Tatra Mt. Math. Publ. **31** (2005), 103–111

TATRA MOUNTAINS
Mathematical Publications

THE DENSITY SET: A SURVEY

GEORGES GREKOS

ABSTRACT. The density set of a given subset A of the natural numbers is the set of pairs of the upper and lower densities of all subsets of A. In this survey, firstly, we summarize the properties of the density set and, secondly, we explain how sets, having the same form as the density set, appear in quite different situations in Analysis: uniform distribution of sequences, mean values of functions.

1. Density

Consider a subset A of $\mathbb{N} = \{1, 2, 3, \dots\}$. The set A can also be visualized as an "integer sequence" $A = \{a_1 < a_2 < \dots\}$. Its *counting function* is defined, for $x \geqslant 1$, by $A(x) = \bigl|A \cap [1, x]\bigr|$. Its *upper* and *lower asymptotic densities* are defined as follows:

$$\overline{d}A = \limsup_{x\to\infty} \frac{A(x)}{x} = \limsup_{k\to\infty} \frac{k}{a_k},$$

$$\underline{d}A = \liminf_{x\to\infty} \frac{A(x)}{x} = \liminf_{k\to\infty} \frac{k}{a_k}.$$

For convenience we put $\alpha = \overline{d}A$, $\alpha' = \underline{d}A$, and we call *density point* of A the pair $D(A) = (\alpha, \alpha')$ belonging to the triangle $\bigl\{(x, y) \in \mathbb{R}^2 \colon 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant x\bigr\}$ of the $\mathbb{R}^2$ plane.

Here are two examples.

1) For the set $\mathbb{P}$ of primes we have $\overline{d}\,\mathbb{P} = 0$ and so $D(\mathbb{P}) = (0, 0)$.

2) Let A_m be the set of positive integers having m as first digit, $1 \leqslant m \leqslant 9$. Then (see, for instance, [FL])

$$D(A_m) = \left(\frac{10}{9(m+1)}, \frac{1}{9m}\right).$$

The last result can be derived from the following property.

2000 Mathematics Subject Classification: 11A05.
Keywords: density, density set, convexity.

Proposition 1. *Let* $0 \leqslant p_1 < q_1 < p_2 < q_2 < \dots$ *be integers. Put* $A = \bigcup_{j \geqslant 1} \left( ]p_j, q_j] \cap \mathbb{N} \right)$. *Then*

$$\overline{d}A = \limsup_{j \to \infty} \frac{A(q_j)}{q_j}, \quad \underline{d}A = \liminf_{j \to \infty} \frac{A(p_j)}{p_j}.$$

For proof see [FG-A, Theorem 8.2] or [MT, Lemma 1].

It will be useful to have in mind that, in general, a density m has the properties:

(i) $0 \leqslant m(A) \leqslant 1$ for all A for which $m(A)$ is defined;

(ii) $m(\emptyset) = 0$, $m(\mathbb{N}) = 1$;

(iii) if $m(A)$ and $m(B)$ are defined and $B \subseteq A$, then $m(B) \leqslant m(A)$.

Moreover, upper and lower densities $\overline{m}$ and $\underline{m}$ are densities, i.e., satisfy (i), (ii) and (iii), are defined **for all** $A \subseteq \mathbb{N}$ and satisfy $\underline{m}(A) \leqslant \overline{m}(A)$.

2. The density set

Suppose that the set A is fixed and so is its (asymptotic) density point $D(A) = (\alpha, \alpha')$. The coordinates of the density point $D(B) = (\beta, \beta')$ of any subset B of A clearly verify $0 \leqslant \beta' \leqslant \alpha'$ and $\beta' \leqslant \beta \leqslant \alpha$. This means that $D(B)$ belongs to the trapezium $\mathrm{Tz} = \mathrm{Tz}(A)$ having vertices $(0,0)$, $(\alpha, 0)$, $D(A)$, (α', α').

A natural question arises: Given any point (x, y) of the above trapezium, is there a set $B \subseteq A$ such that $D(B) = (x, y)$?

