CATEGORY AND MEASURE

N. H. BINGHAM

Colloquium talk, University of York,
21 May 2014

§0. Sources
§1. Regular variation
§2. Baire category
§3. Lebesgue density
§4. Bitopology
§5. Normal numbers
§6. Continued fractions
§7. Non-constructive existence proofs
§8. Random series
§9. Zero-one laws
§10. Dynamical systems
§11. Dichotomy; axioms of set theory
§12. Automatic continuity
§13. BGT and BinO

§*0. Sources*

[O] J. C. OXTOBY, *Measure and category: A survey of the analogies between topological and measure spaces*, GTM 2, Springer, 1980 [1st ed. 1971];

[BGT] N. H. BINGHAM, C. M. GOLDIE and J. L. TEUGELS, *Regular variation*, CUP, 1987;

[BinO] N. H. BINGHAM and A. J. OSTASZEWSKI: A series of papers by myself and Adam Ostaszewski, a topologist, Maths Dept., LSE, 2008 on [21 with more to come; see MathSciNet, arXiv and Adam's homepage];

[Ost] A. J. OSTASZEWSKI: A series of papers by Adam, arising from this [9 solo with more to come, one with Harry [H. I.] Miller, one with Harry and his daughter Leila].

[Kor] J. Korevaar, *Tauberian theorems: A century of developments*. Grundl. math. Wiss. **329**, Springer, 2004.

[Kech] A. S. KECHRIS, *Classical descriptive set theory*. GTM **156**, Springer, 1995.

§*1. Regular variation*

The theory of regular variation is a chapter in the classical theory of functions of a real variable, dating from Karamata in 1930. It is used extensively in probability theory, analysis (particularly Tauberian theory), number theory etc.; see [BGT] for a monograph treatment and [Kor] IV for Tauberian theorems. It concerns relationships of the form

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

Here g satisfies the Cauchy functional equation

$$g(\lambda\mu) = g(\lambda)g(\mu) \qquad \forall \lambda, \mu > 0. \qquad (CFE)$$

Subject to a mild regularity condition, (CFE) forces g to be a power:

$$g(\lambda) = \lambda^{\rho} \qquad \forall \lambda > 0. \qquad (\rho)$$

Then f is said to be *regularly varying* with *index* ρ, written $f \in R_\rho$. The case $\rho = 0$ is basic. A function $f \in R_0$ is called *slowly varying*, written ℓ (for *lente*, or *langsam*). Recall:

f is (Lebesgue) measurable iff inverse images $f^{-1}(U)$ of open sets U are measurable,
f is Baire (has the Baire property) iff inverse images $f^{-1}(U)$ of open sets U have the Baire property, i.e. are a symmetric difference $G\Delta Q$ with G open and Q meagre (of first category – 'small' – §2 below) – are 'nearly open'.
The basic theorem of the subject is the Uniform Convergence Theorem (UCT): if

$$\ell(\lambda x)/\ell(x) \to 1 \quad (x \to \infty) \quad \forall \lambda > 0, \qquad (SV)$$

then the convergence is *uniform* on compact λ-sets in $(0, \infty)$. The basic facts are:
(i) if ℓ is (Lebesgue) measurable, then UCT holds (Korevaar et al. 1949; Karamata 1930 in the continuous case);
(ii) if ℓ is Baire, then UCT holds (Matuszewska 1965);
(iii) in general, UCT need not hold.
Similarly, if f is measurable or Baire, (CFE) implies (ρ), but not in general.

§*2. Baire Category* (René BAIRE (1874-1932) in 1899).

Recall that in a topological space a set is *nowhere dense* if its closure has empty interior, *meagre* (of first category) if it is a countable union of nowhere dense sets, *non-meagre* (second category) otherwise. Baire's Category Theorem: in a complete metric space (even pseudometric), the intersection of countably many dense open sets is itself dense. Similarly for locally compact regular topological spaces. Proof: J. L. KELLEY, *General topology*, Van Nostrand, 1955, 200-201.

