ACTA ARITHMETICA
162.3 (2014)

On the range of Carmichael's universal-exponent function

by

Florian Luca (Santiago de Querétaro and Johannesburg) and
Carl Pomerance (Hanover, NH)

In commemoration of the 100th anniversary of the birth of Paul Erdős

1. Introduction. Let λ denote the Carmichael function. For a natural number n, $\lambda(n)$ is the universal exponent modulo n. Equivalently, it is the largest multiplicative order of elements modulo n. Given the prime factorization $p_1^{\nu_1}\cdots p_k^{\nu_k}$ of $n>1$ we can explicitly give $\lambda(n)$ as follows:

$$\lambda(n)=\operatorname{lcm}[\lambda(p_1^{\nu_1}),\dots,\lambda(p_k^{\nu_k})],$$

where for any prime power $p^\nu>1$ one has

$$\lambda(p^\nu)=\begin{cases}p^{\nu-1}(p-1) & \text{if } p\ge 3 \text{ or } \nu\le 2,\\ 2^{\nu-2} & \text{if } p=2 \text{ and } \nu\ge 3.\end{cases}$$

We also have $\lambda(1)=1$. In this paper we shall be concerned with estimating the counting function for the distinct values of λ. Put

$$V_\lambda(x)=\#\{\lambda(n)\le x : n\ge 1\}.$$

Denote by $V_\varphi(x)$ the corresponding counting function for Euler-function values in $[1,x]$. It has been known now for almost 80 years that

$$V_\varphi(x)=x/(\log x)^{1+o(1)} \quad \text{as } x\to\infty$$

(Erdős [4]). Due to the similarity of λ and φ (they are almost identical on prime powers, and for every n, both $\lambda(n)$ and $\varphi(n)$ share the same prime factors), one might guess that their values are distributed in a similar fashion.

Here is what was known prior to this paper. Since $p-1=\lambda(p)$, it follows that $V_\lambda(x)\ge\pi(x+1)\ge(1+o(1))x/\log x$ as $x\to\infty$. In [2] a somewhat larger lower bound is found, but still of the shape $x/(\log x)^{1+o(1)}$ and similar to the lower bound for $V_\varphi(x)$ in [8] and [13]. It is a simple exercise to see that at least $V_\varphi(x)=o(x)$ as $x\to\infty$, but the corresponding result for $V_\lambda(x)$ is trickier.

2010 *Mathematics Subject Classification*: Primary 11N37.
Key words and phrases: Carmichael function.

DOI: 10.4064/aa162-3-6

In [6], a proof was outlined (using a result from [7]) that $V_\lambda(x) \ll x/(\log x)^{\kappa_0}$ for some constant $\kappa_0 > 0$. This was worked out explicitly in [9], namely κ_0 may be taken as any number smaller than $1 - (e \log 2)/2 = 0.0579153\ldots$.

In this paper, we prove the following bounds on $V_\lambda(x)$.

THEOREM 1.1. *As $x \to \infty$,*

$$V_\lambda(x) \le \frac{x}{(\log x)^{\eta + o(1)}},$$

where $\eta = 1 - (1 + \log\log 2)/\log 2 = .0860713\ldots$ is the Erdős–Tenenbaum–Ford constant.

THEOREM 1.2. *For all large values of x we have*

$$V_\lambda(x) \ge \frac{x}{(\log x)^{.359052}}.$$

In particular, Theorem 1.2 shows that $V_\lambda(x)$ is much larger than $V_\varphi(x)$.

We present a heuristic argument that the "correct" exponent on $\log x$ is that in Theorem 1.1, namely the Erdős–Tenenbaum–Ford constant.

It is perhaps also of interest to estimate $\#\{\lambda(n) : n \le x\}$. Here the count is closer to $V_\varphi(x)$. We prove the following result.

THEOREM 1.3. *As $x \to \infty$,*

$$\#\{\lambda(n) : n \le x\} = \frac{x}{(\log x)^{1+o(1)}}.$$

In an Appendix we present several algorithms for computing $V_\lambda(x)$, discuss their complexity, and give some modest numerical data.

Our proof of Theorem 1.2 depends strongly on careful estimates involving the number of prime factors of numbers 1 less than a prime. This kind of thought has been present since the dawn of this subject in the 1935 paper of Erdős [4], but the argument here is considerably more difficult. It is perhaps of interest that our lower bound actually holds for the smaller set of numbers of the form $\lambda(pq)$, where p, q are primes. It may even sound wrong that there could be so many λ-values of this form, since the number of integers $pq \le x$ is $O(x \log\log x/\log x)$. However, the number of integers pq with $\lambda(pq) \le x$ can be seen to be of magnitude x. Since, in general, the number of integers n with $\lambda(n) \le x$ is an enormous function of x, growing faster than any fixed power of x (see [6, p. 384]), we not only see why it is plausible that there are so many more λ-values to x than there are φ-values, but we also see a possible path to improving Theorem 1.2 and closing the gap with Theorem 1.1. For more in this vein see the discussion at the end of Section 2.1.

