

INTERSECTION THEOREMS FOR SYSTEMS OF SETS

P. ERDÖS and R. RADO†

A version of Dirichlet's box argument asserts that given a positive integer a and any a^2+1 objects x_0, x_1, \dots, x_{a^2} , there are always $a+1$ distinct indices ν ($0 \leq \nu \leq a^2$) such that the corresponding $a+1$ objects x_ν are either all equal to each other or mutually different from each other. This proposition can be restated as follows. Let N be an index set of more than a^2 elements and let, for each element ν of N , X_ν be a one-element set. Then there is a subset N' of N having more than a elements, such that all intersections $X_\mu X_\nu$ corresponding to distinct elements μ, ν of N' have the same value. In this note we investigate extensions of this principle to cases when the sets X_ν are of any prescribed cardinal b . Both a and b are given cardinals, finite or infinite. In the case of finite a and b we obtain estimates for the number which corresponds to a^2 in Dirichlet's case, and we show that when at least one of a and b is infinite then a^{b+1} is the best possible value of that number.

We introduce some definitions‡. A system $\Sigma_1: Y_\nu (\nu \in N)$ of sets Y_ν , where ν ranges over the index set N , is said to contain the system $\Sigma_0: X_\mu (\mu \in M)$ if, for every μ_0 of M , the set X_{μ_0} occurs in Σ_1 at least as often as in Σ_0 , i.e. if

$$|\{\nu: \nu \in N; Y_\nu = X_{\mu_0}\}| \geq |\{\mu: \mu \in M; X_\mu = X_{\mu_0}\}|.$$

If Σ_1 contains Σ_0 and, at the same time, Σ_0 contains Σ_1 , then we do not distinguish between the systems Σ_0 and Σ_1 .

The system Σ_0 is called a (a, b) -system if it consists of a (not necessarily distinct) sets of cardinal b , i.e., if $|M| = a$ and $|X_\mu| = b$ for $\mu \in M$. The system Σ_0 is called a Δ -system if it has the property that the intersections of any two of its sets§ have the same value, i.e. if for

$$\mu_0, \mu_1, \mu_2, \mu_3 \in M; \mu_0 \neq \mu_1; \mu_2 \neq \mu_3$$

we always have $X_{\mu_0} X_{\mu_1} = X_{\mu_2} X_{\mu_3}$. More specifically, Σ_0 is a $\Delta(a)$ -system with kernel K if $|M| = a$ and $X_{\mu_0} X_{\mu_1} = K$ whenever $\mu_0, \mu_1 \in M; \mu_0 \neq \mu_1$. In the special case when $|M| = 1$, say $M = \{\mu_0\}$, we stipulate that $K \subset X_{\mu_0}$, and the empty system Σ_0 , for which $M = \emptyset$, is considered as a $\Delta(0)$ -system with any arbitrary set K as kernel. Expressions such as

$$(> a, \leq b)\text{-system, } \Delta(> a)\text{-system}$$

have their obvious meaning. Trivially, every $(> a, 0)$ -system is a Δ -system, and the box principle stated above asserts that every

† Received 5 December, 1958; read 18 December, 1958.

‡ The cardinal of the set A is denoted by $|A|$, and set union by $A+B$ or $\Sigma(\nu \in N) A_\nu$ and set intersection by AB or $\Pi(\nu \in N) A_\nu$. $A \subset B$ denotes inclusion, in the wide sense. We use the obliteration operator \wedge whose effect consists in removing from a well-ordered series the term above which it is placed. Unless the contrary is stated all sets are allowed to be empty.

§ Not necessarily distinct sets but sets having distinct indices μ .

$(> a^2, 1)$ -system contains a $\Delta(> a)$ -system. In what follows a and b denote arbitrary cardinals, and b^+ is the next larger cardinal to b .

THEOREM I.

- (i) If $a, b \geq 1$ then every $(> b^+ b^b a^{b+1}, \leq b)$ -system contains a $\Delta(> a)$ -system.
- (ii) If $a \geq 2; b \geq 1; a+b \geq \aleph_0$ then every $(\dot{>} a^b, \leq b)$ -system contains a $\Delta(> a)$ -system.