Call a space a *Baire space* if the conclusion of Baire's theorem holds (Baire's original theorem was that $\mathbb{R}$ is a Baire space under the Euclidean topology).

Category gives us a topological way of measuring *size* of a set: the meagre sets are the 'small' sets. Compare the *null sets* (sets of measure zero in Measure Theory).

Call a set with meagre complement *co-meagre* or *residual*.

§*3. Lebesgue Density* (Henri LEBESGUE (1875-1932)).
With $|.|$ Lebesgue measure, recall x is a *density point* of a measurable set A if

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

Recall Lebesgue's Density Theorem (1907): almost all points of a measurable set are density points. Call a set d-*open* if *all* its points are density points. These sets form a topology, the *density topology*, d. Then:
(i) The density topology d is finer than the Euclidean topology $\mathcal{E}$.
(ii) A set has the Baire property in d iff it is measurable.
(iii) A Baire set is d-meagre iff it is (Lebesgue-)null.
(iv) $(\mathbb{R}, d)$ is a Baire space (Lebesgue's density theorem: see e.g. [Kech] 17.47).
(v) A function is d-continuous iff it is approximately continuous in Denjoy's sense (Arnaud DENJOY (1884-1974) in 1916).

§*4. Bitopology*.

Using these ideas, one can handle the category and measure cases of regular variation together, using d for the first and $\mathcal{E}$ for the second. See BinO/Ost, Beyond Lebesgue and Baire, I - IV.

Qualitative v. quantitative measure theory.

Working bitopologically as above, the category case is qualitative measure theory: all that counts in whether the measure of a set is 0 or positive. Quantitative measure theory uses the actual value of the measure of a non-null set. We have been able to reduce the amount of quantitative measure theory needed for regular variation to an irreducible minimum – far less than before, with correspondingly more economical proofs – but not eliminate it altogether.

§*5. Normal numbers.*

Take $[0, 1]$, and look at the binary, decimal, ..., expansions of $x \in [0, 1]$. It is easy to check that if X is a random variable with the uniform distribution in $[0, 1]$, $X \sim U(0, 1)$ (probability = length), and X has dyadic expansion $\sum_1^\infty \epsilon_n/2^n$, then ϵ_n are independent coin-tosses (values 0,1 with prob. half), and conversely. So by the Stong Law of Large Numbers (Borel in 1909 for this case, Kolmogorov in 1933 in general), almost all numbers are normal (equal probabilities for 0 and 1). Similarly for each base $d = 1, 2, \ldots$. Similarly for pairs, triples, ...; similarly for shifts (start the count after 1 place, 2 places, ...). Intersecting countably many sets of full measure gives a set of full measure. So (Borel's Normal Number Theorem, 1909): almost all numbers are strongly normal (in this sense). So, *normality is generic (in the measure sense)*.

A stark contrast between measure and category shows up here: *non-normality is generic in the topological (category) sense*. See e.g. [APT] S. ALBEVERIO, M. PRATSIOVYTI & G. TORBIN: Topological and fractal properties of real numbers which are not normal. *Bull. Sci. Math.* **129** (2005), 615-630.
The set of non-normal numbers is co-meagre. It also has Hausdorff dimension 1.

6. Continued fractions.

From the number-theoretic point of view, the natural way to expand a real is as a *continued fraction*. Thus

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2+\cfrac{2}{3+\cfrac{3}{4+\cfrac{4}{\ddots}}}}}, = 2 + \frac{1}{1+}\ \frac{2}{2+}\ \frac{3}{3+}\cdots$$

is Euler's continued fraction for e of 1737, while Brouncker's continued fraction for π of 1655 is

$$\pi/4 = \cfrac{1}{1 + \cfrac{1^2}{2+\cfrac{3^2}{2+\cfrac{5^2}{\ddots}}}} = \frac{1}{1+}\ \frac{1^2}{2+}\ \frac{3^2}{2+}\ \frac{5^2}{2+}\cdots.$$

Again, normality is generic in the measure sense: Paul LÉVY (1886-1971) in 1929 and subsequently. The limit involves the Gauss law, with density

$$\frac{1}{\log 2(1+x)} \qquad (x \in [0,1])$$

and the Birkhoff-Khinchin ergodic theorem (§10 below). But non-normality is generic in category:

L. OLSEN, Extremely non-normal continued fractions. *Acta Arith.* **108** (2003), 191-202.

See also

M. G. MADRITSCH, Non-normal numbers with respect to infinite Markov partitions, *Discrete Cont. Dyn. Syst. A* **34** (2014), 663-676.

