CHAPTER XIII

Laplace Transforms. Tauberian Theorems. Resolvents

The Laplace transforms are a powerful practical tool, but at the same time their theory is of intrinsic value and opens the door to other theories such as semi-groups. The theorem on completely monotone functions and the basic Tauberian theorem have rightly been considered pearls of hard analysis. (Although the present proofs are simple and elementary, the pioneer work in this direction required originality and power.) Resolvents (sections 9–10) are basic for semi-group theory.

As this chapter must cover diverse needs, a serious effort has been made to keep the various parts as independent of each other as the subject permits, and to make it possible to skip over details. Chapter XIV may serve for collateral reading and to provide examples. The remaining part of this book is entirely independent of the present chapter.

Despite the frequent appearance of regularly varying functions only the quite elementary theorem 1 of VIII,8 is used.

1. DEFINITIONS. THE CONTINUITY THEOREM

Definition 1. *If* F *is a proper or defective probability distribution concentrated on* $\overline{0, \infty}$, *the Laplace transform* φ *of* F *is the function defined for* $\lambda \geq 0$ *by*

$$\varphi(\lambda) = \int_0^\infty e^{-\lambda x}\, F\{dx\}. \tag{1.1}$$

Here and in the sequel it is understood that the *interval of integration is closed* (and may be replaced by $\overline{-\infty, \infty}$). Whenever we speak of the Laplace transform of a distribution F it is tacitly understood that F is concentrated on $\overline{0, \infty}$. As usual we stretch the language and speak of "the

Laplace transform of the random variable X," meaning the transform of its distribution. With the usual notation for expectations we have then

$$\varphi(\lambda) = \mathbf{E}(e^{-\lambda \mathbf{X}}). \tag{1.2}$$

Example. (*a*) Let X assume the values $0, 1, \ldots$ with probabilities $p_0, p_1, \ldots$. Then $\varphi(\lambda) = \sum p_n e^{-n\lambda}$ whereas the generating function is $P(s) = \sum p_n s^n$. Thus $\varphi(\lambda) = P(e^{-\lambda})$ and the Laplace transform differs from the generating function only by the change of variable $s = e^{-\lambda}$. This explains the close analogy between the properties of Laplace transforms and generating functions.

(*b*) *The gamma distribution* with density $f_\alpha(x) = (x^{\alpha-1}/\Gamma(\alpha))e^{-x}$ has the transform

$$\varphi_\alpha(\lambda) = \frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-(\lambda+1)x}\, x^{\alpha-1}\, dx = \frac{1}{(\lambda+1)^\alpha}, \qquad \alpha > 0. \tag{1.3}$$

The next theorem shows that a distribution is recognizable by its transform; without this the usefulness of Laplace transforms would be limited.

Theorem 1. (*Uniqueness.*) *Distinct probability distributions have distinct Laplace transforms.*

First proof. In VIII,(6.4) we have an explicit inversion formula which permits us to calculate F when its transform is known. This formula will be derived afresh in section 4.

Second proof. Put $y = e^{-x}$. As x goes from 0 to ∞ the variable y goes from 1 to 0. We now define a probability distribution G concentrated on $\overline{0, 1]}$ by letting $G(y) = 1 - F(x)$ at points of continuity. Then

$$\varphi(\lambda) = \int_0^\infty e^{-\lambda x}\, F\{dx\} = \int_0^1 y^\lambda\, G\{dy\} \tag{1.4}$$

as is obvious from the fact that the Riemann sums $\sum e^{-\lambda x_k}[F(x_{k+1}) - F(x_k)]$ coincide with the Riemann sums $\sum y_k^\lambda [G(y_k) - G(y_{k+1})]$ when $y_k = e^{-x_k}$. We know from VII,3 that the distribution G is uniquely determined by its moments, and these are given by $\varphi(k)$. Thus the knowledge of $\varphi(1)$, $\varphi(2), \ldots$ determines G, and hence F. This result is stronger than the assertion of the theorem.[1] ▶

The following basic result is a simple consequence of theorem 1.

[1] More generally, a completely monotone function is uniquely determined by its values at a sequence $\{a_n\}$ of points such that $\sum a_n^{-1}$ diverges. However, if the series converges there exist two distinct completely monotone functions agreeing at all points a_n. For an elementary proof of this famous theorem see W. Feller, ***On Müntz' theorem and completely monotone functions***, Amer. Math. Monthly, vol. 75 (1968), pp. 342–350.

Theorem 2. (*Continuity theorem.*) *For* $n = 1, 2, \ldots$ *let* F_n *be a probability distribution with transform* φ_n.

If $F_n \to F$ *where* F *is a possibly defective distribution with transform* φ *then* $\varphi_n(\lambda) \to \varphi(\lambda)$ *for* $\lambda > 0$.

Conversely, if the sequence $\{\varphi_n(\lambda)\}$ *converges for each* $\lambda > 0$ *to a limit* $\varphi(\lambda)$, *then* φ *is the transform of a possibly defective distribution* F, *and* $F_n \to F$.

The limit F *is not defective iff* $\varphi(\lambda) \to 1$ *as* $\lambda \to 0$.

Proof. The first part is contained in the basic convergence theorem of VIII,1. For the second part we use the selection theorem 1 of VIII,6. Let $\{F_{n_k}\}$ be a subsequence converging to the possibly defective distribution F. By the first part of the theorem the transforms converge to the Laplace transform of F. It follows that F is the unique distribution with Laplace transform φ, and so all convergent subsequences converge to the same limit F. This implies the convergence of F_n to F. The last assertion of the theorem is clear by inspection of (1.1). ▶

For clarity of exposition we shall as far as possible reserve the letter F for probability distributions, but instead of (1.1) we may consider more general integrals of the form

$$\omega(\lambda) = \int_0^\infty e^{-\lambda x}\, U\{dx\}. \tag{1.5}$$

where U is a measure attributing a finite mass $U\{I\}$ to the finite interval I, but may attribute an infinite mass to the positive half axis. As usual, we describe this measure conveniently in terms of its improper distribution function defined by $U(x) = U\{\overline{0, x}\}$. In the important special case where U is defined as the integral of a function $u \geq 0$ the integral (1.5) reduces to

$$\omega(\lambda) = \int_0^\infty e^{-\lambda x}\, u(x)\, dx \tag{1.6}$$

Examples. (*c*) If $u(x) = x^a$ with $a > -1$, then $\omega(\lambda) = \Gamma(a+1)/\lambda^{\alpha+1}$ for all $\lambda > 0$.

