Prologue: Regular Variation

P.1 Introduction

Regular variation is a subject both of theoretical interest and of great use in a variety of applications. These include analytic number theory (asymptotics of arithmetic functions, Prime Number Theorem), complex analysis (entire functions – Levin–Pfluger theory) and, particularly, probability theory (limit theorems). The standard work here, covering theory and applications, is BinGoT1987 (BGT below for brevity). As it happens, matters left open in BGT – the foundational question, p. 11 (on measurability and the Baire property), and the contextual question (Appendix 1, on contexts beyond the real analysis to which the bulk of the text is devoted) – motivated our joint work. So this book is motivated by two much earlier and by now well-established texts, Oxtoby (Oxt1980) and BGT. But these serve only as background and motivation here; this book is self-contained and may be read without reference to either.

To make the above more specific, here are some instances of 'what regular variation can do for the mathematician in the street'.

P.2 Probability Theory

The prototypical limit theorem in the subject is Kolmogorov's[1] *strong law of large numbers*: that if (X_n) is a sequence of independent copies of a random variable X drawn from some distribution (or law) F, the averages $S_n/n := \sum_1^n X_k/n$ converge as n increases to some limit c with probability 1 ('almost

[1] In 2023, the postal address of Moscow State University became 1 Kolmogorov Street; cf. 2 Stefan Banach Street, Warsaw, the postal address of the mathematics department of Warsaw University.

surely', a.s.) if and only if X has a *mean* (first moment, *expectation*) (meaning $\mathbb{E}\,[\,|X|\,] < \infty$), and then the limit is the mean $\mu = \mathbb{E}\,[X]$:

$$S_n/n \to \mu \qquad (n \to \infty) \quad a.s., \quad \text{where } \mu = \mathbb{E}[X]. \tag{LLN}$$

Second only to this is the *central limit theorem*: if also one has finite variance, σ^2 say (finite second moment), and centres at means (subtracting $\mathbb{E}[S_n] = n\mu$), the right scaling is by $\sqrt{n}\sigma$, and one then has a limit distribution, the standard normal (or Gaussian) law $\Phi = N(0, 1)$:

$$\mathbb{P}((S_n - n\mu)/\sigma\sqrt{n} \le x) \to \Phi(x) := \int_{-\infty}^{x} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}}\, du \quad (n \to \infty) \text{ for all } x \in \mathbb{R}. \tag{CLT}$$

Because these results are so important, and because one does not always have a (finite) mean and variance, it was of great interest to find versions of them which held under weaker moment conditions. It was realized by Sakovich in 1956 (Sak1956) that regular variation gave the right language here: what one needs for the first is that the *truncated mean is slowly varying*,

$$\int_{-x}^{x} u\,dF(u) \sim \ell_1(x) \qquad (x \to \infty),$$

and for the second that the *truncated variance is slowly varying*,

$$\int_{-x}^{x} u^2\,dF(u) \sim \ell_2(x) \qquad (x \to \infty),$$

where ℓ_1, ℓ_2 are slowly varying (below).

Note. Oddly, despite their importance, these results were overlooked at the time, and they were re-discovered and given prominence in Feller's book (Fel1966). The first author saw them there then ('love at first sight').

More curiously still, although regular variation dates back to 1930 (below), the classic monograph by Gnedenko and Kolmogorov (GneK1954) (the Russian original is from 1949) made no use of it. So its treatment of these and related matters is unnecessarily complicated, and in particular the analysis and the probability are not properly separated. We note that Sakovich's PhD was supervised by Gnedenko.

For more on Gnedenko's work, and his very interesting life, see Bin2014.[2]

Then one has the third member of the trilogy, LLN–CLT–LIL: the *law of the iterated logarithm*. Here the norming, which gives the result its name, is intermediate between those in (LLN) and (CLT), and the conclusion is of a different type:

$$\limsup\,(S_n - n\mu)/\sigma\sqrt{2n\log\log n} = +1 \quad \text{a.s.}, \qquad \liminf \cdots = -1 \quad \text{a.s.}$$

[2] The text of a talk given by the first author at the Gnedenko Centenary Memorial Meeting, Moscow State University, 2012.

Indeed,

$$(S_n - n\mu)/\sigma\sqrt{2n \log\log n} \to\to [-1, 1] \quad \text{a.s.,} \tag{LIL}$$

meaning that all points in $[-1, 1]$, and no others, are limits of subsequences, a.s.

Stable laws (limit laws of centred and normed sums of independent copies) provide another good example. See, e.g., two approaches to the 'domain of attraction' problem here by Pitman and Pitman (PitP2016) and Ostaszewski (Ost2016a).

For more on the early history of regular variation in probability theory, see Bin2007.

Extremes The extreme values in a sample – sample maximum and minimum – have always been of great practical importance (the strength of a chain is that of its weakest link, etc.). The theory here dates back to Fisher and Tippett in 1928, so to before regular variation, though the relevance of regular variation was soon realized. The area is growing in importance nowadays, e.g. because of climate change and global warming. There was pioneering early work by Gnedenko in 1943, but the systematic use of regular variation to study extremes stems from de Haan in 1970 (Haa1970). For background and historical comments, see, e.g., our recent survey BinO2021b and the references there.

While Hardy himself was not interested in probability, the Tauberian theory to which he and Littlewood contributed so much has proved very useful in probability theory; see, e.g., Bin2015b.

P.3 Complex Analysis

Recall (see, e.g., BGT, Ch. 4) that an *Abelian* theorem passes from a stronger mode of convergence to a weaker one (such results are usually easy); a *Tauberian* one gives a converse, under an additional condition (a *Tauberian condition*); *Mercerian* theorems (see, e.g., BGT, Ch. 5) are hybrids, going from a condition on both to a stronger conclusion, under *no* Tauberian condition. A prototypical Abelian result will pass from a function

$$f(x) \sim x^\rho \ell(x) \quad (\ell \in R_0) \quad (x \to \infty)$$

to a Mellin convolution

$$(f * k)(x) := \int_0^\infty k(t) f(x/t) dt/t,$$

where ρ lies in the vertical strip in the complex s-plane where the Mellin transform

$$\hat{k}(s) := \int_0^\infty t^{-s} k(t) dt/t = \int_0^\infty u^s k(1/u) du/u$$

converges absolutely, giving

$$(f * k)(x) \sim \hat{k}(\rho) x^\rho \ell(x) \quad (x \to \infty);$$

the factor $\hat{k}(\rho)$ is to be expected, since if $f(x) = x^\rho$, $(f * k)(x) = \hat{k}(\rho) x^\rho$. The Tauberian converse reverses the implication for kernels satisfying Wiener's condition that $\hat{k}(s)$ be non-vanishing for $Re\ s = \rho$ ((NV) below), under suitable Tauberian conditions on f. The Mercerian (or 'ratio Tauberian', below) statement assumes convergence of the quotient,

$$(f * k)(x)/f(x) \to c \qquad (x \to \infty)$$

for some constant c, and deduces regular variation of both as above, with $c = \hat{k}(\rho)$.