This question was studied in [G1]. It was proved that if (x, y) belongs to the (closed) triangle Tg having vertices $(0,0), (\alpha, 0), D(A)$, then the answer is YES; if (x, y) belongs to the (open below) triangle Tg' with vertices $(0,0)$, $D(A)$, (α', α'), then the answer may be YES or NO **depending on** the set A. Notice that Tg' is empty when $\alpha = \alpha'$ and when $\alpha' = 0$. Note that Tz is the union of Tg and Tg'.

In [G1] the *density set* of a subset A of $\mathbb{N}$ was introduced:

$$S(A) = \left\{ D(B) \colon B \subseteq A \right\}. \tag{1}$$

Given a point $M_0 = (x_0, y_0)$, $0 \leqslant y_0 \leqslant x_0 \leqslant 1$, two sets were constructed. A set $A_1 \subseteq \mathbb{N}$ such that

$$S(A_1) = \mathrm{Trapezium}\left[(0,0), (x_0, 0), (x_0, y_0), (y_0, y_0)\right] =: \mathrm{Tz}_0 \tag{2}$$

and another set $A_2 \subseteq \mathbb{N}$ such that

$$S(A_2) = \mathrm{Triangle}\left[(0,0), (x_0, 0), (x_0, y_0)\right] =: \mathrm{Tg}_0 \,. \tag{3}$$

3. Properties of the density set

With the notations introduced in the preceding sections, the following theorems hold.

THEOREM 1. *Let $A \subseteq \mathbb{N}$ be fixed. Then* $\mathrm{Tg} \subseteq S(A) \subseteq \mathrm{Tz}$ *and* $S(A)$ *is convex and closed.*

THEOREM 2. *Given a point* (x_0, y_0), $0 \leqslant y_0 \leqslant x_0 \leqslant 1$, *and a closed convex region* S *of* $\mathbb{R}^2$ *such that* $\mathrm{Tg}_0 \subseteq S \subseteq \mathrm{Tz}_0$, *there is a set* $A \subseteq \mathbb{N}$ *such that* $S(A) = S$.

The proofs of Theorems 1 and 2 were published in 1978 [G3]. The two theorems were announced and partially proved in [G2], where a proof of Theorem 2 is given in the special case where S is a quadrilateral containing Tg_0 and contained in Tz_0.

In the proof of Theorem 1, crucial role plays the so called *Darboux property.*

PROPOSITION 2. *Given a set* A, $D(A) = (\alpha, \alpha')$ *and* $t \in [0,1]$, *there is a set* $B \subseteq A$ *such that* $D(B) = (t\alpha, t\alpha')$.

This property was proved, among other results, by Tibor Šalát and Robert Tijdeman [ST] in the more general context of "measure" densities. Necessary and sufficient conditions for the validity of the Darboux property for generalized densities are given in the paper [PSV] of this issue.

Another trend, observed in [G1] and investigated in [GV], is the relation between asymptotic densities and the *gap density*

$$\lambda(A) = \limsup_{n\to\infty} \frac{a_{n+1}}{a_n} \in [1, +\infty]$$

of a set $A = \{a_1 < a_2 < \dots\} \subseteq \mathbb{N}$. See also [GST].

PROPOSITION 3. $\alpha'\lambda(A) \leqslant \alpha$ *(with the convention* $0 \times \infty = 0$*).*

Proposition 3 says that $D(A)$ lies below or on the straight line L going through the origin and having slope $\frac{1}{\lambda(A)}$ (convention $\frac{1}{\infty} = 0$). It is proved in [GV] that the whole density set $S(A)$ lies below or on the line L.

Denote by f the function having as graph the upper boundary of the density set:

$$f\colon [0,\alpha] \to [0,\alpha'], \quad \gamma \mapsto \max\{y\colon (\gamma, y) \in S(A)\}.$$

This function is concave. So at each $\gamma \in \,]0,\alpha[$, it has a left and a right derivative $f'_\ell(\gamma) \geqslant f'_r(\gamma)$. Also $f'_r(0)$ and $f'_\ell(\alpha)$ exist, $f'_r(0) \leqslant 1$, $f'_\ell(\alpha) \geqslant 0$. The functions f'_ℓ, defined on $]0,\alpha[$., and f'_r, defined on $[0,\alpha[$, are decreasing. Notice that, for sake of simplicity, one could put $f'_\ell(0) = 1$ and $f'_r(\alpha) = 0$.