THEOREM II. For every a, b such that $a, b \geq 1$ there exists a (a^{b+1}, b) -system which does not contain any $\Delta(> a)$ -system.

THEOREM III. If $1 \leq a, b < \aleph_0$ and

$$c = b! a^{b+1} \left(1 - \frac{1}{2!a} - \frac{2}{3!a^2} - \dots - \frac{b-1}{b!a^{b-1}} \right) \quad (1)$$

then every $(> c, \leq b)$ -system contains a $\Delta(> a)$ -system.

Remarks. 1. It follows from II that I(ii) is best possible, in the sense that, for $a \geq 2; b \geq 1; a+b \geq \aleph_0$ not every $(a^b, \leq b)$ -system contains some $\Delta(> a)$ -system.

2. The (a^{b+1}, b) -system of Theorem II will be constructed explicitly.

3. For $a = 2; b = 2$ the result III is best possible. For we have $c = 12$, and the following $(12, 2)$ -system does not contain any $\Delta(3)$ -system.

$$01, 01, 23, 23, 04, 04, 14, 14, 25, 25, 35, 35, \text{ where } xy = \{x, y\}.$$

However, for $a = 3; b = 2$ Theorem III is not best possible. By II we see that III is best possible except for a factor between 1 and $b!$.

It is not improbable that in (1) the factor $b!$ can be replaced by $c_1 b$, for some absolute positive constant c_1 . Such a sharpened version of III would have some applications in the theory of numbers, and in fact these applications originally gave rise to the present investigations.

Before proving Theorem I we establish a simple lemma which is at the root of a large number of combinatorial arguments.

RAMIFICATION LEMMA. Let α_0 be an ordinal, $c_0, c_1, \dots, \hat{c}_{\alpha_0}$ be cardinals; let S be a set and $M(s_0, s_1, \dots, \hat{s}_\alpha)$ be a subset of S defined for $\alpha < \alpha_0$ and $s_0, \dots, \hat{s}_\alpha \in S$, such that $|M(s_0, \dots, \hat{s}_\alpha)| \leq c_\alpha$. Let V be a set of "vectors" $(s_0, s_1, \dots, \hat{s}_{\alpha_0})$ such that $s_0, \dots, \hat{s}_{\alpha_0} \in S$ and

$$s_\alpha \in M(s_0, \dots, \hat{s}_\alpha) \text{ for } \alpha < \alpha_0. \quad (2)$$

Then

$$|V| \leq c_0 c_1 \dots \hat{c}_{\alpha_0}.$$

Proof. For every subset S' of S choose a representation of the form $S' = \{t_0, t_1, \dots, \hat{t}_k\} \neq \emptyset$, where k is the initial ordinal belonging to $|S'|$.

Define $\phi(S', t)$ for $t \in S'$ by putting $\phi(S', t_\kappa) = \kappa$ ($\kappa < k$). Now let $s = (s_0, \dots, \hat{s}_{\alpha_0}) \in V$. Then, by (2), we can define an ordinal $\kappa(\alpha, s)$ by putting

$$\kappa(\alpha, s) = \phi(M(s_0, \dots, \hat{s}_\alpha, s_\alpha) \quad (\alpha < \alpha_0),$$

and a vector $\psi(s) = (\kappa(0, s), \kappa(1, s), \dots, \hat{\kappa}(\alpha_0, s))$.