In what follows, we use the Vinogradov symbols $\gg$, $\ll$ and $\asymp$ and the Landau symbols O and o with their usual meaning. We use p, q and r with or without subscripts for prime numbers. We write $\Omega(n)$ for the number of prime power divisors (> 1) of n. For a real number $z > 1$, we write $\Omega_z(n)$

for the number of prime powers $p^j \mid n$ with $j > 0$, $p \le z$. Let $P(n)$ denote the largest prime divisor of $n > 1$ (with $P(1) = 1$) and let $P_2(n) = P(n/P(n))$. We write (a, b) for the greatest common divisor of the two integers a, b, and $[a, b]$ for their least common multiple. For integers $k \ge 2$ we put $\log_k x$ for the k-fold iterate of the natural logarithm evaluated at x (and we shall assume that the argument of $\log_k$ is large enough so that this iteration is defined and positive).

2. The upper bound and a heuristic. In this section we prove Theorem 1.1, followed by a heuristic argument that it is best possible.

To begin the proof, we first show that certain sets of integers are negligible. Let x be large and set

$$\mathcal{S}_1 = \{n \le x : P(n) \le x^{1/\log\log x}\}.$$

By a well-known estimate of de Bruijn (see [3]),

$$\#\mathcal{S}_1 \le \frac{x}{(\log x)^{10}} \tag{1}$$

for all sufficiently large values of x.

Next, set $\alpha = 1/\log 2$ and

$$\mathcal{S}_2 = \{n \le x : n \notin \mathcal{S}_1,\ \Omega(n) \ge 1 + \alpha \log\log x\}.$$

Every integer $n \in \mathcal{S}_2$ is of the form pm where $p > x^{1/\log\log x}$ is prime and $\Omega(m) \ge \alpha \log\log x$, and so

$$\#\mathcal{S}_2 \le x \sum_{\substack{m \le x^{1-1/\log\log x} \\ \Omega(m) \ge \alpha\log\log x}} \sum_{p \le x/m} 1 \ll \frac{x \log\log x}{\log x} \sum_{\substack{m \le x^{1-1/\log\log x} \\ \Omega(m) \ge \alpha\log\log x}} \frac{1}{m}.$$

This last sum was estimated in [11, Corollary 2.5] (though this estimate is well-known). One thus gets

$$\#\mathcal{S}_2 \le \frac{x \log\log x}{(\log x)^{\eta}}, \tag{2}$$

where η is the Erdős–Tenenbaum–Ford constant.

Therefore, we may consider the set $\mathcal{S}_3$ of values of λ in $[1, x]$ that are neither in $\mathcal{S}_1$ nor in $\mathcal{S}_2$. Set $k = \lceil \alpha \log\log x \rceil$. Let $n \in \mathcal{S}_3$, so that $\Omega(n) \le k$. Since $P(n) > x^{1/\log\log x}$ and $n = \lambda(N)$ for some integer N, it follows that $P(N) > x^{1/\log\log x}$. Set $p = P(N)$, and so n is of the form $(p-1)m$. Write $\Omega(p-1) = i$, $\Omega(m) = j$, so that $i + j \le k$. We have

$$\#\mathcal{S}_3 \le \sum_{i+j\le k} \sum_{\substack{m \le 2x^{1-1/\log\log x} \\ \Omega(m)=j}} \sum_{\substack{p \le x/m+1 \\ \Omega(p-1)=i}} 1.$$

It follows from Timofeev [14, Theorem 1] that the inner sum here is

$$O\left(\frac{x(\log\log x+\kappa_1)^{i-1}}{(i-1)!m(\log(x/m))^2}\right)=O\left(\frac{x(\log\log x+\kappa_1)^{i+1}}{(i-1)!m(\log x)^2}\right)$$

for an absolute positive constant κ_1. By a suitable adjustment of κ_1 if necessary, it follows from equation (2) of [14] (due to Halász) and partial summation that

$$\begin{aligned}\#\mathcal{S}_3 &\ll \frac{x\log\log x}{(\log x)^2}\sum_{i+j\le k}\frac{(\log\log x+\kappa_1)^{i+j}}{(i-1)!j!}\\ &\ll \frac{x(\log\log x)^2}{(\log x)^2}\sum_{i+j\le k-1}\frac{(\log\log x+\kappa_1)^{i+j}}{i!j!}\\ &= \frac{x(\log\log x)^2}{(\log x)^2}\sum_{l\le k-1}\frac{(2\log\log x+2\kappa_1)^l}{l!}.\end{aligned}$$

Using $1/\log 2<2$ and Stirling's formula, we get

$$\#\mathcal{S}_3\ll\frac{x(\log\log x)^2}{(\log x)^2}\cdot\frac{(2\log\log x+2\kappa_1)^{k-1}}{(k-1)!}\ll\frac{x(\log\log x)^2}{(\log x)^\eta},$$

where η is the Erdős–Tenenbaum–Ford constant. With (1) and (2), this shows that

$$V_\lambda(x)\ll\frac{x(\log\log x)^2}{(\log x)^\eta},$$

completing our proof of Theorem 1.1.