For entire functions of finite order, one can look (in discs centre 0 and large radius r) at growth rates of the function, its maximum modulus $M(r)$ and the zero-counting function $n(r)$ (both in discs centre 0 and radius r). Matters split between integer and non-integer order. Functions with real negative zeros are simplest; write $\mathcal{E}_\rho$ for the class of entire f with order $\rho < \infty$ and negative zeros. For $f \in \mathcal{E}_\rho$, one has the Valiron–Titchmarsh theorem (BGT, §7.2, Th. 7.2.2), proved by Tauberian methods involving regular variation (BGT, Ch. 4), based on the linear integral transform (Stieltjes transform)

$$\log f(z) = \int_0^\infty \frac{zn(t)}{t + z}\, dt/t \qquad (|\arg z| < \pi).$$

For non-integral order, regular variation of either of $n(r)$, $\log f(re^{i\theta})$ implies regular variation of the other, and convergence of the quotient to a non-zero limit. For more on the Valiron–Titchmarsh Theorem, see DrasS1970 and the references cited there.

The question of whether convergence of this quotient implies regular variation of both functions has been called of 'ratio Tauberian' type; it is in fact *Mercerian* (BGT, Ch. 5). The first such results are due to Edrei and Fuchs (EdrF1966) and Shea (She1969), for $f \in \mathcal{E}_\rho$ and the Stieltjes transform above. Such results were extended by Drasin (Dras1968) and Drasin and Shea (DrasS1976) to more general kernels, using Wiener Tauberian theory. Drasin and Shea had *non-negative* kernels k, for which the relevant Mellin transforms converge absolutely in their strip of convergence. Matters are more complicated when the kernel can change sign (as with Fourier sine and cosine transforms, and Hankel transforms), as here there can be strips of conditional convergence (or Abel summability) also; see Jor1974. The Fourier and Hankel cases were considered in detail by Bingham and Inoue (BinI1997; BinI1999).

As may be seen from the Wiener Tauberian Theorem, P.7.1: if regular variation of index ρ (membership of R_ρ) is to appear in the hypothesis and conclusion, the key condition on the kernel is the *non-vanishing* condition

$$\hat{k}(s) \neq 0 \qquad (\operatorname{Re} s = \rho). \tag{NV}$$

In the corresponding Mercerian results, the key condition on k is the *no-repeat* condition

$$\hat{k}(s) = \hat{k}(\rho) \qquad \text{on } \operatorname{Re} s = \rho \text{ only for } s = \rho \tag{NR}$$

(She1969; Jor1974; cf. PalW1934, IV, (18.09)).

One can also consider the *Nevanlinna characteristic*

$$T(r) = T(r, f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \log_+ |f(re^{i\theta})|\, d\theta$$

(see, e.g., Haym1964). For $f \in \mathcal{E}_\rho$, this is given by the *non-linear* integral transform

$$T(r) = \sup \left\{ \int_0^\infty P(r/t, \theta) N(t) dt/t : \theta \in (0, \pi) \right\},$$

where

$$N(r) := \int_0^r n(t) dt/t, \qquad P(t, \theta) := \frac{1}{\pi} \frac{\sin\theta}{t + 2\cos\theta + t^{-1}}.$$

Baernstein (Bae1969) obtained a non-linear Tauberian theorem (the passage from N to T is Abelian, and simple, EdrF1966): for f entire of genus 0, if $T(r) \sim r^\lambda \ell(r)$ as $r \to \infty$ for ℓ slowly varying, then

(a) if $\lambda \in [0, \frac{1}{2}]$, then $N(r) \sim r^\lambda \ell(r)$;
(b) if $\lambda \in [\frac{1}{2}, 1]$, and f has only negative zeros, then $N(r) \sim \sin \pi\lambda\, r^\lambda \ell(r)$.

The corresponding Mercerian (or ratio Tauberian) theorem was proved by Edrei and Fuchs (EdrF1966), for $\lambda \in [\frac{1}{2}, 1]$: if the ratio converges, then both $N(r)$ and $T(r)$ are regularly varying.

Baernstein (Bae1969) conjectured that his results extend to $\mathcal{M}_\rho$, the class of meromorphic functions of finite order ρ, negative zeros and positive poles, but this is not the case. For counterexamples and discussion, see Dras2010. However, they do extend to the subclass $\mathcal{J}_\rho$ of $\mathcal{M}_\rho$ whose zeros a_n and poles b_n are symmetrically related, $a_n = -b_n$ (Will1972). Edrei (Edr1969) removes geometric restrictions on the zeros and poles, but at the cost of obtaining only 'locally Tauberian' results, in which $r \to \infty$ only through the union of a well-chosen sequence of large intervals.

One can extend to real zeros. One can use the language of proximate orders, due to Valiron in 1913, which can be shown to be equivalent to that of regular variation (and thus that Valiron may be credited with initiating the subject). The resulting *Levin–Pfluger theory* (A. Pfluger in 1938, B. Ya. Levin in 1964; BGT, Ch. 7) may be regarded as weakening the severe geometric restriction that all zeros lie on one ray, or one line, as far as possible.

The contrasts between key examples throw light on the theory, which they inspired. To quote BGT (end of §7.6): 'It is instructive to compare the two examples $\sin \pi z$ and $1/\Gamma(z)$. Their rates of growth differ, as above; their zeros differ not so much in their density as in their geometry. An extensive study of the integer-order case has been given by Pfluger (1946), motivated by the contrasts between these examples.'

As well as the maximum modulus, the minimum modulus of an entire function is of interest:

$$M(r) := \sup\{|f(z)| : |z| \le r\}, \qquad m(r) := \inf\{|f(z)| : |z| \le r\}.$$

One has the $\cos \pi\rho$ *theorem* (see BGT, §7.7; Bae1974; Ess1975 and the references there for details): if f is entire of order $\rho \in [0, 1)$,

$$\limsup \frac{\log m(r)}{\log\ M(r)} \ge \cos \pi\rho.$$

Functions extremal here are particularly interesting; see DrasS1969. Here one encounters *exceptional sets*, of logarithmic density 0, which cannot be avoided (Haym1970).

Pólya Peaks. Pólya peaks (of the 'first and second kinds') are a device in real analysis, introduced by Pólya (Poly1923) for the study of entire functions. They were named by Edrei in the 1960s. Their use was extended to meromorphic functions by Hayman (Haym1964, §4.4); for details, see DrasS1972. It turns out that they are intimately linked to the *Matuszewska indices* (BGT, §2.1) $\alpha(f)$, $\beta(f)$ of regular variation: both kinds of peak exist in the interval $[\beta(f), \alpha(f)]$ and nowhere else (the Pólya Peak Theorem; BGT1987, Th. 2.5.2).

