Generalized multiplication tables

Dimitris Koukoulopoulos

University of Illinois at Urbana - Champaign

June 4, 2009

Consider the multiplication table

	1	2	3	4	$\cdots$	N
1	1	2	3	4	$\cdots$	N
2	2	4	6	8	$\cdots$	$2N$
3	3	6	9	12	$\cdots$	$3N$
4	4	8	12	16	$\cdots$	$4N$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
N	N	$2N$	$3N$	$4N$	$\cdots$	N^2

Question (Erdős, 1955)

How many distinct integers does this table contain?

We may ask the same question for products of 3 integers, in which case we have a multiplication ‘box’, or 4 integers, and so on. To this end, define

$$A_{k+1}(N) := |\{n_1 \cdots n_{k+1} : n_i \leq N \ (1 \leq i \leq k+1)\}|.$$

Then the problem is equivalent to estimating $A_{k+1}(N)$.

We may ask the same question for products of 3 integers, in which case we have a multiplication 'box', or 4 integers, and so on. To this end, define

$$A_{k+1}(N) := |\{n_1 \cdots n_{k+1} : n_i \leq N \ (1 \leq i \leq k+1)\}|.$$

Then the problem is equivalent to estimating $A_{k+1}(N)$.

The key to understanding the combinatorics of the multiplication table is the function

$$H_{k+1}(x, \boldsymbol{y}, \boldsymbol{z}) := |\{n \leq x : \exists d_1 \cdots d_k | n \text{ such that } y_i < d_i \leq z_i \ (1 \leq i \leq k)\}|,$$

where $\boldsymbol{y} = (y_1, \ldots, y_k)$ and $\boldsymbol{z} = (z_1, \ldots, z_k)$.

The transition from $H_{k+1}(x, \boldsymbol{y}, \boldsymbol{z})$ to $A_{k+1}(N)$ is achieved via the elementary inequalities

$$H_{k+1}\left(\frac{N^{k+1}}{2^k}, \left(\frac{N}{2}, \dots, \frac{N}{2}\right), (N, \dots, N)\right) \leq A_{k+1}(N)$$

$$\leq \sum_{\substack{2^{m_i} \leq \sqrt{N} \\ 1 \leq i \leq k}} H_{k+1}\left(\frac{N^{k+1}}{2^{m_1+\cdots+m_k}}, \left(\frac{N}{2^{m_1+1}}, \dots, \frac{N}{2^{m_k+1}}\right), \left(\frac{N}{2^{m_1}}, \dots, \frac{N}{2^{m_k}}\right)\right)$$

$$+ N^{k+1/2}.$$

The transition from $H_{k+1}(x, \boldsymbol{y}, \boldsymbol{z})$ to $A_{k+1}(N)$ is achieved via the elementary inequalities

$$H_{k+1}\left(\frac{N^{k+1}}{2^k}, \left(\frac{N}{2}, \dots, \frac{N}{2}\right), (N, \dots, N)\right) \leq A_{k+1}(N)$$
$$\leq \sum_{\substack{2^{m_i} \leq \sqrt{N} \\ 1 \leq i \leq k}} H_{k+1}\left(\frac{N^{k+1}}{2^{m_1+\cdots+m_k}}, \left(\frac{N}{2^{m_1+1}}, \dots, \frac{N}{2^{m_k+1}}\right), \left(\frac{N}{2^{m_1}}, \dots, \frac{N}{2^{m_k}}\right)\right)$$
$$+ N^{k+1/2}.$$

Note that it suffices to study $H_{k+1}(x, \boldsymbol{y}, \boldsymbol{z})$ when

- $\boldsymbol{z} = 2\boldsymbol{y}$;
- the numbers $\log y_1, \dots, \log y_k$ all have the same order of magnitude.

Set $Q(u) := u \log u - u + 1$.

Set $Q(u) := u \log u - u + 1$. Then the following estimate holds.

