)\,\mathrm{d}x, \quad k=0,1,2,\ldots,$$

$\{S_k\}$ is *indeterminate* provided that

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\log\left(\frac{1}{w(x)}\right)\mathrm{d}x < \infty,$$

as we have seen in §C.2.

Both conditions involve integrals of *the same form*, containing, however, different functions. This leads one to think that they might both be reflections of some general *necessary and sufficient condition* expressed in terms of the integral which is the subject of this book. As we shall now see, that turns out to be the case.

1. The criterion with Riesz' function $R(z)$

Given a moment sequence $\{S_k\}$, we take a positive measure μ on $\mathbb{R}$ having the moments S_k, and, for $z\in\mathbb{C}$, put

$$R(z) = \sup\left\{|P(z)|^2\colon P \text{ a } \textit{polynomial} \text{ with } \int_{-\infty}^{\infty}|P(x)|^2\,\mathrm{d}\mu(x)\leqslant 1\right\}.$$

It is only the sequence of S_k which is needed to get $R(z)$ and not the measure μ itself of which they are the moments; indeed, if

$$P(z) = \sum_{k=0}^{N} c_k z^k$$

with the $c_k\in\mathbb{C}$,

$$\int_{-\infty}^{\infty}|P(x)|^2\,\mathrm{d}\mu(x) = \sum_{i=0}^{N}\sum_{j=0}^{N} S_{i+j}c_i\bar{c}_j.$$

Thus, $R(z)$ (which may be infinite at some points) *depends just on the sequence* $\{S_k\}$; it turns out to govern that moment sequence's *determinacy.* Marcel Riesz worked with the *reciprocal* $\rho(z) = 1/R(z)$ instead of with $R(z)$, and the reader should note that, in literature on the moment problem, results are usually stated in terms of $\rho(z)$.

Theorem (M. Riesz). *Given a moment sequence* $\{S_k\}$ *and its associated function* $R(z)$, $\{S_k\}$ *is indeterminate if* $R(x) < \infty$ *on a non-denumerable subset of* $\mathbb{R}$. *Conversely, if* $\{S_k\}$ *is indeterminate,* $R(x) < \infty$ *everywhere on* $\mathbb{R}$ *and*

$$\int_{-\infty}^{\infty} \frac{\log^+ R(x)}{1+x^2} \mathrm{d}x \ < \ \infty.$$

Proof. For the first (and longest) part of the proof, let us suppose that $R(x) < \infty$ for all x belonging to some *non-denumerable* subset E of $\mathbb{R}$. We must establish indeterminacy of $\{S_k\}$.

Take any positive measure μ with $S_k = \int_{-\infty}^{\infty} x^k \,\mathrm{d}\mu(x)$, $k = 0, 1, 2, \ldots$, and let us first show that μ *cannot be supported on a finite set of points.* Suppose, on the contrary, that μ *were* supported on $\{x_1, x_2, \ldots, x_N\}$, say. Put $P_M(x) = M(x - x_1)(x - x_2)\ldots(x - x_N)$; then,

$$\int_{-\infty}^{\infty} |P_M(x)|^2 \,\mathrm{d}\mu(x) = 0,$$

but, if $x \neq x_1, x_2, \ldots$ or x_N, $P_M(x) \to \infty$ as $M \to \infty$, so $R(x) = \infty$. In that case, $R(x)$ could *not* be finite on the non-denumerable set E.

Having established that μ is not supported on a finite set, let us apply Schmidt's orthogonalization procedure to the sequence $1, x, x^2, \ldots$ and the measure μ, obtaining, one after the other, the real polynomials $p_n(x)$, $n \geqslant 0$, with $p_n(x)$ of degree n such that

$$\int_{-\infty}^{\infty} x^k p_n(x) \,\mathrm{d}\mu(x) = 0 \quad \text{for} \quad k = 0, \ldots, n-1,$$

when $n \geqslant 1$. Of course, the construction of the $p_n(x)$ really only depends on the S_k, and not on the particular positive measure μ of which they are the moments. Since no such μ can be supported on a finite set of points, the *orthogonalization process never stops*, and we obtain a non-zero p_n for *each n.* These orthogonal polynomials will be used presently.

Pick any x_0 with $R(x_0) < \infty$. We are going to construct a positive measure ν on $\mathbb{R}$ *having the moments* S_k, *but such that*

$$\nu(\{x_0\}) \geqslant 1/R(x_0) \quad (\textit{sic!}).$$

In order to obtain ν, let us take any large N, and try to find M points $x_1, x_2, \ldots, x_M$ *different from* x_0, with $M = N-1$ or N (it turns out that *either* possibility may occur) such that the *Gauss quadrature formula*

$$(*) \qquad \int_{-\infty}^{\infty} P(x)\,d\mu(x) = \sum_{k=0}^{M} \mu_k P(x_k)$$

holds for all (complex) polynomials P of degree $\leqslant M+N$; here, the μ_k are supposed to be certain coefficients independent of P.

Assume for the time being that we can obtain a quadrature formula $(*)$ for every large N, and consider the situation for *any given fixed* N. In the first place, *the coefficients* μ_k *are all* >0. To see this, pick any k, $0 \leqslant k \leqslant M$, and write

$$Q_k(x) = \prod_{\substack{i \neq k \\ 0 \leqslant i \leqslant M}} (x - x_i).$$

The polynomial Q_k is of degree M, so $P(x) = [Q_k(x)]^2$ is of degree $2M \leqslant M+N$, and we can apply $(*)$ to it, getting

$$(Q_k(x_k))^2 \mu_k = \int_{-\infty}^{\infty} [Q_k(x)]^2\,d\mu(x).$$

The right side is surely >0, for μ is not supported on any finite set of points. Therefore $\mu_k > 0$.

Using the polynomial

$$q(x) = \frac{Q_0(x)}{\sqrt{\mu_0}Q_0(x_0)}$$

of degree M, we have, by $(*)$ applied to $(q(x))^2$,

$$\int_{-\infty}^{\infty} (q(x))^2\,d\mu(x) = 1,$$

whilst $(q(x_0))^2 = 1/\mu_0$. Therefore, since $q(x)$ is a real polynomial, surely

$$R(x_0) \geqslant \frac{1}{\mu_0}$$

by definition of $R(z)$, i.e.,

$$\mu_0 \geqslant 1/R(x_0).$$

Let ν_N be the discrete positive measure supported on the set $x_0, x_1, \ldots, x_M$ defined by the relations

$$\nu_N(\{x_k\}) = \mu_k, \quad k = 0, 1, \ldots, M;$$

according to (*) we will then have

$$(\ast\ast)\qquad S_k = \int_{-\infty}^{\infty} x^k \,\mathrm{d}\mu(x) = \int_{-\infty}^{\infty} x^k \,\mathrm{d}\nu_N(x)$$

for $0 \leqslant k \leqslant M+N$, hence certainly for $k = 0, 1, 2, \ldots, 2N-1$. And, as we have just seen, $\nu_N(\{x_0\}) \geqslant 1/R(x_0)$.

Given any fixed k, the integrals

$$\int_{-\infty}^{\infty} (x^2+1)^k \,\mathrm{d}\nu_N(x)$$

can, according to what has just been shown, be expressed in obvious fashion in terms of the S_n as soon as $N > k$. *They are hence bounded*, and this means we can find *some sequence of* N's *tending to* ∞, and *finite positive measures* $\nu^{(k)}$ on $\mathbb{R}$, $k = 0, 1, 2, \ldots$, with, for *each* k,

$$(x^2+1)^k \,\mathrm{d}\nu_N(x) \longrightarrow \mathrm{d}\nu^{(k)}(x) \quad \mathrm{w}^*$$

as $N \to \infty$ through that sequence. (See Chapter III, §F.1). Let $l \geqslant 0$; then, since $(x^2+1)^{-l}$ is bounded and continuous on $\mathbb{R}$, the w^* convergence just mentioned certainly implies that

$$(x^2+1)^{k-l} \mathrm{d}\nu_N(x) \longrightarrow (x^2+1)^{-l} \mathrm{d}\nu^{(k)}(x),$$

so, if $l = 0, 1, \ldots, k$,

$$\mathrm{d}\nu^{(k-l)}(x) = (x^2+1)^{-l} \mathrm{d}\nu^{(k)}(x)$$

and thus

$$\mathrm{d}\nu^{(k)}(x) = (x^2+1)^{l} \mathrm{d}\nu^{(k-l)}(x).$$

Put $\nu^{(0)} = \nu$. By the preceding relation, $(x^2+1)^k \mathrm{d}\nu(x) = \mathrm{d}\nu^{(k)}(x)$ for $k = 0, 1, 2, \ldots$, so, since the measures $\nu^{(k)}$ are all finite, we have

$$\int_{-\infty}^{\infty} (x^2+1)^k \,\mathrm{d}\nu(x) \;<\; \infty$$

for $k \geqslant 0$. *It is now claimed that the* S_n *are the moments of the measure* ν.

Fixing any n, take a $k > n$. Then

$$\int_{-\infty}^{\infty} x^n \,\mathrm{d}\nu(x) = \int_{-\infty}^{\infty} \frac{x^n}{(x^2+1)^k} \mathrm{d}\nu^{(k)}(x).$$

By the above mentioned w^* convergence, the integral on the right is just the limit of

$$\int_{-\infty}^{\infty} \frac{x^n}{(x^2+1)^k} (x^2+1)^k \,\mathrm{d}\nu_N(x)$$

as N goes to ∞ through its special sequence of values. Each of the latter integrals, however, $= \int_{-\infty}^{\infty} x^n \mathrm{d}\nu_N(x)$ which, by ($\ast\atop\ast$), is just S_n as soon as $2N-1 \geqslant n$. Therefore

$$\int_{-\infty}^{\infty} x^n \mathrm{d}\nu(x) = S_n$$

for any $n \geqslant 0$, as claimed.