(*d*) If $u(x) = e^{ax}$ then $\omega(\lambda) = 1/(\lambda - a)$ for $\lambda > a > 0$, but the integral (1.6) diverges for $\lambda \leq a$.

(*e*) If $u(x) = e^{x^2}$ the integral (1.6) diverges everywhere.

(*f*) By differentiation we get from (1.1)

$$-\varphi'(\lambda) = \int_0^\infty e^{-\lambda x} x\, F\{dx\} \tag{1.7}$$

and this is an integral of the form (1.5) with $U\{dx\} = x\, F\{dx\}$. This example illustrates how integrals of the form (1.5) arise naturally in connection with proper probability distributions. ▶

We shall be interested principally in measures U derived by simple operations from probability distributions, and the integral in (1.5) will generally converge for all $\lambda > 0$. However, nothing is gained by excluding measures for which convergence takes place only for *some* λ. Now $\omega(a) < \infty$ implies $\omega(\lambda) < \infty$ for all $\lambda > a$, and so the values of λ for which the integral in (1.5) converges fill an interval $\overline{a, \infty}$.

Definition 2. *Let U be a measure concentrated on $\overline{0, \infty}$. If the integral in* (1.5) *converges for $\lambda > a$, then the function ω defined for $\lambda > a$ is called the Laplace transform of U.*

If U has a density u, the Laplace transform (1.6) *of U is also called the ordinary Laplace transform of u.*

The last convention is introduced merely for convenience. To be systematic one should consider more general integrals of the form

$$\int_0^\infty e^{-\lambda x}\, v(x)\, U\{dx\} \tag{1.8}$$

and call them "Laplace transform of v with respect to the measure U." Then (1.6) would be the "transform of u with respect to Lebesgue measure" (or ordinary length). This would have the theoretical advantage that one could consider functions u and v of variable signs. For the purposes of this book it is simplest and least confusing to associate Laplace transforms only with measures, and we shall do so.[2]

If U is a measure such that the integral in (1.5) converges for $\lambda = a$, then for all $\lambda > 0$

$$\omega(\lambda+a) = \int_0^\infty e^{-\lambda x}\cdot e^{-ax}\, U\{dx\} = \int_0^\infty e^{-\lambda x}\, U^{\#}\{dx\} \tag{1.9}$$

is the Laplace transform of the bounded measure $U^{\#}\{dx\} = e^{-ax}\, U\{dx\}$, and $\omega(\lambda+a)/\omega(a)$ is the transform of a *probability* distribution. In this way every theorem concerning transforms of probability distributions automatically generalizes to a wider class of measures. Because the graph of the new transform $\omega(\lambda+a)$ is obtained by translation of the graph of ω we shall refer to this extremely useful method as the *translation principle.* For example, since U is uniquely determined by $U^{\#}$, and $U^{\#}$ by $\omega(\lambda+a)$ for $\lambda > 0$, we can generalize theorem 1 as follows.

Theorem 1a. *A measure U is uniquely determined by the values of its Laplace transform* (1.5) *in some interval $a < \lambda < \infty$.*

---

[2] The terminology is not well established, and in the literature the term "Laplace transform of F" may refer either to (1.1) or to (2.6). We would describe (2.6) as the "ordinary Laplace transform of the distribution function F," but texts treating principally such transforms would drop the determinative "ordinary." To avoid ambiguities in such cases the transform (1.1) is then called the *Laplace-Stieltjes* transform.

Corollary. *A continuous function u is uniquely determined by the values of its ordinary Laplace transform* (1.6) *in some interval $a < \lambda < \infty$.*

Proof. The transform determines uniquely the integral U of u, and two distinct continuous[3] functions cannot have identical integrals. ►

[An explicit formula for u in terms of ω is given in VII,(6.6).]

The continuity theorem generalizes similarly to sequences of arbitrary measures U_n with Laplace transforms. The fact that U_n has a Laplace transform implies that $U_n\{I\} < \infty$ for finite intervals I. We recall from VIII,1 and VIII,6 that a sequence of such measures is said to converge to a measure U iff $U_n\{I\} \to U\{I\} < \infty$ for every *finite* interval of continuity of U.

Theorem 2a. (*Extended continuity theorem.*) *For $n = 1, 2, \ldots$ let U_n be a measure with Laplace transform ω_n. If $\omega_n(\lambda) \to \omega(\lambda)$ for $\lambda > a$, then ω is the Laplace transform of a measure U and $U_n \to U$.*

Conversely, if $U_n \to U$ ad the sequence $\{\omega_n(a)\}$ is bounded, then $\omega_n(\lambda) \to \omega(\lambda)$ for $\lambda > a$.

Proof. (*a*) Assume that $U_n \to U$ and that $\omega_n(a) < A$. If $t > 0$ is a point of continuity of U then

$$(1.10) \qquad \int_0^t e^{-(\lambda+a)x}\, U_n\{dx\} \to \int_0^t e^{-(\lambda+a)x}\, U\{dx\}$$

and the left side differs from $\omega_n(\lambda+a)$ by at most

$$(1.11) \qquad \int_t^\infty e^{-(\lambda+a)x}\, U_n\{dx\} < Ae^{-\lambda t}$$

which can be made $<\epsilon$ by choosing t sufficiently large. This means that the upper and lower limits of $\omega_n(\lambda+a)$ differ by less than an arbitrary ϵ, and hence for every $\lambda > 0$ the sequence $\{\omega_n(\lambda+a)\}$ converges to a finite limit.

