Acta Sci. Math., 55 (1991), 53—58

On a problem posed by I. Z. Ruzsa

RICHARD WARLIMONT

In his paper titled "On the small sieve II. Sifting by composite numbers" (*Journal of Number Theory*, **14** (1982), 260—268) I. Z. RUZSA posed the following problem:

For $a \in \mathbf{N}$ and $b \in \mathbf{Z}$ let $R(a, b)$ denote the residue class b mod a. Consider all systems $a_1, \ldots, a_m$ (m not fixed) of natural numbers $1 \leqq a_1 < \ldots < a_m \leqq n$ for which there exist integers $b_1, \ldots, b_m$ such that

$$(*) \qquad \bigcup_{j=1}^{m} R(a_j, b_j) \supset \{1, \ldots, n\}.$$

Put

$$\mu(n) := \min \sum_{j=1}^{m} \frac{1}{a_j}$$

where the minimum has to be taken over all those systems. What can be said abou the behaviour of $\mu(n)$ for $n \to \infty$?

Since $(*)$ implies

$$2n \sum_{j=1}^{m} \frac{1}{a_j} \geqq \sum_{j=1}^{m} \left(\left[\frac{n}{a_j} \right] + 1 \right) \geqq n$$

the lower estimate $\mu(n) \geqq \frac{1}{2}$ follows at once. Ruzsa mentions that he can improve it to

$$\mu(n) \geqq \log \frac{2^5 3^6}{5^2 23^2} = 0.567\,545\,38\ldots$$

and he also gives the upper estimate

$$\mu(n) \leqq \log \frac{5}{2} + O\left(\frac{1}{n}\right).$$

Received October 11, 1988 and in revised form May 29, 1990.

Ruzsa's problem appears to be very delicate but it gives rise to another one which can be solved: Denote by $\mathscr{A}(n)$ the family of all subsets $A\subset\{1, \ldots, n\}$ with the property

$$\sum_{a\in A}\left(\left[\frac{n}{a}\right]+1\right) \geqq n$$

and put

$$\nu(n) := \min_{A\in\mathscr{A}(n)} \sum_{a\in A} \frac{1}{a}.$$

Obviously $\nu(n)\leqq\mu(n)$. I could show

$$\nu(n) = \log\frac{2^5\cdot 3^6}{23^3}+O(n^{-1/3}).$$

Ruzsa simplified my proof. His modifications also led to the better error term $O(1/n)$. With his kind permission I present this simpler version.

Let $Y(n)$ denote the set of all $y=(y_1, \ldots, y_n)\in\mathbf{R}^n$ which fulfil

$$0 \leqq y_j \leqq 1 \ \ (1 \leqq j \leqq n) \quad \text{and} \quad \sum_{j=1}^{n} y_j\left(\left[\frac{n}{j}\right]+1\right) \geqq n.$$

Put

$$\nu^*(n) := \min_{y\in Y(n)} \sum_{j=1}^{n} y_j\frac{1}{j}.$$

Obviously $\nu^*(n)\leqq\nu(n)$. It will be shown that

(1) $$\nu^*(n) = \log\frac{2^5\cdot 3^6}{23^3}+O\left(\frac{1}{n}\right)$$

and

(2) $$\nu(n) \leqq \nu^*(n)+O\left(\frac{1}{n}\right).$$

If we put $\beta_j := \frac{j}{n}\left(\left[\frac{n}{j}\right]+1\right)$ and denote by $Z(n)$ the set of all $z=(z_1, \ldots, z_n)\in\mathbf{R}^n$ which fulfil

$$0 \leqq z_j \leqq \frac{1}{j} \ \ (1 \leqq j \leqq n) \quad \text{and} \quad \sum_{j=1}^{n} z_j\beta_j \geqq 1$$

then

$$\nu^*(n) = \min_{z\in Z(n)} \sum_{j=1}^{n} z_j.$$

Let $(\xi_1, \ldots, \xi_n) \in Z(n)$ be such that $v^*(n) = \sum_{j=1}^{n} \xi_j$. Then

$$0 \leqq \xi_j \leqq \frac{1}{j} \quad (1 \leqq j \leqq n) \quad \text{and} \quad \sum_{j=1}^{n} \xi_j \beta_j = 1.$$

The quantity

$$\gamma = \gamma(n) := \min_{\xi_j > 0} \beta_j$$

turns out to be crucial in the argumentation. One has

(3) $$\beta_j < \gamma \Rightarrow \xi_j = 0,$$

(4) $$\beta_j > \gamma \Rightarrow \xi_j = \frac{1}{j}.$$

Here (3) is evident and (4) can be seen this way: Let m, $1 \leqq m \leqq n$, be such that $\beta_m = \gamma$. Then $\xi_m > 0$. If there existed some k, $1 \leqq k \leqq n$, with $\beta_k > \gamma$ and $\xi_k < \frac{1}{k}$ then $k \neq m$ and

$$\varepsilon := \min \left\{ \frac{\xi_m}{\beta_k}, \frac{1}{\beta_m} \left(\frac{1}{k} - \xi_k \right) \right\} > 0.$$