We have, moreover, $\nu_N(\{x_0\}) \geqslant 1/R(x_0)$ by our construction. Therefore, since the ν_N are positive measures, of which a subsequence tends w^* to ν, we *certainly have*

$$(\dagger) \qquad \nu(\{x_0\}) \geqslant 1/R(x_0).$$

In this way, we have obtained a *positive measure ν having the moments S_k and satisfying* ($\dagger$), where x_0 is *any one of the points in the non-denumerable set E* on which $R(x) < \infty$.

From this fact it follows, however, *that $\{S_k\}$ cannot be determinate.* We have, indeed, a positive measure ν with moments S_k satisfying ($\dagger$) for *each* $x_0 \in E$, and, *in the case of determinacy, those measures ν would have to be all the same.* In other words, there would be a *single measure ν* with $\nu(\{x_0\}) > 0$ *for a non-denumerable set of points* x_0. But that is nonsense. So, if $R(x)$ is finite on a non-denumerable set, we can establish indeterminacy using the quadrature formula ($\ast$). Everything turns, then, on the establishment of that formula, to which we will immediately direct our attention.

There is, however, one remark which should be made at this point, even though it has no bearing on the proof, namely, *that in* ($\dagger$) *we in fact have equality*,

$$\nu(\{x_0\}) = 1/R(x_0).$$

To see this, suppose that $\nu(\{x_0\}) > 1/R(x_0)$. We can get a polynomial P with

$$\int_{-\infty}^{\infty} |P(x)|^2 \,\mathrm{d}\mu(x) = \int_{-\infty}^{\infty} |P(x)|^2 \,\mathrm{d}\nu(x) = 1$$

but $|P(x_0)|^2$ *as close as we like to* $R(x_0)$. Then, however,

$$\int_{-\infty}^{\infty} |P(x)|^2 \,\mathrm{d}\nu(x) \;\geqslant\; |P(x_0)|^2 \nu(\{x_0\})$$

would be $>$ $|P(x_0)|^2/R(x_0)$, *and hence* > 1, a contradiction. We see that the function $R(x)$ *gives the solution to a certain extremal problem*:

$$\boxed{1/R(x_0) = \max\{\mu(\{x_0\}) \colon \mu \text{ a positive measure with the moments } S_k\}.}$$

We have now to prove the quadrature formula $(*)$. For this purpose we use the orthogonal polynomials $p_n(x)$ described at the beginning of the present demonstration; the idea goes back to Gauss. Take any $x_0 \in \mathbb{R}$ and any positive integer N. We can surely find *two real numbers* α and β, *not both zero*, such that

$$Q(x) = \alpha p_N(x) + \beta p_{N+1}(x)$$

vanishes at x_0. The polynomial Q is certainly not identically zero, and in fact it is of degree N or $N+1$, depending on whether $p_N(x_0) = 0$ or not. It is this uncertainty in the degree of Q which forces us to bring in the number M; we take $M = (\text{degree of } Q) - 1$; thus, $M = N - 1$ or N.

$Q(x)$, being of degree $M + 1$, vanishes at x_0 *and at M other points*; it is claimed that these points are *real and distinct*. This statement will be seen to rest entirely on the relation

$$(\S) \qquad \int_{-\infty}^{\infty} P(x)Q(x)\,d\mu(x) = 0,$$

valid for any polynomial P of degree $\leqslant N - 1$, which is an obvious consequence of the formula for Q and the orthogonality property of the polynomials $p_n(x)$.

Suppose, to begin with, that $Q(x)$ *has the real zeros* $x_0, \dots, x_r$ (with repetitions according to multiplicities, as in Chapter III), *and no others, and that* $r < M - 1$. Then, if

$$P(x) = (x - x_0)(x - x_1)\dots(x - x_r),$$

$P(x)Q(x)$ *will not change sign on* $\mathbb{R}$, so, for a suitable constant $c \neq 0$, $cP(x)Q(x) \geqslant 0$ on $\mathbb{R}$. Therefore $\int_{-\infty}^{\infty} cP(x)Q(x)\,d\mu(x) > 0$, for *$\mu$ is not supported on a finite set of points.* However $P(x)$ has degree $r + 1 < M \leqslant N - 1$, so $\int_{-\infty}^{\infty} cP(x)Q(x)\,d\mu(x) = 0$ by (§). We have reached a contradiction, showing that $Q(x)$ must have *at least M real zeros* (counting multiplicities), including x_0. However, $Q(x)$ is of degree $M + 1$. Therefore Q can have *at most one* non-real zero. The coefficients of Q are *real*, however, like those of p_N and p_{N+1}. Hence, *non-real zeros* of Q *must occur in pairs*, and Q *cannot have just one such zero.* This shows that *all the zeros* of Q *are real.*

The real zeros of Q are *distinct*. Suppose, for instance, that Q has at least a *double zero* at a_0; denote the remaining (real) zeros of Q by $a_1, \dots, a_{M-1}$; it is, of course, not excluded that some of them coincide with a_0. Put

$$P(x) = (x - a_1)(x - a_2)\dots(x - a_{M-1});$$

since $Q(x)$ has the factor $(x - a_0)^2$, $P(x)Q(x)$ *does not change sign* on $\mathbb{R}$. Thus, for a suitable constant c,

$$\int_{-\infty}^{\infty} cP(x)Q(x)\,\mathrm{d}\mu(x) > 0$$

as before, and this contradicts (§) since P is of degree $M-1<N$.

Denote now the *real and distinct zeros* of Q by $x_0, x_1, \ldots, x_M$, and let us complete the proof of the quadrature formula (*). Take any polynomial $P(x)$ of degree $\leqslant M+N$. Then, long division of $P(x)$ by $Q(x)$ yields

$$P(x) = D(x)Q(x) + R(x)$$

where, since degree of $Q = M+1$, the degree of R is $\leqslant M$ and the degree of D is $\leqslant N-1$. This last fact implies, by (§), that

$$\int_{-\infty}^{\infty} D(x)Q(x)\,\mathrm{d}\mu(x) = 0,$$

so

$$\int_{-\infty}^{\infty} P(x)\,\mathrm{d}\mu(x) = \int_{-\infty}^{\infty} R(x)\,\mathrm{d}\mu(x).$$

Now, since degree of $R \leqslant M$, Lagrange's interpolation formula gives us

$$R(x) = \sum_{k=0}^{M} \frac{R(x_k)}{Q'(x_k)(x-x_k)}\,Q(x),$$

i.e.

$$R(x) = \sum_{k=0}^{M} \frac{P(x_k)}{Q'(x_k)(x-x_k)}\,Q(x),$$

since $R(x_k) = P(x_k)$ at each zero x_k of Q. Therefore

$$\int_{-\infty}^{\infty} P(x)\,\mathrm{d}\mu(x) = \int_{-\infty}^{\infty} R(x)\,\mathrm{d}\mu(x)$$

$$= \int_{-\infty}^{\infty} \left(\sum_{k=0}^{M} \frac{P(x_k)}{Q'(x_k)(x-x_k)}\,Q(x) \right) \mathrm{d}\mu(x) = \sum_{k=0}^{M} P(x_k)\mu_k,$$

where

$$\mu_k = \int_{-\infty}^{\infty} \frac{Q(x)}{Q'(x_k)(x-x_k)}\,\mathrm{d}\mu(x), \quad k = 0, 1, \ldots, M.$$

Our quadrature formula (*) is thus established, and therewith, the *first part of the theorem.*

Proof of the *second part* of the theorem is quite a bit shorter. Here, we suppose that $\{S_k\}$ is an *indeterminate moment sequence*, and use that property to obtain information about $R(z)$.

We have, then, *two different* positive measures μ and ν with

$$S_k = \int_{-\infty}^{\infty} x^k \, d\mu(x) = \int_{-\infty}^{\infty} x^k \, d\nu(x), \quad k = 0, 1, 2, \ldots.$$

Denote by σ the positive measure $\frac{1}{2}(\mu + \nu)$ and by τ the *real signed measure* $\frac{1}{2}(\mu - \nu)$. Then also

$$S_k = \int_{-\infty}^{\infty} x^k \, d\sigma(x), \quad k = 0, 1, 2, \ldots,$$

so that, if $p(x)$ is *any polynomial*,

$$(\dagger\dagger) \qquad \int_{-\infty}^{\infty} |p(x)|^2 \, d\mu(x) = \int_{-\infty}^{\infty} |p(x)|^2 \, d\sigma(x).$$

For the signed measure τ,

$$\int_{-\infty}^{\infty} x^k \, d\tau(x) = 0, \quad k = 0, 1, 2, \ldots.$$

There is a *trick* based on this identity which, according to M. Riesz, goes back to Markov who used it in studying the moment problem around 1890. The same idea was used by Riesz himself and then, around 1950, by Pollard in the study of weighted polynomial approximation (see next chapter). Take any polynomial $\dot{P}$ and any $z_0 \notin \mathbb{R}$. Then

$$\frac{P(x) - P(z_0)}{x - z_0}$$

is also a polynomial in x, so, by the identity just written,

$$\int_{-\infty}^{\infty} \frac{P(x) - P(z_0)}{x - z_0} \, d\tau(x) = 0.$$

From this we have

$$(\S\S) \qquad P(z) \int_{-\infty}^{\infty} \frac{d\tau(t)}{t - z} = \int_{-\infty}^{\infty} \frac{P(t) \, d\tau(t)}{t - z}$$

whenever $z \notin \mathbb{R}$.

The function

$$F(z) = \int_{-\infty}^{\infty} \frac{d\tau(t)}{t - z}$$

is clearly analytic for $\Im z > 0$; moreover, *it cannot be identically zero there.*

Indeed,

$$\Im F(z) = \int_{-\infty}^{\infty} \frac{\Im z}{|z-t|^2} \mathrm{d}\tau(t),$$

τ being real, so by the remark at the end of §F.1, Chapter III,

$$\Im F(x+\mathrm{i}y)\,\mathrm{d}x \longrightarrow \pi \mathrm{d}\tau(x) \quad \mathrm{w}^*$$

for $y \to 0+$. Therefore $F(z) \equiv 0$ for $\Im z > 0$ would make $\tau = 0$, which is, however, contrary to the initial assumption that $\mu \neq \nu$.

