de Gruyter Studies in Mathematics 37

Editors: Carsten Carstensen · Nicola Fusco
Niels Jacob · Karl-Hermann Neeb

de Gruyter Studies in Mathematics

1 Riemannian Geometry, 2nd rev. ed., *Wilhelm P. A. Klingenberg*
2 Semimartingales, *Michel Métivier*
3 Holomorphic Functions of Several Variables, *Ludger Kaup and Burchard Kaup*
4 Spaces of Measures, *Corneliu Constantinescu*
5 Knots, 2nd rev. and ext. ed., *Gerhard Burde and Heiner Zieschang*
6 Ergodic Theorems, *Ulrich Krengel*
7 Mathematical Theory of Statistics, *Helmut Strasser*
8 Transformation Groups, *Tammo tom Dieck*
9 Gibbs Measures and Phase Transitions, *Hans-Otto Georgii*
10 Analyticity in Infinite Dimensional Spaces, *Michel Hervé*
11 Elementary Geometry in Hyperbolic Space, *Werner Fenchel*
12 Transcendental Numbers, *Andrei B. Shidlovskii*
13 Ordinary Differential Equations, *Herbert Amann*
14 Dirichlet Forms and Analysis on Wiener Space, *Nicolas Bouleau and Francis Hirsch*
15 Nevanlinna Theory and Complex Differential Equations, *Ilpo Laine*
16 Rational Iteration, *Norbert Steinmetz*
17 Korovkin-type Approximation Theory and its Applications, *Francesco Altomare and Michele Campiti*
18 Quantum Invariants of Knots and 3-Manifolds, *Vladimir G. Turaev*
19 Dirichlet Forms and Symmetric Markov Processes, *Masatoshi Fukushima, Yoichi Oshima and Masayoshi Takeda*
20 Harmonic Analysis of Probability Measures on Hypergroups, *Walter R. Bloom and Herbert Heyer*
21 Potential Theory on Infinite-Dimensional Abelian Groups, *Alexander Bendikov*
22 Methods of Noncommutative Analysis, *Vladimir E. Nazaikinskii, Victor E. Shatalov and Boris Yu. Sternin*
23 Probability Theory, *Heinz Bauer*
24 Variational Methods for Potential Operator Equations, *Jan Chabrowski*
25 The Structure of Compact Groups, 2nd rev. and aug. ed., *Karl H. Hofmann and Sidney A. Morris*
26 Measure and Integration Theory, *Heinz Bauer*
27 Stochastic Finance, 2nd rev. and ext. ed., *Hans Föllmer and Alexander Schied*
28 Painlevé Differential Equations in the Complex Plane, *Valerii I. Gromak, Ilpo Laine and Shun Shimomura*
29 Discontinuous Groups of Isometries in the Hyperbolic Plane, *Werner Fenchel and Jakob Nielsen*
30 The Reidemeister Torsion of 3-Manifolds, *Liviu I. Nicolaescu*
31 Elliptic Curves, *Susanne Schmitt and Horst G. Zimmer*
32 Circle-valued Morse Theory, *Andrei V. Pajitnov*
33 Computer Arithmetic and Validity, *Ulrich Kulisch*
34 Feynman-Kac-Type Theorems and Gibbs Measures on Path Space, *József Lörinczi, Fumio Hiroshima and Volker Betz*
35 Integral Representation Theory, *Jaroslas Lukeš, Jan Malý, Ivan Netuka and Jiří Spurný*
36 Introduction to Harmonic Analysis and Generalized Gelfand Pairs, *Gerrit van Dijk*
37 Bernstein Functions, *René Schilling, Renming Song and Zoran Vondraček*

René L. Schilling
Renming Song
Zoran Vondraček

Bernstein Functions

Theory and Applications

De Gruyter

Authors

René L. Schilling
Institut für Stochastik
Technische Universität Dresden
01062 Dresden, Germany

Renming Song
Department of Mathematics
University of Illinois
Urbana, IL 61801, USA

Zoran Vondraček
Department of Mathematics
University of Zagreb
10000 Zagreb, Croatia

Mathematics Subject Classification 2010: 26-02, 30-02, 44-02, 31B, 46L, 47D, 60E, 60G, 60J, 62E

Keywords: Bernstein function, complete Bernstein function, completely monotone function, generalized diffusion, Krein's string theory, Laplace transform, Nevanlinna–Pick function, matrix monotone function, operational calculus, positive and negative definite function, potential theory, Stieltjes transform, subordination

Updates and corrections can be found on www.degruyter.com/bernstein-functions

∞ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

ISBN 978-3-11-021530-4

Printed in Germany.
Typesetting: Da-TeX Gerd Blumenstein, Leipzig, www.da-tex.de.
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen.

Contents

Preface

Bernstein functions and the important subclass of complete Bernstein functions appear in various fields of mathematics—often with different definitions and under different names. Probabilists, for example, know Bernstein functions as Laplace exponents, and in harmonic analysis they are called negative definite functions. Complete Bernstein functions are used in complex analysis under the name Pick or Nevanlinna functions, while in matrix analysis and operator theory, the name operator monotone function is more common. When studying the positivity of solutions of Volterra integral equations, various types of kernels appear which are related to Bernstein functions. There exists a considerable amount of literature on each of these classes, but only a handful of texts observe the connections between them or use methods from several mathematical disciplines.

This book is about these connections. Although many readers may not be familiar with the name *Bernstein function*, and even fewer will have heard of *complete Bernstein functions*, we are certain that most have come across these families in their own research. Most likely only certain aspects of these classes of functions were important for the problems at hand and they could be solved on an *ad hoc* basis. This explains quite a few of the rediscoveries in the field, but also that many results and examples are scattered throughout the literature; the exceedingly rich structure connecting this material got lost in the process. Our motivation for writing this book was to point out many of these connections and to present the material in a unified way. We hope that our presentation is accessible to researchers and graduate students with different backgrounds. The results as such are mostly known, but our approach and some of the proofs are new: we emphasize the structural analogies between the function classes which we believe is a very good way to approach the topic. Since it is always important to know explicit examples, we took great care to collect many of them in the tables which form the last part of the book.

Completely monotone functions—these are the Laplace transforms of measures on the half-line $[0, \infty)$—and Bernstein functions are intimately connected. The derivative of a Bernstein function is completely monotone; on the other hand, the primitive of a completely monotone function is a Bernstein function if it is positive. This observation leads to an integral representation for Bernstein functions: the Lévy–Khintchine formula on the half-line

$$f(\lambda) = a + b\lambda + \int_{(0,\infty)} (1 - e^{-\lambda t})\, \mu(dt), \quad \lambda > 0.$$

Although this is familiar territory to a probabilist, this way of deriving the Lévy–Khintchine formula is not the usual one in probability theory. There are many more connections between Bernstein and completely monotone functions. For example, f is a Bernstein function if, and only if, for all completely monotone functions g the composition $g \circ f$ is completely monotone. Since g is a Laplace transform, it is enough to check this for the kernel of the Laplace transform, i.e. the basic completely monotone functions $g(\lambda) = e^{-t\lambda}$, $t > 0$.

A similar connection exists between the Laplace transforms of completely monotone functions, that is, *double Laplace* or *Stieltjes transforms*, and *complete* Bernstein functions. A function f is a complete Bernstein function if, and only if, for each $t > 0$ the composition $(t + f(\lambda))^{-1}$ of the Stieltjes kernel $(t + \lambda)^{-1}$ with f is a Stieltjes function. Note that $(t + \lambda)^{-1}$ is the Laplace transform of $e^{-t\lambda}$ and thus the functions $(t + \lambda)^{-1}$, $t > 0$, are the basic Stieltjes functions. With some effort one can check that complete Bernstein functions are exactly those Bernstein functions where the measure μ in the Lévy–Khintchine formula has a completely monotone density with respect to Lebesgue measure. From there it is possible to get a surprising geometric characterization of these functions: they are non-negative on $(0, \infty)$, have an analytic extension to the cut complex plane $\mathbb{C} \setminus (-\infty, 0]$ and preserve upper and lower half-planes. A familiar sight for a classical complex analyst: these are the Nevanlinna functions. One could go on with such connections, delving into continued fractions, continue into interpolation theory and from there to operator monotone functions ...

Let us become a bit more concrete and illustrate our approach with an example. The fractional powers $\lambda \mapsto \lambda^\alpha$, $\lambda > 0$, $0 < \alpha < 1$, are easily among the most prominent (complete) Bernstein functions. Recall that

$$f_\alpha(\lambda) := \lambda^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (1 - e^{-\lambda t})\, t^{-\alpha-1}\, dt. \tag{1}$$

Depending on your mathematical background, there are many different ways to derive and to interpret (1), but we will follow probabilists' custom and call (1) the Lévy–Khintchine representation of the Bernstein function f_α. At this point we do not want to go into details, instead we insist that one should read this formula as an integral representation of f_α with the kernel $(1 - e^{-\lambda t})$ and the measure $c_\alpha\, t^{-\alpha-1}\, dt$.

This brings us to negative powers, and there is another classical representation

$$\lambda^{-\beta} = \frac{1}{\Gamma(\beta)} \int_0^\infty e^{-\lambda t} t^{\beta-1}\, dt, \quad \beta > 0, \tag{2}$$

showing that $\lambda \mapsto \lambda^{-\beta}$ is a completely monotone function. It is no accident that the reciprocal of the Bernstein function λ^α, $0 < \alpha < 1$, is completely monotone, nor is it an accident that the representing measure $c_\alpha\, t^{-\alpha-1}\, dt$ of λ^α has a completely monotone density. Inserting the representation (2) for $t^{-\alpha-1}$ into (1) and working out the double integral and the constant, leads to the second important formula for the fractional powers,

$$\lambda^\alpha = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty \frac{\lambda}{\lambda + t}\, t^{\alpha-1}\, dt. \tag{3}$$

We will call this representation of λ^α the Stieltjes representation. To explain why this is indeed an appropriate name, let us go back to (2) and observe that $t^{\alpha-1}$ is a Laplace transform. This shows that $\lambda^{-\alpha}$, $\alpha > 0$, is a double Laplace or Stieltjes transform. Another non-random coincidence is that

$$\frac{f_\alpha(\lambda)}{\lambda} = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty \frac{1}{\lambda+t}\, t^{\alpha-1}\, dt$$

is a Stieltjes transform and so is $\lambda^{-\alpha} = 1/f_\alpha(\lambda)$. This we can see if we replace $t^{\alpha-1}$ by its integral representation (2) and use Fubini's theorem:

$$\frac{1}{f_\alpha(\lambda)} = \lambda^{-\alpha} = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty \frac{1}{\lambda+t}\, t^{-\alpha}\, dt. \tag{4}$$

It is also easy to see that the fractional powers $\lambda \mapsto \lambda^\alpha = \exp(\alpha \log \lambda)$ extend analytically to the cut complex plane $\mathbb{C} \setminus (-\infty, 0]$. Moreover, z^α maps the upper half-plane into itself; actually it contracts all arguments by the factor α. Apart from some technical complications this allows to surround the singularities of f_α—which are all in $(-\infty, 0)$—by an integration contour and to use Cauchy's theorem for the half-plane to bring us back to the representation (3).

