# REGULAR VARIATION: NEW VARIATIONS ON AN OLD THEME

N. H. BINGHAM
Imperial College London

Statistics Department, Columbia University,
New York

22nd September, 2016

# Overview

1. Pre-BGT
2. BGT
3. (Lebesgue) measure and (Baire) category
4. BGT: The gaps
5. BinO: The Bingham-Ostaszewski program
6. The Steinhaus-Weil theorem
7. The Ostrowski theorem
8. Density topologies
9. Infinite combinatorics
10. Analytic sets
11. Sequential and non-sequential aspects
12. Beurling theory
13. Kendall's theorem
14. Scaling and Fechner's law

# 1. Pre-BGT

A function $f : \mathbb{R} \to \mathbb{R}_+$ is *regularly varying* (RV) if

$$f(\lambda x)/f(x) \to g(\lambda) \quad (x \to \infty) \quad \forall \lambda > 0 \qquad (RV)$$

for some $g$. The convergence here is *uniform* on compact $\lambda$-sets in $\mathbb{R}_+ := (0, \infty)$, for $f$ measurable or Baire (below), but not in general (UCT). The limit function $g$ is a power, $g(\lambda) \equiv \lambda^\rho$, again for $f$ measurable/Baire but not in general (Characterization Theorem). Then $\rho$ is called the *index* of regular variation, and one writes $f \in R_\rho$. If

$$f(\lambda x)/f(x) \to 1,$$

i.e. if $\rho = 0$, $f$ is called *slowly varying*, SV, written $\ell$ ('l for lente, or langsam'). There is a Representation Theorem for $\ell \in R_0$, and then each $f \in R_\rho$ has the form $f(x) = x^\rho \ell(x)$.

The above is the *multiplicative* form, useful for *applications*; for proofs, the *additive* form is better:

$$F(x+u) - F(x) \to G(u) \quad (x \to \infty) \quad \forall u \in \mathbb{R}. \qquad (RV_+)$$

This can be usefully generalized to

$$[F(x+u) - F(x)]/\ell(x) \to G(u) \ (x \to \infty) \ \forall u \in \mathbb{R} \qquad (BKdH)$$

for $\ell$ SV (Bojanic-Karamata 1963, de Haan 1970 on).
Regular variation (RV) was introduced by Jovan Karamata (1902-1967) in 1930 (*Mathematica (Cluj)*). For References, see e.g.
Jovan Karamata, *Selected papers* (ed. V. Maric), Beograd, 2009.
His motivation was a short proof of the Hardy-Littlewood Tauberian theorem for Laplace transforms: the *Hardy-Litlewood-Karamata theorem*, HLK. There was earlier work by Landau (1911), Valiron (1913), Pólya (1917), Hardy (1924).

The use of regular variation in probability theory (why we use it!) is due to G. N. Sakovich in 1956 (*Theor. Prob. Appl.* Vol. **1**). Earlier results, in which regular variation was not explicit: domain of attraction for CLT with Gaussian limit (Khinchin, Feller, Lévy in 1935, independently – *truncated variance slowly varying*); domain of attraction for CLT with general (inf. div.) limit: Gnedenko (1939), Doeblin (1940); Gnedenko & Kolmogorov, book, 1949 (Russian), 1954 (English). Validity of weak LLN: *truncated mean slowly varying*.

First textbook account with regular variation explicit: Feller, Vol. II, 1966/1971.

First textbook account of the mathematical theory: Seneta, 1971.

Extensions of the Karamata theory and applications to extreme value theory (EVT): de Haan, 1970 on.

Monograph account of probabilistic theory and applications to EVT: Resnick, 1987.

# 2. Bingham-Goldie-Teugels (BGT)

[BGT] N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular variation*, CUP, 1987/1989.

Chapters 1, Karamata theory; 2, Extensions; 3, De Haan theory; 4, Abelian and Tauberian theorems; 5, Mercerian theorems; 6, Analytic number theory; 7, Complex analysis; 8, Probability theory. Widely cited, still: MathSciNet, 1,288 (1987) + 333 (1989) (Google, 489 (1987) + 4,626 (1989)).