Theorem (Ford (2004), K. (2008))

Let $k \geq 1$, $0 < \delta \leq 1$ and $c \geq 1$. Assume that $x \geq 1$ and $3 \leq y_1 \leq y_2 \leq \cdots \leq y_k \leq y_1^c$ with $\frac{x}{y_1 \cdots y_k} \geq \max\{2^k, y_1^\delta\}$. Then

$$H_{k+1}(x, \mathbf{y}, 2\mathbf{y}) \asymp_{k,\delta,c} \frac{x}{(\log y_1)^{Q(\frac{k}{\log(k+1)})}(\log\log y_1)^{3/2}}.$$

Set $Q(u) := u \log u - u + 1$. Then the following estimate holds.

Theorem (Ford (2004), K. (2008))

Let $k \geq 1$, $0 < \delta \leq 1$ and $c \geq 1$. Assume that $x \geq 1$ and $3 \leq y_1 \leq y_2 \leq \cdots \leq y_k \leq y_1^c$ with $\frac{x}{y_1 \cdots y_k} \geq \max\{2^k, y_1^\delta\}$. Then

$$H_{k+1}(x, \mathbf{y}, 2\mathbf{y}) \asymp_{k,\delta,c} \frac{x}{(\log y_1)^{Q(\frac{k}{\log(k+1)})} (\log \log y_1)^{3/2}}.$$

Consequently

Corollary

$$A_{k+1}(N) \asymp_k \frac{N^{k+1}}{(\log N)^{Q(\frac{k}{\log(k+1)})} (\log \log N)^{3/2}} \quad (N \geq 3).$$

The following heuristic was first given by Ford when $k = 1$.

The following heuristic was first given by Ford when $k = 1$.

Assume that $\log y_1, \ldots, \log y_k$ have the same order of magnitude. For $n \in \mathbb{N}$ write $n = ab$, where

$$a = \prod_{p^e \| n, p \leq 2y_1} p^e.$$

Assume that $\mu^2(a) = 1$ and $a \leq y_1^C$.

The following heuristic was first given by Ford when $k = 1$.

Assume that $\log y_1, \ldots, \log y_k$ have the same order of magnitude. For $n \in \mathbb{N}$ write $n = ab$, where

$$a = \prod_{p^e \| n, p \leq 2y_1} p^e.$$

Assume that $\mu^2(a) = 1$ and $a \leq y_1^C$.

Consider the set

$$D_{k+1}(a) = \{(\log d_1, \ldots, \log d_k) : d_1 \cdots d_k | a\}.$$

The following heuristic was first given by Ford when $k = 1$.

Assume that $\log y_1, \ldots, \log y_k$ have the same order of magnitude. For $n \in \mathbb{N}$ write $n = ab$, where

$$a = \prod_{p^e \| n, p \le 2y_1} p^e.$$

Assume that $\mu^2(a) = 1$ and $a \le y_1^C$.

Consider the set

$$D_{k+1}(a) = \{(\log d_1, \ldots, \log d_k) : d_1 \cdots d_k | a\}.$$

Main assumption: $D_{k+1}(a)$ is well-distributed in $[0, \log a]^k$.

Then we would expect that

$$\begin{aligned}\tau_{k+1}(a,\boldsymbol{y},2\boldsymbol{y}) &:= |\{d_1\cdots d_k|a : y_i < d_i \le 2y_i \ (1\le i\le k)\}| \\ &= \left| D_{k+1}(a) \cap \prod_{i=1}^{k}(\log y_i, \log y_i + \log 2] \right| \\ &\approx |D_{k+1}(a)| \frac{(\log 2)^k}{(\log a)^k} \approx \frac{(k+1)^{\omega(a)}}{(\log y_1)^k}.\end{aligned}$$

Then we would expect that

$$\begin{aligned}\tau_{k+1}(a, \boldsymbol{y}, 2\boldsymbol{y}) &:= |\{d_1 \cdots d_k | a : y_i < d_i \le 2y_i \ (1 \le i \le k)\}| \\ &= \left| D_{k+1}(a) \cap \prod_{i=1}^{k} (\log y_i, \log y_i + \log 2] \right| \\ &\approx |D_{k+1}(a)| \frac{(\log 2)^k}{(\log a)^k} \approx \frac{(k+1)^{\omega(a)}}{(\log y_1)^k}.\end{aligned}$$

This expression is ≥ 1 when

$$\omega(a) \ge m := \left\lfloor \frac{k}{\log(k+1)} \log\log y_1 \right\rfloor + O(1).$$