Coming back to the fractional powers λ^α, $0 < \alpha < 1$, we derive yet another representation formula. First note that $\lambda^\alpha = \int_0^\lambda \alpha s^{-(1-\alpha)}\, ds$ and that the integrand $s^{-(1-\alpha)}$ is a Stieltjes function which can be expressed as in (4). Fubini's theorem and the elementary equality

$$\int_0^\lambda \frac{1}{t+s}\, ds = \log\left(1 + \frac{\lambda}{t}\right)$$

yield

$$\lambda^\alpha = \frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^\infty \log\left(1 + \frac{\lambda}{t}\right)\, t^{\alpha-1}\, dt. \tag{5}$$

This representation will be called the Thorin representation of λ^α. Not every complete Bernstein function has a Thorin representation. The critical step in deriving (5) was the fact that the derivative of λ^α is a Stieltjes function.

What has been explained for fractional powers can be extended in various directions. On the level of functions, the structure of (1) is characteristic for the class $\mathcal{BF}$ of Bernstein functions, (3) for the class $\mathcal{CBF}$ of complete Bernstein functions, and (5) for the Thorin–Bernstein functions $\mathcal{TBF}$. If we consider $\exp(-tf)$ with f from $\mathcal{BF}$, $\mathcal{CBF}$ or $\mathcal{TBF}$, we are led to the corresponding families of completely monotone functions and measures. Apart from some minor conditions, these are the infinitely divisible distributions ID, the Bondesson class of measures BO and the generalized Gamma convolutions GGC. The diagrams in Remark 9.17 illustrate these connections. If we replace (formally) λ by $-A$, where A is a negative semi-definite matrix or a dissipative closed operator, then we get from (1) and (2) the classical formulae for fractional powers, while (3) turns into Balakrishnan's formula. Considering $\mathcal{BF}$ and $\mathcal{CBF}$ we obtain a fully-fledged functional calculus for generators and potential operators. Since complete Bernstein functions are operator monotone functions we can even recover the famous Heinz–Kato inequality.

Let us briefly describe the content and the structure of the book. It consists of three parts. The first part, Chapters 1–10, introduces the basic classes of functions: the positive definite functions comprising the completely monotone, Stieltjes and Hirsch functions, and the negative definite functions which consist of the Bernstein functions and their subfamilies—special, complete and Thorin–Bernstein functions. Two probabilistic intermezzi explore the connection between Bernstein functions and certain classes of probability measures. Roughly speaking, for every Bernstein function f the functions $\exp(-tf)$, $t > 0$, are completely monotone, which implies that $\exp(-tf)$ is the Laplace transform of an infinitely divisible sub-probability measure. This part of the book is essentially self-contained and should be accessible to non-specialists and graduate students.

In the second part of the book, Chapter 11 through Chapter 14, we turn to applications of Bernstein and complete Bernstein functions. The choice of topics reflects our own interests and is by no means complete. Notable omissions are applications in integral equations and continued fractions.

Among the topics are the spectral theorem for self-adjoint operators in a Hilbert space and a characterization of all functions which preserve the order (in quadratic form sense) of dissipative operators. Bochner's subordination plays a fundamental role in Chapter 12 where also a functional calculus for subordinate generators is developed. This calculus generalizes many formulae for fractional powers of closed operators. As another application of Bernstein and complete Bernstein functions we establish estimates for the eigenvalues of subordinate Markov processes. This is continued in Chapter 13 which contains a detailed study of excessive functions of killed and subordinate killed Brownian motion. Finally, Chapter 14 is devoted to two results in the theory of generalized diffusions, both related to complete Bernstein functions through Kreĭn's theory of strings. Many of these results appear for the first time in a monograph.

The third part of the book is formed by extensive tables of complete Bernstein functions. The main criteria for inclusion in the tables were the availability of explicit representations and the appearance in mathematical literature.

In the appendix we collect, for the readers' convenience, some supplementary results.

We started working on this monograph in summer 2006, during a one-month workshop organized by one of us at the University of Marburg. Over the years we were supported by our universities: Institut für Stochastik, Technische Universität Dresden, Department of Mathematics, University of Illinois, and Department of Mathematics, University of Zagreb. We thank our colleagues for a stimulating working environment and for many helpful discussions. Considerable progress was made during the two week Research in Pairs programme at the Mathematisches Forschungsinstitut in Oberwolfach where we could enjoy the research atmosphere and the wonderful library. Our sincere thanks go to the institute and its always helpful staff.

Panki Kim and Hrvoje Šikić read substantial parts of the manuscript. We are grateful for their comments which helped to improve the text. We thank the series editor Niels Jacob for his interest and constant encouragement. It is a pleasure to acknowledge the support of our publisher, Walter de Gruyter, and its editors Robert Plato and Simon Albroscheit.

Writing this book would have been impossible without the support of our families. So thank you, Herta, Jean and Sonja, for your patience and understanding.

Dresden, Urbana and Zagreb
October 2009

René Schilling
Renming Song
Zoran Vondraček

Index of notation

This index is intended to aid cross-referencing, so notation that is specific to a single section is generally not listed. Some symbols are used locally, without ambiguity, in senses other than those given below; numbers following an entry are page numbers.

Unless otherwise stated, binary operations between functions such as $f \pm g$, $f \cdot g$, $f \wedge g$, $f \vee g$, comparisons $f \leqslant g$, $f < g$ or limiting relations $f_j \xrightarrow{j\to\infty} f$, $\lim_j f_j$, $\liminf_j f_j$, $\limsup_j f_j$, $\sup_j f_j$ or $\inf_j f_j$ are always understood pointwise.

Chapter 1

Laplace transforms and completely monotone functions

In this chapter we collect some preliminary material which we need later on in order to study Bernstein functions.

As usual, we define the *(one-sided) Laplace transform* of a function $m : [0, \infty) \to [0, \infty)$ or a measure μ on the half-line $[0, \infty)$ by

$$\mathscr{L}(m; \lambda) := \int_0^\infty e^{-\lambda t} m(t)\, dt \quad \text{or} \quad \mathscr{L}(\mu; \lambda) := \int_{[0,\infty)} e^{-\lambda t}\, \mu(dt), \tag{1.1}$$

respectively, whenever these integrals converge. Obviously, $\mathscr{L}m = \mathscr{L}\mu_m$ if $\mu_m(dt)$ denotes the measure $m(t)\, dt$.

The following real-analysis lemma is helpful in order to show that finite measures are uniquely determined in terms of their Laplace transforms.

Lemma 1.1. *We have for all* $t, x \geqslant 0$

$$\lim_{\lambda\to\infty} e^{-\lambda t} \sum_{k \leqslant \lambda x} \frac{(\lambda t)^k}{k!} = \mathbb{1}_{[0,x]}(t). \tag{1.2}$$

Proof. Let us rewrite (1.2) in probabilistic terms: if X is a Poisson random variable with parameter λt, (1.2) states that

$$\lim_{\lambda\to\infty} \mathbb{P}(X \leqslant \lambda x) = \mathbb{1}_{[0,x]}(t).$$

From the basic formulae for the mean value and the variance of Poisson random variables, $\mathbb{E}X = \lambda t$ and $\operatorname{Var} X = \mathbb{E}((X - \lambda t)^2) = \lambda t$, we find for $t > x$ with Chebyshev's inequality

$$\begin{aligned}
\mathbb{P}(X \leqslant \lambda x) &\leqslant \mathbb{P}\big(|X - \lambda t| \geqslant \lambda(t - x)\big) \\
&\leqslant \frac{\mathbb{E}((X - \lambda t)^2)}{\lambda^2 (t - x)^2} \\
&= \frac{\lambda t}{\lambda^2 (t - x)^2} \xrightarrow{\lambda\to\infty} 0.
\end{aligned}$$

If $t \leqslant x$, a similar calculation yields

$$\begin{aligned}\mathbb{P}(X \leqslant \lambda x) &= 1 - \mathbb{P}\big(X - \lambda t > \lambda(x-t)\big)\\ &\geqslant 1 - \mathbb{P}\big(|X - \lambda t| > \lambda(x-t)\big) \xrightarrow{\lambda\to\infty} 1 - 0,\end{aligned}$$

and the claim follows. □

Proposition 1.2. *A measure μ supported in $[0,\infty)$ is finite if, and only if, $\mathscr{L}(\mu; 0+) < \infty$. The measure μ is uniquely determined by its Laplace transform.*

Proof. The first part of the assertion follows from monotone convergence since we have $\mu[0,\infty) = \int_{[0,\infty)} 1\, d\mu = \lim_{\lambda\to 0} \int_{[0,\infty)} e^{-\lambda t}\,\mu(dt)$.

For the uniqueness part we use first the differentiation lemma for parameter dependent integrals to get

$$(-1)^k \mathscr{L}^{(k)}(\mu;\lambda) = \int_{[0,\infty)} e^{-\lambda t}\, t^k\, \mu(dt).$$

Therefore,

$$\begin{aligned}\sum_{k\leqslant\lambda x} (-1)^k \mathscr{L}^{(k)}(\mu;\lambda)\,\frac{\lambda^k}{k!} &= \sum_{k\leqslant\lambda x} \int_{[0,\infty)} \frac{(\lambda t)^k}{k!} e^{-\lambda t}\,\mu(dt)\\ &= \int_{[0,\infty)} \sum_{k\leqslant\lambda x} \frac{(\lambda t)^k}{k!} e^{-\lambda t}\,\mu(dt)\end{aligned}$$

and we conclude with Lemma 1.1 and dominated convergence that

$$\lim_{\lambda\to\infty} \sum_{k\leqslant\lambda x} (-1)^k \mathscr{L}^{(k)}(\mu;\lambda)\,\frac{\lambda^k}{k!} = \int_{[0,\infty)} \mathbb{1}_{[0,x]}(t)\,\mu(dt) = \mu[0,x]. \tag{1.3}$$

This shows that μ can be recovered from (all derivatives of) its Laplace transform. □

It is possible to characterize the range of Laplace transforms. For this we need the notion of complete monotonicity.

