

Jacek M. Pomykała, Janusz A. Pomykała

## ON THE NORMAL ORDER OF DOMAIN SETS OF BINARY RELATIONS

In the paper we consider any binary relation  $\rho$  with “almost independent” codomain sets. We investigate the influence of the independence condition on the corresponding domain sets of  $\rho$  and the lifting relation  $\hat{\rho}$ .

### 1. Notation

$\mathcal{X}, \mathcal{Y}$  — finite sets

$P(\mathcal{X})$  — family of all subsets of  $\mathcal{X}$

$P_k(\mathcal{Y})$  — family of all  $k$ -element subsets of  $\mathcal{Y}$

$\rho \subset \mathcal{X} \times \mathcal{Y}$  — a binary relation

$\hat{\rho} \subset P(\mathcal{X}) \times P_k(\mathcal{Y})$  — the relation defined by

$$\mathbb{X} \hat{\rho} \mathbb{Y} \Leftrightarrow \forall x \in \mathbb{X} \exists y \in \mathbb{Y} x \rho y \quad (\text{see [3]})$$

$\alpha(x, y)(\hat{\alpha}(x, y))$  — the characteristic function of the relation  $\rho(\hat{\rho})$  respectively

$$U_x = \{y \in Y : x \rho y\}$$

$$V_y = \{x \in X : x \rho y\}$$

$|A|$  — denotes the cardinality or absolute value according to whether  $A$  is a set or a number

$\langle a \rangle$  — the largest integer  $\leq a$ .

The symbol  $\mathcal{O}()$  is employed in the following sense:  $A(X) = \mathcal{O}(B(x))$  where  $B(X)$  is positive, means that

$$\limsup_{X \rightarrow \infty} \frac{|A(X)|}{B(X)} < \infty.$$

### 2. Introduction

Let  $\mathfrak{X} = \{x_1 \prec x_2 \prec \dots\}$ ,  $\mathfrak{Y} = \{y_1 \prec y_2 \prec \dots\}$  be the infinite ordered sets and  $\rho \subset \mathfrak{X} \times \mathfrak{Y}$  any binary relation.

Given any function  $Y = Y(u)$  we denote by  $\rho|_u$  the relation  $\rho$  restricted to the set  $\{x_1, \dots, x_{\langle u \rangle}\} \times \{y_1, \dots, y_{\langle Y(u) \rangle}\}$ .

We investigate the asymptotic behaviour of  $|U_x|$  and  $|V_y|$  for  $\rho = \rho|_u$  as  $u$ -approaches infinity. Let  $\mathcal{X}$  and  $\mathcal{Y}$  denote the domain and codomain respectively of  $\rho|_u$  while  $X$  and  $Y = Y(X)$  their cardinalities. Then any function  $V = V(X, Y)$  (if exists) such that we have

$$(1) \quad \sum_{y \in Y} (V - |V_y|)^2 = \mathcal{O}(YV^2\delta(X))$$

with some  $\delta(X)$  tending to zero as  $X$  approaches infinity we call the normal order of the domain sets  $V_y$  ( $y \in Y$ ). The notion of normal order is traditionally used in the theory of numbers (where  $\rho$  is a divisibility relation). The condition (1) says (roughly speaking) that the value of  $|V_y|$  is asymptotically equal to  $V$  for almost all  $y \in \mathcal{Y}$ .

In the paper we establish a sufficient conditions for the relation  $\rho$  and  $\hat{\rho}$  to have the normal order of the domain sets and we present the explicit formula for  $V(X, Y)$ . In particular taking as  $\rho$  the divisibility relation one obtains the well known conclusions about the normal order of the number theoretical functions  $\omega(n)$  and  $d(n)$  (cf. [1] p. 356).