(*b*) Assume then that $\omega_n(\lambda) \to \omega(\lambda)$ for $\lambda > a$. For fixed $\lambda_0 > a$ the function $\omega_n(\lambda+\lambda_0)/\omega_n(\lambda_0)$ is the Laplace transform of the probability distribution $U_n^\#\{dx\} = (1/\omega_n(\lambda_0))e^{-\lambda_0 x}\, U_n\{dx\}$. By the continuity theorem therefore $U_n^\#$ converges to a possibly defective distribution $U^\#$, and this implies that U_n converges to a measure U such that $U\{dx\} = \omega(\lambda_0)e^{\lambda_0 x}\, U^\#\{dx\}$. ►

The following example shows the necessity of the condition that $\{\omega_n(a)\}$ remain bounded.

[3] The same argument shows that in general u is determined up to values on an arbitrary set of measure zero.

Example. (*g*) Let U_n attach weight e^{n^2} to the point n, and zero to the complement. Since $U_n\{\overline{0, n}\} = 0$ we have $U_n \to 0$, but $\omega_n(\lambda) = e^{n(n-\lambda)} \to \infty$ for all $\lambda > 0$. ▶

One speaks sometimes of the *bilateral transform* of a distribution F with two tails, namely

$$\varphi(\lambda) = \int_{-\infty}^{+\infty} e^{-\lambda x}\, F\{dx\}, \tag{1.12}$$

but this function need not exist for any $\lambda \neq 0$. If it exists, $\varphi(-\lambda)$ is often called the *moment generating function*, but in reality it is the generating function of the sequence $\{\mu_n/n!\}$ where μ_n is the nth moment.

2. ELEMENTARY PROPERTIES

In this section we list the most frequently used properties of the Laplace transforms; the parallel to generating functions is conspicuous.

(i) Convolutions. Let F and G be probability distributions and U their convolution, that is,

$$U(x) = \int_0^x G(x-y)\, F\{dy\}. \tag{2.1}$$

The corresponding Laplace transforms obey the *multiplication rule*

$$\omega = \varphi\gamma. \tag{2.2}$$

This is equivalent to the assertion that for independent random variables $\mathbf{E}(e^{-\lambda(\mathbf{X}+\mathbf{Y})}) = \mathbf{E}(e^{-\lambda\mathbf{X}})\,\mathbf{E}(e^{-\lambda\mathbf{Y}})$, which is a special case of the multiplication rule for expectations.[4]

If F and G have densities f and g, then U has a density u given by

$$u(x) = \int_0^x g(x-y) f(y)\, dy \tag{2.3}$$

and the multiplication rule (2.2) applies to the "ordinary" Laplace transforms (1.6) of f, g, and u.

We now show that the multiplication rule can be extended as follows. *Let F and G be arbitrary measures with Laplace transforms φ and γ converging for $\lambda > 0$. The convolution U has then a Laplace transform ω given by* (2.2). This implies in particular that the multiplication rule applies to the "ordinary" transforms of any two integrable functions f and g and their convolution (2.3).

[4] The converse is false: two variables may be dependent and yet such that the distribution of their sum is given by the convolution formula. [See II,4(*e*) and problem 1 of III,9.]

To prove the assertion we introduce the finite measures F_n obtained by truncation of F as follows: for $x \leq n$ we put $F_n(x) = F(x)$, but for $x > n$ we let $F_n(x) = F(n)$. Define G_n similarly by truncating G. For $x < n$ the convolution $U_n = F_n \star G_n$ does not differ from U, and hence not only $F_n \to F$ and $G_n \to G$, but also $U_n \to U$. For the corresponding Laplace transforms we have $\omega_n = \varphi_n \gamma_n$ and letting $n \to \infty$ we get the assertion $\omega = \varphi\gamma$.

Examples. (*a*) *Gamma distributions.* In example 1(*b*) the familiar convolution rule $f_\alpha \star f_\beta = f_{\alpha+\beta}$ is mirrored in the obvious relation $\varphi_\alpha \varphi_\beta = \varphi_{\alpha+\beta}$.

(*b*) *Powers.* To $u_\alpha(x) = x^{\alpha-1}/\Gamma(\alpha)$ there corresponds the ordinary Laplace transform $\omega_\alpha(\lambda) = \lambda^{-\alpha}$. It follows that the convolution (2.3) of u_α and u_β is given by $u_{\alpha+\beta}$. The preceding example follows from this by the translation principle since $\varphi_\alpha(\lambda) = \omega_\alpha(\lambda+1)$.

(*c*) *If* $a > 0$ *then* $e^{-a\lambda}\omega(\lambda)$ *is the Laplace transform of the measure with distribution function* $U(x-a)$. This is obvious from the definition, but may be considered also as a special case of the convolution theorem inasmuch as $e^{-a\lambda}$ is the transform of the distribution concentrated at the point a. ►

(ii) Derivatives and moments. If F is a probability distribution and φ its Laplace transform (1.1), then φ *possesses derivatives of all orders given by*

$$(2.4) \qquad (-1)^n \varphi^{(n)}(\lambda) = \int_0^\infty e^{-\lambda x} x^n \, F\{dx\}$$

(as always, $\lambda > 0$). The differentiation under the integral is permissible since the new integrand is bounded and continuous.

It follows in particular that F *possesses a finite* nth *moment iff a finite limit* $\varphi^{(n)}(0)$ *exists.* For a random variable $\mathbf{X}$ we can therefore write

$$(2.5) \qquad \mathbf{E}(\mathbf{X}) = -\varphi'(0), \qquad \mathbf{E}(\mathbf{X}^2) = \varphi''(0)$$

with the obvious conventions in case of divergence. The differentiation rule (2.4) remains valid for arbitrary measures F.

(iii) Integration by parts leads from (1.1) to

$$(2.6) \qquad \int_0^\infty e^{-\lambda x} F(x)\, dx = \frac{\varphi(\lambda)}{\lambda}, \qquad \lambda > 0.$$

For probability distributions it is sometimes preferable to rewrite (2.6) in terms of the tail

$$(2.7) \qquad \int_0^\infty e^{-\lambda x}[1 - F(x)]\, dx = \frac{1 - \varphi(\lambda)}{\lambda}.$$

This corresponds to formula **1**; XI,(1.6) for generating functions.