Now put

$$\xi'_m := \xi_m - \varepsilon \beta_k, \quad \xi'_k := \xi_k + \varepsilon \beta_m, \quad \xi'_j := \xi_j \quad \text{for} \quad j \neq m, k.$$

Since $0 \leqq \xi'_j \leqq \frac{1}{j}$ and $\sum_{j=1}^{n} \xi'_j \beta_j = \sum_{j=1}^{n} \xi_j \beta_j = 1$, we have $(\xi'_1, \ldots, \xi'_n) \in Z(n)$. This implies

$$v^*(n) \leqq \sum_{j=1}^{n} \xi'_j = v^*(n) - \varepsilon(\beta_k - \gamma) < v^*(n)$$

which is absurd.

Now put $\delta = \delta(n) := \gamma(n) - 1$. From $\beta_j \leqq 2$ and

$$\beta_j - 1 = \frac{j}{n} \left(1 - \left(\frac{n}{j} - \left[\frac{n}{j} \right] \right) \right) \geqq \frac{j}{n} \left(1 - \frac{j-1}{j} \right) = \frac{1}{n}$$

we see that in particular $\frac{1}{n} \leqq \delta \leqq 1$. If k is an integer, $1 \leqq k \leqq n$, then

$$\beta_j = \frac{j}{n}(k+1) \quad \text{for} \quad \frac{n}{k+1} < j \leqq \frac{n}{k}.$$

We shall show next that

(5) $$\delta(n) \geqq \frac{1}{2500} =: \vartheta \quad \text{for all } n.$$

We put $x := \min\left\{\frac{1}{\delta}, \sqrt{n}\right\}$ and have

$$1 = \sum_{j=1}^{n} \xi_j \beta_j \geqq \sum_{k<x} \sum_{n\gamma/(k+1)<j\leqq n/k} \xi_j \beta_j.$$

Since in the inner sum $\beta_j > \frac{\gamma}{k+1}(k+1) = \gamma$. We infer from (4) that

$$1 \geqq \sum_{k<x} \sum_{n\gamma/(k+1)<j\leqq n/k} \frac{1}{j}\frac{j}{n}(k+1) = \frac{1}{n}\sum_{k<x}(k+1)\left(\left[\frac{n}{k}\right]-\left[\frac{n\gamma}{k+1}\right]\right) \geqq$$

$$\geqq \frac{1}{n}\sum_{k<x}(k+1)\left(\frac{n}{k}-\frac{n\gamma}{k+1}-1\right) = \sum_{k<x}\frac{1}{k} - \delta\sum_{k<x}1 - \frac{1}{n}\sum_{k<x}(k+1) \geqq \sum_{k<x}\frac{1}{k}-3.$$

Therefore $\sum_{k<x}\frac{1}{k} \leqq 4$ which implies $x \leqq 50$.

If $x=\sqrt{n}$ then $\sqrt{n} \leqq 50$ and therefore $\delta \geqq \frac{1}{n} \geqq \frac{1}{2500}$. If $x = \frac{1}{\delta}$ then $\frac{1}{\delta} \leqq 50$, and therefore $\delta \geqq \frac{1}{50}$.

Now comes

$$\delta(n) = \frac{5}{18} + O\left(\frac{1}{n}\right). \tag{6}$$

To establish that we start from

$$1 = \sum_{j=1}^{n}\xi_j\beta_j = \sum_{k<1/\delta}\Big(\sum_{n/(k+1)<j\leqq n\gamma/(k+1)}\xi_j\beta_j + \sum_{n\gamma/(k+1)<j\leqq n/k}\xi_j\beta_j\Big) +$$

$$+\sum_{1/\delta\leqq k\leqq n}\sum_{n/(k+1)<j\leqq n/k}\xi_j\beta_j = S_1+S_2+S_3.$$

We estimate S_3. If $j<\frac{n}{k}$ then $\beta_j < 1+\frac{1}{k} \leqq 1+\delta = \gamma$. If $k > \frac{1}{\delta}$ then $\beta_j \leqq 1+\frac{1}{k} < 1+\delta = \gamma$. Thus by (3) there is at most one term ξ_i, β_i, $i = \frac{n}{1/\delta}$, in S_3 which may not vanish. Therefore, by (5),

$$S_3 = \xi_i\beta_i \leqq 2\frac{1}{i} = \frac{2}{\delta n} \leqq \frac{2}{\vartheta n}$$