# 3. (Lebesgue) measure and (Baire) category

A set is *meagre* (of first (Baire) category) if it is a countable union of nowhere dense sets, *non-meagre* (second category) otherwise; $A$ has the *Baire property* – 'is Baire' – if it is the symmetric difference of an open set and a meagre set ('is nearly open'); $f$ has the *Baire property* – is *Baire* – if $f^{-1}(G)$ is Baire for every open $G$.
The standard work on measure and category:
[Oxt] J. C. Oxtoby, *Measure and category: A survey of the analogies between topological and measure spaces*, 2nd ed., Grad. Texts in Math. **2**, Springer, 1980 (1st ed. 1971).
This explores *measure-category duality* in depth. This works when the measure theory is *qualitative* (is the measure zero or positive?) rather than *quantitative* (what is the measure?). So: it goes as far as the Zero-One Law, but not the Strong Law of Large Numbers.

# 4. BGT: The gaps

The book appeared in the series *Encyclopedia of Mathematics*, and is sometimes referred to as 'the encyclopedia'! What does it miss?

*The foundational gap*.

In Ch. 1, measurability suffices; so does the Baire property (which is topological, not measure-theoretic); neither includes the other. Question: what is the right condition (minimal common generalisation)?

*The contextual gap*.

BGT has 420-odd pages of real analysis, then 4 (Appendix 1) in more general settings; real analysis is not the natural context. What is?

*Hard proofs*.

Some proofs in BGT are hard! E.g., Th. 1.4.3 p. 18-19 is proved (Th. 3.2.5, p.141-143) 120-odd pages later in a more general context; the special case is no easier. How many hard proofs are actually needed? Answer (now!): *zero*.

# 5. BinO: The Bingham-Ostaszewski program

Starting ten years ago, I began work with Adam (A.J.) Ostaszewski (Math. Dept., London School of Economics) to fill the gaps in BGT (§4 above), and extend Oxtoby's work on measure-category duality (§3 above). This is ongoing. In brief: it is *category*, rather than *measure*, that is crucial: we speak of *Category and measure* (the title of our forthcoming book). In particular, one can do RV with just *qualitative* rather than quantitative measure theory. The corpus so far:
BinO1-25 (21 published, the rest on the arXiv), plus Ost1-9 (and MilO, with Harry I. Miller).

# 6. The Steinhaus-Weil theorem

The theory of RV rests on *Steinhaus's theorem* of 1920 (BGT Th. 1.1.1): for a measurable set $A \subset \mathbb{R}$ of positive measure, the difference set $A - A$ contains a neighbourhood of the origin. The same is true for $A$ non-meagre (Piccard 1939; Pettis 1950, 1951).
This was extended to locally compact groups $G$ by Weil in 1940:
A. Weil, *L'intégration dans les groups topologiques*, 1940 (2nd ed. 1965).
Here the sets have positive *Haar* measure (extending *Lebesgue* measure on the line).

# 7. The Ostrowski theorem

The other pillar of RV theory is the *Ostrowski theorem*. A function $f$ is *additive* if

$$f(x+y) \equiv f(x) + f(y), \qquad (CFE)$$

the *Cauchy functional equation*. Obvious examples: $f(x) = cx$: $f$ *linear*. Pathological examples can be constructed using *Hamel bases* of the rationals $\mathbb{Q}$ over the reals $\mathbb{R}$ (G. Hamel, 1905 – this needs the Axiom of Choice, AC, of Zermelo, 1904). But, just a hint of decent behaviour in an additive function guarantees linearity – a *dichotomy* between very good and very bad behaviour. For (Ostrowski, 1929) if an additive function is bounded above or below on a set of positive measure (e.g., any interval), it is linear (BGT, Th. 1.1.7). Similarly for Baire functions (Mehdi, 1964).

# 8. Density topologies

Recall the *Lebesgue Density Theorem*. Writing $|.|$ for Lebesgue measure, call a point $x$ a *density point* of a set $A$ if

$$|A \cap (x - \epsilon, x + \epsilon)|/(2\epsilon) \to 1 \quad (\epsilon \downarrow 0).$$

Then almost all points of a measurable set are density points. Call a set *density-open* if all its points are points of density. Then (Haupt and Pauc, 1952; cf. C. Goffman and collaborators, 1961, Denjoy, 1915) these open sets define a topology, the *density topology*. This is a *fine topology* – it refines the usual (Euclidean) topology. One can deal with the measure and category cases above by working *bitopologically* – passing between the *Euclidean* topology for the category case and the density topology for the measure case (BinO15 (*Colloq. Math.*, 2010), BinO24 (2016+)).