Definition 1.3. A function $f : (0,\infty) \to \mathbb{R}$ is a *completely monotone function* if f is of class C^∞ and

$$(-1)^n f^{(n)}(\lambda) \geqslant 0 \quad \text{for all } n \in \mathbb{N}\cup\{0\} \text{ and } \lambda > 0. \tag{1.4}$$

The family of all completely monotone functions will be denoted by $\mathcal{CM}$.

The conditions (1.4) are often referred to as *Bernstein–Hausdorff–Widder conditions*. The next theorem is known as *Bernstein's theorem*.

The version given below appeared for the first time in [34] and independently in [287]. Subsequent proofs were given in [98] and [86]. The theorem may be also considered as an example of the general integral representation of points in a convex cone by means of its extremal elements. See Theorem 4.8 and [69] for an elementary exposition. The following short and elegant proof is taken from [212].

Theorem 1.4 (Bernstein). *Let $f : (0,\infty) \to \mathbb{R}$ be a completely monotone function. Then it is the Laplace transform of a unique measure μ on $[0,\infty)$, i.e. for all $\lambda > 0$,*

$$f(\lambda) = \mathscr{L}(\mu;\lambda) = \int_{[0,\infty)} e^{-\lambda t}\,\mu(dt).$$

Conversely, whenever $\mathscr{L}(\mu;\lambda) < \infty$ for every $\lambda > 0$, $\lambda \mapsto \mathscr{L}(\mu;\lambda)$ is a completely monotone function.

Proof. Assume first that $f(0+) = 1$ and $f(+\infty) = 0$. Let $\lambda > 0$. For any $a > 0$ and any $n \in \mathbb{N}$, we see by Taylor's formula

$$\begin{aligned} f(\lambda) &= \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(\lambda - a)^k + \int_a^\lambda \frac{f^{(n)}(s)}{(n-1)!}(\lambda - s)^{n-1}\,ds \\ &= \sum_{k=0}^{n-1} \frac{(-1)^k f^{(k)}(a)}{k!}(a-\lambda)^k + \int_\lambda^a \frac{(-1)^n f^{(n)}(s)}{(n-1)!}(s-\lambda)^{n-1}\,ds. \end{aligned} \tag{1.5}$$

If $a > \lambda$, then by the assumption all terms are non-negative. Let $a \to \infty$. Then

$$\begin{aligned} \lim_{a\to\infty} \int_\lambda^a \frac{(-1)^n f^{(n)}(s)}{(n-1)!}(s-\lambda)^{n-1}\,ds &= \int_\lambda^\infty \frac{(-1)^n f^{(n)}(s)}{(n-1)!}(s-\lambda)^{n-1}\,ds \\ &\leqslant f(\lambda). \end{aligned}$$

This implies that the sum in (1.5) converges for every $n \in \mathbb{N}$ as $a \to \infty$. Thus, every term converges as $a \to \infty$ to a non-negative limit. For $n \geqslant 0$ let

$$\rho_n(\lambda) = \lim_{a\to\infty} \frac{(-1)^n f^{(n)}(a)}{n!}(a-\lambda)^n.$$

This limit does not depend on $\lambda > 0$. Indeed, for $\kappa > 0$,

$$\begin{aligned} \rho_n(\kappa) &= \lim_{a\to\infty} \frac{(-1)^n f^{(n)}(a)}{n!}(a-\kappa)^n \\ &= \lim_{a\to\infty} \frac{(-1)^n f^{(n)}(a)}{n!}(a-\lambda)^n \frac{(a-\kappa)^n}{(a-\lambda)^n} = \rho_n(\lambda). \end{aligned}$$

Let $c_n = \sum_{k=0}^{n-1} \rho_k(\lambda)$. Then

$$f(\lambda) = c_n + \int_\lambda^\infty \frac{(-1)^n f^{(n)}(s)}{(n-1)!} (s-\lambda)^{n-1} \, ds.$$

Clearly, $f(\lambda) \geqslant c_n$ for all $\lambda > 0$. Let $\lambda \to \infty$. Since $f(+\infty) = 0$, it follows that $c_n = 0$ for every $n \in \mathbb{N}$. Thus we have obtained the following integral representation of the function f:

$$f(\lambda) = \int_\lambda^\infty \frac{(-1)^n f^{(n)}(s)}{(n-1)!} (s-\lambda)^{n-1} \, ds. \tag{1.6}$$

By the monotone convergence theorem

$$1 = \lim_{\lambda \to 0} f(\lambda) = \int_0^\infty \frac{(-1)^n f^{(n)}(s)}{(n-1)!} s^{n-1} \, ds. \tag{1.7}$$

Let

$$f_n(s) = \frac{(-1)^n}{n!} f^{(n)}\left(\frac{n}{s}\right) \left(\frac{n}{s}\right)^{n+1}. \tag{1.8}$$

Using (1.7) and changing variables according to s/t, it follows that for every $n \in \mathbb{N}$, f_n is a probability density function on $(0, \infty)$. Moreover, the representation (1.6) can be rewritten as

$$\begin{aligned} f(\lambda) &= \int_0^\infty \left(1 - \frac{\lambda}{s}\right)_+^{n-1} \frac{(-1)^n f^{(n)}(s)}{(n-1)!} s^{n-1} \, ds \\ &= \int_0^\infty \left(1 - \frac{\lambda t}{n}\right)_+^{n-1} f_n(t) \, dt. \end{aligned} \tag{1.9}$$

By Helly's selection theorem, Corollary A.8, there exist a subsequence $(n_k)_{k \geqslant 1}$ and a probability measure μ on $(0, \infty)$ such that $f_{n_k}(t)\, dt$ converges weakly to $\mu(dt)$. Further, for every $\lambda > 0$,

$$\lim_{n \to \infty} \left(1 - \frac{\lambda t}{n}\right)_+^{n-1} = e^{-\lambda t}$$

uniformly in $t \in (0, \infty)$. By taking the limit in (1.9) along the subsequence $(n_k)_{k \geqslant 1}$, it follows that

$$f(\lambda) = \int_{(0,\infty)} e^{-\lambda t} \, \mu(dt).$$

Uniqueness of μ follows from Proposition 1.2.

Assume now that $f(0+) < \infty$ and $f(+\infty) = 0$. By looking at $f/f(0+)$ we see that the representing measure for f is uniquely given by $f(0+)\mu$.

Now let f be an arbitrary completely monotone function with $f(+\infty) = 0$. For every $a > 0$, define $f_a(\lambda) := f(\lambda + a)$, $\lambda > 0$. Then f_a is a completely monotone function with $f_a(0+) = f(a) < \infty$ and $f_a(+\infty) = 0$. By what has been already proved, there exists a unique finite measure μ_a on $(0, \infty)$ such that $f_a(\lambda) = \int_{(0,\infty)} e^{-\lambda t}\, \mu_a(dt)$. It follows easily that for $b > 0$ we have $e^{at}\mu_a(dt) = e^{bt}\mu_b(dt)$. This shows that we can consistently define the measure μ on $(0, \infty)$ by $\mu(dt) = e^{at}\mu_a(dt)$, $a > 0$. In particular, the representing measure μ is uniquely determined by f. Now, for $\lambda > 0$,

$$
\begin{aligned}
f(\lambda) = f_{\lambda/2}(\lambda/2) &= \int_{(0,\infty)} e^{(-\lambda/2)t}\, \mu_{\lambda/2}(dt) \\
&= \int_{(0,\infty)} e^{-\lambda t} e^{(\lambda/2)t}\, \mu_{\lambda/2}(dt) = \int_{(0,\infty)} e^{-\lambda t}\, \mu(dt).
\end{aligned}
$$

Finally, if $f(+\infty) = c > 0$, add $c\delta_0$ to μ.

For the converse we set $f(\lambda) := \mathscr{L}(\mu; \lambda)$. Fix $\lambda > 0$ and pick $\epsilon \in (0, \lambda)$. Since $t^n = \epsilon^{-n}(\epsilon t)^n \leqslant n!\epsilon^{-n}e^{\epsilon t}$ for all $t > 0$, we find

$$
\int_{[0,\infty)} t^n\, e^{-\lambda t}\, \mu(dt) \leqslant \frac{n!}{\epsilon^n} \int_{[0,\infty)} e^{-(\lambda-\epsilon)t}\, \mu(dt) = \frac{n!}{\epsilon^n} \mathscr{L}(\mu; \lambda - \epsilon)
$$

and this shows that we may use the differentiation lemma for parameter dependent integrals to get

$$
(-1)^n f^{(n)}(\lambda) = (-1)^n \int_{[0,\infty)} \frac{d^n}{d\lambda^n} e^{-\lambda t}\, \mu(dt) = \int_{[0,\infty)} t^n\, e^{-\lambda t}\, \mu(dt) \geqslant 0. \quad \square
$$

Remark 1.5. The last formula in the proof of Theorem 1.4 shows, in particular, that $f^{(n)}(\lambda) \neq 0$ for all $n \geqslant 1$ and all $\lambda > 0$ unless $f \in \mathcal{CM}$ is identically constant.

Corollary 1.6. *The set $\mathcal{CM}$ of completely monotone functions is a convex cone, i.e.*

$$
sf_1 + tf_2 \in \mathcal{CM} \quad \textit{for all} \;\; s, t \geqslant 0 \;\; \textit{and} \;\; f_1, f_2 \in \mathcal{CM},
$$

which is closed under multiplication, i.e.

$$
\lambda \mapsto f_1(\lambda) f_2(\lambda) \;\; \textit{is in} \;\; \mathcal{CM} \;\; \textit{for all} \;\; f_1, f_2 \in \mathcal{CM},
$$

and under pointwise convergence:

$$
\mathcal{CM} = \overline{\left\{\mathscr{L}\mu \; : \; \mu \textit{ is a finite measure on } [0, \infty)\right\}}
$$

(the closure is taken with respect to pointwise convergence).

Proof. That $\mathcal{CM}$ is a convex cone follows immediately from the definition of a completely monotone function or, alternatively, from the representation formula in Theorem 1.4.