(iv) **Change of scale.** From (1.2) we have $\mathbf{E}(e^{-a\lambda\mathbf{X}}) = \varphi(a\lambda)$ for each fixed $a > 0$, and so $\varphi(a\lambda)$ *is the transform of the distribution* $F\{dx/a\}$ [with distribution function $F(x/a)$]. This relation is in constant use.

Example. (*d*) *Law of large numbers.* Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be independent random variables with a common Laplace transform φ. Suppose $\mathbf{E}(\mathbf{X}j) = \mu$. The Laplace transform of the sum $\mathbf{X}_1 + \cdots + \mathbf{X}_n$ is φ^n, and hence the transform of the average $[\mathbf{X}_1 + \cdots + \mathbf{X}_n]/n$ is given by $\varphi^n(\lambda/n)$. Near the origin $\varphi(\lambda) = 1 - \mu\lambda + o(\lambda)$ [see (2.5)] and so as $n \to \infty$

$$\lim \varphi^n\left(\frac{\lambda}{n}\right) = \lim\left(1 - \frac{\mu\lambda}{n}\right)^n = e^{-\mu\lambda}. \tag{2.8}$$

But $e^{-\mu\lambda}$ is the transform of the distribution concentrated at μ, and so the distribution of $[\mathbf{X}_1 + \cdots + \mathbf{X}_n]/n$ tends to this limit. This is the weak law of large numbers in the Khintchine version, which does not require the existence of a variance. True, the proof applies directly only to positive variables, but it illustrates the elegance of Laplace transform methods. ►

3. EXAMPLES

(*a*) *Uniform distribution.* Let F stand for the uniform distribution concentrated on $\overline{0, 1}$. Its Laplace transform is given by $\varphi(\lambda) = (1 - e^{-\lambda})/\lambda$. Using the binomial expansion it is seen that the n-fold convolution $F^{n\star}$ has the transform

$$\varphi^n(\lambda) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} e^{-\lambda k} \lambda^{-n}. \tag{3.1}$$

As λ^{-n} is the transform corresponding to $U(x) = x^n/n!$ example 2(*c*) shows that $e^{-k\lambda}\lambda^{-n}$ corresponds to $(x-k)_+^n/n!$ where x_+ denotes the function that equals 0 for $x \leq 0$ and x for $x \geq 0$. Thus

$$F^{n\star}(x) = \frac{1}{n!} \sum_{k=0}^{n} (-1)^k \binom{n}{k} (x-k)_+^n. \tag{3.2}$$

This formula was derived by direct calculation in I,(9.5) and by a passage to the limit in problem 20 of **1**; XI.

(*b*) *Stable distributions with exponent* $\frac{1}{2}$. The distribution function

$$G(x) = 2[1 - \mathfrak{N}(1/\sqrt{x})], \qquad x > 0 \tag{3.3}$$

(where $\mathfrak{N}$ is the standard normal distribution) has the Laplace transform

$$\gamma(\lambda) = e^{-\sqrt{2\lambda}}. \tag{3.4}$$

This can be verified by elementary calculations, but they are tedious and we prefer to derive (3.4) from the limit theorem 3 in **1**; III,7 in which the distribution G was first encountered. Consider a simple symmetric random walk (coin tossing), and denote by $\mathbf{T}$ the epoch of the first return to the origin. The cited limit theorem states that G is the limit distribution of the normalized sums $(\mathbf{T}_1+\cdots+\mathbf{T}_n)/n^2$, where $\mathbf{T}_1, \mathbf{T}_2, \ldots$ are independent random variables distributed like $\mathbf{T}$. According to **1**; XI,(3.14) the generating function of $\mathbf{T}$ is given by $f(s) = 1 - \sqrt{1-s^2}$, and therefore

$$(3.5) \qquad \gamma(\lambda) = \lim\,[1-\sqrt{1-e^{-2\lambda/n^2}}]^n = \lim\left[1 - \frac{\sqrt{2\lambda}}{n}\right]^n = e^{-\sqrt{2\lambda}}.$$

We have mentioned several times that G *is a stable distribution*, but again the direct computational verification is laborious. Now obviously $\gamma^n(\lambda) = \gamma(n^2\lambda)$ which is the same as $G^{n\star}(x) = G(n^{-2}x)$ and proves the stability without effort.

(*c*) *Power series and mixtures*. Let F be a probability distribution with Laplace transform $\varphi(\lambda)$. We have repeatedly encountered distributions of the form

$$(3.6) \qquad G = \sum_{k=0}^{\infty} p_k F^{k\star}$$

where $\{p_k\}$ is a probability distribution. If $P(s) = \sum p_k s^k$ stands for the generating function of $\{p_k\}$, the Laplace transform of G is obviously given by

$$(3.7) \qquad \gamma(\lambda) = \sum_{k=0}^{\infty} p_k \varphi^k(\lambda) = P(\varphi(\lambda)).$$

This principle can be extended to arbitrary power series with positive coefficients. We turn to specific applications.

(*d*) *Bessel function densities*. In example II,7(*c*) we saw that for $r = 1, 2, \ldots$ the density

$$(3.8) \qquad v_r(x) = e^{-x}\frac{r}{x}I_r(x)$$

corresponds to a distribution of the form (3.6) where F is exponential with $\varphi(\lambda) = 1/(\lambda+1)$, and $\{p_k\}$ is the distribution of the first-passage epoch through the point $r > 0$ in an ordinary symmetric random walk. The generating function of this distribution is

$$(3.9) \qquad P(s) = \left(\frac{1-\sqrt{1-s^2}}{s}\right)^r$$

[see **1**; XI,(3.6)]. Substituting $s = (1+\lambda)^{-1}$ we conclude that *the ordinary Laplace transform of the probability density* (3.8) *is given by*

$$(3.10) \qquad [\lambda + 1 - \sqrt{(\lambda+1)^2 - 1}\,]^r.$$

That v_r *is a probability density and* (3.10) *its transform* has been proved only for $r = 1, 2, \ldots$. However, the statement is true[5] for *all* $r > 0$. It is of probabilistic interest because it implies the convolution formula $v_r \star v_s = v_{r+s}$ and thus the *infinite divisibility of* v_r. (See section 7.)