Since $F(z) \not\equiv 0$ in $\{\Im z > 0\}$, we can use (§§) to get a formula for $P(z)$ in that half plane:

$$P(z) = \frac{1}{F(z)} \int_{-\infty}^{\infty} \frac{P(t)\mathrm{d}\tau(t)}{t-z}.$$

In particular, if $z = \xi + \mathrm{i}$ with ξ real,

$$\begin{aligned} |P(\xi+\mathrm{i})| &\leqslant \frac{1}{|F(\xi+\mathrm{i})|} \int_{-\infty}^{\infty} |P(t)|\,|\mathrm{d}\tau(t)| \\ &\leqslant \frac{1}{|F(\xi+\mathrm{i})|} \int_{-\infty}^{\infty} |P(t)|\mathrm{d}\sigma(t), \end{aligned}$$

since $|\mathrm{d}\tau(t)| \leqslant \mathrm{d}\sigma(t)$. Let now $P = p^2$, where p is any polynomial. By the preceding relation and (††), $|p(\xi+\mathrm{i})|^2 \leqslant (1/|F(\xi+\mathrm{i})|)\int_{-\infty}^{\infty}|p(t)|^2\mathrm{d}\mu(t)$, so, by definition of $R(z)$,

$$R(\xi+\mathrm{i}) \leqslant \frac{1}{|F(\xi+\mathrm{i})|}.$$

The analytic function $F(z)$ is clearly *bounded* in $\{\Im z \geqslant 1\}$ and continuous up to the line $\Im z = 1$. Since, as we have seen, $F(z) \not\equiv 0$ there, we have, applying the first theorem of §G.2, Chapter III to the half plane $\{\Im z \geqslant 1\}$,

$$\int_{-\infty}^{\infty} \frac{\log^- |F(\xi+\mathrm{i})|}{\xi^2+1} \mathrm{d}\xi < \infty.$$

Combined with the previous inequality, this yields

$$(\ddagger) \qquad \int_{-\infty}^{\infty} \frac{\log^+ R(\xi+\mathrm{i})}{\xi^2+1} \mathrm{d}\xi < \infty.$$

Using this result, we can now estimate $R(x)$ *on the real axis*.

Let $p(z)$ be any polynomial with

$$\int_{-\infty}^{\infty} |p(t)|^2 \mathrm{d}\mu(t) \leqslant 1;$$

then, *by definition* (!),

$$|p(\xi + \mathrm{i})|^2 \leqslant R(\xi + \mathrm{i}).$$

On the other hand, by the theorem of §E, Chapter III, applied to the half plane $\Im z \leqslant 1$,

$$\log|p(x)| \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\log^+|p(\xi+\mathrm{i})|}{(x-\xi)^2+1}\,\mathrm{d}\xi$$

for $x \in \mathbb{R}$. (Note that $p(z)$ is an *entire function of exponential type zero*! The reader who does not wish to resort to the result from Chapter III may of course easily verify the inequality for *polynomials* $p(z)$ directly.) These two relations yield

$$2\log|p(x)| \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\log^+ R(\xi+\mathrm{i})}{(x-\xi)^2+1}\,\mathrm{d}\xi, \quad x \in \mathbb{R},$$

whence, taking the supremum of $2\log|p(x)|$ for such polynomials p,

$$\log R(x) \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\log^+ R(\xi+\mathrm{i})}{(x-\xi)^2+1}\,\mathrm{d}\xi, \quad x \in \mathbb{R},$$

(by definition again!). We can, of course, replace $\log R(x)$ by $\log^+ R(x)$ in this inequality, since the right-hand side is $\geqslant 0$.

We see from (‡) that the integral on the right in the relation just obtained is $< \infty$ for each $x \in \mathbb{R}$. That is, $R(x) < \infty$ *for every real* x *if our moment sequence is indeterminate*, this is part of what we wanted to prove. Again,

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\log^+ R(x)}{1+x^2}\,\mathrm{d}x \;&\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\log^+ R(\xi+\mathrm{i})}{(x-\xi)^2+1}\cdot\frac{1}{x^2+1}\,\mathrm{d}\xi\,\mathrm{d}x \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \log^+ R(\xi+\mathrm{i}) \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{((\xi-x)^2+1)(x^2+1)}\,\mathrm{d}\xi \\ &= \int_{-\infty}^{\infty} \frac{2\log^+ R(\xi+\mathrm{i})}{\xi^2+4}\,\mathrm{d}\xi,\end{aligned}$$

and the last integral is finite by (‡). This shows that

$$\int_{-\infty}^{\infty} \frac{\log^+ R(x)}{x^2+1}\,\mathrm{d}x \;<\; \infty,$$

and the *second part* of our theorem is completely proved.

Corollary. *The moment sequence* $\{S_k\}$ *is determinate iff, for the function* $R(z)$ *associated with it,*

$$\int_{-\infty}^{\infty} \frac{\log^+ R(x)}{1+x^2}\,\mathrm{d}x \;=\; \infty.$$

Remark. The corollary does not give the full story. What the theorem really says is that there is an *alternative* for the function $R(x)$: *either* $R(x) = \infty$ *everywhere on* $\mathbb{R}$ *save, perhaps, on a countable set of points, or else* $R(x) < \infty$ *everywhere on* $\mathbb{R}$ *and*

$$\int_{-\infty}^{\infty} \frac{\log^+ R(x)}{x^2+1} \mathrm{d}x \;<\; \infty.$$

Scholium. Take the *normalized orthogonal polynomials* $P_n(x)$ corresponding to a positive measure μ with the moments S_k. Like the $p_n(x)$ used in the first part of the proof of the above theorem, the P_n are gotten by applying Schmidt's orthogonalization procedure (with the measure μ) to the successive powers $1, x, x^2, x^3, \ldots$; here, however, one also imposes the supplementary conditions

$$\int_{-\infty}^{\infty} [P_n(x)]^2 \,\mathrm{d}\mu(x) = 1,$$

making each $P_n(x)$ a *constant multiple* of $p_n(x)$. One of course needs only the S_k to compute the successive P_n.

It is easy to express $R(x)$ in terms of these P_n; we have, in fact,

$$\boxed{R(x) \;=\; \sum_{n=0}^{\infty} (P_n(x))^2.}$$

Proof of this relation may be left to the reader – first work out

$$R_N(x) = \max\Big\{|p(x)|^2\colon\ p \text{ a polynomial of degree } \leqslant N \text{ with } \int_{-\infty}^{\infty} |p(t)|^2 \,\mathrm{d}\mu(t) = 1\Big\}$$

by writing $p(t) = \sum_{n=0}^{N} \alpha_n P_n(t)$ and using Lagrange's method; then make $N \to \infty$. The boxed formula seems at first sight to break down if any μ with the moments S_k is supported on a finite number of points, say M. In that case, the formula can, however, be saved by taking $P_n(x) = \infty$ for $n \geqslant M$ and x lying outside the support of μ. This makes sense, because the *only polynomial of degree* $n \geqslant M$ orthogonal to the powers $1, x, \ldots, x^{M-1}$ with respect to a measure supported on M points is ***zero****, hence can't be normalized.* The vain attempt to normalize it gives us the form 0/0, which we are of course at liberty to take as ∞ outside the support of that measure.

2. Derivation of the results in §C from the above one

Let us first deduce Carleman's theorem in §C.1 from that of M. Riesz. According to the *second* theorem of §C.1 (in the scholium of that article), Carleman's theorem is *equivalent* to the following statement: *the moment sequence* $\{S_k\}$ *is determinate provided that*

$$(*) \qquad \int_0^\infty \frac{\log S(r)}{1+r^2}\,\mathrm{d}r \;=\; \infty,$$

where

$$S(r) \;=\; \sup_{k\geqslant 0} \frac{r^{2k}}{S_{2k}} \quad \text{for } r>0.$$

To verify this, observe that, if the S_k are the moments of a positive measure μ, the polynomials

$$q_k(x) \;=\; x^k/\sqrt{S_{2k}}$$

satisfy

$$\int_{-\infty}^{\infty} (q_k(x))^2\,\mathrm{d}\mu(x) = 1,$$

so surely $R(x) \geqslant (q_k(x))^2$ for each k, by definition of $R(z)$. Therefore

$$R(x) \geqslant S(|x|).$$

Also, $S(|x|) \;\geqslant\; 1/S_0 \;>\; 0$, so $\log S(|x|)$ is *bounded below*. It is thus clear that $(*)$ implies

$$\int_{-\infty}^{\infty} \frac{\log^+ R(x)}{1+x^2}\,\mathrm{d}x \;=\; \infty.$$

The moment sequence $\{S_k\}$ is therefore *determinate* by the corollary to Riesz' theorem.

Consider now the theorem of §C.2. We are given a positive integrable function $w(x)$ with

$$\int_{-\infty}^{\infty} \frac{\log w(x)}{1+x^2}\,\mathrm{d}x \;>\; -\infty,$$

and want to prove that the moment sequence $S_k = \int_{-\infty}^{\infty} x^k w(x)\,\mathrm{d}x$ is *indeterminate* using Riesz' theorem.

Observe that the integrability of $w(x)$ makes $\int_{-\infty}^{\infty}(w(x)/(1+x^2))\mathrm{d}x < \infty$, so, surely, $\int_{-\infty}^{\infty}(\log^+ w(x)/(1+x^2))\mathrm{d}x < \infty$. Our other assumption on w therefore implies that

$$(\ast\!\ast) \qquad \int_{-\infty}^{\infty} \frac{\log^- w(x)}{1+x^2}\,\mathrm{d}x \;<\; \infty.$$

Take any polynomial p with $\int_{-\infty}^{\infty}|p(x)|^2 w(x)\mathrm{d}x \leqslant 1$. Then, surely,

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{|p(x)|^2 w(x)}{(x-\xi)^2+1}\mathrm{d}x \;\leqslant\; \frac{1}{\pi}$$

for any real ξ, whence, by the inequality between arithmetic and geometric means,

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{2\log|p(x)| + \log w(x)}{(x-\xi)^2+1}\mathrm{d}x \;\leqslant\; \log\frac{1}{\pi} \;\leqslant\; 0,$$

i.e.