If μ_j denotes the representing measure of f_j, $j = 1, 2$, the convolution

$$\mu[0,u] := \mu_1 \star \mu_2[0,u] := \iint_{[0,\infty)\times[0,\infty)} \mathbb{1}_{[0,u]}(s+t)\,\mu_1(ds)\mu_2(dt)$$

is the representing measure of the product $f_1 f_2$. Indeed,

$$\int_{[0,\infty)} e^{-\lambda u}\,\mu(du) = \int_{[0,\infty)}\int_{[0,\infty)} e^{-\lambda(s+t)}\,\mu_1(ds)\mu_2(dt) = f_1(\lambda)f_2(\lambda).$$

Write $M := \{\mathscr{L}\mu \ :\ \mu \text{ is a finite measure on } [0,\infty)\}$. Theorem 1.4 shows that $M \subset \mathcal{CM} \subset \overline{M}$. We are done if we can show that $\mathcal{CM}$ is closed under pointwise convergence. For this choose a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{CM}$ such that $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ exists for every $\lambda > 0$. If μ_n denotes the representing measure of f_n, we find for every $a > 0$

$$\mu_n[0,a] \leqslant e^{a\lambda}\int_{[0,a]} e^{-\lambda t}\,\mu_n(dt) \leqslant e^{a\lambda} f_n(\lambda) \xrightarrow{n\to\infty} e^{a\lambda} f(\lambda)$$

which means that the family of measures $(\mu_n)_{n\in\mathbb{N}}$ is bounded in the vague topology, hence vaguely sequentially compact, see Appendix A.1. Thus, there exist a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ and some measure μ such that $\mu_{n_k} \to \mu$ vaguely. For $\chi \in C_c[0,\infty)$ with $0 \leqslant \chi \leqslant 1$, we find

$$\int_{[0,\infty)} \chi(t)e^{-\lambda t}\,\mu(dt) = \lim_{k\to\infty}\int_{[0,\infty)} \chi(t)e^{-\lambda t}\,\mu_{n_k}(dt) \leqslant \liminf_{k\to\infty} f_{n_k}(\lambda) = f(\lambda).$$

Taking the supremum over all such χ, we can use monotone convergence to get

$$\int_{[0,\infty)} e^{-\lambda s}\,\mu(dt) \leqslant f(\lambda).$$

On the other hand, we find for each $a > 0$

$$\begin{aligned} f_{n_k}(\lambda) &= \int_{[0,a)} e^{-\lambda t}\,\mu_{n_k}(dt) + \int_{[a,\infty)} e^{-\frac{1}{2}\lambda t}e^{-\frac{1}{2}\lambda t}\,\mu_{n_k}(dt) \\ &\leqslant \int_{[0,a)} e^{-\lambda t}\,\mu_{n_k}(dt) + e^{-\frac{1}{2}\lambda a} f_{n_k}\left(\frac{1}{2}\lambda\right). \end{aligned}$$

If we let $k \to \infty$ and then $a \to \infty$ along a sequence of continuity points of μ we get $f(\lambda) \leqslant \int_{[0,\infty)} e^{-\lambda t}\,\mu(dt)$ which shows that $f \in \mathcal{CM}$ and that the measure μ is actually independent of the particular subsequence. In particular, $\mu = \lim_{n\to\infty}\mu_n$ vaguely in the space of measures supported in $[0,\infty)$. $\square$

The seemingly innocuous closure assertion of Corollary 1.6 actually says that on the set $\mathcal{CM}$ the notions of *pointwise convergence*, *locally uniform convergence*, and even *convergence in the space* $C^\infty(0,\infty)$ coincide. This situation reminds remotely of the famous Montel's theorem from the theory of analytic functions, see e.g. Berenstein and Gay [21, Theorem 2.2.8].

Corollary 1.7. *Let* $(f_n)_{n\in\mathbb{N}}$ *be a sequence of completely monotone functions such that the limit* $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ *exists for all* $\lambda \in (0,\infty)$*. Then* $f \in \mathcal{CM}$ *and* $\lim_{n\to\infty} f_n^{(k)}(\lambda) = f^{(k)}(\lambda)$ *for all* $k \in \mathbb{N}\cup\{0\}$ *locally uniformly in* $\lambda \in (0,\infty)$.

Proof. From Corollary 1.6 we know already that $f \in \mathcal{CM}$. Moreover, we have seen that the representing measures μ_n of f_n converge vaguely in $[0,\infty)$ to the representing measure μ of f. By the differentiation lemma for parameter dependent integrals we infer

$$\begin{aligned} f_n^{(k)}(\lambda) &= (-1)^k \int_{[0,\infty)} t^k\, e^{-\lambda t}\, \mu_n(dt) \\ &\xrightarrow{n\to\infty} (-1)^k \int_{[0,\infty)} t^k\, e^{-\lambda t}\, \mu(dt) = f^{(k)}(\lambda), \end{aligned}$$

since $t \mapsto t^k e^{-\lambda t}$ is a function that vanishes at infinity, cf. (A.3) in Appendix A.1.

Finally, assume that $|\lambda - \kappa| \leqslant \delta$ for some $\delta > 0$. Using the elementary estimate $|e^{-\lambda t} - e^{-\kappa t}| \leqslant |\lambda-\kappa|\, t\, e^{-(\kappa\wedge\lambda)t}$, $\lambda,\kappa,t > 0$, we conclude that for $\kappa,\lambda \geqslant \epsilon$ and all $\epsilon > 0$

$$\begin{aligned} \left|f_n^{(k)}(\lambda) - f_n^{(k)}(\kappa)\right| &\leqslant \int_{(0,\infty)} \left|e^{-\lambda t} - e^{-\kappa t}\right| t^k\, \mu_n(dt) \\ &\leqslant \delta \int_{(0,\infty)} e^{-(\kappa\wedge\lambda)t}\, t^{k+1}\, \mu_n(dt) \\ &= \delta\left|f_n^{(k+1)}(\kappa\wedge\lambda)\right|. \end{aligned}$$

Using that $\lim_{n\to\infty} f_n^{(k+1)}(\kappa\wedge\lambda) = f^{(k+1)}(\kappa\wedge\lambda)$, we find for sufficiently large values of n

$$\left|f_n^{(k)}(\lambda) - f_n^{(k)}(\kappa)\right| \leqslant 2\delta \sup_{\gamma\geqslant\epsilon}\left|f^{(k+1)}(\gamma)\right|.$$

This proves that the functions $f_n^{(k)}$ are uniformly equicontinuous on $[\epsilon,\infty)$. Therefore, the convergence $\lim_{n\to\infty} f_n^{(k)}(\lambda) = f^{(k)}(\lambda)$ is locally uniform on $[\epsilon,\infty)$ for every $\epsilon > 0$. Since $\epsilon > 0$ was arbitrary, we are done. □

Remark 1.8. The representation formula for completely monotone functions given in Theorem 1.4 has an interesting interpretation in connection with the Kreĭn–Milman theorem and the Choquet representation theorem. The set

$$\left\{f \in \mathcal{CM} \,:\, f(0+) = 1\right\}$$

is a basis of the convex cone $\mathcal{CM}_b$, and its extremal points are given by

$$e_t(\lambda) = e^{-\lambda t}, \quad 0 \leqslant t < \infty, \quad \text{and} \quad e_\infty(\lambda) = \mathbb{1}_{\{0\}}(\lambda),$$

see Phelps [234, Lemma 2.2], Lax [199, p. 139] or the proof of Theorem 4.8. These extremal points are formally defined for $\lambda \in [0, \infty)$ with the understanding that $e_\infty|_{(0,\infty)} \equiv 0$. Therefore, the representation formula from Theorem 1.4 becomes a Choquet representation of the elements of $\mathcal{CM}_b$,

$$\int_{[0,\infty)} e^{-\lambda t}\, \mu(dt) = \int_{[0,\infty]} e^{-\lambda t}\, \mu(dt), \quad \lambda \in (0, \infty).$$

In particular, the functions

$$e_t|_{(0,\infty)}$$

are prime examples of completely monotone functions. Theorem 1.4 and Corollary 1.6 tell us that every $f \in \mathcal{CM}$ can be written as an 'integral mixture' of the extremal $\mathcal{CM}$-functions $\{e_t|_{(0,\infty)} : 0 \leqslant t < \infty\}$.

It was pointed out in [256] that the conditions (1.4) are redundant. The following proof of this fact is from [109].

Proposition 1.9. *Let $f : (0, \infty) \to \mathbb{R}$ be a C^∞ function such that $f \geqslant 0$, $f' \leqslant 0$ and $(-1)^n f^{(n)} \geqslant 0$ for infinitely many $n \in \mathbb{N}$. Then f is a completely monotone function.*

Proof. Let $n \geqslant 2$ be such that $(-1)^n f^{(n)}(\lambda) \geqslant 0$ for all $\lambda > 0$. By Taylor's formula, for every $a > 0$

$$f(\lambda) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (\lambda - a)^k + (-1)^n \int_a^\lambda \frac{(-1)^n f^{(n)}(s)}{(n-1)!} (\lambda - s)^{n-1}\, ds.$$

If n is even

$$f(\lambda) \geqslant \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (\lambda - a)^k,$$

while for n odd

$$f(\lambda) \leqslant \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (\lambda - a)^k.$$

Dividing by $(\lambda - a)^{n-1}$, letting $\lambda \to \infty$, and by using that $f(\lambda)$ is non-increasing, we arrive at $f^{(n-1)}(a) \leqslant 0$ in case n is even, and $f^{(n-1)}(a) \geqslant 0$ in case n is odd. Thus, $(-1)^{n-1} f^{(n-1)}(a) \geqslant 0$. It follows inductively that $(-1)^k f^{(k)}(a) \geqslant 0$ for all $k = 0, 1, \ldots, n$. Since n can be taken arbitrarily large, and $a > 0$ is arbitrary, the proof is completed. $\square$

Comments 1.10. Standard references for Laplace transforms include D. V. Widder's monographs [289, 290] and Doetsch's treatise [80]. For a modern point of view we refer to Berg and Forst [29] and Berg, Christensen and Ressel [28]. The most comprehensive tables of Laplace transforms are the Bateman manuscript project [91] and the tables by Prudnikov, Brychkov and Marichev [239].

The *concept* of complete monotonicity seems to go back to S. Bernstein [32] who studied functions on an interval $I \subset \mathbb{R}$ having positive derivatives of all orders. If $I = (-\infty, 0]$ this is, up to a change of sign in the variable, complete monotonicity. In later papers, Bernstein refers to functions enjoying this property as *absolument monotone*, see the appendix *première note,* [33, pt. IV, p. 190], and in [34] he states and proves Theorem 1.4 for functions on the negative half-axis.