(*e*) *Another Bessel density.* In (3.6) choose for F the exponential distribution with $\varphi(\lambda) = 1/(\lambda+1)$ and for $\{p_k\}$ the Poisson distribution with $P(s) = e^{-t+ts}$. It is easy to calculate G explicitly, but fortunately this task was already accomplished in example II,7(*a*). We saw there that the density

$$w_\rho(x) = e^{-t-x}\sqrt{(x/t)^\rho}\, I_\rho(2\sqrt{tx}) \tag{3.11}$$

defined in II,(7.2) is the convolution of our distribution G with a gamma density $f_{1,\rho+1}$. It follows that the ordinary Laplace transform of w_ρ is the product of our γ with the transform of $f_{1,\rho+1}$, namely $(\lambda+1)^{\rho+1}$. Accordingly, *the probability density* (3.11) *has the Laplace transform*

$$\frac{1}{(\lambda+1)^{\rho+1}}\, e^{-t+t/(\lambda+1)}. \tag{3.12}$$

For $t = 1$ we see using the translation rule (1.9) that $\sqrt{x^\rho}\, I_\rho(2\sqrt{x})$ *has the ordinary transform* $\lambda^{-\rho-1}e^{1/\lambda}$.

(*f*) *Mixtures of exponential densities.* Let the density f be of the form

$$f(x) = \sum_{k=1}^{n} p_k a_k e^{-a_k x}, \quad p_k > 0, \qquad \sum_{k=1}^{n} p_k = 1 \tag{3.13}$$

where for definiteness we assume $0 < a_1 < \cdots < a_n$. The corresponding Laplace transform is given by

$$\varphi(\lambda) = \sum_{k=1}^{n} p_k \frac{a_k}{\lambda + a_k} = \frac{Q(\lambda)}{P(\lambda)} \tag{3.14}$$

where P is a polynomial of degree n with roots $-a_k$, and Q is a polynomial of degree $n-1$. Conversely, for any polynomial Q of degree $n-1$ the ratio $Q(\lambda)/P(\lambda)$ admits of a partial fraction expansion of the form (3.14) with

$$a_r p_r = \frac{Q(-a_r)}{P'(-a_r)} \tag{3.15}$$

[see 1; XI,(4.5)]. For (3.14) to correspond to a mixture (3.13) it is necessary and sufficient that $p_r > 0$ and that $Q(0)/P(0) = 1$. From the graph of P it is clear that $P'(-a_r)$ and $P'(-a_{r+1})$ are of opposite signs, and hence the same must be true of $Q(-a_r)$ and $Q(-a_{r+1})$. In other words, it is necessary that Q has a root $-b_r$ between $-a_r$ and $-a_{r+1}$. But as Q

[5] This result is due to H. Weber. The extremely difficult analytic proof is now replaced by an elementary proof in J. Soc. Industr. Appl. Math., vol. 14 (1966) pp. 864–875.

cannot have more than $n-1$ roots $-b_r$ we conclude that these must satisfy

$$0 < a_1 < b_1 < a_2 < b_2 < \cdots < b_{n-1} < a_n. \tag{3.16}$$

This guarantees that all p_r are of the same sign, and we reach the following conclusion: Let P and Q be polynomials of degree n and $n-1$, respectively, and $Q(0)/P(0) = 1$. *In order that* $Q(\lambda)/P(\lambda)$ *be the Laplace transform of a mixture* (3.13) *of exponential densities it is necessary and sufficient that the roots* $-a_r$ *of* P *and* $-b_r$ *of* Q *be distinct and* (*with proper numbering*) *satisfy* (3.16). ▶

4. COMPLETELY MONOTONE FUNCTIONS. INVERSION FORMULAS

As we saw in VII,2 a function f in $\overline{0, 1}$ is a generating function of a positive sequence $\{f_n\}$ iff f is absolutely monotone, that is, if f possesses positive derivatives $f^{(n)}$ of all orders. An analogous theorem holds for Laplace transforms, except that now the derivatives alternate in sign.

Definition 1. *A function* φ *on* $\overline{0, \infty}$ *is completely monotone if it possesses derivatives* $\varphi^{(n)}$ *of all orders and*

$$(-1)^n \varphi^{(n)}(\lambda) \geq 0, \qquad \lambda > 0. \tag{4.1}$$

As $\lambda \to 0$ the values $\varphi^{(n)}(\lambda)$ approach finite or infinite limits which we denote by $\varphi^{(n)}(0)$. Typical examples are $1/\lambda$ and $1/(1+\lambda)$.

The following beautiful theorem due to S. Bernstein (1928) was the starting point of much research, and the proof has been simplified by stages. We are able to give an extremely simple proof because the spade work was laid by the characterization of generating functions derived in theorem 2 of VII,2 as a consequence of the law of large numbers.

Theorem 1. *A function* φ *on* $\overline{0, \infty}$ *is the Laplace transform of a probability distribution* F, *iff it is completely monotone, and* $\varphi(0) = 1$.

We shall prove a version of this theorem which appears more general in form, but can actually be derived from the restricted version by an appeal to the translation principle explained in connection with (1.9).

Theorem 1a. *The function* φ *on* $\overline{0, \infty}$ *is completely monotone iff it is of the form*

$$\varphi(\lambda) = \int_0^\infty e^{-\lambda x}\, F\{dx\}, \qquad \lambda > 0, \tag{4.2}$$

where F *is not a necessarily finite measure on* $\overline{0, \infty}$.

(By our initial convention the interval of integration is *closed:* a possible atom of F at the origin has the effect that $\varphi(\infty) > 0$.)