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{2\log|p(x)|}{(x-\xi)^2+1}\mathrm{d}x \;\leqslant\; -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log w(x)}{(x-\xi)^2+1}\mathrm{d}x.$$

The left-hand integral is, however, $\geqslant 2\log|p(\xi+\mathrm{i})|$ by the second theorem of §G.2, Chapter III. ($p(z)$ is entire, of exponential type *zero*. For polynomials, the fact in question may also be easily verified directly.) We therefore have

$$\log|p(\xi+\mathrm{i})|^2 \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log^- w(x)}{(x-\xi)^2+1}\mathrm{d}x,$$

and, taking the supremum over such polynomials p,

$$\log R(\xi+\mathrm{i}) \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log^- w(x)}{(x-\xi)^2+1}\mathrm{d}x.$$

Here, one may, of course, replace $\log R(\xi+\mathrm{i})$ by $\log^+ R(\xi+\mathrm{i})$ on the left.

For $x_0 \in \mathbb{R}$,

$$\log R(x_0) \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log^+ R(\xi+\mathrm{i})}{(x_0-\xi)^2+1}\mathrm{d}\xi,$$

just as in the proof of the second part of Riesz' theorem. Substituting in the previous inequality on the right and changing the order of integration, we get finally

$$\log R(x_0) \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{2\log^- w(x)}{(x_0-x)^2+4}\mathrm{d}x, \quad x_0 \in \mathbb{R}.$$

But the integral on the right is *finite* by (*). Therefore $R(x_0) < \infty$ for each $x_0 \in \mathbb{R}$, so the moment sequence $\{S_k\}$ is *indeterminate* by Riesz' theorem. We are done.

VI

Weighted approximation on the real line

In the study of weighted approximation on $\mathbb{R}$, we start with a function $W(x) \geqslant 1$, henceforth called a *weight*, defined for $-\infty < x < \infty$. We usually suppose that $W(x) \to \infty$ for $x \to \pm\infty$, but *do not always assume W continuous*, and frequently allow it to be *infinite* on some large sets.

Given a weight W, we take the space $\mathscr{C}_W(\mathbb{R})$ consisting of continuous functions $\varphi(x)$ defined on $\mathbb{R}$ with $\varphi(x)/W(x) \longrightarrow 0$ for $x \to \pm\infty$, and write $\|\varphi\|_W = \sup_x |\varphi(x)/W(x)|$ for $\varphi \in \mathscr{C}_W(\mathbb{R})$. Being presented with a certain subset $\mathscr{E}$ of $\mathscr{C}_W(\mathbb{R})$, we then ask whether $\mathscr{E}$ is *dense* in $\mathscr{C}_W(\mathbb{R})$ in the norm $\| \ \|_W$ – this is the so-called *weighted approximation problem.*

The following preliminary observation will be used continually.

Lemma. *$\mathscr{E}$ is $\| \ \|_W$-dense in $\mathscr{C}_W(\mathbb{R})$ iff, for some $c \notin \mathbb{R}$, all the functions*

$$\frac{1}{(x-c)^n W(x)} \quad \textit{and} \quad \frac{1}{(x-\bar{c})^n W(x)}, \quad n = 1, 2, 3, \ldots,$$

can be approximated uniformly on $\mathbb{R}$ by functions of the form $f(x)/W(x)$ with $f \in \mathscr{E}$.

Proof. *Only if* is manifest. For *if*, take any function $\varphi \in \mathscr{C}_W(\mathbb{R})$ and *first* construct a continuous function ψ of *compact support* such that $|(\varphi(x) - \psi(x))/W(x)| < \varepsilon/3$ on $\mathbb{R}$. We can, for instance, put

$$\psi(x) = \begin{cases} \varphi(x), & |x| \leqslant A, \\ \dfrac{2A - |x|}{A}\varphi(x), & A \leqslant |x| \leqslant 2A, \\ 0, & |x| \geqslant 2A; \end{cases}$$

the desired relation will then hold if A is taken *large enough*, since $\varphi(x)/W(x) \longrightarrow 0$ for $x \to \pm\infty$.

By the appropriate version of Weierstrass' theorem, linear combinations of the functions $(x-c)^{-n}$, $(x-\bar{c})^{-n}$, $n = 1,2,3,\ldots$, can now be used to approximation $\psi(x)$ *uniformly* on $\mathbb{R}$, so we can get such a linear combination $\sigma(x)$ with $|\psi(x)-\sigma(x)| < \varepsilon/3$ on $\mathbb{R}$, whence (since $W(x) \geqslant 1$!), $|(\psi(x)-\sigma(x))/W(x)| < \varepsilon/3$ there.

If, now, we can find an $f \in \mathscr{E}$ with $|\sigma(x)/W(x) - f(x)/W(x)| < \varepsilon/3$ for $x \in \mathbb{R}$, we'll have $\|\sigma - f\|_W \leqslant \varepsilon/3$. Then, since $\|\varphi - \psi\|_W \leqslant \varepsilon/3$ and $\|\psi - \sigma\|_W \leqslant \varepsilon/3$, we obtain $\|\varphi - f\|_W \leqslant \varepsilon$. This establishes the *if* part of the lemma.

In the first weighted approximation problem we consider, the so-called *Bernstein approximation problem*, $\mathscr{E}$ consists of *polynomials*. *Then, of course, we must impose on the weight W the supplementary requirement that*

$$\frac{x^n}{W(x)} \longrightarrow 0 \quad \textit{as } x \to \pm\infty$$

for all $n \geqslant 0$. In another problem, whose treatment is similar to that of Bernstein's, $\mathscr{E}$ consists of all finite linear combinations of the exponentials $e^{i\lambda x}$ with $-a \leqslant \lambda \leqslant a$ for some given positive a. If a weight W satisfies the supplementary condition, it is natural to compare the solution of Bernstein's problem with those of the latter one for different values of a. One may also study approximation using *weighted L_p norms* instead of the weighted uniform norm $\| \quad \|_W$.

These questions are taken up in the present chapter. Some of the methods applied in studying them resemble closely the one used to prove the *second* part of Riesz' theorem (previous chapter, §D.1). There is indeed a relation between the material of this chapter and the determinacy problem discussed in the preceding one, and results obtained in the study of either subject may sometimes be applied to the other.

I know of no book entirely devoted to the matters mentioned above; the one by Nachbin has very little concerning them and is really about something else. One who wishes to go into the subject should first read Mergelian's *Uspekhi* paper and then Akhiezer's; both have been translated into English. There is material in the complements at the end of the second edition of Akhiezer's book on approximation theory, and also in de Branges' book. It is worthwhile to study S. Bernstein's original investigations on weighted polynomial approximation; most of his papers on this are in volume two of his collected works. Some of the results given near the end of the present chapter are from a paper of mine published around 1964.

A. Mergelian's treatment of weighted polynomial approximation

Let W be a weight such that

$$\frac{x^n}{W(x)} \longrightarrow 0 \quad \text{as } x \to \pm\infty \quad \text{for } n = 1, 2, 3, \ldots;$$

the solution of Bernstein's approximation problem for W turns out to be governed by the quantity

$$\Omega(z) = \sup\left\{ |P(z)|\colon\ P \text{ a polynomial and } \left|\frac{P(t)}{(t-\mathrm{i})W(t)}\right| \leqslant 1 \quad \text{on } \mathbb{R} \right\}$$

introduced by Mergelian.

Note the similarity between the definition of this quantity and that of the Riesz function $R(z)$ given in §D.1 of the preceding chapter.

Note especially that the condition $|P(t)/(t-\mathrm{i})W(t)| \leqslant 1$ *and not the seemingly more natural one* $|P(t)/W(t)| \leqslant 1$ *is used in defining* $\Omega(z)$. About this, more later.

1. Criterion in terms of finiteness of $\Omega(z)$

Theorem (Mergelian). *The polynomials are* $\|\ \|_W$*-dense in* $\mathscr{C}_W(\mathbb{R})$ *iff* $\Omega(z_0) = \infty$ *for one non-real* z_0, *and, if this happens, then* $\Omega(z) = \infty$ *for all non-real* z.

Proof (Mergelian).

Only if: Suppose the polynomials are dense in $\mathscr{C}_W(\mathbb{R})$. Then, given any $z_0 \notin \mathbb{R}$, we can find polynomials $Q_n(t)$ such that the quantities

$$\delta_n = \sup_{t\in\mathbb{R}} \left| \frac{(t-z_0)^{-1} - Q_n(t)}{W(t)} \right|$$

tend to *zero* as $n \to \infty$. Put

$$P_n(t) = \frac{1 - (t-z_0)Q_n(t)}{\delta_n};$$

$P_n(t)$ is for each n a polynomial, $P_n(z_0) = 1/\delta_n \xrightarrow[n]{} \infty$, and $|P_n(t)/(t-z_0)W(t)| \leqslant 1$ for $t \in \mathbb{R}$. There is obviously a number $K(z_0) > 0$ *depending only on* z_0 such that

$$\frac{1}{K(z_0)} \leqslant \left|\frac{t-z_0}{t-\mathrm{i}}\right| \leqslant K(z_0) \quad \text{for } t \in \mathbb{R}.$$

Therefore $|P_n(t)/(t-\mathrm{i})W(t)| \leqslant K(z_0)$ on $\mathbb{R}$, so, since $P_n(z_0) \xrightarrow[n]{} \infty$, we see that $\Omega(z_0) = \infty$, establishing the *only if* part.