Following Schur (probably [258]), Hausdorff [123, p. 80] calls a sequence $(\mu_n)_{n\in\mathbb{N}}$ *total monoton*—the literal translation *totally monotone* is only rarely used, e.g. in Hardy [118]; the modern terminology is *completely monotone* and appears for the first time in [287]—if all iterated differences $(-1)^k \Delta^k \mu_n$ are non-negative where $\Delta\mu_n := \mu_{n+1} - \mu_n$. Hausdorff focusses in [123, 124] on the *moment problem*: the μ_n are of the form $\int_{(0,1]} t^n \, \mu(dt)$ for some measure μ on $(0, 1]$ if, and only if, the sequence $(\mu_n)_{n\in\mathbb{N}\cup\{0\}}$ is *total monoton*; moreover, he introduces the moment function $\mu_\lambda := \int_{(0,1]} t^\lambda \, \mu(dt)$ which he also calls *total monoton*. A simple change of variables $(0, 1] \ni t \rightsquigarrow e^{-u}, u \in [0, \infty)$, shows that $\mu_\lambda = \int_{[0,\infty)} e^{-\lambda u} \, \tilde{\mu}(du)$ for some suitable image measure $\tilde{\mu}$ of μ. This means that every moment sequence gives rise to a unique completely monotone function. The converse is much easier since

$$\int_{[0,\infty)} e^{-\lambda u} \, \tilde{\mu}(du) = \sum_{n=0}^{\infty} \frac{(-1)^n \lambda^n}{n!} \int_{[0,\infty)} u^n \, \tilde{\mu}(du).$$

Many historical comments can be found in the second part, pp. 29–44, of Ky Fan's memoir [94]. For an up-to-date survey we recommend the scholarly commentary by Chatterji [63] written for Hausdorff's collected works [125]. More general *higher monotonicity properties* of the type that f satisfies (1.4) only for $n \in \{0, 1, \dots, N\}$, $N \in \mathbb{N} \cup \{0\}$, were used by Hartman [119] in connection with Bessel functions and solutions of second-order ordinary differential equations. Further applications of $\mathcal{CM}$ and related functions to ordinary differential equations can be found, e.g. in Lorch *et al.* [204] and Mahajan and Ross [207], see also the study by van Haeringen [282] and the references given there. The connection between integral equations and $\mathcal{CM}$ are extensively covered in the monographs by Gripenberg *et al.* [112] and Prüss [240].

A by-product of the proof of Proposition 1.2 is an example of a so-called real inversion formula for Laplace transforms. Formula (1.3) is due to Dubourdieu [86] and Feller [98], see also Pollard [238] and Widder [289, p. 295] and [290, Chapter 6]. Our presentation follows Feller [100, VII.6].

The proof of Bernstein's theorem, Theorem 1.4, also contains a real inversion formula for the Laplace transform: (1.8) coincides with the operator $L_{k,y}(f(\lambda))$ of Widder [290, p. 140] and, up to a constant, also [289, p. 288]. Since the proof of Theorem 1.4 relies on a compactness argument using subsequences, the weak limit $f_{n_k}(t)\,dt \to \mu(dt)$ might depend on the actual subsequence $(n_k)_{k\in\mathbb{N}}$. If we combine this argument with Proposition 1.2, we get at once that all subsequences lead to the same μ and that, therefore, the weak limit of the full sequence $f_n(t)\,dt \to \mu(dt)$ exists.

The representation (1.6) was also obtained in [87] in the following way: because of $f(+\infty) = 0$ we can write $f(\lambda) = \int_\lambda^\infty (-1) f'(t_1)\,dt_1$. Since $-f'$ is again completely monotone and satisfies $-f'(+\infty) = 0$, the same argument proves that $f(\lambda) = \int_\lambda^\infty \int_{t_1}^\infty f''(t_2)\,dt_2\,dt_1$. By induction, for every $n \in \mathbb{N}$,

$$f(\lambda) = \int_\lambda^\infty \int_{t_1}^\infty \cdots \int_{t_{n-1}}^\infty (-1)^n f^{(n)}(t_n)\,dt_n \cdots dt_2\,dt_1.$$

The representation (1.6) follows by using Fubini's theorem and reversing the order of integration. The rest of the proof is now similar to our presentation.

It is possible to avoid the compactness argument in Theorem 1.4 and to give an 'intuitionistic' proof, see van Herk [284, Theorem 33] who did this for the class $\mathcal{S}$ (denoted by $\{F\}$ in [284]) of Stieltjes functions which is contained in $\mathcal{CM}$; his arguments work also for $\mathcal{CM}$.

A proof of Theorem 1.4 using Choquet's theorem or the Kreĭn–Milman theorem can be found in Kendall [170], Meyer [214] or Choquet [69, 68]. A modern textbook version is contained in Lax [199, Chapter 14.3, p. 138], Phelps [234, Chapter 2], Becker [19] and also in Theorem 4.8 below. Gneiting's short note [109] contains an example showing that one cannot weaken the Bernstein–Hausdorff–Widder conditions beyond what is stated in Proposition 1.9.

There is a deep geometric connection between completely monotone functions and the problem when a metric space can be embedded into a Hilbert space $\mathfrak{H}$. The basic result is due to Schoenberg [255] who proves that a function f on $[0,\infty)$ with $f(0) = f(0+)$ is completely monotone if, and only if, $\xi \mapsto f(|\xi|^2)$, $\xi \in \mathbb{R}^d$, is positive definite for all dimensions $d \geqslant 0$, cf. Theorem 12.14.

Chapter 2
Stieltjes functions

Stieltjes functions are a subclass of completely monotone functions. They will play a central role in our study of complete Bernstein functions. In Theorem 7.3 we will see that f is a Stieltjes function if, and only if, $1/f$ is a complete Bernstein function. This allows us to study Stieltjes functions via the set of all complete Bernstein functions which is the focus of this tract. Therefore we restrict ourselves to the definition and a few fundamental properties of Stieltjes functions.

Definition 2.1. A (non-negative) *Stieltjes function* is a function $f : (0,\infty) \to [0,\infty)$ which can be written in the form

$$f(\lambda) = \frac{a}{\lambda} + b + \int_{(0,\infty)} \frac{1}{\lambda + t}\,\sigma(dt) \tag{2.1}$$

where $a, b \geqslant 0$ are non-negative constants and σ is a measure on $(0,\infty)$ such that $\int_{(0,\infty)} (1+t)^{-1}\sigma(dt) < \infty$. We denote the family of all Stieltjes functions by $\mathcal{S}$.

The integral appearing in (2.1) is also called the *Stieltjes transform* of the measure σ. Using the elementary relation $(\lambda + t)^{-1} = \int_0^\infty e^{-tu}e^{-\lambda u}\,du$ and Fubini's theorem one sees that it is also a double Laplace transform. In view of the uniqueness of the Laplace transform, see Proposition 1.2, a, b and σ appearing in the representation (2.1) are uniquely determined by f. Since some authors consider measures σ on the compactification $[0,\infty]$, there is no marked difference between Stieltjes transforms and Stieltjes functions in the sense of Definition 2.1.

It is sometimes useful to rewrite (2.1) in the following form

$$f(\lambda) = \int_{[0,\infty]} \frac{1+t}{\lambda + t}\,\bar{\sigma}(dt) \tag{2.2}$$

where $\bar{\sigma} := a\delta_0 + (1+t)^{-1}\sigma(dt) + b\delta_\infty$ is a finite measure on the compact interval $[0,\infty]$.

Since for $z = \lambda + i\kappa \in \mathbb{C} \setminus (-\infty, 0]$ and $t \geqslant 0$

$$\left|\frac{1}{z+t}\right| = \frac{1}{\sqrt{(\lambda+t)^2 + \kappa^2}} \asymp \frac{1}{t+1},$$

i.e. there exist two positive constants $c_1 < c_2$ (depending on λ and κ) such that

$$\frac{c_1}{t+1} \leqslant \frac{1}{\sqrt{(\lambda+t)^2 + \kappa^2}} \leqslant \frac{c_2}{t+1},$$

we can use (2.2) to extend $f \in \mathcal{S}$ uniquely to an analytic function on $\mathbb{C} \setminus (-\infty, 0]$. Note that

$$\operatorname{Im} z \cdot \operatorname{Im} \frac{1}{z+t} = \operatorname{Im} z \cdot \frac{-\operatorname{Im} z}{|z+t|^2} = -\frac{(\operatorname{Im} z)^2}{|z+t|^2}$$

which means that the mapping $z \mapsto f(z)$ swaps the upper and lower complex half-planes. We will see in Corollary 7.4 below that this property is also sufficient to characterize $f \in \mathcal{S}$.

Theorem 2.2. (i) *Every* $f \in \mathcal{S}$ *is of the form*

$$f(\lambda) = \mathscr{L}(a \cdot dt; \lambda) + \mathscr{L}\big(b \cdot \delta_0(dt); \lambda\big) + \mathscr{L}\big(\mathscr{L}(\sigma; t)\, dt; \lambda\big)$$

for the measure σ *appearing in* (2.1). *In particular,* $\mathcal{S} \subset \mathcal{CM}$ *and* $\mathcal{S}$ *consists of all completely monotone functions having a representation measure with completely monotone density on* $(0, \infty)$.

(ii) *The set* $\mathcal{S}$ *is a convex cone: if* $f_1, f_2 \in \mathcal{S}$*, then* $s f_1 + t f_2 \in \mathcal{S}$ *for all* $s, t \geqslant 0$.

(iii) *The set* $\mathcal{S}$ *is closed under pointwise limits: if* $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}$ *and if the limit* $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ *exists for all* $\lambda > 0$*, then* $f \in \mathcal{S}$.