7. Non-constructive existence proofs.
Note that these methods enable one to prove behaviour is *generic*, and thus does occur, without being able to construct any specific example. For instance, we have no explicit example of a number strongly normal to all bases simultaneously, even though almost all of them are. The best result known is the strong normality to all bases that are powers of 10 of *Champernowne's number*

$$0.12345678910111213....99100101...$$

Similarly for functions. Brownian motion $B := (B_t)$ is Brownian motion (BM) has paths $t \mapsto B_t$ which are a.s. *everywhere continuous but nowhere differentiable*. This can be proved for BM by probabilistic methods – but such behaviour is *generic* in $C(0,1)$ topologically, as was shown by Banach by category methods in 1931 [O, Ch. 11].

§*8. Random series.*
The prototypical random series is that in the classic Paley-Wiener-Zygmund (PWZ) construction of BM of 1933: if

$$B_t = \sum_0^\infty \lambda_n Z_n \Delta_n(t), \qquad t \in [0, 1],$$

where the Z_n are independent standard normal random variables, the Δ_n are the *Schauder functions* (wavelet system, obtained from integrating the Haar functions) and the λ_n are suitable constants, then $B = (B_t)$ is Brownian motion (the expansion above is of *spectral* type: it splits the time-dependence in the $\Delta_n(t)$ from the ω-dependence (randomness) in the $Z_n = Z_n(\omega)$). The standard work here is Jean-Pierre KAHANE, *Some random series of functions*, 2nd ed., CUP, 1985 [1st ed. 1968]. There, Kahane considers series with "random signs", $\sum \pm a_n$, where the $\pm$ are independent coin-tosses, + or - with prob. 1/2. By Kolmogorov's zero-one law (§9 below), such a series will converge with probability 0 or 1. It is classical that the condition for a.s. convergence is $a = (a_n) \in \ell_2$:

$$\sum |a_n|^2 < \infty \quad \Leftrightarrow \sum \pm a_n \text{ converges } a.s.$$

By contrast, the condition for convergence off a meagre set is much stronger, $a \in \ell_1$. Writing q.e. for *quasi-everywhere* (everywhere off a meagre set):

$$\sum |a_n| < \infty \quad \Leftrightarrow \sum \pm a_n \text{ converges q.e.}$$

(interpretation: "it doesn't converge until it has to"). Similarly for random Fourier series, Gaussian series (using standard normals instead of random signs), etc. See e.g.

J.-P. KAHANE, Baire's category and trigonometric series. *J. Analyse Math.* **80** (2000), 143-182,

R. KAUFMAN, Thin sets, differentiable functions and the category method. *J. Fourier Analysis* (1995), 311-6.

Similarly for random fractals (Falconer; Barlow, Bass, ...).

§*9. Zero-one laws*.
Given a sequence of events (measurable sets in a probability space), their *tail σ-field* is the σ-field generated by sets invariant under deletion of finitely many events. *Kolmogorov's zero-one law* of 1933 states that for independent events the probability of a tail event (event in the tail σ-field) is 0 or 1. Example (§8): the probability a random series (with independent terms, understood) converges is 0 or 1.
For (X_n) independent and identically distributed random variables, call an event *symmetric* if it is invariant under finite permutations of the X_n. The *Hewitt-Savage zero-one law* of 1955 says that such a symmetric event has prob. 0 or 1. There is a *Lévy* 0-1 law, the first martingale convergence theorem, and a number of others; see e.g.
A. N. SHIRYAEV, *Probability*, 2nd ed., Springer, 1996 [1st ed. 1984], Ch. IV.
Zero-one laws extend to the category case [O, Ch. 21]: a tail event with the Baire property is meagre or co-meagre.