If: When $z_0 \notin \mathbb{R}$, it is convenient to work with the quantity

$$M(z_0) = \sup\left\{ |P(z_0)|:\ P \text{ a polynomial and } \left|\frac{P(t)}{(t-z_0)W(t)}\right| \leqslant 1 \quad \text{on } \mathbb{R} \right\}.$$

One sees by using the number $K(z_0)$ brought in during the above argument that $\Omega(z_0) = \infty$ iff $M(z_0) = \infty$, as long as $z_0 \notin \mathbb{R}$.

One advantage of introducing $M(z)$ lies in its *continuity property*:

$$(*) \qquad \left|\frac{1}{M(z)} - \frac{1}{M(\zeta)}\right| \leqslant \frac{|\zeta - z|}{|\Im z|\,|\Im \zeta|}.$$

To verify this, take any polynomial P with $|P(t)/(t-\zeta)W(t)| \leqslant 1$ on $\mathbb{R}$ and $|P(\zeta)|$ *close to* $M(\zeta)$. For the polynomial in t

$$Q(t) \;=\; \frac{P(t)-P(\zeta)}{t-\zeta}(t-z) + P(\zeta)$$

we have $Q(z) = P(\zeta)$, whilst, for $t \in \mathbb{R}$, $|Q(t)/(t-z)W(t)| \leqslant$ $|P(t)/(t-\zeta)W(t)| + |P(\zeta)|\,|(z-\zeta)/(t-z)(t-\zeta)W(t)|$ and this is $\leqslant 1 + |P(\zeta)|\,|z-\zeta|/|\Im z|\,|\Im\zeta|$, since, as we are always assuming, $W(x) \geqslant 1$. Put now

$$R(t) \;=\; \left(1 + |P(\zeta)|\frac{|z-\zeta|}{|\Im z|\,|\Im\zeta|}\right)^{-1} \cdot Q(t);$$

R is a polynomial in t, and $|R(t)/(t-z)W(t)| \leqslant 1$ on $\mathbb{R}$. Therefore $|R(z)| \leqslant M(z)$; however,

$$|R(z)| \;=\; \frac{|P(\zeta)|}{1 + |P(\zeta)|\dfrac{|z-\zeta|}{|\Im z|\,|\Im\zeta|}}.$$

Thence,

$$\frac{1}{M(z)} \leqslant \frac{1}{|R(z)|} = \frac{1}{|P(\zeta)|} + \frac{|z-\zeta|}{|\Im z|\,|\Im\zeta|},$$

so, since we can have $|P(\zeta)|$ as close as we like to $M(\zeta)$, we get

$$\frac{1}{M(z)} \leqslant \frac{1}{M(\zeta)} + \frac{|z-\zeta|}{|\Im z|\,|\Im\zeta|}$$

This relation and the similar one obtained by reversing the rôles of z and ζ in the argument just made give us $(*)$.

Armed with $(*)$, we proceed with the *if* part of our proof. Suppose then that $z_0 \notin \mathbb{R}$ and that $\Omega(z_0) = \infty$; this means that $M(z_0) = \infty$, so we can find polynomials $P_n(t)$ with $|P_n(t)/(t-z_0)W(t)| \leqslant 1$ on $\mathbb{R}$ whilst $|P_n(z_0)| \xrightarrow[n]{} \infty$. For the polynomials

$$Q_n(t) \;=\; \frac{P_n(z_0) - P_n(t)}{(t-z_0)P_n(z_0)},$$

we will have, for $t \in \mathbb{R}$,

$$\sup_{t\in\mathbb{R}} \left| \frac{1}{(t-z_0)W(t)} - \frac{Q_n(t)}{W(t)} \right| \;\leqslant\; \frac{1}{|P_n(z_0)|},$$

a quantity tending to 0 as $n \to \infty$.

Therefore the function $(t-z_0)^{-1}$ can be approximated as closely as we like, in the norm $\|\ \|_W$, by polynomials.

It is now claimed that

$$(\ast\ast) \qquad \int_0^{2\pi} \Omega(z_0 + \rho e^{i\vartheta})\,d\vartheta = \infty$$

for each $\rho > 0$. Since $\Omega(z_0) = \infty$, there are polynomials $q_n(t)$ with $|q_n(t)/(t-i)W(t)| \leqslant 1$ on $\mathbb{R}$ (N.B. Here it *is* $t - i$ in the denominator and *not* $t - z_0$!), and $|q_n(z_0)| \xrightarrow[n]{} \infty$. For each n, $|q_n(z)| \leqslant \Omega(z)$ *by definition*, therefore

$$\int_0^{2\pi} \Omega(z_0 + \rho e^{i\vartheta})\,d\vartheta \;\geqslant\; \int_0^{2\pi} |q_n(z_0 + \rho e^{i\vartheta})|\,d\vartheta \;\geqslant\; 2\pi |q_n(z_0)|.$$

Since the quantity on the right $\to \infty$ with n, we have $(\ast\ast)$.

If $0 < \rho < |\Im z_0|$, there is a z_ρ, $|z_\rho - z_0| = \rho$, with $\Omega(z_\rho) = \infty$. Indeed, $(\ast\ast)$ implies the existence of a sequence of points ζ_k with $|\zeta_k - z_0| = \rho$ and $\Omega(\zeta_k) \xrightarrow[k]{} \infty$. Suppose, wlog, that $\zeta_k \xrightarrow[k]{} z_\rho$. Comparison of the definitions of $\Omega(\zeta)$ and $M(\zeta)$ shows immediately that

$$\Omega(\zeta) \;\leqslant\; \sup_{t\in\mathbb{R}} \left| \frac{t-i}{t-\zeta} \right| \cdot M(\zeta).$$

Here, the supremum is *clearly bounded above* for $|\zeta - z_0| = \rho$ since $|\Im z_0| - \rho > 0$; therefore our choice of the sequence $\{\zeta_k\}$ makes $M(\zeta_k) \xrightarrow[k]{} \infty$. Because $\zeta_k \xrightarrow[k]{} z_\rho$, *we then get* $M(z_\rho) = \infty$ *by* $(*)$, since $|\Im\zeta| \geqslant |\Im z_0| - \rho > 0$ on the circle $|\zeta - z_0| = \rho$. Thus, $\Omega(z_\rho) = \infty$. (A mistake I made here while lecturing was pointed out to me by Dr Raymond Couture.)

We thus have points z_ρ for which $\Omega(z_\rho) = \infty$ with $|z_\rho - z_0| = \rho$, when $\rho > 0$ is sufficiently small.

For each such z_ρ, $1/(t - z_\rho)$ can be approximated in $\|\ \|_W$-norm by polynomials in t; this is shown by the argument used above for $1/(t-z_0)$. We can therefore obtain a sequence of points $z_n \neq z_0$ tending to z_0, such that each of the functions

$$\frac{1}{(t-z_n)W(t)}$$

is the *uniform limit*, on $\mathbb{R}$, of polynomials in t divided by $W(t)$. This fact makes it possible for us to show (by taking limits of difference quotients of successively higher order) that *each* of the expressions

$$\frac{1}{(t-z_0)^m W(t)}, \quad m = 1, 2, 3, \ldots,$$

can be uniformly approximated on $\mathbb{R}$ by polynomials in t divided by $W(t)$ $(\geqslant 1)$.

Now we have $\Omega(\bar{z}_0) = \Omega(z_0)$, for, if P is a polynomial with $|P(t)/(t-\mathrm{i})W(t)| \leqslant 1$ on $\mathbb{R}$, the polynomial $P^*(t)$ whose coefficients are the *complex conjugates* of the corresponding ones of $P(t)$ also satisfies $|P^*(t)/(t-\mathrm{i})W(t)| \leqslant 1$, $t \in \mathbb{R}$. Therefore in the present case $\Omega(\bar{z}_0) = \infty$, so, by the above discussion, each of the functions $1/(t-\bar{z}_0)^m W(t)$, $m = 1, 2, 3, \ldots$, can be uniformly approximated on $\mathbb{R}$ by polynomials in t divided by $W(t)$. As we just saw, the same is true for the functions $1/(t-z_0)^m W(t)$. *According to the general lemma given at the very beginning of this chapter, polynomials must hence be* $\| \ \|_W$*-dense in* $\mathscr{C}_W(\mathbb{R})$. The *if* part of our theorem is thus established.

We are done.

2. A computation

In the next article and later on, as well, we will need a *formula* for

$$\sup_{t \in \mathbb{R}} \left| \frac{t-\mathrm{i}}{t-z} \right|.$$

Lemma. *When* $\Im z > 0$,

$$\boxed{\sup_{t \in \mathbb{R}} \left| \frac{t-\mathrm{i}}{t-z} \right| = \frac{|z+\mathrm{i}| + |z-\mathrm{i}|}{2\Im z}}$$

Proof. $|(t-\mathrm{i})/(t-z)| = |1-(z-\mathrm{i})/(t-\mathrm{i})|^{-1}$. In order to simplify the writing, put $z - \mathrm{i} = \zeta$; then we have to calculate $\inf_{t \in \mathbb{R}} |1 - \zeta/(t-\mathrm{i})|$. The linear fractional transformation $t \to 1/(t-\mathrm{i})$ takes the real axis into the *circle* having the segment $[0, \mathrm{i}]$ as *diameter*:

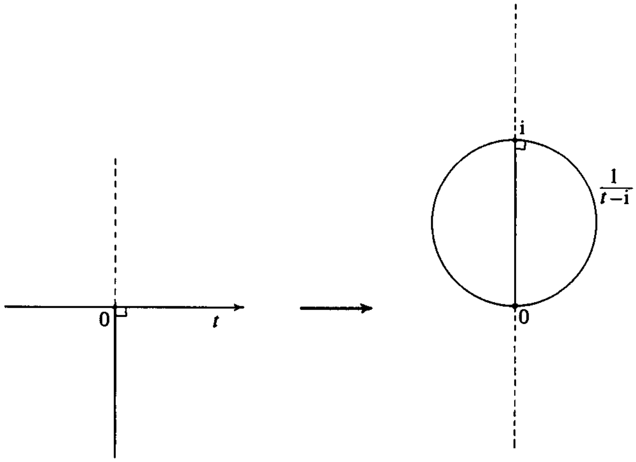


Figure 32

Therefore, as t ranges over the real axis, $\zeta/(t-\mathrm{i})$ ranges over the circle γ_ζ with segment $[0, \zeta\mathrm{i}]$ as diameter:

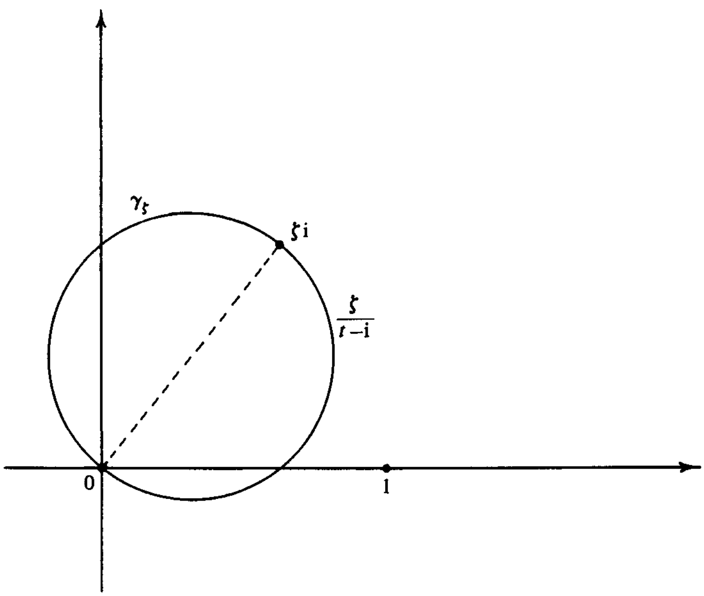


Figure 33

Since $\Im z > 0$, $\Im \zeta > -1$, so $\Re(\zeta i) < 1$. Therefore the point 1 must lie *outside* the circle γ_ζ:

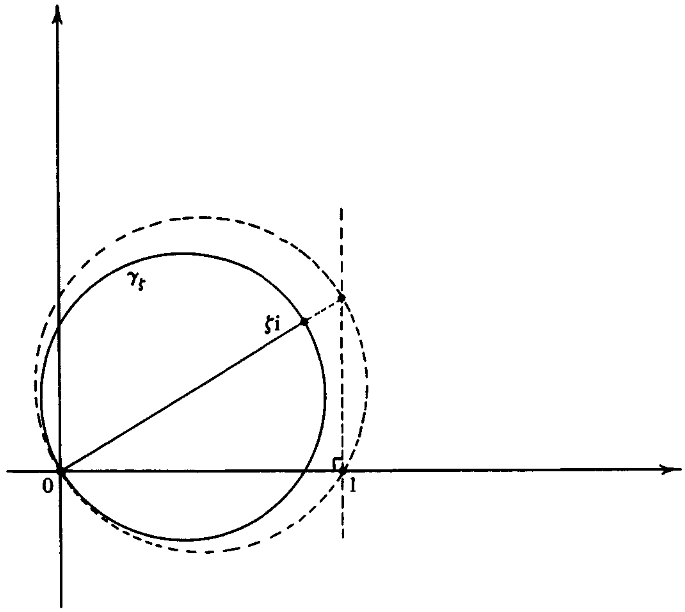


Figure 34

Our quantity $\inf_{t\in\mathbb{R}}|1 - \zeta/(t - i)|$, which is simply the distance from 1 to γ_ζ, can thus be read off from the diagram:

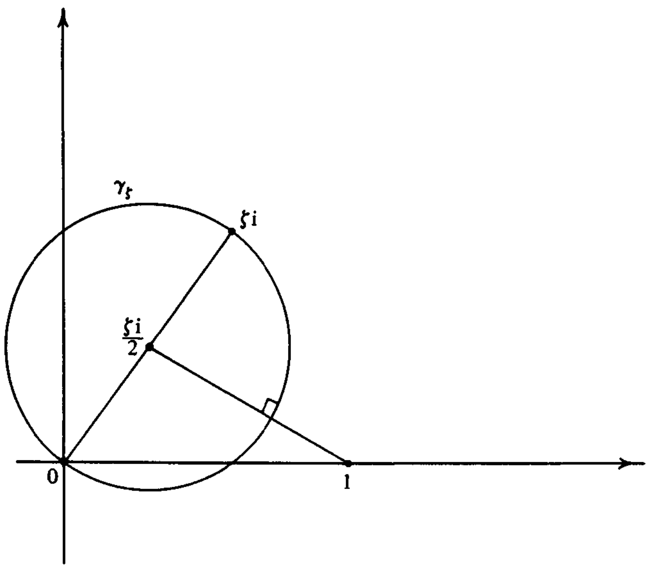


Figure 35

We see that

$$\inf_{t\in\mathbb{R}}\left|1-\frac{\zeta}{t-\mathrm{i}}\right| = \left|1-\frac{\zeta\mathrm{i}}{2}\right| - \text{radius of } \gamma_\zeta$$

$$= \left|1-\frac{\zeta\mathrm{i}}{2}\right| - \left|\frac{\zeta\mathrm{i}}{2}\right| = \left|\frac{1}{2}-\frac{\mathrm{i}z}{2}\right| - \left|\frac{1}{2}+\frac{\mathrm{i}z}{2}\right|.$$

Finally,

$$\sup_{t\in\mathbb{R}}\left|\frac{t-\mathrm{i}}{t-z}\right| = \frac{1}{\left|\frac{1}{2}-\frac{\mathrm{i}z}{2}\right| - \left|\frac{1}{2}+\frac{\mathrm{i}z}{2}\right|} = \frac{|z+\mathrm{i}|+|z-\mathrm{i}|}{2\Im z},$$

proving the lemma.

Corollary. *For* $\Im z > 0$,

$$\sup_{t\in\mathbb{R}}\left|\frac{t-\mathrm{i}}{t-z}\right| \leqslant \frac{1+|z|}{\Im z}.$$

This inequality will be sufficient for our purposes.

3. Criterion in terms of $\int_{-\infty}^{\infty}(\log\Omega(t)/(1+t^2))\,\mathrm{d}t$

We return to the consideration of the quantity $\Omega(t)$ introduced at the beginning of this §, and to its connection with weighted polynomial approximation.

Theorem (Mergelian). *Polynomials are dense in* $\mathscr{C}_W(\mathbb{R})$ *iff*

$$(*) \qquad \int_{-\infty}^{\infty}\frac{\log\Omega(t)}{1+t^2}\mathrm{d}t = \infty.$$

Remark. Since $W(t)\geqslant 1$ we *always have* $\log\Omega(z)\geqslant 0$, i.e., $\Omega(z)\geqslant 1$, because 1 is a *polynomial* (!), and $|1/(t-\mathrm{i})W(t)|\leqslant 1$ on $\mathbb{R}$.

Proof of theorem

Only if: We must show that, if $(*)$ *fails*, the polynomials *can't be dense* in $\mathscr{C}_W(\mathbb{R})$.

Assume, then, that

$$\int_{-\infty}^{\infty}\frac{\log\Omega(t)}{1+t^2}\mathrm{d}t < \infty,$$

and take *any polynomial* $P(t)$ with $|P(t)/(t-\mathrm{i})W(t)|\leqslant 1$ on $\mathbb{R}$. Then, by a very simple version of the second theorem of §G.2, Chapter III (which, for *polynomials*, can be easily verified directly),

$$\log|P(\mathrm{i})| \leqslant \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log|P(t)|}{1+t^2}\mathrm{d}t.$$

Here, *by definition* (compare with the proof of the second part of M. Riesz' theorem in §D.1, Chapter V),

$$|P(t)| \leqslant \Omega(t),$$

so

$$\log|P(\mathrm{i})| \leqslant \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log\Omega(t)}{1+t^2}\mathrm{d}t,$$

and, taking the supremum of $\log|P(\mathrm{i})|$ over all polynomials P subject to the condition given above, we get

$$\log\Omega(\mathrm{i}) \leqslant \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log\Omega(t)}{1+t^2}\mathrm{d}t < \infty.$$

The quantity $\Omega(\mathrm{i})$ is thus finite. Therefore polynomials *cannot be dense* in $\mathscr{C}_W(\mathbb{R})$ by the first Mergelian theorem of article 1.

If: Supposing that polynomials *are not dense* in $\mathscr{C}_W(\mathbb{R})$, we must show that $(*)$ *is false.*

If polynomials are *not dense* in $\mathscr{C}_W(\mathbb{R})$, the Hahn–Banach theorem (whose validity does *not*, by the way, depend on $\mathscr{C}_W(\mathbb{R})$'s being complete!) furnishes us with a *bounded linear functional* on $\mathscr{C}_W(\mathbb{R})$ which is *not identically zero thereon*, but *is zero at each of the polynomials.* It is convenient to denote the *value* of this linear functional at a member φ of $\mathscr{C}_W(\mathbb{R})$ by the expression

$$L\left(\frac{\varphi(t)}{W(t)}\right);$$

the reason for this is that *then* we will simply have

$$(\ast\!\ast)\qquad \left|L\left(\frac{\varphi(t)}{W(t)}\right)\right| \leqslant C\sup_{t\in\mathbb{R}}\left|\frac{\varphi(t)}{W(t)}\right|$$

with some constant C, for $\varphi\in\mathscr{C}_W(\mathbb{R})$.

► (N.B. We are NOT writing $L(\varphi(t)/W(t))$ as $\int_{-\infty}^{\infty}(\varphi(t)/W(t))\mathrm{d}\mu(t)$ with a Radon measure μ. That's because we are *not assuming any continuity* of $W(t)$ here, so the existence of such a measure μ is problematical.)

Let us continue with this part of the proof. We have our linear functional L, such that

$$(\dagger)\qquad L\left(\frac{P(t)}{W(t)}\right) = 0$$

for every polynomial P, whilst

$$(\S)\qquad L\left(\frac{\varphi_0(t)}{W(t)}\right)\neq 0$$

for some $\varphi_0\in\mathscr{C}_W(\mathbb{R})$.