### 3. Almost independence of codomain sets

Having endowed the discrete measure on  $\mathcal{Y}$  one regards two sets  $U_x, U_{x'}$  as independent iff

$$|U_x \cap U_{x'}| = |U_x||U_{x'}||\mathcal{Y}|^{-1}.$$

In many examples (see below) the above quantity is valid only approximately. We shall therefore replace the above condition by the average one, replacing the corresponding equality by the asymptotic equality as  $X \rightarrow \infty$ . More precisely we shall postulate the following approximate formulas (for any distinct  $x_1, x_2 \in \mathcal{X}$ )

$$(A_1) \quad \sum_{y \in \mathcal{Y}} \alpha(x_1, y) = Y\mu(x_1) + r(x_1; \mathcal{Y}),$$

$$(A_2) \quad \sum_{y \in \mathcal{Y}} \alpha(x_1, y)\alpha(x_2, y) = Y\mu(x_1)\mu(x_2) + r'(x_1, x_2; \mathcal{Y}),$$

where  $\mu(x)$  is some nonnegative function such that  $0 \leq \mu(x) \leq 1$  and  $r, r'$  are to be regarded as the remainder terms.

Assume that the following conditions hold true

$$(E1) \quad \sum_{x \in \mathcal{X}} \mu(x) = F(X)\{1 + \mathcal{O}(\delta(X))\},$$

$$(E2) \quad \sum_{x \in \mathcal{X}} r(x; \mathcal{Y}) = \mathcal{O}(YF(X)[\delta(X) + F(X)^{-1}]),$$

$$(E3) \quad \sum_{\substack{x_1, x_2 \in \mathcal{X} \\ x_1 \neq x_2}} r'(x_1, x_2; \mathcal{Y}) = \mathcal{O}(YF(X)^2[\delta(X) + F(X)^{-1}]),$$

with some  $F(X)$ ,  $\delta(X)$  satisfying

$$\lim_{X \rightarrow \infty} F(X) = \infty; \quad \lim_{X \rightarrow \infty} \delta(X) = 0.$$

Then we say that the codomain sets  $U_x$  ( $x \in \mathcal{X}$ ) are almost independent in  $\mathcal{Y}$ .

#### 4. The principal result

**THEOREM.** *Under the assumptions  $(A_1)$ – $(A_2)$ ,  $(E1)$ – $(E3)$  we have that*

$$(*) \quad \sum_{y \in \mathcal{Y}} \left( \sum_{x \in \mathcal{X}} \alpha(x, y) - F(X) \right)^2 = \mathcal{O}\{YF(X)^2[\delta(X) + F(X)^{-1}]\} \quad \text{as } X \rightarrow \infty.$$

**Proof.** Squaring out the left-hand-side of  $(*)$  we obtain that it is equal to

$$\sum_y \left\{ \sum_{x_1, x_2} \alpha(x_1, y) \alpha(x_2, y) - 2 \sum_x \alpha(x, y) F(X) + F(X)^2 \right\} = W - 2V + U, \text{ say}$$

where the variables in the above sums range over  $y \in \mathcal{Y}$ ,  $x_1, x_2, x \in \mathcal{X}$ . Changing the order of summation we obtain

$$W = \sum_{x_1, x_2} \sum_y \alpha(x_1, y) \alpha(x_2, y) = \sum_{x_1 \neq x_2} \sum_y \alpha(x_1, y) \alpha(x_2, y) + \sum_x \sum_y \alpha(x, y)^2.$$

The second term contributes to the error term, since

$$\begin{aligned} \sum_x \sum_y \alpha(x, y)^2 &\leq \sum_x \sum_y \alpha(x, y) = \sum_x |Y\mu(x) + r(x; \mathcal{Y})| \\ &\leq \{YF(X)[1 + \mathcal{O}(\delta(X))] + \mathcal{O}(YF(X)[\delta(X) + F(X)^{-1}])\} \\ &= \{YF(X)[1 + \mathcal{O}(\delta(X) + F(X)^{-1})]\} = \mathcal{O}\{YF(X)^2[\delta(X) + F(X)^{-1}]\}. \end{aligned}$$