The fact, proved in [GV], that $S(A)$ lies below or on L, can be stated as

$$f_r'(0) \leqslant \frac{1}{\lambda(A)} \leqslant 1\,.$$

In [GV] we also proved a stronger version of Theorem 2 (see its formulation): given a "convenient" convex closed region S, f being the function corresponding to the upper boundary of S, and given any real number $u \in \left[f_r'(0), 1\right]$, there is a set $A \subseteq \mathbb{N}$ such that $S(A) = S$ and $\frac{1}{\lambda(A)} = u$.

As for the closure of S, it is proved in [G3] that given $\gamma \in \,]0, \alpha[$, the set $B \subseteq A$ defined inductively by

$$B = A \cap \left\{n \in \mathbb{N} \colon B(n-1) + 1 \leqslant \gamma n\right\}$$

has $\overline{d}B = \gamma$ and $\underline{d}B$ is maximal: if $C \subseteq A$ and $\overline{d}C = \gamma$, then $\underline{d}C \leqslant \underline{d}B$. It follows that $S(A)$ is closed and $f(\gamma) = \underline{d}B$.

Two more results were proved in [GV]. One of them is the following theorem.

Theorem 3. *Given two "convenient" (see Theorem 2) convex closed regions S_1, S_2 of $\mathbb{R}^2$ with $S_1 \subseteq S_2$, there are sets $A_1 \subseteq A_2 \subseteq \mathbb{N}$ such that $S(A_i) = S_i$, $i = 1, 2$.*

The other result is Theorem 4 below which characterizes the slope at the origin of the upper boundary of the density set in terms of the following variant of the gap density. Let A be a set with positive upper density: $\alpha > 0$. For $n \in \mathbb{N}$ and $\varepsilon \in [0, \alpha[$, let

$$E(A, \varepsilon, n) = \left\{\ell \in \mathbb{N} \cup \{0\} \colon A(n+q) - A(n) \leqslant \varepsilon q\,,\ q = 0, 1, \dots, \ell\right\}.$$

Notice that this set is

- non empty, because it contains 0;
- finite, otherwise the upper asymptotic density of A would be less than or equal to ε.

Put $\ell(A, \varepsilon, n) = \max E(A, \varepsilon, n)$ and

$$\lambda(A, \varepsilon) = \limsup_{n \to \infty} \frac{n + \ell(A, \varepsilon, n)}{n}\,.$$

Notice that $\lambda(A, 0) = \lambda(A)$, previously defined in this section. The function $\varepsilon \mapsto \lambda(A, \varepsilon)$ is increasing and so $\lim\limits_{\varepsilon \to 0^+} \lambda(A, \varepsilon)$ exists and is $\geqslant \lambda(A)$.

Theorem 4. *Let $A \subseteq \mathbb{N}$ be a set having positive upper asymptotic density. The letter f denotes the function corresponding to the upper boundary of $S(A)$. We have*

$$f_r'(0) = \frac{1}{\lim\limits_{\varepsilon \to 0^+} \lambda(A, \varepsilon)}$$

(with the convention $\frac{1}{\infty} = 0$).

4. Other occurrences and generalizations

4.1. Let $A^c = \mathbb{N} \setminus A$. Due to the relations

$$\overline{d}A^c = 1 - \underline{d}A\,, \quad \underline{d}A^c = 1 - \overline{d}A$$

the set $T(A) = \{D(C)\colon C \supseteq A\}$, included in the trapezium $[(\alpha,\alpha'),(1,\alpha'),(1,1),(\alpha,\alpha)]$, has a similar form: it contains the triangle $[(\alpha,\alpha'),(1,\alpha'),(1,1)]$ and it is convex and closed.