Proof. The necessity of the condition follows by formal differentiation as in (2.4). Assuming φ to be completely monotone consider $\varphi(a-as)$ for fixed $a > 0$ and $0 < s < 1$ as a function of s. Its derivatives are evidently positive and by theorem 2 of VII,2 the Taylor expansion

$$\varphi(a-as) = \sum_{n=0}^{\infty} \frac{(-a)^n \varphi^{(n)}(a)}{n!} s^n \tag{4.3}$$

is valid for $0 \leq s < 1$. Thus

$$\varphi_a(\lambda) = \varphi(a-ae^{-\lambda/a}) = \sum_{n=0}^{\infty} \frac{(-a)^n \varphi^{(n)}(a)}{n!} e^{-n\lambda/a} \tag{4.4}$$

is the Laplace transform of an arithmetic measure attributing mass $(-a)^n\varphi^{(n)}(a)/n!$ to the point n/a (where $n = 0, 1, \ldots$). Now $\varphi_a(\lambda) \to \varphi(\lambda)$ as $a \to \infty$. By the extended continuity theorem there exists therefore a measure F such that $F_a \to F$ and φ is its Laplace transform. ▶

We have not only proved theorem 1*a*, but the relation $F_a \to F$ may be restated in the form of the important

Theorem 2. (*Inversion formula.*) *If* (4.2) *holds for* $\lambda > 0$, *then at all points of continuity*[6]

$$F(x) = \lim_{a\to\infty} \sum_{n \leq ax} \frac{(-a)^n}{n!} \varphi^{(n)}(a). \tag{4.5}$$

This formula is of great theoretical interest and permits various conclusions. The following boundedness criterion may serve as an example of particular interest for semi-group theory. (See problem 13.)

Corollary. *For* φ *to be of the form*

$$\varphi(\lambda) = \int_0^\infty e^{-\lambda x} f(x)\, dx \qquad \textit{where} \quad 0 \leq f \leq C \tag{4.6}$$

it is necessary and sufficient that

$$0 \leq \frac{(-a)^n \varphi^{(n)}(a)}{n!} \leq \frac{C}{a} \tag{4.7}$$

for all $a > 0$.

[6] The inversion formula (4.5) was derived in VII,(6.4) as a direct consequence of the law of large numbers. In VII,(6.6) we have an analogous *inversion formula for integrals of the form* (4.6) *with continuous* f (not necessarily positive).

Proof. Differentiating (4.6) under the integral we get (4.7) [see (2.4)]. Conversely, (4.7) implies that φ is completely monotone and hence the transform of a measure F. Substituting from (4.7) into (4.5) we conclude that

$$F(x_2) - F(x_1) \leq C(x_2 - x_1)$$

for any pair $x_1 < x_2$. This means that F has bounded difference ratios and hence F is the integral of a function $f \leq C$ (see V,3). ▶

Theorem 1 leads to simple *tests* that a given function is the Laplace transform of a probability distribution. The standard technique is illustrated by the proof of

Criterion 1. *If* φ *and* ψ *are completely monotone so is their product* $\varphi\psi$.

Proof. We show by induction that the derivatives of $\varphi\psi$ alternate in sign. Assume that for *every* pair φ, ψ of completely monotone functions the first n derivatives of $\varphi\psi$ alternate in sign. As $-\varphi'$ and $-\psi'$ are completely monotone the induction hypothesis applies to the products $-\varphi'\psi$ and $-\varphi\psi'$, and we conclude from $-(\varphi\psi)' = -\varphi'\psi - \varphi\psi'$ that in fact the first $n + 1$ derivatives of $\varphi\psi$ alternate in sign. Since the hypothesis is trivially true for $n = 1$ the criterion is proved. ▶

The same proof yields the useful

Criterion 2. *If* φ *is completely monotone and* ψ *a positive function with a completely monotone derivative then* $\varphi(\psi)$ *is completely monotone.* (*In particular*, $e^{-\psi}$ *is completely monotone.*)

Typical applications are given in section 6 and in the following example, which occurs frequently in the literature with unnecessary complications.

Example. (*a*) *An equation occurring in branching processes.* Let φ be the Laplace transform of a probability distribution F with expectation $0 < \mu \leq \infty$, and let $c > 0$. We prove that *the equation*

$$\beta(\lambda) = \varphi(\lambda + c - c\beta(\lambda)) \tag{4.8}$$

has a unique root $\beta(\lambda) \leq 1$ *and* β *is the Laplace transform of a distribution* B *which is proper iff* $\mu c \leq 1$, *defective otherwise.*

(See XIV,4 for applications and references.)

Proof. Consider the equation

$$\varphi(\lambda + c - cs) - s = 0 \tag{4.9}$$

for fixed $\lambda > 0$ and $0 \leq s \leq 1$. The left side is a convex function which assumes a negative value at $s = 1$ and a positive value at $s = 0$. It follows that there exists a unique root.

To prove that the root $\beta(\lambda)$ is a Laplace transform put $\beta_0 = 0$ and recursively $\beta_{n+1} = \varphi(\lambda+c-c\beta_n)$. Then $\beta_0 \leq \beta_1 \leq 1$ and since φ is decreasing this implies $\beta_1 \leq \beta_2 \leq 1$, and by induction $\beta_n \leq \beta_{n+1} \leq 1$. The limit of the bounded monotone sequence $\{\beta_n\}$ satisfies (4.8) and hence $\beta = \lim \beta_n$. Now $\beta_1(\lambda) = \varphi(\lambda+c)$ is completely monotone and criterion 2 shows recursively that $\beta_2, \beta_3, \ldots$ are completely monotone. By the continuity theorem the same is true of the limit β, and hence β is the Laplace transform of a measure B. Since $\beta(\lambda) \leq 1$ for all λ the total mass of B is $\beta(0) \leq 1$. It remains to decide under what conditions $\beta(0) = 1$.