In fact, the use of Pólya peaks in the results above (Edrei, Drasin and Shea, Jordan) may be avoided; see Bingham and Inoue (BinI2000a). This may be preferred on thematic grounds in complex analysis, as well as to simplify the proofs.

Recently, essential use of O-regular variation has been made in the theory of *ultraholomorphic functions*; see JimSS2019 for background and details.

P.4 Analytic Number Theory

For background on Abelian, Tauberian and Mercerian theorems as mentioned above, see, e.g., BGT, §4.5; Kor2004.

Tauberian theorems such as the Hardy–Littlewood–Karamata theorem are extensively used in analytic number theory (see, e.g., Ten2015; BGT, Ch. 4). So too is e.g. Kohlbecker's Tauberian Theorem on asymptotics of partitions (BGT, Th. 4.12.1). Tauberian (and Mercerian) theorems can be used to give short proofs of the Prime Number Theorem (BGT, §6.2).

The *prime divisor functions*,

$$\omega(n) := \#\text{ distinct prime divisors of } n,$$
$$\Omega(n) := \#\text{ prime divisors of } n \qquad \text{(counted with multiplicity)}$$

(Ten2015, I.2.2), illustrate our approach well. The classical estimates are (Hardy and Ramanujan in 1917; Ten2015, I.3.6,7)

$$\frac{1}{x}\sum_{n\le x}\omega(n) = \log\log x + c_1 + O(1/\log x) \qquad (x\to\infty),$$

with c_1 a known constant, and similarly for $\Omega(n)$ with a different known constant $c_2 > c_1$. One also has the classical Erdős–Kac central limit theorem of 1939,

$$\frac{1}{x}\,|\{n\le x : \omega(n)\le \log\log x + \lambda\sqrt{\log\log x}\}| \to \Phi(\lambda) \quad (x\to\infty \text{ for all } \lambda\in\mathbb{R}),$$

the beginning of probabilistic number theory, and its refinement of Berry–Esseen type, due to Rényi and Turán in 1958, with error term $O(1/\sqrt{\log\log x})$ uniform in λ (Ten2015, III.4.15). Our methods give (BinI2000b)

$$\frac{1}{\lambda x}\sum_{n\le\lambda x}\omega(n) - \frac{1}{x}\sum_{n\le x}\omega(n) \sim \frac{\log\lambda}{\log x} \quad (x\to\infty \text{ for all } \lambda\in\mathbb{R}).$$

This is a statement of regular-variation type, so it has a representation theorem, namely

$$\frac{1}{x}\sum_{n\le x}\omega(n) = C + \int_2^x (1+o(1))\frac{dt}{t\log t} + o(1/\log x)$$

(note the *two* error terms, one inside the integral, one outside). This is not comparable to the classical results. There, it is the size of the error terms that counts, but there is no information on behaviour under differencing; here, matters are reversed. Similarly for results of Mertens (Ten2015, I.1.4; HarW2008, Th. 425, 427) on sums over primes p (BinI2000b),

$$\sum_{p \le x} \frac{\log p}{p}, \qquad \sum_{p \le x} 1/p.$$

P.5 Regular Variation: Preliminaries

The subject of regular variation originates with the Yugoslav mathematician Jovan Karamata (1902–1967) in 1930 (Kar2009). It concerns limiting relations of the form

$$f(\lambda x)/f(x) \to g(\lambda) \qquad (x \to \infty) \quad \forall\, \lambda > 0, \tag{K}$$

for positive functions f on $\mathbb{R}_+$. Relevant here is the multiplicative group of positive reals, $(\mathbb{R}_+, \times)$, with Haar measure dx/x. While this formulation is the one useful for applications, for theory it is more convenient to pass to the additive group of reals, $(\mathbb{R}, +)$, Haar measure Lebesgue measure dx, where we write this as

$$h(u + x) - h(x) \to k(u) \qquad (x \to \infty) \quad \forall\, u \in \mathbb{R}. \tag{K$_+$}$$

We can pass at will between these two formulations via the exp/log isomorphism. The core of the resulting theory is treated in full in Chapter 1 of BGT, with further results (e.g. with lim replaced by lim sup – where one may lose measurability, by 'character degradation') in Chapter 2 of BGT.

The limit function g in (K) satisfies the *Cauchy functional equation*

$$g(\lambda\mu) = g(\lambda)g(\mu) \qquad (\lambda,\ \mu > 0). \tag{CFE}$$

For background on functional equations, see AczD1989; the classic context is Ban1920.

P.6 Topological Regular Variation

Solutions to (CFE), as is typical with functional equations, exhibit a sharp dichotomy: they are either *very nice* (continuous, here) or *very nasty* (pathological – unbounded above and below on every interval, or even on any non-meagre Baire set or non-null measurable set). Since (as we shall see below) such 'bad' solutions can be manufactured at will from a Hamel basis (of the reals, as a vector space over the rationals), we will call this the *Hamel pathology*.

Under mild regularity conditions (such as measurability, the Baire property, No Trumps **NT**, etc.), this gives

$$g(\lambda) \equiv \lambda^\rho$$

for some $\rho \in \mathbb{R}$. Then f is called *regularly varying* with *index* ρ, $f \in R_\rho$. Functions in class R_0 are called *slowly varying*, written ℓ (for *lente*, or *langsam*).

By taking logarithms, (K) may be written in terms of the limit of the difference of a function at arguments λx and x. It turns out that this may be fruitfully generalized by introducing a denominator $\ell \in R_0$:

$$[f(\lambda x) - f(x)]/\ell(x) \to k(\lambda) \qquad (x \to \infty) \qquad \forall \lambda > 0 \qquad \text{(BK/DH)}$$

(using a denominator in R_ρ for $\rho \neq 0$ gives nothing new; see, e.g., BGT, §3.2). This study goes back to Bojanic and Karamata (BojK1963), and independently to de Haan (Haa1970), whence the name (BK/DH); see BGT, Chapter 3 for a full account.

There are three key theorems that underlie any form of regular variation (there are two forms above; more will follow). These are (under mild conditions):

the *Uniform Convergence Theorem*, **UCT:** the convergence in (K), (BK/DH) takes place *uniformly* on compact λ-sets in $\mathbb{R}_+$;

the *Representation Theorem***:** giving that $\ell \in R_0$ if and only if it may be written in the form

$$\ell(x) = c(x) \exp\{\int_1^x \epsilon(u)\, du/u\} \qquad (x \geq 1) \qquad \text{(RT)}$$

where

$$c(x) \to c \in \mathbb{R}_+, \qquad \epsilon(x) \to 0 \qquad (x \to \infty)$$

(here $\epsilon(.)$ may be taken as smooth as desired, while $c(.)$ may be as rough as the mild regularity conditions allow);

the *Characterization Theorem***:** giving the form of $g(\lambda)$ as λ^ρ as above and that of k in (BK/DH) as

$$k(\lambda) = c h_\rho(\lambda),\ c \in \mathbb{R}_+; \quad h_\rho(\lambda) := \int_1^\lambda u^{\rho-1}\, du = (\lambda^\rho - 1)/\rho \ (\lambda > 0),$$

with the usual ‘l’Hospital convention’ that the right-hand side above is taken as $\log \lambda$ when $\rho = 0$.