4.2. One can define (asymptotic) densities and prove similar results for discrete monoids like $\mathbb{N}^2$ or the set $\mathcal{F}$ of all finite subsets of $\mathbb{N}$. Densities are defined as the upper and the lower limits, when $N \to \infty$, of:

$$N^{-2}\big|\{(m,n) \in A\colon m \leqslant N,\ n \leqslant N\}\big| \ \text{ for } A \subseteq \mathbb{N}^2\,,$$

$$2^{-N}\big|\{E \in \mathcal{A}\colon E \subseteq \{1,2,\dots,N\}\}\big| \ \text{ for } \mathcal{A} \subseteq \mathcal{F}\,.$$

4.3. Theorems 1 and 2 were proved [G0] in the context of *generalized densities.* These densities are encountered in [RV], [A] or [PSV] and are defined as the upper and lower limits $(n \to \infty)$ of

$$\left(\sum_{j=1}^{n} w(j)\right)^{-1} \sum_{a\in A,\ a\leqslant n} w(a)\,,$$

where $w\colon \mathbb{N} \to \mathbb{R}_+ = [0,+\infty[$ is a *weight* function satisfying the conditions

$$\sum_{n=1}^{\infty} w(n) = \infty$$

and

$$\lim_{n\to\infty} \frac{w(n) + \sum\limits_{j=1}^{n-1} \big|w(j) - w(j+1)\big|}{\sum\limits_{j=1}^{n} w(j)} = 0\,.$$

5. Uniform distribution modulo 1

Let $u = (u_n)_{n\in\mathbb{N}}$ be a real sequence. In 1971, Michel Mendès-France (see [M-F], [B]) was interested in subsequences $u_A = (u_n)_{n\in A}$ of u which are uniformly distributed modulo 1 (abbreviated: u.d.) and such that A has a density $dA = \overline{d}A = \underline{d}A$. He proved that the set

$$\{\alpha\colon \exists A \subseteq \mathbb{N},\ dA = \alpha,\ u_A \text{ u.d.}\}$$

is either empty or an interval of the type $[0,x]$ or $[0,x[$. He conjectured that it is a closed one. This was proved by Yves Dupain and Jacques Lesca [DL].

In [DG] we prove the following two results.

1) Given a real sequence $u=(u_n)$, the set

$$E(u)=\left\{(\overline{d}A,\, \underline{d}A)\in \mathbb{R}^2 \colon u_A \text{ u.d.}\right\},$$

if it is not empty, is a convex closed region of the plane and has the same form as the density set. It is worthwhile to notice here that there is no maximal set A such that u_A be u.d. (unless u itself is u.d.).

2) Given a "convenient" (as in Theorem 2) closed convex region S, there is a real sequence $u=(u_n)$ such that $E(u)=S$.

6. Mean values and convexity

Upper and lower densities are limit mean values. It seems that the reason of the appearance of—non trivial—convexity is the fact that we *take the mean.* This point is developed in [G4] and [GH]. Before presenting these results, let me say what I mean by "non trivial convexity".

Fix a bounded real sequence $a=(\alpha_n)_{n\in\mathbb{N}}$ with non negative terms and set

$$\alpha=\limsup \alpha_n\,, \qquad \alpha'=\liminf \alpha_n\,.$$

Take all non negative sequences $b=(\beta_n)_{n\in\mathbb{N}}$ "dominated" by α, that is satisfying $0\leqslant \beta_n\leqslant \alpha_n$ for all n. One can verify that the set

$$\left\{(\beta,\beta')\colon b \text{ is dominated by } a\right\}$$

is *equal* to the trapezium $\left[(0,0),(\alpha,0),(\alpha,\alpha'),(\alpha',\alpha')\right]$. But we saw that the particular case $\alpha_n=A(n)/n$, $\beta_n=B(n)/n$ with $B\subseteq A$, A fixed and B variable, may give rise to a proper convex subset of the above trapezium. Now, $A(n)/n$ and $B(n)/n$ are mean values of the truncated characteristic functions of A and B, respectively.

In [G4] I proved the following theorem.

Theorem 5. *Let $\mathcal{F}$ be the set of bounded locally integrable functions $\mathbb{R}_+\to\mathbb{R}_+$. For $f\in\mathcal{F}$, let*

$$L(f)=\left(\limsup_{x\to\infty}\frac{1}{x}\int_0^x f(t)\,dt,\ \liminf_{x\to\infty}\frac{1}{x}\int_0^x f(t)\,dt\right).$$

If f is fixed, then the set

$$R(f) = \{L(g): g \in \mathcal{F},\ g \leqslant f\}$$

is convex, closed and of the same form as the density set (cf. Theorem 1). Inversely, given a convenient (cf. Theorem 2) region R_0 of $\mathbb{R}^2$, there is $f \in \mathcal{F}$ such that $R(f) = R_0$.