Consider the function

$$F(z) \;=\; L\left(\frac{1}{(t-z)W(t)}\right),$$

defined whenever $z\notin\mathbb{R}$. In the first place, $F(z)$ *cannot vanish identically for both* $\Im z>0$ and $\Im z<0$. Suppose it *did*. A simple modification of the general lemma given at the beginning of this chapter (whose verification is left to the reader) guarantees, for each $\varepsilon>0$, the existence of a finite linear combination $\varphi_\varepsilon(t)$ of the fractions $1/(t-c)$, $c\notin\mathbb{R}$, such that $\|\varphi_0-\varphi_\varepsilon\|_W<\varepsilon$ for the function φ_0 figuring in (§). If, then, $F(c)=L(1/(t-c)W(t))=0$ for every $c\notin\mathbb{R}$, we'd have $L(\varphi_\varepsilon(t)/W(t))=0$, whence $|L(\varphi_0(t)/W(t))|<C\varepsilon$ by (⁑). Squeezing ε, we get a contradiction with (§).

Wlog, $F(z)$ is not identically zero in $\{\Im z>0\}$. *It is analytic there.* To see this, observe that if $z\notin\mathbb{R}$, the difference quotient

$$\frac{(t-z-\Delta z)^{-1}-(t-z)^{-1}}{\Delta z} \;=\; \frac{1}{(t-z)(t-z-\Delta z)}$$

tends to $(t-z)^{-2}$ *uniformly for* $t\in\mathbb{R}$ as $\Delta z\to 0$. Therefore, by the linearity of L and (⁑),

$$\frac{F(z+\Delta z)-F(z)}{\Delta z} \;\longrightarrow\; L\left(\frac{1}{(t-z)^2W(t)}\right)$$

as $\Delta z\to 0$, since $W(t)\geqslant 1$ on $\mathbb{R}$. This shows that $F'(z)$ exists at every $z\notin\mathbb{R}$ and establishes analyticity of $F(z)$ in $\{\Im z>0\}$.

From (⁑), we get

$$|F(z)|\leqslant C \quad\text{for } \Im z\geqslant 1.$$

Since $F(z)\not\equiv 0$ in the upper half plane, the first theorem of §G.2, Chapter III, shows that

$$(\dagger\dagger)\qquad \int_{-\infty}^{\infty}\frac{\log^-|F(x+\mathrm{i})|}{1+x^2}\mathrm{d}x \;<\; \infty.$$

We can now bring in the Markov–Riesz–Pollard trick already used in proving the second part of Riesz' theorem in §D.1 of the previous chapter.

Take any polynomial $P(t)$ and any fixed z, $\Im z > 0$. Then

$$\frac{P(t) - P(z)}{t - z}$$

is *also* a polynomial in t, so, applying (†) to *it*, we get

$$L\left(\frac{P(t) - P(z)}{(t - z)W(t)}\right) = 0,$$

i.e., in terms of $F(z) = L(1/(t - z)W(t))$,

$$F(z)P(z) = L\left(\frac{P(t)}{(t - z)W(t)}\right).$$

We can thus write

$$(\S\S) \qquad P(z) = \frac{1}{F(z)}L\left(\frac{P(t)}{(t - z)W(t)}\right)$$

for $\Im z > 0$, provided z is not a zero of the analytic function $F(z)$. The idea now is to use (§§) together with (††) in order to show that

$$\int_{-\infty}^{\infty} \frac{\log \Omega(t)}{1 + t^2} dt < \infty.$$

Take any polynomial $P(t)$ such that

$$\left|\frac{P(t)}{(t - \mathrm{i})W(t)}\right| \leqslant 1 \quad \text{on} \quad \mathbb{R}.$$

Then, $|P(t)/(t - z)W(t)| \leqslant \sup_{t\in\mathbb{R}}|(t - \mathrm{i})/(t - z)|$ *which, by the previous article, is* $\leqslant (1 + |z|)/\Im z$ for $\Im z > 0$ (see the corollary there). Putting $z = x + \mathrm{i}$, $x \in \mathbb{R}$, we thus get, from ($\ast\atop\ast$),

$$\left|L\left(\frac{P(t)}{(t - x - \mathrm{i})W(t)}\right)\right| \leqslant C(1 + \sqrt{(x^2 + 1)}).$$

Referring to (§§), we see that

$$|P(x + \mathrm{i})| \leqslant \frac{C(1 + \sqrt{(x^2 + 1)})}{|F(x + \mathrm{i})|}$$

for any polynomial P with $|P(t)/(t - \mathrm{i})W(t)| \leqslant 1$ on $\mathbb{R}$. Taking the supremum of $|P(x + \mathrm{i})|$ over such P, we find that

$$\Omega(x + \mathrm{i}) \leqslant \frac{C(1 + \sqrt{(x^2 + 1)})}{|F(x + \mathrm{i})|},$$

that is, writing $C' = \log C$,

$$(\ddagger) \qquad \log\Omega(x+\mathrm{i}) \;\leqslant\; C' + \log(1+\sqrt{(x^2+1)}) + \log^-|F(x+\mathrm{i})|.$$

We use the last relation in conjunction with (††) in order to get a grip on $\log\Omega(t)$ for real t. The procedure being followed here is like the one used in proving the second part of Riesz' theorem (§D.1, previous chapter). I call it a *hall of mirrors* argument because it consists in our *first going up to the line* $\Im z = 1$ from the *real axis* and *then going back down to the real axis again.* Our reason for engaging in this roundabout manoeuvre is that we do not have any simple way of controlling $|F(z)|$ when z gets near $\mathbb{R}$ (unless we bring in H_p-spaces, whose use we are avoiding as much as possible!). Let P be a polynomial. By the second theorem of §G.2, Chapter III,

$$\log|P(t)| \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log|P(x+\mathrm{i})|}{(x-t)^2+1}\mathrm{d}x.$$

If also $|P(t)/(t-\mathrm{i})W(t)| \leqslant 1$, we have, of course, $|P(x+\mathrm{i})| \leqslant \Omega(x+\mathrm{i})$, so, taking the supremum of $\log|P(t)|$ for such P,

$$\log\Omega(t) \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log\Omega(x+\mathrm{i})}{(t-x)^2+1}\mathrm{d}x.$$

Plug in (‡) on the right, multiply by $1/(t^2+1)$, and integrate t from $-\infty$ to ∞. After changing the order of integration and using the identity

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{((t-x)^2+1)(t^2+1)} = \frac{2}{x^2+4},$$

we obtain

$$\int_{-\infty}^{\infty}\frac{\log\Omega(t)}{1+t^2}\mathrm{d}t \;\leqslant\; \pi C' + \int_{-\infty}^{\infty}\frac{2\log(1+\sqrt{(x^2+1)})}{x^2+4}\mathrm{d}x + \int_{-\infty}^{\infty}\frac{2\log^-|F(x+\mathrm{i})|}{x^2+4}\mathrm{d}x.$$

The first integral on the right is obviously finite. The *second is also finite by* (††). The integral

$$\int_{-\infty}^{\infty}\frac{\log\Omega(t)}{1+t^2}\mathrm{d}t$$

is therefore *finite*, contradicting (∗). This completes the proof of the *if* part of the theorem, and we are done.

B. Akhiezer's method, based on use of $W_*(z)$

The function $\Omega(z)$ introduced by Mergelian, which, as we have seen, indicates by its size whether or not the polynomials are dense in $\mathscr{C}_W(\mathbb{R})$, is equal to $\sup\{|P(z)|: P$ a polynomial and $|P(t)| \leqslant |t-\mathrm{i}|W(t)$ for $t \in \mathbb{R}\}$. The presence of the multiplier $|t-\mathrm{i}|$ in front of $W(t)$ is disconcerting, and it would seem more natural to work with the quantity

$$W_*(z) = \sup\{|P(z)|: P \text{ a polynomial and } |P(t)| \leqslant W(t) \text{ on } \mathbb{R}\}.$$

On the real axis, $W_*(x) \leqslant W(x)$, and so $W_*(x)$ is a kind of *lower regularization* of $W(x)$ by *polynomials*. (Recall that the idea of using *some* kind of lower regularization occurred already in the study of quasianalyticity (Chapter IV); the *convex logarithmic regularization* which turned out to be useful there is *not the same*, however, as the *regularization by polynomials* dealt with here.)

We are always assuming that $W(x) \geqslant 1$. Therefore, since 1 is a *polynomial* (!), we *certainly* have $W_*(z) \geqslant 1$.

1. Criterion in terms of $\int_{-\infty}^{\infty}(\log W_*(x)/(1+x^2))\mathrm{d}x$

The following theorem, due to Akhiezer, is implicity contained in the work of S. Bernstein, who was in possession of all the elements of the proof. Bernstein, who devoted much effort to the study of the problem bearing his name, was apparently unable to see that a solution was within his reach, and never formulated this next result.

Theorem (Akhiezer). *Let $W(x)$ be continuous. Then the polynomials are dense in $\mathscr{C}_W(\mathbb{R})$ iff*

$$\int_{-\infty}^{\infty} \frac{\log W_*(x)}{1+x^2}\mathrm{d}x = \infty.$$

Remark. As we shall see from the proof, the *continuity* requirement on $W(x)$ can be much relaxed. What is really needed here is that $W(x)$ be *finite on a set of points which is not too sparse.*

Proof of Theorem.

If: Comparison of the definitions of $W_*(z)$ and $\Omega(z)$ shows that $W_*(x) \leqslant \Omega(x)$. Therefore, if $\int_{-\infty}^{\infty}(\log W_*(x)/(1+x^2))\mathrm{d}x = \infty$, we certainly have $\int_{-\infty}^{\infty}(\log \Omega(x)/(1+x^2))\mathrm{d}x = \infty$, so polynomials *are* dense in $\mathscr{C}_W(\mathbb{R})$ by Mergelian's second theorem (§A.3). Note that the continuity of W plays no role here.