Even with the simplest functional equation that arises here (the Cauchy), some mild regularity condition is required. There is a dichotomy: as above solutions are either very nice (powers λ^ρ or the $h_\rho(\lambda)$ as above) or very nasty – pathological (e.g. unbounded above and below on any non-negligible set).

Exceptional Sets. There are situations in which the passage to the limit in slow and regular variation needs to avoid some *exceptional set*; see BGT, §2.9. Examples arise in complex analysis: BGT, Th. 7.2.4 (a result of Titchmarsh in 1927), and in work of Drasin and Shea (DrasS1976) on functions extremal for the $\cos \pi\rho$ theorem of Wiman and Valiron mentioned above. We will need such exceptional sets below, in dealing with sequential aspects of regular variation.

Thinning (Quantifier Weakening). Another key question, going back to a conjecture of Karamata, concerns *weakening the quantifier*, $\forall$, in (K), (BK/DH): requiring the convergence to take place for *some but not all* $\lambda > 0$. Rather than having a continuum of conditions to check, one may be able to reduce this to finitely many, or even to just *two* (of course, one could not expect just one to suffice!). Results of this kind – which one might call *thinning* results, as they involve thinning of the λ-set on which convergence is required (cf. BinO2010a) – go back to Heiberg (Hei1971) and Seneta (Sen1973). Interestingly, given the side-condition of Heiberg–Seneta type, one no longer needs to impose the regularity condition needed above to eliminate pathology.

Matters were taken further in Bingham and Goldie (BinGo1982a) and Bingham and Ostaszewski (BinO2018a; BinO2020a): with

$$g^*(\lambda) := \limsup_{x\to\infty} f(\lambda x)/f(x),$$

assume

$$\limsup_{\lambda\downarrow 1} g^*(\lambda) \le 1.$$

Then the following are equivalent (for positive f):

(i) there exists $\rho \in \mathbb{R}$ such that

$$f(\lambda x)/f(x) \to \lambda^\rho \qquad (x \to \infty) \qquad \forall \lambda > 0;$$

(ii) $g(\lambda) := \lim_{x\to\infty} f(\lambda x)/f(x)$ exists, finite, for a λ-set of positive measure [a non-meagre Baire set];

(iii) $g(\lambda)$ exists, finite, in a λ-set dense in $\mathbb{R}_+$;

(iv) $g(\lambda)$ exists, finite, for $\lambda = \lambda_1,\ \lambda_2$ with $(\log \lambda_1)/\log \lambda_2$ irrational.

The reader may recognize that *Kronecker's Theorem* (HarW2008, Ch. XXIII) lies behind (iv) here.

There are corresponding thinning results for (BK/DH); see BGT, Th. 3.2.5, Th. 1.4.3. As remarked there, the result for (K) is no easier, despite its simpler context. This is because the thinning takes place in the quantifier over λ, which affects only the numerator in (BK/DH). The effect is that the general ℓ in the denominator is no harder to handle than $\ell \equiv 1$, when (BK/DH) reduces to the logarithmic form of (K).

Remark The proofs of these two thinning results from BGT are the hardest in that book on the theory of regular variation as such (that of the Drasin–Shea–Jordan theorem, BGT, §5.2, is harder, but belongs rather to Mercerian theory). The search for simpler proofs was the motivation behind several of our papers, on thinning (BinO2010a) and quantifier weakening (BinO2018a; BinO2020a).

Frullani Integrals. The *Frullani integral* (G. Frullani, 1828; see Ostr1976 for the history) of a locally integrable function ψ on $\mathbb{R}_+$ is the improper integral

$$I = I(\psi; a, b) := \int_{0+}^{\infty-} \{\psi(\lambda t) - \psi(t)\}dt/t = \lim_{\epsilon\downarrow 0, X\uparrow\infty} \int_{\epsilon}^{X} \{\psi(\lambda t) - \psi(t)\}dt/t.$$

Writing $bt = u$, we see that $I(\psi; a, b) = I(\psi; a/b, 1) = I(\psi, a/b)$, say. So we may restrict attention to

$$I = I(\psi; \lambda) := \int_{0+}^{\infty-} \{\psi(\lambda t) - \psi(t)\}dt/t = \lim_{\epsilon\downarrow 0, X\uparrow\infty} \int_{\epsilon}^{X} \{\psi(\lambda t) - \psi(t)\}dt/t.$$

Now

$$\int_{\epsilon}^{X} \{\psi(\lambda t) - \psi(t)\}dt/t = \int_{X}^{\lambda X} \psi(t)dt/t - \int_{\epsilon}^{\lambda\epsilon} \psi(t)dt/t.$$

So the two limits, concerning behaviour at ∞ and at 0, may be handled separately. One obtains (BinGo1982b, §6; BGT Th. 1.6.6):

Theorem *For ψ and its Frullani integral $I(\psi; \lambda)$ as above, the following are equivalent:*

(i) $I(\psi; \lambda)$ *exists for all* $\lambda \in \mathbb{R}_+$*;*
(ii) $I(\psi; \lambda)$ *exists for* λ *in a set of positive measure [a non-meagre Baire set];*
(iii) $I(\psi; \lambda)$ *exists for* λ *in a dense set in* $\mathbb{R}_+$ *(or for* $\lambda = \lambda_1$, λ_2 *with* $(\log \lambda_1)/\log \lambda_2$ *irrational), and*

$$\liminf_{\lambda\downarrow 1} \liminf_{x\to\infty} \int_{x}^{\lambda x} \psi(t)dt/t \ge 0,$$
$$\liminf_{\lambda\downarrow 1} \liminf_{x\to\infty} \int_{x}^{\lambda x} \psi(1/t)dt/t \ge 0.$$

Each of (i)–(iii) *holds if and only if both of the following finite limits exist for some (all)* $\sigma > 0$*:*

$$M = M(\psi) = \lim_{x\to\infty} \sigma x^{-\sigma} \int_1^x u^\sigma \psi(u) du/u,$$
$$m = m(\psi) = \lim_{x\to\infty} \sigma x^{-\sigma} \int_1^x u^\sigma \psi(1/u) du/u.$$

Then the Frullani integral is given by

$$I(\psi, \lambda) = (M - m) \log \lambda \qquad (\lambda \in \mathbb{R}_+).$$

This result extends those of Hardy and Littlewood (HarL1924), where $\sigma = 1$.

Frullani integrals occur in probability theory, e.g. in the fluctuation theory of Lévy processes; see, e.g., Bert1996, III.1, p. 73.