Then we would expect that

$$\begin{aligned}\tau_{k+1}(a, \boldsymbol{y}, 2\boldsymbol{y}) &:= |\{d_1 \cdots d_k | a : y_i < d_i \le 2y_i \ (1 \le i \le k)\}| \\ &= \left| D_{k+1}(a) \cap \prod_{i=1}^{k} (\log y_i, \log y_i + \log 2] \right| \\ &\approx |D_{k+1}(a)| \frac{(\log 2)^k}{(\log a)^k} \approx \frac{(k+1)^{\omega(a)}}{(\log y_1)^k}.\end{aligned}$$

This expression is ≥ 1 when

$$\omega(a) \ge m := \left\lfloor \frac{k}{\log(k+1)} \log\log y_1 \right\rfloor + O(1).$$

We have

$$|\{n \le x : \omega(a) = r\}| \approx \frac{x}{\log y_1} \frac{(\log\log y_1)^r}{r!}.$$

Therefore

$$H_{k+1}(x,\mathbf{y},2\mathbf{y}) \approx \frac{x}{\log y_1}\sum_{r\geq m}\frac{(\log\log y_1)^r}{r!}$$
$$\asymp \frac{x}{(\log y_1)^{Q(\frac{k}{\log(k+1)})}(\log\log y_1)^{1/2}}.$$

Therefore

$$H_{k+1}(x, \mathbf{y}, 2\mathbf{y}) \approx \frac{x}{\log y_1} \sum_{r \geq m} \frac{(\log\log y_1)^r}{r!}$$
$$\asymp \frac{x}{(\log y_1)^{Q(\frac{k}{\log(k+1)})} (\log\log y_1)^{1/2}}.$$

This is slightly too big. The problem arises from the fact that $D_{k+1}(a)$ is usually not well-distributed, but instead it has many clumps.

Therefore

$$H_{k+1}(x, \mathbf{y}, 2\mathbf{y}) \approx \frac{x}{\log y_1} \sum_{r \geq m} \frac{(\log\log y_1)^r}{r!}$$
$$\asymp \frac{x}{(\log y_1)^{Q(\frac{k}{\log(k+1)})} (\log\log y_1)^{1/2}}.$$

This is slightly too big. The problem arises from the fact that $D_{k+1}(a)$ is usually not well-distributed, but instead it has many clumps.

To see this define

$$L_{k+1}(a) := \mathrm{Vol}\Big(\bigcup_{d_1 \cdots d_k | a} [\log(d_1/2), \log d_1) \times \cdots \times [\log(d_k/2), \log d_k) \Big),$$

which is a quantitative measure of how well-distributed $D_{k+1}(a)$ is.

If $a = p_1 \cdots p_m$, where $p_1 < \cdots < p_m \le 2y_1$, then we expect that

$$\log\log p_j \sim \frac{j}{m} \log\log y_1 = j \frac{\log(k+1)}{k} + O(1).$$

If $a = p_1 \cdots p_m$, where $p_1 < \cdots < p_m \le 2y_1$, then we expect that

$$\log\log p_j \sim \frac{j}{m} \log\log y_1 = j\frac{\log(k+1)}{k} + O(1).$$

But with probability tending to 1 there is a j so that

$$\log\log p_j \le j\frac{\log(k+1)}{k} - (\log\log y_1)^{1/3},$$

If $a = p_1 \cdots p_m$, where $p_1 < \cdots < p_m \le 2y_1$, then we expect that

$$\log\log p_j \sim \frac{j}{m} \log\log y_1 = j\frac{\log(k+1)}{k} + O(1).$$

But with probability tending to 1 there is a j so that

$$\log\log p_j \le j\frac{\log(k+1)}{k} - (\log\log y_1)^{1/3},$$

which implies that

$$\begin{aligned} L_{k+1}(a) &\le \tau_{k+1}(p_{j+1}\cdots p_m) L_{k+1}(p_1\cdots p_j) \\ &\lesssim (k+1)^m \exp\{-(\log\log y_1)^{1/3}\}. \end{aligned}$$