Then $|\kappa(\alpha, s)| < c_\alpha$, and $s \neq s'$ implies $\psi(s) \neq \psi(s')$ as can be seen by considering the least α with $s_\alpha \neq s'_\alpha$. We conclude that

$$|V| = |\{s : s \in V\}| = |\{\psi(s) : s \in V\}| \leq c_0 c_1 \dots \hat{c}_{\alpha_0}.$$

Proof of Theorem I. We suppose that the system

$$\Sigma : X_\nu \quad (\nu \in N)$$

is a $(|N|, \leq b)$ -system which does not contain any $\Delta(> a)$ -system, and our aim is to deduce that

$$|N| \leq b + b^b a^{b+1}. \tag{3}$$

Throughout the proof the letters α, β denote ordinals such that $|\alpha|, |\beta| \leq b$, and μ, ν denote elements of N . A subset N' of N is called a Δ -set with kernel K if the system $X_\nu (\nu \in N')$ is a Δ -system with kernel K . Put $X = \Sigma (\nu \in N) X_\nu$, and choose an object θ such that $\theta \notin X$. Well-order the sets X and N as well as the set of all subsets of N .

We define elements $f_\alpha(\nu)$ as follows. Let α_0 be fixed, and suppose that $f_\alpha(\nu)$ has already been defined for $\alpha < \alpha_0$ and for all ν , and that

$$f_\alpha(\nu) \in X_\nu + \{\theta\} \quad (\alpha < \alpha_0; \nu \in N).$$

Let $\nu_0 \in N$. We now proceed to define $f_{\alpha_0}(\nu_0)$. Put, for any functions $g_0(\nu), \dots, \hat{g}_{\alpha_0}(\nu)$ defined for $\nu \in N$ and for any $x_0, \dots, \hat{x}_{\alpha_0} \in X + \{\theta\}$

$$F(g_0, \dots, \hat{g}_{\alpha_0}; x_0, \dots, \hat{x}_{\alpha_0}) = N \Pi (\alpha < \alpha_0) \{\nu : g_\alpha(\nu) = x_\alpha\}.$$

Let $N' \subset N; K \subset X + \{\theta\}$. Put

$$H(N') = \Sigma (\nu \in N') X_\nu$$

and define $\Gamma(N', K)$ to be the first subset N'' of N such that (i) $N'' \subset N'$, (ii) N'' is a Δ -set with kernel K , (iii) N'' is maximal such that (i), (ii) hold. Then, by hypothesis about Σ ,

$$|\Gamma(N', K)| \leq a.$$

Put $N(\alpha_0, \nu_0) = F(f_0, \dots, \hat{f}_{\alpha_0}; f_0(\nu_0), \dots, \hat{f}_{\alpha_0}(\nu_0))$,

$$K(\alpha_0, \nu_0) = \{f_\alpha(\nu_0) : \alpha < \alpha_0\},$$

$$N^*(\alpha_0, \nu_0) = \Gamma(N(\alpha_0, \nu_0), K(\alpha_0, \nu_0)).$$

Then $\nu_0 \in N(\alpha_0, \nu_0); N^*(\alpha_0, \nu_0) \subset N(\alpha_0, \nu_0); |N^*(\alpha_0, \nu_0)| \leq s$.

Case 1. $\theta \in K(\alpha_0, \nu_0)$. Then put $f_{\alpha_0}(\nu_0) = \theta$.

Case 2. $\theta \notin K(\alpha_0, \nu_0)$.

Case 2a. $\nu_0 \in N^*(\alpha_0, \nu)$. Then put $f_{\alpha_0}(\nu_0) = \theta$.

Case 2b. $\nu_0 \notin N^*(\alpha_0, \nu_0)$. Then, by (iii) above, $N^*(\alpha_0, \nu_0) + \{\nu_0\}$ is not a Δ -set with kernel $K(\alpha, \nu_0)$. If we now assume that $N^*(\alpha_0, \nu_0) = \emptyset$ then this last fact implies that $K(\alpha_0, \nu_0) \not\subset X_{\nu_0}$ which, however, is false. Hence $N^*(\alpha_0, \nu_0) \neq \emptyset$, and there is a first element ν_1 of $N^*(\alpha_0, \nu_0)$ such that $X_{\nu_0} X_{\nu_1} \neq K(\alpha_0, \nu_0)$. Since we are in Case 2, we have $K(\alpha_0, \nu_0) \subset X_{\nu_0}$. Since $\nu_1 \in N^*(\alpha_0, \nu_0)$ we have $K(\alpha_0, \nu_0) \subset X_{\nu_1}$. Hence $K(\alpha_0, \nu_0) \subsetneq X_{\nu_0} X_{\nu_1}$, and we may define $f_{\alpha_0}(\nu_0)$ to be the first element of the set $X_{\nu_0} X_{\nu_1} - K(\alpha_0, \nu_0)$. This completes the definition of $f_{\alpha}(\nu)$. We have, in Case 2b,