Proof. Since $(\lambda + t)^{-1} = \int_0^\infty e^{-(\lambda+t)u}\, du$, assertion (i) follows from (2.1) and Fubini's theorem; (ii) is obvious. For (iii) we argue as in the proof of Corollary 1.6: assume that f_n is given by (2.2) where we denote the representing measure by $\bar\sigma_n = a_n \delta_0 + (1+t)^{-1}\sigma_n(dt) + b_n \delta_\infty$. Since

$$\bar\sigma_n[0, \infty] = f_n(1) \xrightarrow{n\to\infty} f(1) < \infty,$$

the family $(\bar\sigma_n)_{n\in\mathbb{N}}$ is uniformly bounded. By the Banach–Alaoglu theorem, Corollary A.6, we conclude that $(\bar\sigma_n)_{n\in\mathbb{N}}$ has a weak* convergent subsequence $(\bar\sigma_{n_k})_{k\in\mathbb{N}}$ such that $\bar\sigma := \text{vague-}\lim_{k\to\infty} \bar\sigma_{n_k}$ is a bounded measure on the compact space $[0, \infty]$. Since $t \mapsto (1+t)/(\lambda+t)$ is in $C[0, \infty]$, we get

$$f(\lambda) = \lim_{k\to\infty} f_{n_k}(\lambda) = \lim_{k\to\infty} \int_{[0,\infty]} \frac{1+t}{\lambda+t}\, \bar\sigma_{n_k}(dt) = \int_{[0,\infty]} \frac{1+t}{\lambda+t}\, \bar\sigma(dt),$$

i.e. $f \in \mathcal{S}$. Since the limit $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ exists—independently of any subsequence—and since the representing measure is uniquely determined by the function f, $\bar\sigma$ does not depend on any subsequence. In particular, $\sigma = \lim_{n\to\infty} \sigma_n$ vaguely in the space of measures supported in $(0, \infty)$. □

Remark 2.3. Let $f, f_n \in \mathcal{S}$, $n \in \mathbb{N}$, where we write a, b, σ and a_n, b_n, σ_n for the constants and measures appearing in (2.1) and $\bar\sigma_n, \bar\sigma$ for the corresponding representation measures from (2.2). If $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$, the proof of Theorem 2.2 shows that

$$\text{vague-}\lim_{n\to\infty} \sigma_n = \sigma \quad \text{and} \quad \text{vague-}\lim_{n\to\infty} \bar\sigma_n = \bar\sigma.$$

Combining this with the portmanteau theorem, Theorem A.7, it is possible to show that

$$a = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \left(a_n + \int_{(0,\epsilon)} \frac{\sigma_n(dt)}{1+t} \right)$$

and

$$b = \lim_{R \to \infty} \liminf_{n \to \infty} \left(b_n + \int_{(R,\infty)} \frac{\sigma_n(dt)}{1+t} \right);$$

in the above formulae we can replace $\liminf_n$ by $\limsup_n$. We do not give the proof here, but we refer to a similar situation for Bernstein functions which is worked out in Corollary 3.8.

In general, it is not true that $\lim_{n\to\infty} a_n = a$ or $\lim_{n\to\infty} b_n = b$. This is easily seen from the following examples: $f_n(\lambda) = \lambda^{-1+1/n}$ and $f_n(\lambda) = \lambda^{-1/n}$.

Remark 2.4. Just as in the case of completely monotone functions, see Remark 1.8, we can understand the representation formula (2.2) as a particular case of the Kreĭn–Milman or Choquet representation. The set

$$\{f \in \mathcal{S} \, : \, f(1) = 1\}$$

is a basis of the convex cone $\mathcal{S}$, and its extremal points are given by

$$e_0(\lambda) = \frac{1}{\lambda}, \quad e_t(\lambda) = \frac{1+t}{\lambda+t}, \ \ 0 < t < \infty, \quad \text{and} \quad e_\infty(\lambda) = 1.$$

To see that the functions e_t, $0 \leqslant t \leqslant \infty$, are indeed extremal, note that the equality

$$e_t(\lambda) = \epsilon f(\lambda) + (1-\epsilon) g(\lambda), \quad f, g \in \mathcal{S},$$

and the uniqueness of the representing measures in (2.2) imply that

$$\delta_t = \epsilon \bar{\sigma}_f + (1-\epsilon)\bar{\sigma}_g$$

($\bar{\sigma}_f, \bar{\sigma}_g$ are the representing measures) which is only possible if $\bar{\sigma}_f = \bar{\sigma}_g = \delta_t$. Conversely, since every $f \in \mathcal{S}$ is given by (2.2), the family $\{e_t\}_{t\in[0,\infty]}$ contains all extremal points. In particular,

$$1, \quad \frac{1}{\lambda}, \quad \frac{1}{\lambda+t}, \quad \frac{1+t}{\lambda+t}, \quad t > 0,$$

are examples of Stieltjes functions and so are their integral mixtures, e.g.

$$\lambda^{\alpha-1} \ \ (0 < \alpha < 1), \quad \frac{1}{\sqrt{\lambda}} \arctan \frac{1}{\sqrt{\lambda}}, \quad \frac{1}{\lambda} \log(1+\lambda)$$

which we obtain if we choose $\frac{1}{\pi}\sin(\alpha\pi)\,t^{\alpha-1}\,dt$, $\mathbb{1}_{(0,1)}(t)\frac{dt}{2\sqrt{t}}$ or $\frac{1}{t}\mathbb{1}_{(1,\infty)}(t)\,dt$ as the representing measures $\sigma(dt)$ in (2.1).

The closure assertion of Theorem 2.2 says, in particular, that on the set $\mathcal{S}$ the notions of *pointwise convergence*, *locally uniform convergence*, and even *convergence in the space C^∞* coincide, cf. also Corollary 1.7.

Comments 2.5. The Stieltjes transform appears for the first time in the famous papers [268, 269] where T. J. Stieltjes investigates continued fractions in order to solve what we nowadays call the *Stieltjes Moment Problem*. For an appreciation of Stieltjes' achievements see the contributions of W. van Assche and W. A. J. Luxemburg in [270, Vol. 1].

The name Stieltjes transform for the integral (2.1) was coined by Doetsch [79] and, independently, Widder [288]. Earlier works, e.g. Perron [231], call $\int(\lambda+t)^{-1}\,\sigma(dt)$ *Stieltjes integral* (German: *Stieltjes'sches Integral*) but terminology has changed since then. Sometimes the name *Hilbert–Hankel transform* is also used in the literature, cf. Lax [199, p. 185]. A systematic account of the properties of the Stieltjes transform is given in [289, Chapter VIII].

Stieltjes functions are discussed by van Herk [284] as class $\{F\}$ in the wider context of moment and complex interpolation problems, see also the Comments 6.12, 7.16. Van Herk uses the integral representation

$$\int_{[0,1]} \frac{1}{1-s+sz}\,\chi(ds)$$

which can be transformed by the change of variables $t = s^{-1}-1 \in [0,\infty]$ for $s \in [0,1]$ into the form (2.2); $\bar{\sigma}(dt)$ and $\chi(ds)$ are image measures under this transformation. Van Herk only observes that his class $\{F\}$ contains $\mathcal{S}$, but comparing [284, Theorem 7.50] with our result (7.2) in Chapter 7 below shows that $\{F\} = \mathcal{S}$.

In his paper [135] F. Hirsch introduces Stieltjes transforms into potential theory and identifies $\mathcal{S}$ as a convex cone operating on the abstract potentials, i.e. the densely defined inverses of the infinitesimal generators of C_0-semigroups. Hirsch establishes several properties of the cone $\mathcal{S}$, which are later extended by Berg [22, 23]. A presentation of this material from a potential theoretic point of view is contained in the monograph [29, Chapter 14] by Berg and Forst, the connections to the moment problem are surveyed in [25].

Theorem 2.2 appears in Hirsch [135, Proposition 1], but with a different proof.

In contrast to completely monotone functions, $\mathcal{S}$ is not closed under multiplication. This follows easily since the (necessarily unique) representing measure of $\lambda \mapsto (\lambda+a)^{-1}(\lambda+b)^{-1}$, $0<a<b$, is $(b-t)^{-1}\delta_a(dt) + (a-t)^{-1}\delta_b(dt)$ which is a signed measure. We will see in Proposition 7.10 that $\mathcal{S}$ is still logarithmically convex. It is known, see Hirschman and Widder [140, VII 7.4], that the product $\int_{(0,\infty)}(\lambda+t)^{-1}\sigma_1(dt)\int_{(0,\infty)}(\lambda+t)^{-1}\sigma_2(dt)$ is of the form $\int_{(0,\infty)}(\lambda+t)^{-2}\,\sigma(dt)$ for some measure σ. The latter integral is often called a *generalized Stieltjes transform*. The following related result is due to Srivastava and Tuan [264]: if $f \in L^p(0,\infty)$ and $g \in L^q(0,\infty)$ with $1<p,q<\infty$ and $r^{-1} = p^{-1}+q^{-1} < 1$, then there is some $h \in L^r(0,\infty)$ such that $\mathscr{L}^2(f;\lambda)\mathscr{L}^2(g;\lambda) = \mathscr{L}^2(h;\lambda)$ holds. Since

$$h(t) = f(t)\cdot\text{p.v.}\int_0^\infty \frac{g(u)}{u-t}\,du + g(t)\cdot\text{p.v.}\int_0^\infty \frac{f(u)}{u-t}\,du,$$

h will, in general, change its sign even if f, g are non-negative.

Chapter 3
Bernstein functions

We are now ready to introduce the class of *Bernstein functions* which are closely related to completely monotone functions. The notion of Bernstein functions goes back to the potential theory school of A. Beurling and J. Deny and was subsequently adopted by C. Berg and G. Forst [29], see also [25]. S. Bochner [50] calls them *completely monotone mappings* (as opposed to completely monotone functions) and probabilists still prefer the term *Laplace exponents*, see e.g. Bertoin [36, 38]; the reason will become clear from Theorem 5.2.

Definition 3.1. A function $f : (0,\infty) \to \mathbb{R}$ is a *Bernstein function* if f is of class C^∞, $f(\lambda) \geqslant 0$ for all $\lambda > 0$ and

$$(-1)^{n-1} f^{(n)}(\lambda) \geqslant 0 \quad \text{for all } n \in \mathbb{N} \text{ and } \lambda > 0. \tag{3.1}$$

The set of all Bernstein functions will be denoted by $\mathcal{BF}$.

It is easy to see from the definition that, for example, the fractional powers $\lambda \mapsto \lambda^\alpha$, are Bernstein functions if, and only if, $0 \leqslant \alpha \leqslant 1$.

The key to the next theorem is the observation that a non-negative C^∞-function $f : (0,\infty) \to \mathbb{R}$ is a Bernstein function if, and only if, f' is a completely monotone function.

Theorem 3.2. *A function $f : (0,\infty) \to \mathbb{R}$ is a Bernstein function if, and only if, it admits the representation*

$$f(\lambda) = a + b\lambda + \int_{(0,\infty)} (1 - e^{-\lambda t})\, \mu(dt), \tag{3.2}$$

where $a, b \geqslant 0$ and μ is a measure on $(0,\infty)$ satisfying $\int_{(0,\infty)} (1 \wedge t)\, \mu(dt) < \infty$.