The first one is equal to

$$\begin{aligned} \sum_{x_1 \neq x_2} \{Y\mu(x_1)\mu(x_2) + r'(x_1, x_2; \mathcal{Y})\} &= \sum_{x_1, x_2} Y\mu(x_1)\mu(x_2) - \sum_{x_1} Y\mu(x_1)^2 \\ &\quad + \sum_{x_1 \neq x_2} r(x_1, x_2; \mathcal{Y}) = YF(X)^2\{1 + \mathcal{O}(\delta(X))\}^2 \end{aligned}$$

$$\begin{aligned}
& + \mathbb{O}\left[Y \sum_x \mu(x)\right] + \mathbb{O}\{Y F(X)^2[\delta(X) + F(X)^{-1}]\} \\
& = Y F(X)^2\{1 + \mathbb{O}(\delta(X)) + \mathbb{O}(F(X)^{-1}[1 + \delta(X)] + \mathbb{O}([\delta(X) + F(X)^{-1}])\} \\
& = Y F(X)^2\{1 + O[\delta(X) + F(X)^{-1}]\}.
\end{aligned}$$

Similarly we show

$$\begin{aligned}
V &= F(X) \sum_x \sum_y \alpha(x, y) = F(X) \sum_x (Y \mu(x) + r(x; \mathcal{Y})) \\
&= F(X) \{Y F(X) \{1 + \mathbb{O}(\delta(X))\} + \mathbb{O}(Y F(X)^2[\delta(X) + F(X)^{-1}])\} \\
&= Y F(X)^2 \{1 + \mathbb{O}(\delta(X) + F(X)^{-1})\}, \\
U &= \sum_y F(X)^2 = Y F(X)^2.
\end{aligned}$$

Therefore we have

$$W - 2V + U = O\{Y F(X)^2[\delta(X) + F(X)^{-1}]\}$$

which completes the proof of the Theorem.

**COROLLARY.** *Under the assumptions of the Theorem there exists an absolute constant  $C > 0$  such that*

$$\begin{aligned}
(2) \quad \text{card} \left\{ y \in \mathcal{Y} : \sum_{x \in \mathcal{X}} \alpha(x, y) > F(X) \{1 + [\delta(X) + F(X)^{-1}]^{\frac{1}{3}}\} \right. \\
\left. \text{or } \sum_{x \in \mathcal{X}} \alpha(x, y) < F(X) \{1 - [\delta(X) + F(X)^{-1}]^{\frac{1}{3}}\} \right\} \leq C Y [\delta(X) + F(X)^{-1}]^{\frac{1}{3}}.
\end{aligned}$$

**PROOF.** Denoting the above set by  $S(\mathcal{Y}, \mathcal{X})$  we obtain in view of (\*)

$$\sum_{y \in S(\mathcal{Y}, \mathcal{X})} [\delta(X) + F(X)^{-1}]^{\frac{2}{3}} F(X)^2 = \mathbb{O}\{Y F(X)^2[\delta(X) + F(X)^{-1}]\}$$

hence for some constant  $C > 0$  it holds

$$\text{card } S(\mathcal{Y}, \mathcal{X}) \leq C Y [\delta(X) + F(X)^{-1}]^{\frac{1}{3}}$$

as asserted.

**PROPOSITION.** *Under the assumptions of the above corollary we have*

$$\begin{aligned}
(3) \quad \text{card} \left\{ \mathbb{Y} \in P_k(\mathcal{Y}) : \sum_{\mathbb{X} \in P(\mathcal{X})} \hat{\alpha}(\mathbb{X}, \mathbb{Y}) > 2^{kF(X)(1+[\delta(X)+F(X)^{-1}]^{\frac{1}{3}})} \right\} \\
\leq C Y^k [\delta(X) + F(X)^{-1}] \\
(4) \quad \text{card} \left\{ \mathbb{Y} \in P_k(\mathcal{Y}) : \sum_{\mathbb{X} \in P(\mathcal{X})} \hat{\alpha}(\mathbb{X}, \mathbb{Y}) < 2^{F(X)(1-[\delta(X)+F(X)^{-1}]^{\frac{1}{3}})} \right\}
\end{aligned}$$