In [GH] we unify the results obtained in the discrete and in the continuous cases, we also consider non bounded regions produced by non bounded functions (cf. Theorem 5) and we give a more comprehensive proof of the "inverse" theorem (analogue to the Theorem 2 in Section 3 of this survey) by introducing a geometrically defined auxiliary function.

Here is a more detailed summary of the content of [GH]. Let T be a boundless subset of $\mathbb{R}_+$ containing 0. (N.B. T should be $\mathbb{N} \cup \{0\}$ in the case of asymptotic densities; $\mathbb{R}_+$ in the case of limit mean values). Let $\mathcal{C}$ denote the set of all increasing functions $h: T \to \mathbb{R}_+$ such that $h(0) = 0$. We call *representative point* of a function h in $\mathcal{C}$, the ordered pair

$$r(h) = \big(s(h), i(h)\big) \in \overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+ ,$$

where

$$s(h) = \limsup_{t \to +\infty} \frac{h(t)}{t} \quad \text{and} \quad i(h) = \liminf_{t \to +\infty} \frac{h(t)}{t} .$$

Given two real functions g and h, defined on T, we say that g *is subordinated to* h if and only if the function $h - g$ is increasing.

We study the *representative set* defined as

$$R(h) = \{r(g): g \in \mathcal{C},\ g \text{ subordinated to } h\} .$$

If $s(h) < +\infty$, we obtain a general result whose special cases are the Theorems 1 and 2, Section 3, or the Theorem 5, Section 6, of this survey.

We also study the case where $s(h) = +\infty$ and $i(h) < +\infty$ and the case where both $s(h)$ and $i(h)$ are infinite.

The proof of the inverse theorem is completely recast. In the finite case ($s(h) < +\infty$) we proceed as follows. Let a, b be real numbers, $0 \leqslant b \leqslant a$. Let S be a closed convex region containing $\big[(0,0),(a,0),(a,b)\big]$ and contained in $\big[(0,0),(a,0),(a,b),(b,b)\big]$. Define $f: [0,a] \to [0,b]$ by $f(u) = \max\{v: (u,v) \in S\}$. Because of the convexity of S, for each $x \in [0,1]$, there is one and only one straight line $\Delta(x)$ of slope x tangent to the upper boundary of S. That is:

- the slope of $\Delta(x)$ is equal to x,
- $\Delta(x) \cap (\text{ interior of } S) = \emptyset$, and
- there is at least one $u \in [0, \alpha]$ such that $\big(u, f(u)\big) \in \Delta(x)$.

In fact, for each x, there is only one such $u \in [0,a]$, except if the graph of f contains segments and x is the slope of such a segment. We introduce the auxiliary function

$$\mu\colon [0,1] \to [b,a]\,, \quad x \mapsto \mu(x) = f(u) + (a-u)x\,,$$

where $\bigl(u, f(u)\bigr)$ is *anyone* of the common points of $\Delta(x)$ and S. One can verify by simple affine transformation that $\mu(x)$ depends only on x, not on u. The equation of $\Delta(x)$ in variables X, Y is now $Y - \mu(x) = (X-a)x$. Geometrically, $\mu(x)$ is the ordinate of the point of intersection of $\Delta(x)$ with the line of equation $X = a$. The function μ is increasing, $\mu(0) = b$, $\mu(1) = a$. We prove that $\mu(x)/x$ is decreasing for $0 < x \leqslant 1$.

When $T = \mathbb{R}_+$, a function $g\colon T \to \mathbb{R}_+$ such that $R(g) = S$ is constructed by successive homothetics of the function μ in the following way. Let $t_0 = 0 < t_1 < t_2 < \dots$ be a sequence of real numbers such that $t_{n+1}/t_n \to \infty$ ($t_n = n!$, for instance). Let $I(n) = \bigl[t_n, t_{n+1}\bigl[$. We define

$$g(t) = 0 \quad \text{for } t \in I(0)\,, \qquad g(t) = t\mu\left(\frac{t_n}{t}\right) \quad \text{for } t \in I_n$$

and we prove that $R(g) = S$.