$$f_{\alpha_0}(\nu_0) \in X_{\nu_1} \subset H(N^*(\alpha_0, \nu_0)). \quad (4)$$

Let $\nu \in N$. If for some α , we have $f_{\alpha}(\nu) \neq \theta$ then Case 2b applies to α , and hence also to each $\beta \leq \alpha$; the elements $f_{\beta}(\nu)$ ($\beta \leq \alpha$) are therefore distinct elements of X_{ν} . Hence, in view of $|X_{\nu}| \leq b$, there is β_{ν} such that $f_{\alpha}(\nu) \in X_{\beta_{\nu}}$ for $\alpha < \beta_{\nu}$, and $f_{\beta_{\nu}}(\nu) = \theta$. Then Case 2a applies to β_{ν} , and we have $\nu \in N^*(\beta_{\nu}, \nu)$. This shows that†

$$N = \Sigma(\nu \in N) N^*(\beta_{\nu}, \nu), \quad |N| \leq a |\{N^*(\beta_{\nu}, \nu) : \nu \in N\}|.$$

We now prove that on the right hand side N^* may be replaced by N . Let $N(\beta_{\mu}, \mu) = N(\beta_{\nu}, \nu)$. Then $\mu \in N(\beta_{\mu}, \mu) = N(\beta_{\nu}, \nu)$; $f_{\alpha}(\mu) = f_{\alpha}(\nu) \in X_{\nu}$ for $\alpha < \beta_{\nu}$; $\beta_{\mu} \geq \beta_{\nu}$ and hence, by symmetry, $\beta_{\nu} \geq \beta_{\mu}$. Therefore $\beta_{\mu} = \beta_{\nu}$,

$$K(\beta_{\mu}, \mu) = K(\beta_{\nu}, \nu); \quad N^*(\beta_{\mu}, \mu) = N^*(\beta_{\nu}, \nu).$$

Thus $|\{N^*(\beta_{\nu}, \nu) : \nu \in N\}| \leq |\{N(\beta_{\nu}, \nu) : \nu \in N\}|$.

For any α and any $x_0, \dots, \hat{x}_{\alpha} \in X$ put

$$G(x_0, \dots, \hat{x}_{\alpha}) = F(f_0, \dots, \hat{f}_{\alpha}; x_0, \dots, \hat{x}_{\alpha}),$$

$$M(x_0, \dots, \hat{x}_{\alpha}) = H\left(\Gamma\left(G(x_0, \dots, \hat{x}_{\alpha}), \{x_0, \dots, \hat{x}_{\alpha}\}\right)\right).$$

Then $N(\alpha, \nu) = G(f_0(\nu), \dots, \hat{f}_{\alpha}(\nu))$.

Let α_0 be fixed such that $\alpha_0 \in \{\beta_{\nu} : \nu \in N\}$. Choose any ν with $\beta_{\nu} = \alpha_0$.

Then $N(\beta_{\nu}, \nu) = G(f_0(\nu), \dots, \hat{f}_{\alpha_0}(\nu))$.

Hence $N(\beta_{\nu}, \nu)$ is determined if the vector $s = (f_0(\nu), \dots, \hat{f}_{\alpha_0}(\nu))$ is known. Denote by S the set of all such vectors s , i.e. the set of all s which correspond to choices of ν such that $\beta_{\nu} = \alpha_0$. Let now $s = (s_0, \dots, \hat{s}_{\alpha_0}) \in S$; $\alpha < \alpha_0$. We proceed to show that $s_{\alpha} \in M(s_0, \dots, \hat{s}_{\alpha})$.