$$\leq CY^k[\delta(X) + F(X)^{-1}]^{\frac{1}{3}}.$$

**Proof.** Letting  $S(\mathcal{Y}, \mathcal{X})$ ,  $S_k(\mathcal{Y}, \mathcal{X})$ ,  $S'_k(\mathcal{Y}, \mathcal{X})$  be the corresponding sets in (2), (3), (4) respectively and  $\mathbb{Y} = \{y_1, \dots, y_k\}$ , we will prove the implication

$$(5) \quad \{y_1, \dots, y_k\} \in S_k(\mathcal{Y}, \mathcal{X}) \cup S'_k(\mathcal{Y}, \mathcal{X}) \Rightarrow \exists_{1 \leq j \leq k} \quad y_j \in S(\mathcal{Y}, \mathcal{X}).$$

This condition completes the proof since the cardinality of such  $\{y_1, \dots, y_k\}$  is bounded by

$$\binom{Y}{k-1} CY[\delta(X) + F(X)^{-1}]^{\frac{1}{3}} \leq CY^k[\delta(X) + F(X)^{-1}]^{\frac{1}{3}}$$

as it is required in (3) and (4). To prove (5) let

$$\mathbb{X}' = \mathbb{X}'(k) = \{x \in \mathcal{X} : \hat{\alpha}(\{x\}, \{y_1, \dots, y_k\}) = 1\}$$

and assume on the contrary that  $\forall_{1 \leq j \leq k} y_j \in Y \setminus S(\mathcal{Y}, \mathcal{X})$ .

Then we have

$$\sum_{\mathbb{X} \in P(\mathcal{X})} \hat{\alpha}(\mathbb{X}, \{y_1, \dots, y_k\}) = 2^{|\mathbb{X}'|}$$

where in view of (2)

$$\begin{aligned} |\mathbb{X}'| &= \sum_{x \in \mathcal{X}} \hat{\alpha}(\{x\}, \{y_1, \dots, y_k\}) \\ &\leq \sum_{j=1}^k \sum_{x \in \mathcal{X}} \alpha(x, y_j) \leq kF(X)\{1 + [\delta(X) + F(X)^{-1}]^{\frac{1}{3}}\} \end{aligned}$$

and

$$|\mathbb{X}'| \geq \max_{1 \leq j \leq k} \sum_{x \in \mathcal{X}} \alpha(x, y_j) \geq F(X)\{1 - [\delta(X) + F(X)^{-1}]^{\frac{1}{3}}\}$$

on the other hand. Hence in view of (3) and (4) we obtain that  $\{y_1, \dots, y_k\}$  is neither in  $S_k(\mathcal{Y}, \mathcal{X})$  nor  $S'_k(\mathcal{Y}, \mathcal{X})$ . This contradiction proves that (5) holds true and therefore the proof of Proposition is established.

**Remark.** If we attach with any pair  $(x, y)$  such that  $xpy$  the “arrow” with the endpoints  $x$  and  $y$  respectively then we can express the conditions (E1)–(E3) by saying that

- (i) the number of all arrows is asymptotically equal to  $YF(X)$
- (ii) the ends of arrows form the “independent” sets in codomain of the relation.

As the conclusion we obtain that the almost all coimages of points have asymptotically  $F(X)$  elements.

### 5. Examples

EXAMPLE 1. Let  $U$  be any set of cardinality  $n$  ( $n \rightarrow \infty$ ),  $\rho$  — the relation of inclusion. We set  $\mathcal{X} = P_1(U)$ ,  $\mathcal{Y} = P(U)$ . We have

$$|U_x| = 2^{n-1}, \quad |U_{x_1} \cap U_{x_2}| = 2^{n-2}, \quad (x_1 \neq x_2).$$