If $a = p_1 \cdots p_m$, where $p_1 < \cdots < p_m \le 2y_1$, then we expect that

$$\log\log p_j \sim \frac{j}{m}\log\log y_1 = j\frac{\log(k+1)}{k} + O(1).$$

But with probability tending to 1 there is a j so that

$$\log\log p_j \le j\frac{\log(k+1)}{k} - (\log\log y_1)^{1/3},$$

which implies that

$$\begin{aligned} L_{k+1}(a) &\le \tau_{k+1}(p_{j+1}\cdots p_m)L_{k+1}(p_1\cdots p_j) \\ &\lesssim (k+1)^m \exp\{-(\log\log y_1)^{1/3}\}. \end{aligned}$$

This is much less than $\tau_{k+1}(a) = (k+1)^m$ and so most of the time $D_{k+1}(a)$ contains large clusters of points.

We must focus on abnormal a's for which

$$\log\log p_j \geq j\frac{\log(k+1)}{k} - O(1) \quad (1 \leq j \leq m).$$

We must focus on abnormal *a*'s for which

$$\log\log p_j \geq j\frac{\log(k+1)}{k} - O(1) \quad (1 \leq j \leq m).$$

The probability that *a* has this property is about

$$\frac{1}{m} \asymp \frac{1}{\log\log y_1} \quad \text{(Ford)}.$$

We must focus on abnormal *a*'s for which

$$\log\log p_j \geq j\frac{\log(k+1)}{k} - O(1) \quad (1 \leq j \leq m).$$

The probability that *a* has this property is about

$$\frac{1}{m} \asymp \frac{1}{\log\log y_1} \quad \text{(Ford)}.$$

This leads to the refined heuristic estimate

$$H_{k+1}(x, \boldsymbol{y}, 2\boldsymbol{y}) \approx \frac{x}{(\log y_1)^{Q(\frac{1}{\log \rho})}(\log\log y_1)^{3/2}},$$

which is the correct one.

The natural generalization of the multiplication table problem is the estimation of

$$A_{k+1}(N_1, \ldots, N_{k+1}) := |\{n_1 \cdots n_{k+1} : n_i \le N_i \ (1 \le i \le k+1)\}|.$$

The natural generalization of the multiplication table problem is the estimation of

$$A_{k+1}(N_1, \ldots, N_{k+1}) := |\{n_1 \cdots n_{k+1} : n_i \leq N_i \ (1 \leq i \leq k+1)\}|.$$

As before, to understand this question we study $H_{k+1}(x, \mathbf{y}, 2\mathbf{y})$.

The difference is that we now drop the assumption that the numbers $\log y_1, \ldots, \log y_k$ are of the same order of magnitude.

The natural generalization of the multiplication table problem is the estimation of

$$A_{k+1}(N_1, \ldots, N_{k+1}) := |\{n_1 \cdots n_{k+1} : n_i \leq N_i \ (1 \leq i \leq k+1)\}|.$$

As before, to understand this question we study $H_{k+1}(x, \mathbf{y}, 2\mathbf{y})$.

The difference is that we now drop the assumption that the numbers $\log y_1, \ldots, \log y_k$ are of the same order of magnitude.

When $k = 1$, Ford's result immediately implies that

Corollary

Let $3 \leq N_1 \leq N_2$. Then

$$A_2(N_1, N_2) \asymp \frac{N_1 N_2}{(\log N_1)^{Q(\frac{1}{\log 2})} (\log \log N_1)^{3/2}}.$$

For $n \in \mathbb{N}$ write $n = a_1 \cdots a_k b$, where

$$a_i = \prod_{p^e \| n, 2y_{i-1} < p \le 2y_i} p^e.$$

Assume that $\mu^2(a_i) = 1$ and $a_i \le y_i^C$ for $1 \le i \le k$.

For $n \in \mathbb{N}$ write $n = a_1 \cdots a_k b$, where

$$a_i = \prod_{p^e \| n, 2y_{i-1} < p \le 2y_i} p^e.$$

Assume that $\mu^2(a_i) = 1$ and $a_i \le y_i^C$ for $1 \le i \le k$.

Set

$$D_{k+1}(\boldsymbol{a}) := \{(\log d_1, \dots, \log d_k) : d_1 \cdots d_i | a_1 \cdots a_i \ (1 \le i \le k)\}$$

and assume that $D_{k+1}(\boldsymbol{a})$ is well-distributed in $[0, \log y_1] \times \cdots \times [0, \log y_k]$.