§*10. Dynamical systems*.

As in the theory of Markov chains, call a point x of an open set G *recurrent* for G w.r.t. a volume-preserving homeomorphism T if infinitely many points of the orbit $\{x, Tx, T^2x, \ldots, T^nx, \ldots\}$ are in G. Poincaré's recurrence theorem (work on celestial mechanics of 1899 – [O, Ch. 17]) says that almost all and quasi-all points of G are recurrent (category and measure behave the same here). This work was extended to the Birkhoff-Khinchin ergodic theorem, on existence of limits of the form

$$\lim \frac{1}{n} \sum_{0}^{n-1} I_E(T^k x).$$

Such limits exist almost everywhere, but may exist only on a meagre set: measure and category differ here, perhaps drastically.

Generic behaviour for dynamical systems is of great interest; see e.g.

S. ALPERN & V. S.PRASAD, *Typical dynamics of volume-preserving homeomorphism*, Cambridge Tracts Math. **139**, CUP, 2000.

§*11. Dichotomy; axioms of set theory*

The reals $\mathbb{R}$ form a vector space over the rationals $\mathbb{Q}$, and so (by the Axiom of Choice AC, in the form of Zorn's Lemma) have a basis – a *Hamel basis* (G. Hamel, 1905). Hence additive functions $k(x)$ not of the form cx can be constructed. By Ostrowski's theorem ([BGT, Th. 1.1.7]; §13 below), such a k is *pathological* (unbounded above and below on every non-meagre/non-null set, etc.).

This *dichotomy* between the very nice and the very nasty runs through functional equations (such as (CFE) here), regular variation, etc., cf. 0-1 laws (§9).

If we replace limits in §1 by limsup and liminf (often needed), we may lose measurability/Baire property. To study the extent of the degradation here, one needs *descriptive set theory* [Kech].

Recall that in some axioms of set theory, e.g. the Axiom of Determinacy AD, *all* sets are measurable and Baire.

§*12. Automatic continuity*

Because of the basic role of additive functions (or multiplicative – one has a choice here), it is important to know when one can say such a function is continuous (so one is in the nice case of the dichotomy of §11). Such functions are homomorphisms (of $(\mathbb{R}, +)$ or $(\mathbb{R}_+, \times)$). Much is known about *automatic continuity* of such homomorphisms, in various contexts:
Banach algebras (Dales, Woodin, ...);
harmonic analysis (Gelfand, Helson, ...);
topological groups, analytic spaces, (Hoffmann-Jorgensen,...).
See BinO (Aequat. Math. 2009, PAMS 2010) for results on automatic continuity and analytic sets (Lusin, Souslin, ...). For background on analytic sets (highly relevant here), see [Kech], and C. A. Rogers et al., *Analytic sets*, AP, 1980 (Proc. LMS Conference, UCL, 1978 – the birth of the Bingham-Ostaszewski collaboration).

§*13. BGT and BinO*.

Regular variation rests on two results:

(i) *Steinhaus' Theorem*. If A is measurable and non-null/Baire and non-meagre, then $A-A$ contains a neighbourhood of the origin.

(ii) Ostrowski's Theorem. A solution of the Cauchy functional equation (CFE) which is measurable/Baire is a power, as in (ρ) of §1.

See our paper Dichotomy and infinite combinatorics: the theorems of Steinhaus and Ostrowski. *MPCPS* **150** (2011), 1-22.

BGT dates from 1987, and has lasted well. But it left two main gaps:

(i) *The foundational question*. Measurability and the Baire property both work. What is the appropriate common generalisation?

(ii) *The contextual question*. BGT is set in $\mathbb{R}$, but not restricted to it. What is the appropriate setting? See BinO/Ost for answers, and for a reduction of the number of hard proofs in the area to *zero*. NHB