Hence  $(A_1)$  and  $(A_2)$  are satisfied with  $\mu(x_1) = \mu(x_2) = \frac{1}{2}$ ,  $r(x_1; \mathcal{Y}) = r'(x_1, x_2; \mathcal{Y}) = 0$ . Therefore (E1) holds with  $F(X) = F(n) = \frac{n}{2}$ ,  $\delta(n) = 0$ , while (E2), (E3) hold trivially. By the Corollary we obtain that the number of  $\mathcal{Y} \in P(U)$  having more than  $\frac{n}{2}(1 + (\frac{n}{2})^{-\frac{1}{3}})$  (or less than  $\frac{n}{2}(1 - (\frac{n}{2})^{-\frac{1}{3}})$ ) elements is at most  $C2^n n^{-\frac{1}{3}}$  with some constant  $C$  not depending on  $n$ .

EXAMPLE 2.  $\mathcal{X} = P_1(U)$ ,  $\mathcal{Y}$  — set all subsets of  $U$  with even number of elements (including the empty set). Since

$$\sum_{m=0}^n (-1)^m \binom{n}{m} = (1-1)^n = 0$$

we have that

$$|\mathcal{Y}| = \frac{1}{2}2^n = 2^{n-1}; \quad |U_x| = \frac{1}{2}2^{n-1} = \frac{1}{2}|\mathcal{Y}|$$

$$|U_{x_1} \cap U_{x_2}| = \frac{1}{2}2^{n-2} = \frac{1}{4}|\mathcal{Y}|, \quad (x_1 \neq x_2).$$

Therefore we obtain the same conclusion as in Example 1.

EXAMPLE 3.  $\mathcal{Y}$  — family of sets with the number of elements divisible by 3. Using the formulas

$$\sum_{k=0}^{\langle \frac{n}{3} \rangle} \binom{n}{3k} = \frac{1}{3} \left( 2^n + 2 \cos \frac{n\pi}{3} \right),$$

$$\sum_{k=0}^{\langle \frac{n}{3} \rangle} \binom{n}{3k+1} = \frac{1}{3} \left( 2^n + 2 \cos \frac{(n-2)\pi}{3} \right),$$

$$\sum_{k=0}^{\langle \frac{n}{3} \rangle} \binom{n}{3k+2} = \frac{1}{3} \left( 2^n + 2 \cos \frac{(n-4)\pi}{3} \right),$$

which can be proved by induction (see ex. 1.17.18 [2]), we infer that

$$|Y| = \frac{1}{3} \left( 2^n + 2 \cos \frac{n\pi}{3} \right),$$

$$|U_x| = \frac{1}{3} \left( 2^{n-1} + 2 \cos \frac{(n-5)\pi}{3} \right),$$

$$|U_{x_1} \cap U_{x_2}| = \frac{1}{3} \left( 2^{n-2} + 2 \cos \frac{(n-4)\pi}{3} \right), \quad (x_1 \neq x_2).$$

Hence  $(A_1)$ ,  $(A_2)$  hold with

$$\mu(x_1) = \mu(x_2) = \frac{1}{2}, \quad |r(x_1; \mathcal{Y})| \leq \frac{2}{3}; \quad |r(x_1, x_2; \mathcal{Y})| \leq \frac{2}{3}.$$

Therefore (E1) is satisfied with  $F(n) = \frac{n}{2}$ ,  $\delta(n) = 0$ . Also for  $n \geq 8$  the conditions (E2) and (E3) hold true, since

$$\sum_{x \in \mathcal{X}} 1 \leq \sum_{x_1, x_2 \in \mathcal{X}} 1 \leq n^2 \leq \frac{2^n}{3} - 1 \leq |\mathcal{Y}|.$$

The conclusion is therefore the same as in Example 1.