Convergence and Cesàro Convergence. The mathematics of the Frullani integral above yields as a by-product the results below (BinGo1982b, §6), showing exactly what is needed for a Cesàro convergent function or sequence to converge. For functions: for ϕ locally integrable on $[0, \infty)$,

$$\frac{1}{x} \int_0^x \phi(t) dt \to c \qquad (x \to \infty)$$

if and only if

$$\phi(x) = a(x) + b(x), \qquad \text{where } a(x) \to c, \ \int_1^\infty b(t) dt/t \text{ is convergent.}$$

For sequences:

$$\frac{1}{n} \sum_1^n s_k \to c \qquad (n \to \infty)$$

if and only if

$$s_n = a_n + b_n, \qquad \text{where } a_n \to c, \ \sum_1^\infty b_n/n \text{ is convergent.}$$

We include the proof (due to G. E. H. Reuter) as it is so short.

Proof If $\sum b_n/n$ converges, $(b_1 + \cdots + b_n)/n \to 0$ by Kronecker's Lemma. So if $s_n = a_n + b_n$ as above, $s_n \to c$ (C_1).

Conversely, if $s_n \to c$ (C_1), set $a_{n+1} := (s_1 + \cdots + s_n)/n$. Then $s_n = a_n + n(a_{n+1} - a_n)$, and this is the required decomposition, since if $b_n := n(a_{n+1} - a_n)$, $\sum b_n/n$ converges. □

Beurling Slow Variation. For $\phi\colon \mathbb{R} \to \mathbb{R}_+$ Baire or measurable, ϕ is called *Beurling slowly varying* if

$$\phi(x) = o(x) \quad \text{and} \quad \phi(x + t\phi(x))/\phi(x) \to 1 \text{ for all } t \in \mathbb{R},\ (x \to \infty). \quad \text{(BSV)}$$

This originated in unpublished lecture notes of Beurling in 1957 on his Tauberian theorem (below); see Kor2004, IV.11.

If also the convergence in (BSV) is locally uniform in t, ϕ is called *self-neglecting*, written $\phi \in \mathrm{SN}$. The term and the concept arose in probability theory; see, e.g., BinO2021b and the references there.

It was shown by Bloom (Blo1976) that for ϕ *continuous*, the convergence in (BSV) is indeed locally uniform. In fact Bloom's proof needs only the *Darboux property*, or *intermediate-value property*, that ϕ takes every value between any two values it attains. For this and other results, see BinO2014; Ost2015a. Here it is enough to require that ϕ takes a dense set of values between any two values attained, but the question of whether a Darboux-like property can be dropped altogether remains open.

The Representation Theorem gives the Beurling slowly varying ϕ as those *positive* functions of the form

$$\phi(x) = c(x) \int_0^x \epsilon(u)du \qquad (x \in \mathbb{R}),$$

where ϵ is C^∞ with $\epsilon(x) \to 0$ as $x \to \infty$, and c is Baire/measurable with $c(x) \to c \in (0,\infty)$ as $x \to \infty$ (BGT Th. 2.11.3; BinO2014, §9; Ost2015a, p. 731).

P.7 Tauberian Theorems

We first recall *Wiener's Tauberian Theorem* (see, e.g., Har1949, XII; Kor2004, II):

Theorem P.7.1 (Wiener's Tauberian Theorem) *Suppose $K \in L_1(\mathbb{R})$ with Fourier transform $\hat{K}$ non-vanishing on $\mathbb{R}$, and $H \in L_\infty(\mathbb{R})$. If*

$$\int K(x-y)H(y)\,dy \;\to\; c\int K(y)\,dy \qquad (x \to \infty),$$

then, for all $G \in L_1(\mathbb{R})$,

$$\int G(x-y)H(y)\,dy \;\to\; c\int G(y)\,dy \qquad (x \to \infty).$$

Here integrals are over $\mathbb{R}$ and we use additive notation; one may work multiplicatively with $\int_0^\infty K(x/y)H(y)dy/y$, etc. The *Tauberian condition* here is of O-type: $H \in L_\infty(\mathbb{R})$, or $H = O(1)$.

Beurling's Tauberian Theorem generalizes Wiener's Tauberian Theorem, to which it reduces in the special case $\phi \equiv 1$:

Theorem P.7.2 (Beurling's Tauberian Theorem) *Suppose $K \in L_1(\mathbb{R})$ with Fourier transform $\hat{K}$ non-vanishing on $\mathbb{R}$, with ϕ Beurling slowly varying and $H \in L_\infty(\mathbb{R})$. If*

$$\int K\left(\frac{x-y}{\phi(x)}\right) H(y)\,dy/\phi(x) \;\to\; c\int K(y)\,dy \qquad (x \to \infty),$$

then, for all $G \in L_1(\mathbb{R})$,

$$\int G\left(\frac{x-y}{\phi(x)}\right) H(y)\,dy/\phi(x) \;\to\; c\int G(y)\,dy \qquad (x \to \infty).$$

Note that the arguments $x - y$ in Theorem P.7.1 involve the additive *group* $(\mathbb{R}, +)$, and the x/y of its multiplicative version that of the multiplicative *group* $(\mathbb{R}_+, \times)$, while the $(x - y)/\phi(x)$ of Theorem P.7.2 involve the *ring* structure of $\mathbb{R}$. The two results are thus structurally distinct.

For a short and elegant reduction of Beurling's Tauberian Theorem to Wiener's, see Kor2004, IV, Th. 11.1. For an early use of Beurling's Tauberian Theorem in probability theory, see Bin1981.

The Borel–Tauber Theorem

The two most basic families of summability methods are the *Cesàro* $C_\alpha (\alpha > 0)$ and *Abel* methods A; see, e.g., Har1949, V–VII; Kor2004, I. Perhaps next in importance, though harder, are the *Euler* E_p $(p \in (0, 1))$ and *Borel* methods; see, e.g., Har1949, VIII, IX; Kor2004, VI. The Euler and Borel methods (plus those of Taylor and Meyer–König) belong to the family of *circle methods* (German: Kreisverfahrung; see MeyK1949). The name derives from the circle of convergence of a power series; such methods were used for analytic continuation by power series.

The key Tauberian theorem for the Borel method – 'Borel–Tauber Theorem' – is:

Theorem P.7.3 (Borel–Tauber Theorem) *For $s_n := \sum_0^n a_k$: if*

$$e^{-x}\sum_0^\infty s_n x_n/n! \;\to\; s \qquad (x \to \infty)$$

and $a_n = O(1/\sqrt{n})$, then

$$s_n \;\to\; c \qquad (n \to \infty).$$

The setting of Theorem P.7.3 is discrete, involving a sum, while that of Theorems P.7.1 and P.7.2 is continuous, involving an integral. To pass between them, one may either use 'Wiener's second theorem': see, e.g., Har1949, 12.7, which demands less of the integrand (so $H(y)dy$ becomes a Stieltjes integral, $dU(y)$, say) but more of the kernel K; or use an auxiliary approximation argument. There is much more to be said here, but we must refer for further detail to, e.g., Kor2004, VI.