By construction $s = \beta(0)$ is the *smallest* root of the equation

$$\varphi(c-cs) - s = 0. \tag{4.10}$$

Considered as a function of s the left side is convex; it is positive for $s = 0$ and vanishes for $s = 1$. A second root $s < 1$ exists therefore iff at $s = 1$ the derivative is positive, that is iff $-c\varphi'(0) > 1$. Otherwise $\beta(0) = 1$ and β is the Laplace transform of a proper probability distribution B. Hence B is proper iff $-c\varphi'(0) = c\mu \leq 1$. ►

5. TAUBERIAN THEOREMS

Let U be a measure concentrated on $\overline{0, \infty}$ and such that its Laplace transform

$$\omega(\lambda) = \int_0^\infty e^{-\lambda x}\, U\{dx\} \tag{5.1}$$

exists for $\lambda > 0$. It will be convenient to describe the measure U in terms of its improper distribution function defined for $x \geq 0$ by $U\{\overline{0, x}\}$. We shall see that under fairly general conditions the behavior of ω near the origin uniquely determines the asymptotic behavior of $U(x)$ as $x \to \infty$ and vice versa. Historically any relation describing the asymptotic behavior of U in terms of ω is called a Tauberian theorem, whereas theorems describing the behavior of ω in terms of U are usually called Abelian. We shall make no distinction between these two classes because our relations will be symmetric.

To avoid unsightly formulas involving reciprocals we introduce two positive variables t and τ related by

$$t\tau = 1. \tag{5.2}$$

Then $\tau \to 0$ when $t \to \infty$.

To understand the background of the Tauberian theorems note that for fixed t the change of variables $x = ty$ in (5.1) shows that $\omega(\tau\lambda)$ is the Laplace transform corresponding to the improper distribution function

$U(ty)$. Since ω decreases it is possible to find a sequence $\tau_1, \tau_2, \ldots \to 0$ such that as τ runs through it

$$\frac{\omega(\tau\lambda)}{\omega(\tau)} \to \gamma(\lambda) \tag{5.3}$$

with $\gamma(\lambda)$ finite at least for $\lambda > 1$. By the extended continuity theorem the limit γ is the Laplace transform of a measure G and as t runs through the reciprocals $t_k = 1/\tau_k$

$$\frac{U(tx)}{\omega(\tau)} \to G(x) \tag{5.4}$$

at all points of continuity of G. For $x = 1$ it is seen that the asymptotic behavior of $U(t)$ as $t \to \infty$ is intimately connected with the behavior of $\omega(t^{-1})$.

In principle we could formulate this fact as an all-embracing Tauberian theorem, but it would be too clumsy for practical use. To achieve reasonable simplicity we consider only the case where (5.3) is valid for *any* approach $\tau \to 0$, that is, when ω varies regularly at 0. The elementary lemma[7] 1 of VIII,8 states that the limit γ is necessarily of the form $\gamma(\lambda) = \lambda^{-\rho}$ with $\rho \geq 0$. The corresponding measure is given by $G(x) = x^\rho/\Gamma(\rho+1)$, and (5.4) implies that U varies regularly and the exponents of ω and U are the same in absolute value. We formulate this important result together with its converse in

Theorem 1. *Let U be a measure with a Laplace transform ω defined for $\lambda > 0$. Then each of the relations*

$$\frac{\omega(\tau\lambda)}{\omega(\tau)} \to \frac{1}{\lambda^\rho}, \qquad \tau \to 0 \tag{5.5}$$

and

$$\frac{U(tx)}{U(t)} \to x^\rho, \qquad t \to \infty \tag{5.6}$$

implies the other as well as

$$\omega(\tau) \sim U(t)\,\Gamma(\rho+1). \tag{5.7}$$

Proof. (*a*) Assume (5.5). The left side is the Laplace transform corresponding to $U(tx)/\omega(\tau)$, and by the extended continuity theorem this implies

$$\frac{U(tx)}{\omega(\tau)} \to \frac{x^\rho}{\Gamma(\rho+1)}. \tag{5.8}$$

For $x = 1$ we get (5.7), and substituting this back into (5.8) we get (5.6).

[7] This lemma is used *only* to justify the otherwise artificial form of the relations (5.5) and (5.6). The theory of regular variation is *not* used in this section [except for the side remark that (5.18) implies (5.16)].

(*b*) Assume (5.6). Taking Laplace transforms we get

$$\frac{\omega(\tau\lambda)}{U(t)} \to \frac{\Gamma(\rho+1)}{\lambda^\rho} \tag{5.9}$$

provided the extended continuity theorem is applicable, that is, *provided* the left-side remains bounded for some λ. As under (*a*) it is seen that (5.9) implies (5.7) and (5.5), and to prove the theorem it suffices to verify that $\omega(\tau)/U(t)$ remains bounded.

On partitioning the domain of integration by the points $t, 2t, 4t, \ldots$ it is clear that

$$\omega(\tau) \le \sum_0^\infty e^{-2^{n-1}}\, U(2^n t). \tag{5.10}$$

In view of (5.7) there exists a t_0 such that $U(2t) < 2^{\rho+1}\, U(t)$ for $t > t_0$. Repeated application of this inequality yields

$$\frac{\omega(\tau)}{U(t)} \le \sum_0^\infty 2^{n(\rho+1)} e^{-2^{n-1}} \tag{5.11}$$

and so the left side indeed remains bounded as $t \to \infty$. ▶

Examples. (*a*) $U(x) \sim \log^2 x$ as $x \to \infty$ iff $\omega(\lambda) \sim \log^2 \lambda$ as $\lambda \to 0$. Similarly $U(x) \sim \sqrt{x}$ iff $\omega(\lambda) \sim \frac{1}{2}\sqrt{\pi/\lambda}$.