EXAMPLE 4. Let  $\rho$  be the divisibility relation and  $T$  — be a large positive integer ( $T \rightarrow \infty$ ). We define

$$\mathcal{Y} = \{f(t) = t^2 + 1, \quad t \text{ — positive integer } t \leq T\},$$

$$\mathcal{X} = \{p : p \text{ — prime number, } p \leq \sqrt{T}\}.$$

Then we have the approximation formula

$$\begin{aligned} \sum_{y \in \mathcal{Y}} \alpha(p, y) &= \sum_{\substack{t \leq T \\ f(t) \equiv 0(p)}} 1 = \sum_{\substack{\vartheta \pmod{p} \\ f(\vartheta) \equiv 0 \pmod{p}}} \sum_{\substack{t \leq T \\ t \equiv \vartheta \pmod{p}}} 1 \\ &= \sum_{\substack{\vartheta \pmod{p} \\ f(\vartheta) \equiv 0 \pmod{p}}} \frac{T - \vartheta}{p} + \left( \frac{t - \vartheta}{p} - \left\langle \frac{T - \vartheta}{p} \right\rangle \right) \\ &= \frac{T}{p} \omega(p) + \omega(p) \left( \frac{T - \vartheta}{p} - \left\langle \frac{T - \vartheta}{p} \right\rangle - \frac{\vartheta}{p} \right), \end{aligned}$$

where

$$\omega(p) = \begin{cases} 1 & \text{if } p = 2 \\ 2 & \text{if } p \equiv 1 \pmod{4} \\ 0 & \text{if } p \equiv 3 \pmod{4}. \end{cases}$$

The choice  $\mu(p) = \frac{\omega(p)}{p}$  assures that  $|r(p; Y)| \leq 4$ . The same argument shows that

$$\sum_{y \in \mathcal{Y}} \alpha(p_1, y) \alpha(p_2, y) = T \mu(p_1) \mu(p_2) + r'(p_1, p_2; \mathcal{Y})$$

with  $|r'(p_1, p_2; \mathcal{Y})| \leq 4$ .

The assertion (E1) follows from the well known formula (see e.g. [1]) for

the prime reciprocal series

$$\sum_{p \leq \sqrt{T}} \frac{1}{p} = \log \log \sqrt{T} \{1 + \mathcal{O}(\log \log T)^{-1}\}.$$

Since  $X = \pi(\sqrt{T})$  — number of primes  $\leq \sqrt{T}$ , we obtain (manipulating with the error term  $\mathcal{O}(\cdot)$ ) that

$$\sum_{p \leq \sqrt{T}} \frac{1}{p} = \log \log X \{1 + \mathcal{O}[(\log \log X)^{-1}]\}.$$

Moreover we have

$$\sum_{p_1, p_2 \leq \sqrt{T}} |r'(p_1, p_2; \mathcal{Y})| \leq 8T = \mathcal{O}(X \log \log X)$$

hence the conditions (E2), (E3) are satisfied with  $Y = T$ ,  $X = \pi(\sqrt{T})$ ,  $F(X) = \log \log X$ ;  $\delta(X) = (\log \log X)^{-1}$ .

Therefore by the Corollary we infer that almost every value  $f(t)$ , ( $t \leq T$ ) has

$$\log \log X \{1 + \mathcal{O}[(\log \log X)^{-\frac{1}{3}}]\} = \log \log T \{1 + \mathcal{O}[(\log \log T)^{-\frac{1}{3}}]\}$$

prime divisors not exceeding  $\sqrt{T}$ . Since  $y = f(t)$  has at most 4 prime divisors in the interval  $[\sqrt{T}, T^2 + 1]$  the same conclusion holds for the number of all prime divisors of  $f(t)$ .

## References

- [1] G. H. Hardy, E. M. Wright, *An introduction to the theory of numbers*, (second edition) Oxford At the Clarendon Press, 1945.
- [2] L. Jeśmanowicz, J. Łoś, *Zbiór zadań z algebry*, PWN, (wyd. V) Warszawa 1972.
- [3] C. Brink, *Powering structures*, the Lecture given in a Semester on Algebra at International Mathematical Banach Center, Warsaw 1991.

Jacek M. Pomykała  
WARSAW UNIVERSITY  
ul. Banacha 2  
02-097 WARSZAWA, POLAND

Janusz A. Pomykała  
DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF PEDAGOGICS  
00-110 SIEDLCE, POLAND

*Received March 12, 1993.*