The Tauberian Condition. Here the Tauberian condition is $a_n = O(1/\sqrt{n})$, and this is best-possible in that no weaker O-condition would suffice here. But because the weights $e^{-x}x^n/n!$ (of course those of the *Poisson* distribution $P(x)$ with parameter x) are non-negative, a *one-sided* Tauberian condition suffices: $a_n = O_L(1/\sqrt{n})$, meaning that $\sqrt{n}a_n$ is bounded below (or with O_R and bounded above). In fact such 'pointwise' conditions on the individual a_n are not needed, but rather 'averaged' forms of them involving differences of the s_n. The classical one is of 'slow-decrease' type, due to R. Schmidt in 1925 (see Kor2004, VI.12):

$$\liminf\, (s_m - s_n) \geq 0 \qquad (m, n \to \infty,\ 0 \leq \sqrt{m} - \sqrt{n} \ \to\ 0).$$

Such one-sided Tauberian conditions are studied at length in Bingham and Goldie (BinGo1983).

Valiron Methods. For $\beta \in (0, 1)$, write V_β for the *Valiron* summability method (Bin1984b), given by writing

$$\frac{1}{x^\beta\sqrt{2\pi}}\sum_0^\infty s_k \exp\left\{-\frac{1}{2}(x-k)^2/x^{2\beta}\right\} \ \to\ s \qquad (x \to \infty)$$

as

$$s_n \ \to\ s. \tag{V_β}$$

Our principal concern is with the case $\beta = \frac{1}{2}$ (see BinT1986). One sees that the sum above is a discrete form of the condition in Theorem P.7.2 with $K(x) = e^{-x^2/2}/\sqrt{2\pi}$. This is the standard normal probability density Φ or $N(0, 1)$, with Fourier transform (characteristic function) $\exp\{-\frac{1}{2}t^2\}$, which is non-vanishing as in Theorems P.7.1 and P.7.2. This K may thus serve as a Wiener kernel. In the notation of Theorems P.7.1 and P.7.2, one can obtain boundedness of $H(.)$ from the other conditions; see Har1949, p. 220 for the pointwise Tauberian condition $a_n = O(1/\sqrt{n})$ and Har1949, p. 225 for a reference to Vijayaraghan's method of monotone minorants for the slow-decrease Tauberian condition. This allows an easy proof of Theorem P.7.3 from Theorem P.7.2.

That the methods $V_{\frac{1}{2}}$ and B are intimately linked has been known since Hardy and Littlewood in 1916 (HarL1916). In probabilistic language, this link reflects the *central limit theorem*: the Poisson law $P(x)$ with large parameter x is an n-fold convolution of $P(x/n)$ with itself, and so approaches normality. Much more is true: the rapid tail-decay of the Poisson laws allows the use of large-deviation methods. See Kor2004, VI, Th. 6.1, where the range $|n - x| < x^\gamma$ occurs, where $1/2 < \gamma < 2/3$. As Korevaar remarks, this parameter range is the 'signature' of large deviations.

Jakimovski and Karamata–Stirling Methods. There are other summability methods whose weights exhibit central-limit behaviour. We consider independent random variables X_n, integer-valued (so that the weights will form a matrix, below), with partial sums $S_n = \sum_1^n X_k$; write

$$a_{nk} := \mathbb{P}(S_n = k),$$

and write $A = (a_{nk})$ for the summability matrix. The classical case is of *Jakimovski methods* (Jak1959; ZelB1970, §70); here the X_n are Bernoulli (0–1 valued), with

$$\mathbb{P}(X_n = 1) = p_n, \qquad P(X_n = 0) = q_n := 1 - p_n.$$

Writing $p_n = 1/(1 + d_n)$ $(d_n \geq 0)$, this gives

$$\prod_{j=1}^{n} \left(\frac{x + d_j}{1 + d_j} \right) = \sum_{k=0}^{n} a_{nk} x^k,$$

and the *Jakimovski method* $[F, d_n]$. The motivating examples are:

(i) the Euler methods, with $d_n = 1/\lambda$, say written $E(\lambda)$;

(ii) the *Karamata–Stirling methods* $\mathrm{KS}(\lambda)$, with $d_n = (n-1)/\lambda$. Here

$$a_{nk} = \lambda^k S_{nk}/(\lambda)_n,$$

with (S_{nk}) the Stirling numbers of the second kind and

$$(\lambda)_n := \lambda(\lambda + 1) \cdots (\lambda + n - 1).$$

See Bin1988 for their Tauberian theory and BinS1990 for LLN and LIL results.

Turning from the non-identically distributed Bernoulli case to the identically distributed general integer-valued case gives the *random-walk* methods (Bin1984a).

All the summability methods considered here are closely enough linked to be *equivalent for bounded sequences* (as are Euler and Borel, and indeed as are Cesàro and Abel).

Riesz Means and Moving Averages. With K as above, taking $H(x) = H_a(x) := a^{-1}I_{[0,a]}(x)$ gives conclusions of the form

$$\frac{1}{a\sqrt{n}} \sum_{n \le k < n+a\sqrt{n}} s_k \to s \qquad (n \to \infty),$$

passing back from integrals to sums as above. These are *Riesz means* (HarR1915, IV; Har1949, §4.16, §5.16); 'typical means' there and in ChanM1952, or *moving averages* in the language of probability and statistics. For more on Riesz means and Beurling moving averages, see Bin1981; Bin2019. For related moving averages, see BinG2015; BinO2016a; BinG2017.

The Fourier transform of H_a here is $\hat{H}_a(t) = (\exp(iat) - 1)/(iat)$, which has real zeros, so H cannot be used as a Wiener kernel. But two such H_a with a_1/a_2 irrational may be used, as their Fourier transforms have no common zeros (see, e.g., Wie1933, §10 Th. 6; BinI2000b).

In addition to Riesz means and moving averages, there is a third mode of convergence relevant here, 'perturbed Cesàro convergence with rate'. For $\beta \in (0, 1)$, one has (BinT1986, Th. 3) the equivalence as $n \to \infty$ of

$$s_n \to \sigma \qquad R(\exp\left(n^{1-\beta}\right), 1),$$

$$\frac{1}{un^{\beta}} \sum_{n \le k < n+un^{\beta}} s_k \to s, \quad \text{for some (all) } u > 0,$$

$$\frac{1}{n+1} \sum_{0}^{n} (s_k + \epsilon_k) = s + o\left(1/n^{1-\beta}\right) \quad \text{for some } \epsilon_n \to 0.$$

The most important case, $\beta = \frac{1}{2}$, is in Bin1981, Th. 2 and BinGo1983, Th. 3. It has distinguished antecedents. That the third statement is *sufficient* for Borel (and so Euler) convergence without the ϵ_n terms (so is clearly sufficient with them) is due to Hardy (Har1904, p. 55; Har1949, Th. 149: the first predates Karamata's work, the second does not). An approach via regular variation gives sufficiency of the general result: the relevant Representation Theorem gives the ϵ_n, which plays the role of the error term within the sum or integral there.