Only if: Assuming that $\int_{-\infty}^{\infty}(\log W_*(x)/(1+x^2))\mathrm{d}x$ is *finite*, we show that *any collection of polynomials P with $\|P\|_W \leqslant 1$ forms a normal family in the complex plane.* For this, a *hall of mirrors* argument like the one at the end of §A.3 is used. If P is any polynomial with $\|P\|_W \leqslant 1$, the second theorem of

§G.2, Chapter III and the very definition of W_* give, for real ξ,

$$\log|P(\xi+\mathrm{i})| \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log|P(t)|}{(t-\xi)^2+1}\mathrm{d}t \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log W_*(t)}{(t-\xi)^2+1}\mathrm{d}t.$$

Taking the supremum of $\log|P(\xi+\mathrm{i})|$ for such P, we find, as usual,

$$\log W_*(\xi+\mathrm{i}) \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log W_*(t)}{(t-\xi)^2+1}\mathrm{d}t.$$

Now suppose that $\Im z\leqslant 0$. Using again the second theorem of §G.2, Chapter III, but this time in the half plane $\Im z\leqslant 1$, we see that for any polynomial P,

$$\log|P(z)| \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{(1+|\Im z|)\log|P(\xi+\mathrm{i})|}{|\xi+\mathrm{i}-z|^2}\mathrm{d}\xi.$$

If also $\|P\|_W\leqslant 1$, we have $|P(\xi+\mathrm{i})|\leqslant W_*(\xi+\mathrm{i})$, so, by the inequality just found for the latter function (which, by the way, is $\geqslant 1$),

$$\log|P(z)| \;\leqslant\; \frac{1}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{(1+|\Im z|)\log W_*(t)}{|\xi+\mathrm{i}-z|^2|\xi+\mathrm{i}-t|^2}\mathrm{d}t\,\mathrm{d}\xi.$$

Changing the order of integration, and using the identity

$$\frac{2+|\Im z|}{|t+2\mathrm{i}-z|^2} \;=\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{(1+|\Im z|)\mathrm{d}\xi}{|\xi+\mathrm{i}-z|^2|t-\xi-\mathrm{i}|^2},$$

valid for $\Im z\leqslant 0$, we find that

$$\log|P(z)| \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{(2+|\Im z|)\log W_*(t)}{|t+2\mathrm{i}-z|^2}\mathrm{d}t.$$

Apply now the corollary from §A.2. In the present situation, where $\Im z\leqslant 0$, we get

$$\sup_{t\in\mathbb{R}}\left|\frac{t-\mathrm{i}}{t+2\mathrm{i}-z}\right|^2 \;\leqslant\; \left(\frac{1+|z-2\mathrm{i}|}{2+|\Im z|}\right)^2,$$

whence, by the preceding, for $\Im z\leqslant 0$,

$$(\ast)\qquad \log|P(z)| \;\leqslant\; \frac{(3+|z|)^2}{2+|\Im z|}\cdot\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\log W_*(t)}{1+t^2}\mathrm{d}t$$

whenever P is a polynomial with $\|P\|_W\leqslant 1$.

For such polynomials P, however, $(\ast)$ is also valid for $\Im z\geqslant 0$. This is seen by an argument *just* like the above one, working *first* with $\log|P(\xi-\mathrm{i})|$ instead of $\log|P(\xi+\mathrm{i})|$, and *then* using the second theorem of §G.2, Chapter III in the half plane $\Im z\geqslant -1$ instead of the half plane $\Im z\leqslant 1$.

The polynomials P with $\|P\|_W\leqslant 1$ thus satisfy $(\ast)$ in the whole complex plane. Since $\int_{-\infty}^{\infty}(\log W_*(t)/(1+t^2))\mathrm{d}t$ is *finite*, such polynomials *form* a *normal family* in $\mathbb{C}$.

Once we know that the polynomials P with $\|P\|_W \leqslant 1$ *do* form a *normal family* in $\mathbb{C}$, it is *manifest* that polynomials *cannot* be $\|\ \ \|_W$-dense in $\mathscr{C}_W(\mathbb{R})$. Suppose, indeed, that $\varphi \in \mathscr{C}_W(\mathbb{R})$ and that we have polynomials P_n with $\|P_n - \varphi\|_W \underset{n}{\longrightarrow} 0$. We may wlog take $\|\varphi\|_W < 1$, then $\|P_n\|_W \leqslant 1$ for all sufficiently large n, hence, wlog, for all n.

These $P_n(z)$ therefore form a normal family in $\mathbb{C}$, so a *subsequence of them* must tend u.c.c. in $\mathbb{C}$ to *some entire function* $\Phi(z)$. *At every* $x \in \mathbb{R}$, $\Phi(x)$ *and* $\varphi(x)$ *must coincide, since* $W(x)$, *being continuous, must be finite on* $\mathbb{R}$ (!). The function $\varphi \in \mathscr{C}_W(\mathbb{R})$ which is $\|\ \ \|_W$-approximable by polynomials *must thus coincide on* $\mathbb{R}$ *with some entire function.* Since *lots* of continuous $\varphi \in \mathscr{C}_W(\mathbb{R})$ *don't do that*, we see that polynomials *cannot be* $\|\ \ \|_W$-dense in $\mathscr{C}_W(\mathbb{R})$.

We have finished the *only if* part of Akhiezer's theorem, which is now completely proved.

Remark. We see already from the argument at the very end of the above proof that *we need merely assume* $W(x) < \infty$ *on some closed subset of* $\mathbb{R}$ *with a finite limit point, instead of the continuity* of W on $\mathbb{R}$, and *then* the property

$$\int_{-\infty}^{\infty} \frac{\log W_*(t)}{1+t^2}\,\mathrm{d}t \;<\; \infty$$

will surely imply that the polynomials are not $\|\ \ \|_W$-dense in $\mathscr{C}_W(\mathbb{R})$. Even this assumption on $W(x)$ can be very much weakened, as we shall see in the next article.

2. Description of $\|\ \ \|_W$ limits of polynomials when $\int_{-\infty}^{\infty}(\log W_*(t)/(1+t^2))\mathrm{d}t < \infty$

A small refinement of the calculations made in proving the *only if* part of the previous theorem yields an elegant result.

Theorem (Akhiezer). *If* $\int_{-\infty}^{\infty}(\log W_*(t)/(1+t^2))\mathrm{d}t < \infty$, *every function in* $\mathscr{C}_W(\mathbb{R})$ *which can be* $\|\ \ \|_W$*-approximated by polynomials coincides, on the subset of* $\mathbb{R}$ *where* $W(x) < \infty$, *with some entire function of zero exponential type.*

Proof. We start from the estimate of $\log|P(z)|$ found in the preceding article for polynomials P with $\|P\|_W \leqslant 1$. As we saw there, if $\Im z \leqslant 0$ and P is a polynomial with $\|P\|_W \leqslant 1$,

$$\log|P(z)| \;\leqslant\; \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{(2+|\Im z|)\log W_*(t)}{|t+2\mathrm{i}-z|^2}\,\mathrm{d}t.$$

Take any $\varepsilon > 0$. Since $\int_{-\infty}^{\infty}(\log W_*(t)/(1+t^2))\mathrm{d}t < \infty$, there is a finite M_ε

such that

$$(\dagger) \qquad \frac{1}{\pi}\int_{\{\log W_*(t) > M_\varepsilon\}} \frac{\log W_*(t)}{1+t^2}\,dt \;<\; \varepsilon;$$

we then break up the integral of the preceding relation into the sum of *two*, *one* over

$$\{t \in \mathbb{R}:\quad \log W_*(t) \leqslant M_\varepsilon\}$$

and the *other* over the set where $\log W_*(t) > M_\varepsilon$. We obtain in this way

$$\log|P(z)| \;\leqslant\; M_\varepsilon + \frac{1}{\pi}\int_{\{\log W_*(t) > M_\varepsilon\}} \frac{(2+|\Im z|)\log W_*(t)}{|t+2i-z|^2}\,dt.$$

Apply now the corollary from §A.2. We find, by virtue of (†), that

$$\log|P(z)| \;\leqslant\; M_\varepsilon + \frac{(1+|z-2i|)^2}{(2+|\Im z|)}\varepsilon;$$

this holds whenever $\Im z \leqslant 0$ if P is a polynomial with $\|P\|_W \leqslant 1$.

One can, of course, use exactly the same kind of reasoning for the half plane $\Im z \geqslant 0$. We see in this way that if P is any polynomial with $\|P\|_W \leqslant 1$, the relation

$$(\dagger\dagger) \qquad \log|P(z)| \;\leqslant\; M_\varepsilon + \frac{(3+|z|)^2}{2+|\Im z|}\varepsilon$$

holds in the entire complex plane.

This inequality we refine still more by use of a Phragmén–Lindelöf argument.

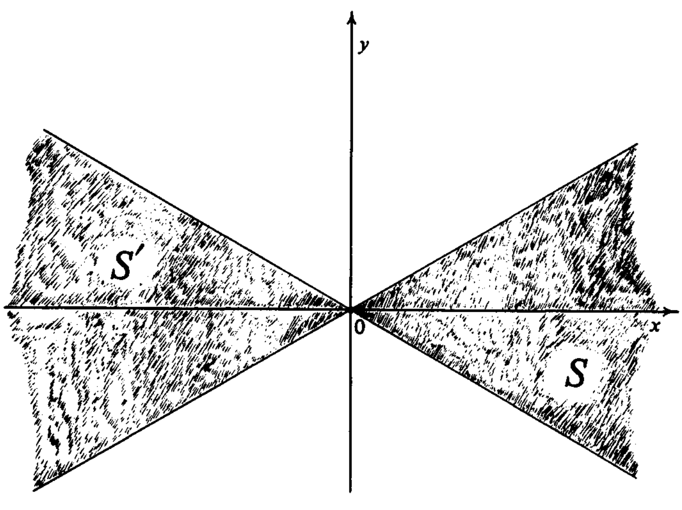


Figure 36