† We remind the reader that $|\{N^*(\beta_{\nu}, \nu) : \nu \in N\}|$ denotes the number of distinct sets $N^*(\beta_{\nu}, \nu)$.

We can choose $\nu \in N$ such that $\beta_\nu = \alpha_0$, and $s_\beta = f_\beta(\nu)$ for all $\beta < \alpha_0$. Then

$$\begin{aligned} s_\alpha &= f_\alpha(\nu) \in H(N^*(\alpha, \nu)) = H(\Gamma(N(\alpha, \nu), K(\alpha, \nu))) \\ &= H\left(\Gamma\left(G(f_0(\nu), \dots, \hat{f}_\alpha(\nu)), \{f_0(\nu), \dots, \hat{f}_\alpha(\nu)\}\right)\right) \\ &= M(s_0, \dots, \hat{s}_\alpha). \end{aligned}$$

In addition, we have

$$|M(s_0, \dots, \hat{s}_\alpha)| = |H(N^*(\alpha, \nu))| \leq b |N^*(\alpha, \nu)| \leq ba.$$

Hence, by the ramification lemma, when ν ranges through all values for which β_ν has the fixed value α_0 , there arise at most $(ba)^{|\alpha_0|}$ distinct vectors $(f_0(\nu), \dots, \hat{f}_{\alpha_0}(\nu))$. We deduce that

$$\begin{aligned} |N| &\leq a |\{N^*(\beta_\nu, \nu) : \nu \in N\}| \leq a |\{N(\beta_\nu, \nu) : \nu \in N\}| \\ &= a \Sigma(|\alpha_0| \leq b) |\{N(\beta_\nu, \nu) : \nu \in N; \beta_\nu = \alpha_0\}| \\ &\leq a \Sigma(|\alpha_0| \leq b) (ba)^{|\alpha_0|} \leq a(ba)^b b^+, \end{aligned}$$

which proves (3) and so establishes I(i).

Part (ii) of Theorem I follows from (i). For if $a \geq 2$; $b \geq 1$; $a+b \geq \aleph_0$; then $b^+ b^b a^{b+1} = a^b$.

Proof of Theorem II. Choose sets A, B such that $|A| = a$; $|B| = b$, and let F be the set of all mappings of B into A . Consider the system

$$\Sigma : X(t, f) = \{x, f(x) : x \in B\} \quad (t \in A; f \in F).$$

We consider the members of Σ as indexed by the pairs (t, f) . In fact, they do not depend on t . Then Σ is a (a^{b+1}, b) -system. Let us assume that Σ contains a $\Delta(> a)$ -system Σ' with kernel K , say the system

$$\Sigma' : X_\rho = X(t_\rho, f_\rho) \quad (\rho \in R).$$

Then $|R| > a$ and

$$(t_\rho, f_\rho) \neq (t_\sigma, f_\sigma) \quad \text{for } \{\rho, \sigma\} \neq \emptyset \subset R. \tag{5}$$

Let $x \in B$. Then $|\{f_\rho(x) : \rho \in R\}| \leq |A| = a < |R|$, and hence there is $\{\rho_x, \sigma_x\} \neq \emptyset \subset R$ with $f_{\rho_x}(x) = f_{\sigma_x}(x)$. Then, for any $\rho \in R$,

$$(x, f_{\rho_x}(x)) \in X_{\rho_x} X_{\sigma_x} = K \subset X_\rho = \{(y, f_\rho(y)) : y \in B\},$$

$f_{\rho_x}(x) = f_\rho(x)$, so that f_ρ is independent of ρ . Since $|\{t_\rho : \rho \in R\}| \leq |A| < |R|$ there is $\{\rho_1, \sigma_1\} \neq \emptyset \subset R$ with $t_{\rho_1} = t_{\sigma_1}$. But then $(t_{\rho_1}, f_{\rho_1}) = (t_{\sigma_1}, f_{\sigma_1})$ which contradicts (5). This proves Theorem II.