In particular, the triplet (a, b, μ) determines f uniquely and vice versa.

Remark 3.3. (i) The representing measure μ and the characteristic triplet (a, b, μ) from (3.2) are often called the *Lévy measure* and the *Lévy triplet* of the Bernstein function f. The formula (3.2) is called the *Lévy–Khintchine representation* of f.

(ii) A useful variant of the representation formula (3.2) can be obtained by an application of Fubini's theorem. Since

$$\begin{aligned}\int_{(0,\infty)} (1-e^{-\lambda t})\,\mu(dt) &= \int_{(0,\infty)} \int_{(0,t)} \lambda e^{-\lambda s}\,ds\,\mu(dt)\\ &= \int_0^\infty \int_{(s,\infty)} \lambda e^{-\lambda s}\mu(dt)\,ds\\ &= \int_0^\infty \lambda e^{-\lambda s}\mu(s,\infty)\,ds\end{aligned}$$

we get that any Bernstein function can be written in the form

$$f(\lambda) = a + b\lambda + \lambda \int_{(0,\infty)} e^{-\lambda s}\,M(s)\,ds \tag{3.3}$$

where $M(s) = M_\mu(s) = \mu(s,\infty)$ is a non-increasing, right-continuous function. Integration by parts and the observation that $\int_0^\infty s e^{-\lambda s}\,ds = \lambda^{-2}$ and $\int_0^\infty e^{-\lambda s}\,ds = \lambda^{-1}$ yield

$$f(\lambda) = \lambda^2 \int_{(0,\infty)} e^{-\lambda s}\,k(s)\,ds = \lambda^2 \mathscr{L}(k;\lambda) \tag{3.4}$$

with $k(s) = as + b + \int_0^s M(t)\,dt$, compare with Theorem 6.2(iii). Note that k is positive, non-decreasing and concave.

(iii) The integrability condition $\int_{(0,\infty)}(1\wedge t)\,\mu(dt) < \infty$ ensures that the integral in (3.2) converges for some, hence all, $\lambda > 0$. This is immediately seen from the convexity inequalities

$$\frac{t}{1+t} \leqslant 1 - e^{-t} \leqslant 1 \wedge t \leqslant 2\,\frac{t}{1+t}, \quad t > 0$$

and the fact that for $\lambda \geqslant 1$ [respectively for $0 < \lambda < 1$] and all $t > 0$

$$1\wedge t \leqslant 1\wedge(t\lambda) \leqslant \lambda(1\wedge t) \qquad \big[\text{respectively } \lambda(1\wedge t) \leqslant 1\wedge(t\lambda) \leqslant 1\wedge t\big].$$

(iv) A useful consequence of the above estimate and the representation formula (3.2) are the following formulae to calculate the coefficients a and b:

$$a = f(0+) \quad \text{and} \quad b = \lim_{\lambda\to\infty} \frac{f(\lambda)}{\lambda}.$$

The first formula is obvious while the second follows from (3.2) and the dominated convergence theorem: $1 - e^{-\lambda t} \leqslant 1 \wedge (\lambda t)$ and $\lim_{\lambda\to\infty}(1-e^{-\lambda t})/\lambda = 0$.

(v) Formula (3.3) shows, in particular, that

$$\int_0^1 \mu(t,\infty)\,dt = \int_0^1 M(t)\,dt < \infty. \tag{3.5}$$

Since a non-increasing function which is integrable near zero is $o(1/t)$ as $t \to 0$, we conclude from (3.5) that

$$\lim_{t\to 0+} t\mu(t,\infty) = \lim_{t\to 0+} tM(t) = 0. \tag{3.6}$$

Proof of Theorem 3.2. Assume that f is a Bernstein function. Then f' is completely monotone. According to Theorem 1.4, there exists a measure ν on $[0,\infty)$ such that for all $\lambda > 0$

$$f'(\lambda) = \int_{[0,\infty)} e^{-\lambda t}\,\nu(dt).$$

Let $b := \nu\{0\}$. Then

$$\begin{aligned} f(\lambda) - f(0+) = \int_0^\lambda f'(y)\,dy &= b\lambda + \int_0^\lambda \int_{(0,\infty)} e^{-yt}\,\nu(dt)\,dy \\ &= b\lambda + \int_{(0,\infty)} \frac{1-e^{-\lambda t}}{t}\,\nu(dt). \end{aligned}$$

Write $a := f(0+)$ and define $\mu(dt) := t^{-1}\nu|_{(0,\infty)}(dt)$. Then the calculation from above shows that (3.2) is true. That $a, b \geqslant 0$ is obvious, and from the elementary (convexity) estimate

$$(1-e^{-1})(1\wedge t) \leqslant 1 - e^{-t}, \quad t \geqslant 0,$$

we infer

$$\int_{(0,\infty)} (1\wedge t)\,\mu(dt) \leqslant \frac{e}{e-1}\int_{(0,\infty)} (1-e^{-t})\,\mu(dt) = \frac{e}{e-1}\,f(1) < \infty.$$

Conversely, suppose that f is given by (3.2) with (a,b,μ) as in the statement of the theorem. Since $te^{-\lambda t} \leqslant t \wedge (e\lambda)^{-1}$, we can apply the differentiation lemma for parameter-dependent integrals for all λ from $[\epsilon, \epsilon^{-1}]$ and all $\epsilon > 0$. Differentiating (3.2) under the integral sign yields

$$f'(\lambda) = b + \int_{(0,\infty)} e^{-\lambda t}\,t\,\mu(dt) = \int_{[0,\infty)} e^{-\lambda t}\,\nu(dt),$$

where $\nu(dt) := t\mu(dt) + b\delta_0(dt)$. This formula shows that f' is a completely monotone function. Therefore, f is a Bernstein function.

Because $f(0+) = a$ and because of the uniqueness assertion of Theorem 1.4 it is clear that (a,b,μ) and $f \in \mathcal{BF}$ are in one-to-one correspondence. $\square$

The derivative of a Bernstein function is completely monotone. The converse is only true, if the primitive of a completely monotone function is positive. This fails, for example, for the completely monotone function λ^{-2} whose primitive, $-\lambda^{-1}$, is not a Bernstein function. The next proposition characterizes the image of $\mathcal{BF}$ under differentiation.

Proposition 3.4. *Let $g(\lambda) = b + \int_{(0,\infty)} e^{-\lambda t}\,\nu(dt)$ be a completely monotone function. It has a primitive $f \in \mathcal{BF}$ if, and only if, the representing measure ν satisfies $\int_{(0,\infty)}(1+t)^{-1}\,\nu(dt) < \infty$.*

Proof. Assume that f is a Bernstein function given in the form (3.2). Then

$$f'(\lambda) = b + \int_{(0,\infty)} e^{-\lambda t}\, t\, \mu(dt)$$

is completely monotone and the measure $\nu(dt) := t\mu(dt)$ satisfies

$$\int_{(0,\infty)} \frac{1}{1+t}\,\nu(dt) = \int_{(0,\infty)} \frac{t}{1+t}\,\mu(dt) \leqslant \int_{(0,\infty)} (1 \wedge t)\,\mu(dt) < \infty.$$

Retracing the above steps reveals that $\int_{(0,\infty)}(1+t)^{-1}\,\nu(dt) < \infty$ is also sufficient to guarantee that $g(\lambda) := b + \int_{(0,\infty)} e^{-\lambda t}\,\nu(dt)$ has a primitive which is a Bernstein function. □

Theorem 3.2 allows us to extend Bernstein functions onto the right complex half-plane $\overrightarrow{\mathbb{H}} := \{z \in \mathbb{C} : \operatorname{Re} z > 0\}$.

Proposition 3.5. *Every $f \in \mathcal{BF}$ has an extension $f : \overline{\overrightarrow{\mathbb{H}}} \to \overline{\overrightarrow{\mathbb{H}}}$ which is continuous for $\operatorname{Re} z \geqslant 0$ and holomorphic for $\operatorname{Re} z > 0$.*

Proof. The function $\lambda \mapsto 1 - e^{-\lambda t}$ appearing in (3.2) has a unique holomorphic extension. If $z = \lambda + i\kappa$ is such that $\lambda = \operatorname{Re} z > 0$ we get

$$|1 - e^{-zt}| = \left| \int_0^{zt} e^{-\zeta}\, d\zeta \right| \leqslant t|z| \quad \text{and} \quad |1 - e^{-zt}| \leqslant 1 + |e^{-zt}| \leqslant 2.$$

This means that (3.2) converges uniformly in $z \in \overrightarrow{\mathbb{H}}$ and $f(z)$ is well defined and holomorphic on $\overrightarrow{\mathbb{H}}$. Moreover,

$$\begin{aligned}
\operatorname{Re} f(z) &= a + b\operatorname{Re} z + \int_{(0,\infty)} \operatorname{Re}\,(1 - e^{-zt})\,\mu(dt) \\
&= a + b\lambda + \int_{(0,\infty)} \big(1 - e^{-\lambda t}\cos(\kappa t)\big)\,\mu(dt)
\end{aligned}$$

which is positive since $\lambda = \operatorname{Re} z \geqslant 0$ and $1 - e^{-\lambda t}\cos(\kappa t) \geqslant 1 - e^{-\lambda t} \geqslant 0$.

Continuity up to the boundary follows from the estimate

$$
\begin{aligned}
|f(z) - f(w)| &\leqslant b|z-w| + \int_{(0,\infty)} |e^{-wt} - e^{-zt}|\,\mu(dt) \\
&\leqslant b|z-w| + \int_{(0,\infty)} 2 \wedge (t|w-z|)\,\mu(dt)
\end{aligned}
$$

for all $z, w \in \overrightarrow{\mathbb{H}}$ and the dominated convergence theorem. □

The following structural characterization comes from Bochner [50, pp. 83–84] where Bernstein functions are called completely monotone mappings.

Theorem 3.6. *Let f be a positive function on $(0,\infty)$. Then the following assertions are equivalent.*

(i) $f \in \mathcal{BF}$.

(ii) $g \circ f \in \mathcal{CM}$ *for every* $g \in \mathcal{CM}$.

(iii) $e^{-uf} \in \mathcal{CM}$ *for every* $u > 0$.