P.8 General Regular Variation

One can usefully combine and generalize all three forms of regular variation (Karamata, Bojanic–Karamata/de Haan, Beurling) encountered above. In BinO2020a we study *general regular variation*, in which one has

$$[f(x + t\phi(x)) - f(x)]/h(x) \to K(t) \quad \text{locally uniformly in } t. \qquad \text{(GRV)}$$

Here f is the function under study, $\phi \in BSV$ and h are auxiliary, and the limit K is called the kernel. By using the algebraic machinery of *Popa groups*, one can substantially reduce the theory to those of the earlier three. In addition, one encounters a number of *functional equations*: Cauchy, Gołąb–Schinzel, Chudziak–Jabłońska, Beurling–Goldie, Goldie. See BinO2020a for further detail (and the planned sequel to this book).

Sequential Results: Kendall's Theorem. As above, regular variation is a *continuous-variable* theory, while our preferred tool, the Baire Category Theorem, is a *discrete-variable* theorem about sequences. But it has long been recognized that sequential results are possible and useful; see, e.g., BGT, §1.9. One finds there reference to early work by Croft (Cro1957), Kingman (Kin1964) and Kendall (in particular Ken1968, Th. 16):

Theorem (Kendall's Theorem) *If*

$$\limsup_{x\to\infty} x_n = \infty, \qquad \limsup_{x\to\infty} x_{n+1}/x_n = 1$$

and, for some continuous positive functions f and g, interval $I = (a, b)$, $0 < a < b < \infty$ and sequence (a_n),

$$a_n f(\lambda x_n) \to g(\lambda) \in \mathbb{R}_+ \quad (n \to \infty) \quad \textit{for all} \quad \lambda \in (a, b),$$

then f varies regularly.

If then $f(x) \sim x^\rho \ell(x)$, one has (BinO2020b)

$$a_n \sim c x_n^{-\rho} \ell(x_n).$$

Because of the importance of Kendall's Theorem in applications, we should thus generalize this result as far as possible, in the light of what is now known. It turns out that one can generalize all three of f, g, I above, but at the cost of introducing an *exceptional set* (BGT, §2.9; DrasS1976). For a function f, say that $f(x)$ has *essential limit* $L = L(f)$, finite, as $x \to \infty$,

$$\text{ess-lim} f(.) = L,$$

if for all $\epsilon > 0$ there exist $X = X(\epsilon, f) \in \mathbb{R}$ and meagre $M = M(\epsilon, f)$ such that

$$|f(x) - L| < \epsilon \quad \text{for all } x > X, \ x \notin M.$$

Then (BinO2020b, Th. 2.3) one can weaken continuity of f to being Baire, continuity of g to being positive, and I an interval to being a non-meagre Baire set. The weakened conclusion is that

$$K(s) := \text{ess-lim}_{x\to\infty} f(s\lambda)/f(\lambda)$$

exists, finite and multiplicative. One calls such f *weakly quasi-regularly varying*. If further g is Baire, then (BinO2020b, Th. 2.5)

$$K(s) \equiv s^\kappa \quad \text{for some } \kappa \in \mathbb{R};$$

one calls such f *strongly quasi-regularly varying*.

There is also a character-degradation theorem (BinO2020b, Th. 8.1): if $k = k(.,.)$ is Borel, then K, where

$$K(s) := \text{ess-lim}_{x\to\infty}\, k(s,x)$$

is of ambiguous analytic class Δ_2^1 (see Chapter 7).

Functional Equations: Hamel Bases. The definition

$$f(\lambda x)/f(x) \to g(\lambda) \quad (x \to \infty) \quad \text{for all } \lambda \in (0,\infty)$$

leads immediately to

$$g(\lambda\mu) = g(\lambda)g(\mu) \quad \text{for all } \lambda, \mu \in (0,\infty)$$

(BGT, 1.4.1). This is the *Cauchy functional equation*, in multiplicative form. While this is the form preferred for applications, for theory it is better to change from this multiplicative setting in $(\mathbb{R}_+, \times)$ to the corresponding additive setting in $(\mathbb{R}, +)$ by writing $h(x) := \log f(e^x)$, $k(x) := \log g(e^x)$, giving

$$k(u+v) = k(u) + k(v) \quad \text{for all } u, v \in \mathbb{R}, \tag{CFE}$$

the Cauchy functional equation on the line. Such functions k are called *additive*. From (CFE), one obtains

$$k(mu) = mk(u), \quad k(u/n) = k(u)/n \quad \text{for all } u \in \mathbb{R},\ m \in \mathbb{N},\ n \in \mathbb{N}\setminus 0,$$

so

$$k(qu) = q\,k(u) \quad \text{for all } u \in \mathbb{R},\ q \in \mathbb{Q}.$$

Thus, writing $c := k(1)$,

$$k(x) = c\,x \quad \text{for all } x \in \mathbb{R}$$

if k is continuous, by approximation. So, continuous additive functions are linear.

One can easily extend this result vastly beyond continuity (BGT1987, 1.1.3). One obtains (Ostr1929, for measurable k; Meh1964, for the Baire case) that if an additive function k is bounded above or below on a non-null measurable set [a non-meagre Baire set], k is linear. Thus, an additive function k is linear or (highly) pathological.

To proceed, we need to invoke the Axiom of Choice, AC, in some form (e.g. Zorn's Lemma); that is, to extend the axiom system we work with from ZF (Zermelo–Fraenkel) to ZFC (i.e. ZF + AC). One can now prove easily that every vector space has a basis (see, e.g., Jec1973, 2.2.2). Conversely, it was shown by Blass in 1984 that existence of bases implies AC (Bla1984).

Regarding the real line $\mathbb{R}$ as a vector space over the rationals $\mathbb{Q}$ as ground field, $\mathbb{R}(\mathbb{Q})$ say, if we work in ZFC this shows (G. Hamel in 1905, Ham1905) that we have a *basis*, H say ('H for Hamel', below) for $\mathbb{R}$ over Q. Of course, H is uncountable; indeed, it has the power $\mathfrak{c}$ of the continuum (Kucz1985, Th. IV.2.3, p. 82).

We may now define, at will, *any* function $g: H \to \mathbb{R}$. This may be extended uniquely to a homomorphism $f: \mathbb{R} \mapsto \mathbb{R}$: each $x \in \mathbb{R}$ may be written uniquely as a finite linear combination

$$x = \sum \alpha_i b_i \qquad (c_i \in \mathbb{Q},\ b_i \in H).$$

Then

$$f(x) := \sum \alpha_i g(b_i).$$

If $x \in H$, the above representation of x reduces to $x = x$, so

$$f \mid H = g.$$

Also, if $y \in \mathbb{R}$ has the representation

$$y = \sum \beta_i b_i,$$

$$f(y) = \sum \beta_i g(b_i).$$

Then $x + y$ has the representation $\sum_i (\alpha_i + \beta_i) b_i$ (the range of summation here being the union of those in the two finite sums for x and y), so

$$f(x+y) = \sum (\alpha_i + \beta_i) g(b_i) = \sum \alpha_i g(b_i) + \sum \beta_i g(b_i) = f(x) + f(y):$$

f is additive. Were f continuous, we could make it discontinuous by changing its value at one point. But then by the Ostrowski and Mehdi results, f would be unbounded above and below on every interval, say. As no such change can be induced in a continuous function by changing its value at one point, we conclude that f is already discontinuous. Thus a Hamel basis gives us a way of manufacturing pathological (discontinuous) additive functions at will. We call this behaviour the *Hamel pathology*. Such pathological functions – or, identifying a function with its graph, functions with graph a Hamel basis in the plane – are called *Hamel functions*.