Proof of Theorem III. Let $1 \leq a, b < \aleph_0$. By Theorem I there exists a least number d , where $d < \aleph_0$, such that every $(> d, \leq b)$ -system contains

a $\Delta(> a)$ -system. Denote this number by $f(a, b)$. We have to show that

$$f(a, b) \leq c,$$

where c is defined by (1).

There is a least number $\phi(a, b)$ such that every $(> \phi, \leq b)$ -system

$$\Sigma : X_\mu \quad (\mu \in M)$$

which satisfies $X_\mu \neq X_\nu$ for $\{\mu, \nu\} \neq \emptyset \subset M$, contains a $\Delta(> a)$ -system. Clearly, $\phi \leq f$. Also, $\phi(a, 1) = a$. We first show that

$$f(a, b) \leq a\phi(a, b). \quad (6)$$

Let

$$\Sigma' : X_\nu \quad (\nu \in N)$$

be a $(> a\phi(a, b), \leq b)$ -system which does not contain any $\Delta(> a)$ -system.

We have to deduce a contradiction. Let

$$\nu_0 \in N; K(\nu_0) = \{\nu : \nu \in N; X_\nu = X_{\nu_0}\}.$$

Then X_ν ($\nu \in K(\nu_0)$) is a Δ -system, and therefore $|K(\nu_0)| \leq a$. Hence, if

$$\{X_\nu : \nu \in N\} = \{X_\mu : \mu \in M\}, \quad X_\mu \neq X_\nu \text{ for } \{\mu, \nu\} \neq \emptyset \subset M,$$

then $|M| > \phi(a, b)$, and it follows from the definition of ϕ that the system X_μ ($\mu \in M$) contains a $\Delta(> a)$ -system. This is the required contradiction.

There is a $(\phi(a, b), \leq b)$ -system

$$\Sigma : X_\nu \quad (\nu \in N),$$

where $X_\mu \neq X_\nu$ for $\mu \neq \nu$, which does not contain any $\Delta(> a)$ -system.

Let N_0 be a maximal subset of N such that $X_\mu X_\nu = \emptyset$ for $\{\mu, \nu\} \neq \emptyset \subset N_0$.

Then $|N_0| \leq a$, since X_ν ($\nu \in N_0$) is a Δ -system. Put $X^* = \Sigma(\nu \in N_0) X_\nu$.

Then we can choose elements

$$x_\mu \in X_\mu X^* \quad (\mu \in N - N_0).$$

Let $\xi \in X^*$. Then there is $\nu_0(\xi) \in N_0$ with $\xi \in X_{\nu_0(\xi)}$. Then the system (of sets of at most $b-1$ elements)

$$X_{\nu_0(\xi)} - \{\xi\}; X_\mu - \{\xi\} \quad (\mu \in N - N_0; x_\mu = \xi)$$

does not contain any $\Delta(> a)$ -system since any such system Σ'' would yield a $\Delta(> a)$ -system contained in Σ if we add to each member of Σ'' the element ξ .

Hence $1 + |\{\mu : \mu \in N - N_0; x_\mu = \xi\}| \leq \phi(a, b-1)$,

$$\begin{aligned} \phi(a, b) &= |N| = |N_0| + |N - N_0| \leq a + \Sigma(\xi \in X^*) |\{\mu : \mu \in N - N_0; x_\mu = \xi\}| \\ &\leq a + (\phi(a, b-1) - 1)ba = -a(b-1) + ab\phi(a, b-1), \end{aligned}$$

$$\frac{\phi(a, b)}{b! a^b} \leq -\frac{b-1}{b! a^{b-1}} + \frac{\phi(a, b-1)}{(b-1)! a^{b-1}}.$$

By means of $b-1$ successive applications of this inequality we obtain

$$\frac{\phi(a, b)}{b! a^b} \leq -\frac{b-1}{b! a^{b-1}} - \frac{b-2}{(b-1)! a^{b-2}} - \dots - \frac{1}{2! a} + \frac{\phi(a, 1)}{1! a}.$$

In view of (6) and $\phi(a, 1) = a$ this is the desired result.

The University,
Reading.