Then

$$\begin{aligned}\tau(n, \boldsymbol{y}, 2\boldsymbol{y}) &= \{(d_1, \dots, d_k) : d_1 \cdots d_i | a_1 \cdots a_i, \\ &\qquad y_i < d_i \le 2y_i \ (1 \le i \le k)\} \\ &= \left| D_{k+1}(\boldsymbol{a}) \cap \prod_{i=1}^{k} (\log y_i, \log y_i + \log 2] \right| \\ &\approx \frac{\prod_{i=1}^{k} (k - i + 2)^{\omega(a_i)}}{\prod_{i=1}^{k} \log y_i}.\end{aligned}$$

Then

$$\begin{aligned}\tau(n, \boldsymbol{y}, 2\boldsymbol{y}) &= \{(d_1, \ldots, d_k) : d_1 \cdots d_i | a_1 \cdots a_i, \\ &\qquad\qquad y_i < d_i \leq 2y_i \ (1 \leq i \leq k)\} \\ &= \left| D_{k+1}(\boldsymbol{a}) \cap \prod_{i=1}^{k} (\log y_i, \log y_i + \log 2] \right| \\ &\approx \frac{\prod_{i=1}^{k} (k - i + 2)^{\omega(a_i)}}{\prod_{i=1}^{k} \log y_i}.\end{aligned}$$

Set $\ell_i := \log 3 \frac{\log y_i}{\log y_{i-1}}$ and

$$\mathscr{H} := \left\{ (r_1, \ldots, r_k) \in (\mathbb{N} \cup \{0\})^k : \sum_{i=1}^{k} r_i \log(k-i+2) \geq \sum_{i=1}^{k} \ell_i (k-i+1) \right\}.$$

Then

$$\tau(n, \boldsymbol{y}, 2\boldsymbol{y}) = \{(d_1, \dots, d_k) : d_1 \cdots d_i | a_1 \cdots a_i, \\ y_i < d_i \le 2y_i \ (1 \le i \le k)\} \\ = \left| D_{k+1}(\boldsymbol{a}) \cap \prod_{i=1}^{k} (\log y_i, \log y_i + \log 2] \right| \\ \approx \frac{\prod_{i=1}^{k} (k-i+2)^{\omega(a_i)}}{\prod_{i=1}^{k} \log y_i}.$$

Set $\ell_i := \log 3 \frac{\log y_i}{\log y_{i-1}}$ and

$$\mathscr{H} := \left\{ (r_1, \dots, r_k) \in (\mathbb{N} \cup \{0\})^k : \sum_{i=1}^{k} r_i \log(k-i+2) \ge \sum_{i=1}^{k} \ell_i (k-i+1) \right\}.$$

Then, heuristically, $\tau(n, \boldsymbol{y}, 2\boldsymbol{y}) \ge 1$ if-f $(\omega(a_1), \dots, \omega(a_k)) \in \mathscr{H}$.

We have that

$$|\{n \le x : \omega(a_i) = r_i \ (1 \le i \le k)\}| \approx \frac{x}{\log y_k} \prod_{i=1}^{k} \frac{\ell_i^{r_i}}{r_i!}.$$

We have that

$$|\{n \le x : \omega(a_i) = r_i \ (1 \le i \le k)\}| \approx \frac{x}{\log y_k} \prod_{i=1}^{k} \frac{\ell_i^{r_i}}{r_i!}.$$

So using Taylor's theorem and Lagrange multipliers we find that

$$\begin{aligned} H_{k+1}(x, \boldsymbol{y}, 2\boldsymbol{y}) &\approx \frac{x}{\log y_k} \sum_{\boldsymbol{r} \in \mathscr{H}} \prod_{i=1}^{k} \frac{\ell_i^{r_i}}{r_i!} \\ &\asymp \frac{x}{\sqrt{\log\log y_k} \prod_{i=1}^{k} \left(\frac{\log y_i}{\log y_{i-1}}\right)^{Q((k-i+2)^{\alpha})}}, \end{aligned}$$

where $\alpha = \alpha(k, \boldsymbol{y})$ satisfies

$$\sum_{i=1}^{k} (k-i+2)^{\alpha} \log(k-i+2)\ell_i = \sum_{i=1}^{k} (k-i+1)\ell_i.$$

Indeed, we may prove the following estimate.