# 9. Infinite combinatorics

Combinatorics is often thought of as a part of *finite* mathematics (enumerative combinatorics – counting things): see e.g.
R. P. Stanley, *Enumerative combinatorics*, CUP, Vol. 1, 1997, Vol. 2, 1999.
But one of the great Paul Erdős's contributions was *infinite combinatorics*. See e.g.
Terence Tao and Van H. Vu, *Additive combinatorics*, CUP, 2006,
R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey theory*, 2nd ed., Wiley, 1990 (1st ed. 1980),
N. Alon and J. H. Spencer, *The probabilistic method*, 3rd ed., Wiley, 2008
(non-constructive existence proofs: behaviour is possible because it is generic).

We make systematic use of results from infinite combinatorics, including the *Kestelman-Borwein-Ditor theorem* (KBD: Kestelman, 1947; Borwein and Ditor, 1978). Call null and meagre sets *negligible*, their complements non-negligible, and say a result holds for *quasi all* points if it holds off a negligible set.

**Th. KBD**. If $z_n \to 0$ and $T \subset \mathbb{R}$ is non-negligible, then for quasi all $t \in T$ there is an infinite set $M_t$ with

$$\{t + z_m : m \in M_t\} \subset T.$$

This is the key to short proofs of the Uniform Convergence Theorem (UCT) for RV (BGT, Th. 1.2.1), the main result of the subject.

# 10. Analytic sets

For $X$ a Polish (complete separable metric) space, recall that the *Borel* sets $\mathcal{B}(X)$ are those in the $\sigma$-field generated by the open (or closed) sets. The *analytic* sets are the continuous images of Polish spaces, denoted $\Sigma_1^1(X)$ (or $A(X)$). This notation is taken from the descriptive theory of sets; see e.g.
A. S. Kechris, *Classical descriptive set theory*, Grad. Texts in Math. **156**, Springer, 1995.
These generalize the Borel sets: by Souslin's theorem, a set is Borel iff it is analytic and co-analytic.
The use of analytic sets in probability theory goes back to:
C. Dellacherie, *Capacités et processus stochastiques*, Erg. Math. **67**, Springer, 1972,
C. Dellacherie, *Ensembles analytiques, capacités, mesures de Hausdorff*, LNM **295**, Springer, 1972.

It was developed further in an LMS conference *Analytic sets* organised by C. A. Rogers, University College London, 1978, where my collaboration with Adam Ostaszewski was born. A recent application is extensions to Ostrowski's theorem (§7). F. Burton Jones, 1942: an additive function $f$ *continuous* on a set $T$ which is analytic and contains a Hamel basis (= spans $\mathbb{R}$) is linear, so $f(x) = cx$. Kominek, 1981: similarly with *bounded* in place of continuous. For a unified proof of a common extension, to $T$ analytic and spanning a non-negligible set, see BinO8 (PAMS, 2010).

# 11. Sequential and non-sequential aspects

RV is a *continuous-variable* theory ('$(x \to \infty)$' in $(RV)$). But, BGT §1.6 on Sequences looks like ordinary RV. Also, of the several proofs of UCT in BGT §1.2, all but one are by contradiction – via a *sequence* 'witnessing' to the contradiction. We now know that *category*, rather than measure, is decisive for RV. This depends on the Baire Category Theorem. This in turn is *sequential*: its proofs are sequential. Also, Baire category arguments only need the Axiom of Dependent Choice, DC ('what is needed to make mathematical induction work'), rather than the full Axiom of Choice, AC. The *positive* statements in measure theory also only need DC. But, full measure theory includes 'negative' statements: e.g. existence of non-measurable sets. The classic Vitali example here explicitly uses AC. In brief: category needs only DC; measure theory needs more. So measure-category duality breaks down at the level of axiomatic assumptions. For a full account, with many references, see the arXiv version (Appendix) of BinO23, *Category-measure duality: Jensen convexity and Berz sublinearity*.