We estimate S_1. If $j < \frac{n\gamma}{k+1}$ then $\beta_j < \frac{\gamma}{k+1}(k+1) = \gamma$. Thus by (3) there is at most one term $\xi_i\beta_i$ in the innermost sum which may not vanish and

$$\xi_i\beta_i \leqq \frac{1}{i}\frac{i}{n}(k+1) = \frac{k+1}{n}.$$

Therefore, by (5),

$$S_1 \leqq \sum_{k<1/\delta} \frac{k+1}{n} \leqq \frac{2}{\delta^2 n} \leqq \frac{2}{\vartheta^2 n}.$$

Finally we evaluate S_2. In S_2 we have $\beta_j > \frac{\gamma}{k+1}(k+1) = \gamma$ which by (4) implies $\xi_j = \frac{1}{j}$. Therefore

$$S_2 = \frac{1}{n}\sum_{k<1/\delta}(k+1)\left(\left[\frac{n}{k}\right]-\left[\frac{n\gamma}{k+1}\right]\right) =$$

$$= \frac{1}{n}\sum_{k<1/\delta}(k+1)\left(\frac{n}{k}-\frac{n\gamma}{k+1}+r\right) \text{ where } |r| \leqq 1$$

$$= \sum_{k<1/\delta}\left(\frac{1}{k}-\delta\right)+R \text{ where } |R| \leqq \frac{2}{\vartheta^2 n}.$$

Thus we obtain

$$1-\frac{6}{\vartheta^2 n} \leqq \sum_{k<1/\delta}\left(\frac{1}{k}-\delta\right) \leqq 1+\frac{2}{\vartheta^2 n}.$$

If we put

$$f(t) := \sum_{k<1/t}\left(\frac{1}{k}-t\right) \quad \text{for} \quad 0 < t \leqq 1$$

then one easily verifies that the following inequality holds true:

$$|t_1-t_2| \leqq |f(t_1)-f(t_2)| \quad \text{for} \quad 0 < t_1, \quad t_2 \leqq 1.$$

Since $f\left(\frac{5}{18}\right) = 1$ we obtain

$$\left|\delta-\frac{5}{18}\right| \leqq |f(\delta)-1| \leqq \frac{6}{\vartheta^2 n}.$$

Proof of (1).

$$\nu^*(n) = \sum_{j=1}^{n}\xi_j = \sum_{k<1/\delta}\Big(\sum_{n/(k+1)<j\leqq n\gamma/(k+1)}\xi_j+\sum_{n\gamma/(k+1)<j\leqq n/k}\xi_j\Big)+$$

$$+\sum_{1/\delta\leqq k\leqq n}\ \sum_{n/(k+1)<j\leqq n/k}\xi_j = T_1+T_2+T_3.$$

Because of $\beta_j>1$ we have

$$T_1+T_3 \leqq S_1+S_3 \ll \frac{1}{n}.$$

Further for $n \geqq n_0$ by (6) one has

$$T_2 = \sum_{k=1}^{3} \sum_{n\gamma/(k+1)<j\leqq n/k} \frac{1}{j} = \sum_{k=1}^{3} \left(\log \frac{k+1}{\gamma k} + O\left(\frac{1}{n}\right) \right) = \log \frac{4}{\gamma^3} + O\left(\frac{1}{n}\right).$$

But (6) implies

$$\gamma^3 = (1+\delta)^3 = \left(\frac{23}{18}\right)^3 + O\left(\frac{1}{n}\right).$$

Proof of (2). Since $\frac{1}{\delta}$ is no integer we may write $k > \frac{1}{\delta}$ in T_3. Therefore $\xi_j = 0$ in T_3. In T_1 there are at most 3 terms with $\xi_j < \frac{1}{j}$. These are replaced by $\frac{1}{j} \leqq \frac{4}{n}$. If we denote the new system by ξ'_j then we have

$$\xi'_j = 0 \quad \text{or} \quad \frac{1}{j} \quad (1 \leqq j \leqq n) \quad \text{and} \quad \sum_{j=1}^{n} \xi'_j \beta_j \geqq \sum_{j=1}^{n} \xi_j \beta_j \geqq 1.$$

Therefore

$$v(n) \leqq \sum_{j=1}^{n} \xi'_j \leqq \sum_{j=1}^{n} \xi_j + \frac{12}{n} = v^*(n) + \frac{12}{n}.$$

UNIVERSITÄT REGENSBURG
INSTITUT FÜR MATHEMATIK
D—8400 REGENSBURG
UNIVERSITÄTSSTRASSE 31