(*b*) Let F be a probability distribution with Laplace transform φ. The measure $U\{dx\} = x\, F\{dx\}$ has the transform $-\varphi'$. Hence if $-\varphi'(\lambda) \sim \mu\lambda^{-\rho}$ as $\lambda \to \infty$ then

$$U(x) = \int_0^x y\, F\{dy\} \sim \frac{\mu}{\Gamma(\rho+1)}\, x^\rho, \qquad x \to \infty$$

and vice versa. This generalizes the differentiation rule (2.4) which is contained in (5.7) for $\rho = 0$. ▶

It is sometimes useful to know to what extent the theorem remains valid in the limit $\rho \to \infty$. We state the result in the form of a

Corollary. *If for some* $a > 1$ *as* $t \to \infty$

$$\text{either} \quad \frac{\omega(\tau a)}{\omega(\tau)} \to 0 \quad \text{or} \quad \frac{U(ta)}{U(t)} \to \infty \tag{5.12}$$

then

$$\frac{U(t)}{\omega(\tau)} \to 0. \tag{5.13}$$

Proof. The first relation in (5.12) implies that $\omega(\tau\lambda)/\omega(\tau) \to 0$ for $\lambda > a$ and by the extended continuity theorem $U(tx)/\omega(\tau) \to 0$ for all $x > 0$. The second relation in (5.12) entails (5.13) because

$$\omega(\tau) \geq \int_0^{at} e^{-x/t}\, U\{dx\} \geq e^{-a}\, U(ta). \qquad \blacktriangleright$$

In applications it is more convenient to express theorem 1 in terms of slow variation. We recall that a positive function L defined on $\overline{0, \infty}$ *varies slowly at* ∞ if for every fixed x

$$\frac{L(tx)}{L(t)} \to 1, \qquad t \to \infty. \tag{5.14}$$

L varies slowly at 0 if this relation holds as $t \to 0$, that is, if $L(1/x)$ varies slowly at ∞. Evidently U satisfies (5.6) iff $U(x)/x^\rho$ varies slowly at ∞ and similarly (5.5) holds iff $\lambda^\rho \omega(\lambda)$ varies slowly at 0. Consequently theorem 1 may be rephrased as follows.

Theorem 2. *If L is slowly varying at infinity and $0 \leq \rho < \infty$, then each of the relations*

$$\omega(\tau) \sim \tau^{-\rho} L\left(\frac{1}{\tau}\right), \qquad \tau \to 0, \tag{5.15}$$

and

$$U(t) \sim \frac{1}{\Gamma(\rho+1)}\, t^\rho L(t), \qquad t \to \infty \tag{5.16}$$

implies the other.

Theorem 2 has a glorious history. The implication (5.16) → (5.15) (from the measure to the transform) is called an Abelian theorem; the converse (5.15) → (5.16) (from transform to measure), a Tauberian theorem. In the usual setup, the two theorems are entirely separated, the Tauberian part causing the trouble. In a famous paper G. H. Hardy and J. E. Littlewood treated the case $\omega(\lambda) \sim \lambda^{-\rho}$ by difficult calculations. In 1930, J. Karamata created a sensation by a simplified proof for this special case. (This proof is still found in texts on complex variables and Laplace transforms.) Soon afterwards he introduced the class of regularly varying functions and proved theorem 2; the proof was too complicated for textbooks, however. The notion of slow variation was introduced by R. Schmidt about 1925 in the same connection. Our proof simplifies and unifies the theory and leads to the little-known, but useful, corollary.

A great advantage of our proof is that it applies without change when the roles of infinity and zero are interchanged, that is, if $\tau \to \infty$ while $t \to 0$. In this way we get the dual theorem connecting the behavior of ω at infinity with that of U at the origin. [It will not be used in this book except to derive (6.2).]

Theorem 3. *The last two theorems and the corollary remain valid when the roles of the origin and infinity are interchanged, that is, for $\tau \to \infty$ and $t \to 0$.*

Theorem 2 represents the main result of this section, but for completeness we derive two useful complements. First of all, when U has a density $U' = u$ it is desirable to obtain estimates for u. This problem cannot be treated in full generality, because a well-behaved distribution U can have an extremely ill-behaved density u. In most applications, however, the density u will be *ultimately monotone*, that is, monotone in some interval $\overline{x_0, \infty}$. For such densities we have

Theorem 4.[8] *Let* $0 < \rho < \infty$. *If* U *has an ultimately monotone derivative* u *then as* $\lambda \to 0$ *and* $x \to \infty$, *respectively,*

$$\omega(\lambda) \sim \frac{1}{\lambda^\rho} L\left(\frac{1}{\lambda}\right) \quad \textit{iff} \quad u(x) \sim \frac{1}{\Gamma(\rho)} x^{\rho-1} L(x). \tag{5.17}$$

(For a formally stronger version see problem 16.)

Proof. The assertion is an immediate consequence of theorem 2 and the following

Lemma. *Suppose that* U *has an ultimately monotone density* u. *If* (5.16) *holds with* $\rho > 0$ *then*

$$u(x) \sim \rho U(x)/x, \qquad x \to \infty. \tag{5.18}$$

[Conversely, (5.18) implies (5.16) even if u is not monotone. This is contained in VIII,(9.6) with $Z = u$ and $p = 0$.]

Proof. For $0 < a < b$

$$\frac{U(tb) - U(ta)}{U(t)} = \int_a^b \frac{u(ty)t}{U(t)}\, dy. \tag{5.19}$$

As $t \to \infty$ the left side tends to $b^\rho - a^\rho$. For sufficiently large t the integrand is monotone, and then (5.16) implies that it remains bounded as $t \to \infty$. By the selection theorem of VIII,6 there exists therefore a sequence $t_1, t_2, \ldots \to \infty$ such that as t runs through it

$$\frac{u(ty)t}{U(t)} \to \psi(y) \tag{5.20}$$

at all points of continuity. It follows that the integral of ψ over $\overline{a, b}$ equals $b^\rho - a^\rho$, and so $\psi(y) = \rho y^{\rho-1}$. This limit being independent of the sequence $\{t_k\}$ the relation (5.20) is true for an arbitrary approach $t \to \infty$, and for $y = 1$ it reduces to (5.18). ▶

[8] This includes the famous Tauberian theorem of E. Landau. Our proof serves as a new example of how the selection theorem obviates analytical intricacies.