The argument above can be presented for additive functions $k\colon \mathbb{R}^d \to \mathbb{R}$ (Kucz1985, V.2); we take $d = 1$ here for simplicity.

Additive functions thus have the property that even a little regularity forces great regularity (the form $c\ x$). Ostaszewski (Ost2015a, p. 729) lists ways in which this can happen: additive functions are continuous if they are:

- Baire (Ban1932, I §3 Th.4);
- measurable (Fre1913; Fre1914);
- bounded on a non-null measurable set (Ostr1929);
- bounded on a non-meagre Baire set (Meh1964).

See BinO2011a for details and references.

P.9 Hamel Bases

Despite the pathological behaviour of the Hamel *functions* above, Hamel *bases* as sets may not themselves be pathological. In 1920 Sierpiński (Sie1920) showed that:

- a Hamel basis H can be (Lebegue-)measurable;
- (Th. I) any measurable Hamel basis has measure 0;
- any Hamel basis has inner measure 0;
- a Hamel basis can be non-measurable.

Thus the classes $\mathcal{H}_1$, $\mathcal{H}_2$ of measurable and non-measurable Hamel bases are both non-empty. Sierpiński also showed (Th. II) that no Hamel base can be an analytic set – indeed, it cannot even be a Borel set. He ends with a corollary of his proofs: There exist two measurable sets $X, Y \subseteq \mathbb{R}$ such that the set of sums $\{x + y\colon\ x \in X, y \in Y\}$ is non-measurable.

Being a Hamel basis is a purely algebraic concept, while we can switch between the measure and category cases by switching between the density topology (Chapter 7) and the Euclidean topology. We conclude that the classes of Hamel bases with and without the Baire property are both non-empty:

- a Hamel basis may or may not have the Baire property, both cases being possible.
- Sierpiński (Sie1935) also showed this, assuming the Continuum Hypothesis, CH, for part of it.

F. Burton Jones showed in 1942 that an additive function continuous on a set T which is analytic and contains a Hamel basis is continuous (Jon1942b); see also Jon1942b. Kominek proved in 1981 the analogous result with ‘continuous’ replaced by 'bounded' (Kom1981). Motivated by the analogy between these two results, the present authors (BinO2010a) gave a result with both the Jones and Kominek theorems as corollaries, using Choquet's Capacitability Theorem. They also deduced Jones' theorem from Kominek's and gave another proof of the Uniform Convergence Theorem for slowly varying functions.

Płotka (Plo2003) showed that *every* function $f : \mathbb{R} \to \mathbb{Q}$ can be represented as the pointwise sum of two Hamel functions.

Recall that a *perfect* set is a non-empty closed set with no isolated points. A subset A of a Polish space X is called *Marczewski measurable* if for every perfect set $P \subseteq X$ either $P \cap A$ or $P \setminus A$ contains a perfect set. If every perfect set P contains a perfect subset which misses A, then A is called *Marczewski null*. Marczewski (Mar1935) (writing as E. Szpilrajn) showed that the Marczewski measurable sets form a σ-field, and the Marczewski null sets form a σ-ideal.

Miller and Popvassilev (MillP2000) show:

- (Th. 10) There exists a Hamel basis H for $\mathbb{R}$ which is Marczewski null.
- (Th. 8) There exists a Hamel basis H for $\mathbb{R}^2$ which is Marczewski null.
- (Th. 14) There exists a Hamel basis H for $\mathbb{R}$ which is Marczewski measurable and perfectly dense.

Dorais, Filipów and Natkaniec (DorFN2013) show (Th. 4.2) 'deep differences between Lebesgue or Baire measurability and Marczewski measurability by constructing a discontinuous additive function that is Marczewski measurable'. They also show (Ex. 4.1) that there exist additive (discontinuous) functions that are not Marczewski measurable. For further background, see Kha2004.

P.10 Scaling and Fechner's Law

Fechner's Law (Gustav Fechner (1801–1887) in 1860) may be viewed as stating that, when two related physically meaningful functions f and g have no natural scale in which to measure their units, and are reasonably smooth, then their relationship is given by a power law:

$$f = cg^{\alpha}. \tag{F}$$

For background, see, e.g., Bin2015a; Han2004, §5.6.

Fechner's Law emerges naturally from regular variation, as follows (we restrict attention to the basic case, with f, g positive, increasing and unbounded). They satisfy some unknown functional relationship, say,

$$f(x) = \phi(g(x)) : \qquad f = \phi \circ g.$$

As there is no natural scale, then at least asymptotically this relationship should be scale-independent regarding x. So changing scale by λ,

$$f(\lambda x) \sim \psi(\lambda) f(x) \quad \text{for all } \lambda > 0 \tag{RV}$$

for some function $\psi(.) > 0$. Under a minimal smoothness assumption (the Baire property or measurability suffice), f is regularly varying with index $\alpha > 0$, and ψ is a power:

$$f \in R_\alpha \subseteq R := \bigcup_{\alpha > 0} R_\alpha.$$

Similarly from $g = \phi^{\leftarrow} \circ f$, $g \in R$, and from $\psi = f \circ g^{\leftarrow}$ with $f, g \in R$, $\phi \in R$ also:

$$\phi(\lambda) = \lambda^\alpha \ell(\lambda) \in R_\alpha,$$

with $\ell \in R_0$.

The classically important special case is the simplest one, ℓ constant, $\ell \equiv c$:

$$\phi(x) = cx^\alpha; \qquad f(x) = cg(x)^\alpha : \qquad f = cg^\alpha,$$

giving Fechner's Law.

Illustrative Example: Athletics Times. For aerobic running below ultra distances (800 m to the marathon, say), time t and distance d show Fechner dependence:

$$t = cd^\alpha.$$

Here c (time per unit distance) reflects the quality of the athlete, while α is approximately constant between athletes. This is illustrated on a real data set (the first author's half-marathon and marathon times) in Bingham and Fry (BinF2010, §8.2.3).

The statistics needed (regression) extends to the study of ageing also. The Rule of Thumb for ageing athletes (over 40, say) is: expect to lose a minute a year on your marathon time through ageing alone. It is well borne out by this data set (BinF2010, Ex. 1.3, Ex. 9.6).