# 12. Beurling theory

Call $\phi : \mathbb{R} \to \mathbb{R}_+$ *Beurling slowly varying*, $\phi \in BSV$, if

$$\phi(x + t\phi(x))/\phi(x) \to 1 \ (x \to \infty) \ \forall t \in \mathbb{R},$$

and *self-neglecting*, $\phi \in SN$, if this holds uniformly on compact $t$-sets. See BGT §2.11 for *BSV* (motivation: Beurling's extension of Wiener's Tauberian theorem, unpublished by Beurling, 1957, published by Peterson and by Moh, 1971, and for *SN*, work of the Netherlands school on EVT). Bloom, 1976: continuity suffices for UCT. B & Ostaszewski (BinO19, Trans. LMS 2014; BinO20, Aequat Math. 2015; BinO21, Indag. Math., 2016; NHB, Ragnar Norberg Festschrift, to appear) give a thorough treatment of this area. Applications:
(i) Beurling moving averages: for $\phi \in SN$,

$$[U(x + t\phi(x)) - U(x)]/\phi(x) \to ct \ (x \to \infty) \ \forall t \geq 0. \qquad (BMA)$$

The Wiener Tauberian theorem deals with integral transforms of *convolution* type,
$\int f(t)k(x/t)dt/t$ or $\int F(u)K(x-u)du$. The Beurling Tauberian theorem (designed for the Tauberian theory of Borel summability, but very useful in probability theory – see e.g. NHB, *Ann. Prob.* **9** (1981)) is 'convolution-like':

$$(f *_\phi k)(x) := \int f\Big(\frac{x-t}{\phi(x)}\Big)\frac{dt}{\phi(x)}$$

(reducing to ordinary convolution for $\phi \equiv 1$). These were algebraically messy, till Adam discovered a clever algebraic approach in BinO21.

(ii) Stable laws directly (i.e., not by specialising from the Lévy-Khintchine formula). See two recent papers on this:
E. J. G. Pitman and J. G. Pitman, A direct approach to the stable distributions
(an extension by Jim Pitman of a handwritten manuscript by his late father Edwin),
A. J. Ostaszewski, Stable laws and Beurling kernels
(suggested to Adam by refereeing this paper for my Festschrift, in the light of BinO21). For these (and much else!), see
[PANT] *Probability, Analysis and Number Theory*. Papers in Honour of N. H. Bingham (ed. C. M. Goldie and A. Mijatovic), *Advances in Applied Probability* Special Volume **48A**, 2016.

# 13. Kendall's theorem

*Kendall's theorem* (1968; BGT, Th. 1.9.2): if $x_n \to \infty$, $\limsup x_{n+1}/x_n = 1$ (e.g. $x_n = n$), and for some $a_n \to \infty$ and $f$, $T$ as above

$$a_n f(\lambda x_n) \to g(\lambda) \quad (n \to \infty) \quad \forall \lambda \in T,$$

then $f$ RV (and $a_n$ RV, and $g$ a power).
This important result underlies the definition of *regularly varying measures* in higher dimensions. For $\mathbf{X}$ a random vector, say $\mathbf{X}$ has *regularly varying distribution* (or law) if

$$nP(\mathbf{X}/a_n \in .) \to \mu(.) \ (n \to \infty)$$

for some measure $\mu$ (again, $a_n$ is RV). See e.g.
H. Hult, F. Lindskog, T. Mikosch and G. Samorodnitsky, *Ann. Appl. Prob.*, 2005.

# 14. Scaling and Fechner's law

Fechner's law (1860): if $f$, $g$ are related, physically meaningful, have no natural scale and are (reasonably) smooth, they obey a *power law*: $f = cg^{\alpha}$. Example (my motivation): athletics times: for, e.g., 5k and 10k, or half-marathon and marathon, etc. Here $c$ is a measure of the athlete's quality (speed), while $\alpha$ seems approximately constant over people.

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

No scale: asymptotically scale-independent:

$f(\lambda x) \sim \psi(\lambda) f(x)$: $f$ RV.

So $\psi$ is a power.

Similarly, from $g = \phi^{\leftarrow} \circ f$, $g$ RV.

So as $\phi = f \circ g^{\leftarrow}$, $\phi$ RV: $\phi(x) = \ell(x) x^{\alpha}$, $\ell$ SV. Simplest case, $\ell \equiv c$: $f = cg^{\alpha}$: Fechner's law.

See e.g. NHB, PIMB (2015).