Proof. The proof relies on the following formula for the n-th derivative of the composition $h = g \circ f$ due to Faa di Bruno [93], see also [111, formula 0.430]:

$$
h^{(n)}(\lambda) = \sum_{(m,i_1,\dots,i_\ell)} \frac{n!}{i_1!\cdots i_\ell!}\, g^{(m)}\big(f(\lambda)\big) \prod_{j=1}^{\ell} \left(\frac{f^{(j)}(\lambda)}{j!}\right)^{i_j} \tag{3.7}
$$

where $\sum_{(m,i_1,\dots,i_\ell)}$ stands for summation over all $\ell \in \mathbb{N}$ and all $i_1,\dots,i_\ell \in \mathbb{N}\cup\{0\}$ such that $\sum_{j=1}^{\ell} j \cdot i_j = n$ and $\sum_{j=1}^{\ell} i_j = m$.

(i)⇒(ii) Assume that $f \in \mathcal{BF}$ and $g \in \mathcal{CM}$. Then $h(\lambda) = g(f(\lambda)) \geqslant 0$. Multiply formula (3.7) by $(-1)^n$ and observe that $n = m + \sum_{j=1}^{\ell}(j-1)\cdot i_j$. The assumptions $f \in \mathcal{BF}$ and $g \in \mathcal{CM}$ guarantee that each term in the formula multiplied by $(-1)^n$ is non-negative. This proves that $h = g \circ f \in \mathcal{CM}$.

(ii)⇒(iii) This follows from the fact that $g(\lambda) := g_u(\lambda) := e^{-\lambda u}$, $u > 0$, is completely monotone.

(iii)⇒(i) The series $e^{-uf(\lambda)} = \sum_{j=0}^{\infty} \frac{(-1)^j u^j}{j!}\,[f(\lambda)]^j$ and all of its formal derivatives (w.r.t. λ) converge uniformly, so we can calculate $\frac{d^n}{d\lambda^n} e^{-uf(\lambda)}$ by termwise differentiation. Since e^{-uf} is completely monotone, we get

$$
0 \leqslant (-1)^n \frac{d^n}{d\lambda^n} e^{-uf(\lambda)} = \sum_{j=1}^{\infty} \frac{u^j}{j!}\,(-1)^{n+j} \frac{d^n}{d\lambda^n}\big[f(\lambda)\big]^j.
$$

Dividing by $u > 0$ and letting $u \to 0$ we see

$$
0 \leqslant (-1)^{n+1} \frac{d^n}{d\lambda^n} f(\lambda).
$$

□

Theorems 3.2 and 3.6 have a few important consequences.

Corollary 3.7. (i) *The set $\mathcal{BF}$ is a convex cone: if $f_1, f_2 \in \mathcal{BF}$, then $sf_1+tf_2 \in \mathcal{BF}$ for all $s,t \geqslant 0$.*

(ii) *The set $\mathcal{BF}$ is closed under pointwise limits: if $(f_n)_{n\in\mathbb{N}} \subset \mathcal{BF}$ and if the limit $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ exists for every $\lambda > 0$, then $f \in \mathcal{BF}$.*

(iii) *The set $\mathcal{BF}$ is closed under composition: if $f_1, f_2 \in \mathcal{BF}$, then $f_1 \circ f_2 \in \mathcal{BF}$. In particular, $\lambda \mapsto f_1(c\lambda)$ is in $\mathcal{BF}$ for any $c > 0$.*

(iv) *For all $f \in \mathcal{BF}$ the function $\lambda \mapsto f(\lambda)/\lambda$ is in $\mathcal{CM}$.*

(v) *$f \in \mathcal{BF}$ is bounded if, and only if, in* (3.2) *$b = 0$ and $\mu(0,\infty) < \infty$.*

(vi) *Let $f_1, f_2 \in \mathcal{BF}$ and $\alpha, \beta \in (0,1)$ such that $\alpha + \beta \leqslant 1$. Then $\lambda \mapsto f_1(\lambda^\alpha) f_2(\lambda^\beta)$ is again a Bernstein function.*

Proof. (i) This follows immediately from Definition 3.1 or, alternatively, from the representation formula (3.2).

(ii) For every $u > 0$ we know that e^{-uf_n} is a completely monotone function and that $e^{-uf(\lambda)} = \lim_{n\to\infty} e^{-uf_n(\lambda)}$. Since $\mathcal{CM}$ is closed under pointwise limits, cf. Corollary 1.6, e^{-uf} is completely monotone and $f \in \mathcal{BF}$.

(iii) Let $f_1, f_2 \in \mathcal{BF}$. For any $g \in \mathcal{CM}$ we use the implication (i)$\Rightarrow$(ii) of Theorem 3.6 to get $g \circ f_1 \in \mathcal{CM}$, and then $g \circ (f_1 \circ f_2) = (g \circ f_1) \circ f_2 \in \mathcal{CM}$. The converse direction (ii)$\Rightarrow$(i) of Theorem 3.6 shows that $f_1 \circ f_2 \in \mathcal{BF}$.

(iv) Note that $(1 - e^{-\lambda t})/\lambda = \int_0^1 e^{-\lambda s}\, ds$ is completely monotone. Therefore,

$$\frac{f(\lambda)}{\lambda} = \frac{a}{\lambda} + b + \int_{(0,\infty)} \frac{1 - e^{-\lambda t}}{\lambda}\, \mu(dt)$$

is the limit of linear combinations of completely monotone functions which is, by Corollary 1.6, completely monotone.

(v) That $b = 0$ and $\mu(0,\infty) < \infty$ imply the boundedness of f is clear from the representation (3.2). Conversely, if f is bounded, $b = 0$ follows from Remark 3.3(iv), and $\mu(0,\infty) < \infty$ follows from (3.2) and Fatou's lemma.

(vi) We know that the fractional powers $\lambda \mapsto \lambda^\alpha$, $0 \leqslant \alpha \leqslant 1$, are Bernstein functions. Since $h(\lambda) := f_1(\lambda^\alpha) f_2(\lambda^\beta)$ is positive, it is enough to show that the derivative h' is completely monotone. We have

$$\begin{aligned} h'(\lambda) &= \alpha f_1'(\lambda^\alpha)\lambda^{\alpha-1} f_2(\lambda^\beta) + \beta f_2'(\lambda^\beta)\lambda^{\beta-1} f_1(\lambda^\alpha) \\ &= \lambda^{\alpha+\beta-1} \left(\alpha f_1'(\lambda^\alpha)\, \frac{f_2(\lambda^\beta)}{\lambda^\beta} + \beta f_2'(\lambda^\beta)\, \frac{f_1(\lambda^\alpha)}{\lambda^\alpha} \right). \end{aligned}$$

Note that $f_1'(\kappa)$, $f_2'(\kappa)$ and, by part (iv), $\kappa^{-1} f_1(\kappa)$, $\kappa^{-1} f_2(\kappa)$ are completely monotone. By Theorem 3.6(ii) the functions $f_1'(\lambda^\alpha)$, $f_2'(\lambda^\beta)$, $f_1(\lambda^\alpha)/\lambda^\alpha$ and $f_2(\lambda^\beta)/\lambda^\beta$

are again completely monotone. Since $\alpha + \beta \leqslant 1$, $\lambda \mapsto \lambda^{\alpha+\beta-1}$ is completely monotone. As sums and products of completely monotone functions are in $\mathcal{CM}$, see Corollary 1.6, h' is completely monotone. $\square$

Just as for completely monotone functions, the closure assertion of Corollary 3.7 says that on the set $\mathcal{BF}$ the notions of *pointwise convergence*, *locally uniform convergence*, and even *convergence in the space C^∞* coincide.

Corollary 3.8. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of Bernstein functions such that the limit $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ exists for all $\lambda \in (0,\infty)$. Then $f \in \mathcal{BF}$ and for all $k \in \mathbb{N} \cup \{0\}$ the convergence $\lim_{n\to\infty} f_n^{(k)}(\lambda) = f^{(k)}(\lambda)$ is locally uniform in $\lambda \in (0,\infty)$.*

If (a_n, b_n, μ_n) and (a, b, μ) are the Lévy triplets for f_n and f, respectively, see (3.2)*, we have*

$$\lim_{n\to\infty} \mu_n = \mu \quad \textit{vaguely in } (0,\infty),$$

and

$$a = \lim_{R\to\infty} \liminf_{n\to\infty} \big(a_n + \mu_n[R,\infty)\big), \quad b = \lim_{\epsilon\to 0} \liminf_{n\to\infty} \left(b_n + \int_{(0,\epsilon)} t\, \mu_n(dt)\right).$$

In both formulae we may replace $\liminf_n$ by $\limsup_n$.

Proof. From Corollary 3.7(ii) we know that $f \in \mathcal{BF}$. Obviously, $\lim_{n\to\infty} e^{-f_n} = e^{-f}$; by Theorem 3.6, the functions e^{-f_n}, e^{-f} are completely monotone and we can use Corollary 1.7 to conclude that

$$e^{-f_n} \xrightarrow{n\to\infty} e^{-f} \quad \text{and} \quad (-f_n')\, e^{-f_n} \xrightarrow{n\to\infty} (-f')\, e^{-f}$$

locally uniformly on $(0,\infty)$. In particular, $\lim_{n\to\infty} f_n'(\lambda) = f'(\lambda)$ for each $\lambda \in (0,\infty)$. Again by Corollary 1.7 and the complete monotonicity of f_n', f' we see that for $k \geqslant 1$ the derivatives $f_n^{(k)}$ converge locally uniformly to $f^{(k)}$. By the mean value theorem,

$$|f_n(\lambda) - f(\lambda)| = |\log e^{-f(\lambda)} - \log e^{-f_n(\lambda)}| \leqslant C\, |e^{-f(\lambda)} - e^{-f_n(\lambda)}|$$

with $C \leqslant e^{f(\lambda)+f_n(\lambda)}$. The locally uniform convergence of e^{-f_n} ensures that C is bounded for $n \in \mathbb{N}$ and λ from compact sets in $(0,\infty)$; this proves locally uniform convergence of f_n to f on $(0,\infty)$.

Differentiating the representation formula (3.2) we get

$$f_n'(\lambda) = b_n + \int_{(0,\infty)} t\, e^{-\lambda t}\, \mu_n(dt) = \int_{[0,\infty)} e^{-\lambda t} \big(b_n \delta_0(dt) + t\, \mu_n(dt)\big),$$

implying that $b_n\delta_0(dt) + t\, \mu_n(dt)$ converge vaguely to $b\delta_0(dt) + t\, \mu(dt)$. This proves at once that $\mu_n \to \mu$ vaguely on $(0,\infty)$ as $n \to \infty$.