Theorem (K, 2008)

Let $3 = y_0 \le y_1 \le \cdots \le y_k$ *with* $\frac{x}{y_1 \cdots y_k} \ge \max\{2^k, y_k\}$. *Then*

$$\frac{H_{k+1}(x, \mathbf{y}, 2\mathbf{y})}{x} \ll_k \frac{\min\left\{1, \frac{(\log\log 3y_{i_0-1})(\log 3\frac{\log y_k}{\log y_{i_0}})}{\ell_{i_0}}\right\}}{\sqrt{\log\log y_k}\prod_{i=1}^{k}\left(\frac{\log y_i}{\log y_{i-1}}\right)^{Q((k-i+2)^{\alpha})}},$$

where i_0 *is such that* $\ell_{i_0} = \max_{1 \le i \le k} \ell_i$.

When $k = 2$, this upper bound is the correct order of $H_3(x, \mathbf{y}, 2\mathbf{y})$.

Theorem (K, 2008)

Let $3 \le y_1 \le y_2$ *so that* $\frac{x}{y_1 y_2} \ge \max\{4, y_2\}$. *Then*

$$\frac{H_3(x, \mathbf{y}, 2\mathbf{y})}{x} \asymp \frac{(\log\log 3y_1)(\log 3\frac{\log y_2}{\log y_1})}{(\log\log y_2)^{5/2}(\log y_1)^{Q(3^\alpha)}\left(\frac{\log y_2}{\log y_1}\right)^{Q(2^\alpha)}},$$

where $\mathbf{y} = (y_1, y_2)$.

As a direct corollary we obtain the order of magnitude of the number of integers in a 3-dimensional multiplication table.

Corollary

For $3 \leq N_1 \leq N_2 \leq N_3$ we have that

$$\frac{A_3(N_1, N_2, N_3)}{N_1 N_2 N_3} \asymp \frac{(\log\log N_1)(\log 3\frac{\log N_2}{\log N_1})}{(\log\log N_2)^{5/2}(\log N_1)^{Q(3^\alpha)}\left(\frac{\log N_2}{\log N_1}\right)^{Q(2^\alpha)}}.$$

In general, we may reduce the counting in $H_{k+1}(x, \boldsymbol{y}, 2\boldsymbol{y})$ (local distribution of factorizations) to estimating averages of

$$L_{k+1}(\boldsymbol{a}) := \mathrm{Vol}\Big(\bigcup_{\substack{d_1 \cdots d_i | a_1 \cdots a_i \\ 1 \le i \le k}} [\log(d_1/2), \log d_1) \times \cdots [\log(d_k/2), \log d_k) \Big)$$

(global distribution of factorizations).

In general, we may reduce the counting in $H_{k+1}(x,\boldsymbol{y},2\boldsymbol{y})$ (local distribution of factorizations) to estimating averages of

$$L_{k+1}(\boldsymbol{a}) := \mathrm{Vol}\Big(\bigcup_{\substack{d_1\cdots d_i|a_1\cdots a_i\\ 1\le i\le k}} [\log(d_1/2),\log d_1)\times\cdots[\log(d_k/2),\log d_k)\Big)$$

(global distribution of factorizations).

Proposition (Ford (2004), K. (2008))

Let $k\ge 1$, $x\ge 1$ and $3\le y_1\le\cdots\le y_k$ with $\frac{x}{y_1\cdots y_k}\ge\max\{2^k,y_k\}$. Then

$$\frac{H_{k+1}(x,\boldsymbol{y},2\boldsymbol{y})}{x} \asymp_k \prod_{i=1}^{k}\Big(\frac{\log y_i}{\log y_{i-1}}\Big)^{-(k-i+2)} \sum_{\substack{a_i\in\mathscr{P}(y_{i-1},y_i)\\ 1\le i\le k}} \frac{L_{k+1}(\boldsymbol{a})}{a_1\cdots a_k}.$$

Thank you for your attention!