REFERENCES

[A] ALEXANDER, R.: *Density and multiplicative structure of sets of integers*, Acta. Arith. **12** (1976), 321–332.

[B] BÉSINEAU, J.: *Ensembles d'entiers à caractères presque périodiques et équirépartition*, in: Répartition modulo 1 (Actes Colloq., Marseille-Luminy, 1974), Lecture Notes in Math., Vol. 475, Springer-Verlag, Berlin, 1975, pp. 1–12.

[DL] DUPAIN, Y.—LESCA, J.: *Répartition des sous-suites d'une suite donnée*, Acta Arith. **23** (1973), 307–314.

[DG] DUPAIN, Y.—GREKOS, G.: *Densités des sous-suites equiréparties d'une suite donnée*, in: Conference on Additive Number Theory (Bordeaux, 1977), Univ. Bordeaux I, Talence, 1978, pp. 43–56.

[FG-A] FUCHS, A.—GIULIANO-ANTONINI, R.: *Théorie générale des densités*, Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. **14** (1990), 253–294.

[FL] FUCHS, A.—LETTA, G.: *Le problème du premier chiffre décimal pour les nombres premiers*, J. Combin. **3** (1996), 601–607. Electron. J. Comb. **3**, (1996), Research paper R 25, 7.p.

[G0] GREKOS, G.: *Doctorat de 3ème cycle, Université de Paris VI, 22.6.1976, "Sur la répartition des densités des sous-suites d'une suite donnée"*, Jury: Charles Pisot (Président), Jean-Marc Deshouillers (Directeur de thèse), Michel Waldschmidt.

[G1] GREKOS, G.: *Suites d'entiers de densités données*, Séminaire Delange-Pisot-Poitou, 17e année (1975–1976), Théor. des Nombres: Fasc. 2, Exp. No. G15, 8 p., Secrétariat Math., Paris, 1977.

[G2] GREKOS, G.: *Densités asymptotiques de sous-suites*, Séminaire de Théorie des Nombres (1976–1977), Lab. Théor. des Nombres, Exp. No. 7, 10 p., CNRS, Talence, 1977.

[G3] GREKOS, G.: *Répartition des densités des sous-suites d'une suite d'entiers*, J. Number Theory **10** (1978), 177–191.

[G4] GREKOS, G.: *Moyennes limites et convexité*, Period. Math. Hungar. **22** (1990), 159–166.

[GH] GREKOS, G.—HADDAD, L.: *Densités, moyennes et ensemble représentatif: unification du cas discret et du cas continu*, Ann. Faculté Sci. Toulouse **10** (2001), 35–68.

[GST] GREKOS, G.—ŠALÁT, T.—TOMANOVA, J.: *Gaps and densities*, Bulletin Mathématique Soc. Sci. Math. Roumaine (N.S.) **46** (2003–2004), 121–141.

[GV] GREKOS, G.—VOLKMANN, B.: *On densities and gaps*, J. Number Theory **26** (1987), 129–148.

[M-F] MENDÈS-FRANCE, M.: *Suites et sous-suites équiréparties modulo* 1, in: J. Arithm. Françaises (Univ. Provence, Marseille, 1971), Exp. No. 4, 5 p., Univ. Provence, Marseille, 1971.

[MT] MIŠÍK, L.—TÓTH, J. T.: *Logarithmic density of a sequence of integers and density of its ratio set*, J. Théor. Nombres Bordeaux **15** (2003), 309–318.

[PSV] PAŠTÉKA, M.—ŠALÁT, T.—VISNYAI, T.: *Remarks on Buck's measure density and a generalization of asymptotic density*, Tatra Mt. Math. Publ. **31** (2005), 87–101.

[RV] ROHRBACH, H.—VOLKMANN, B.: *Verallgemeinerte asymptotische Dichten*, J. Reine Angew. Math. **194** (1955), 195–209.

[ST] ŠALÁT, T.—TIJDEMAN, R.: *Asymptotic densities of sets of positive integers*, Math. Slovaca **33** (1983), 199–207.

Received November 12, 2004

Département de Mathématiques
Faculté des Sciences et Techniques
23, rue du Dr. Paul Michelon
F–42023 Saint-Etienne Cédex 2
FRANCE

E-mail: grekos@univ-st-etienne.fr