2.1. A heuristic lower bound. Fix a positive integer k. We consider numbers $n=\lambda(p_1\dots p_k)\le x$ where the primes $p_1,\dots,p_k$ are distinct, and each p_i-1 is squarefree. Thus, n is squarefree. Consider the set S_i of primes dividing p_i-1. The Venn diagram of these k sets has 2^k-1 subset intersections (we do not consider the empty subset), and so we have an ordered factorization of n into 2^k-1 factors. We ask when an ordered factorization of an even squarefree number n into 2^k-1 factors corresponds to an equation $n=\lambda(p_1\dots p_k)$ in this way. Suppose $n\le x$ and $\Omega(n)>\beta\log\log x$. Then n has more than

$$(2^k-1)^{\beta\log\log x}=(\log x)^{\beta\log(2^k-1)}$$

ordered factorizations into 2^k-1 factors. There are k different subsets of these factors, each of size 2^{k-1}, that we wish to express as p_i-1. The "chance" that all of these are shifted primes is about $(\log x)^{-k}$. Thus, we "expect" n to be a λ value in this way if

$$\beta\log(2^k-1)>k.$$

If $\beta > 1/\log 2 + \epsilon$, this last inequality holds for all large values of k, so we "expect" an even squarefree n with $\Omega(n) > (1/\log 2 + \epsilon)\log\log x$ to be a λ value. The number of such $n \le x$ is greater than $x/(\log x)^{\eta+\delta}$ for all sufficiently large values of x, where η is the Erdős–Tenenbaum–Ford constant and $\delta \to 0^+$ as $\epsilon \to 0^+$. This argument then suggests that Theorem 1.1 is best possible.

Let $V_\lambda^{(k)}(x)$ denote the number of integers in $[1, x]$ of the form $\lambda(n)$, where $\Omega(n) = k$. Our proof of Theorem 1.2 actually shows that

$$V_\lambda^{(2)}(x) \ge \frac{x}{(\log x)^{.359052}}$$

for all sufficiently large values of x. On the other hand, the above heuristic argument suggests that

$$V_\lambda^{(2)}(x) = \frac{x}{(\log x)^{\beta_2+o(1)}}, \quad \text{where} \quad \beta_2 := 1 - \frac{2}{\log 3}(1 - \log 2 + \log\log 3).$$

Note that $\beta_2 = .270169\ldots$. To make further progress, it seems reasonable to try and prove this estimate.

3. Some sieve estimates. We shall use some standard upper bound sieve methods such as are found in [10], and also some non-standard ones that follow as consequences. We have the following results.

LEMMA 3.1. *Uniformly for all positive integers a and real numbers $t > 1$, the number of integers $n \le t$ with $an + 1$ prime is*

$$O\bigg(\frac{a}{\varphi(a)} \cdot \frac{t}{\log t}\bigg).$$

In addition, the number of primes $p \le t$ with $ap + 1$ also prime is

$$O\bigg(\frac{a}{\varphi(a)} \cdot \frac{t}{(\log t)^2}\bigg).$$

The first part follows from the Brun–Titchmarsh inequality. Both parts can be found in [10, Chapter 2].

LEMMA 3.2. *Uniformly for all pairs of positive integers $a < b$ and real numbers $t > 1$, the number of integers $n \le t$ with both $an + 1$ and $bn + 1$ prime is*

$$O\bigg(\frac{ab(b-a)(a,b)}{\varphi(ab(b-a))\varphi((a,b))} \cdot \frac{t}{(\log t)^2}\bigg).$$

If we ask in addition that n be prime, the estimate is the same but with $(\log t)^3$ instead of $(\log t)^2$.

These too are standard results, see [10].

We consider variants of these two lemmas where we restrict the number of primes dividing n. Before stating these results we consider the following useful fact.

LEMMA 3.3. *Uniformly for $0 < \alpha \le 1$ we have*

$$\sum_{\substack{P(n)\le x\\ \Omega(n)\le \alpha \log\log x}} \frac{1}{\varphi(n)} \ll (\log x)^{\alpha-\alpha\log\alpha}.$$

Proof. By essentially the same proof as [11, Lemma 2.4], for $0 < z < 2$ we have

$$\sum_{P(n)\le x} \frac{z^{\Omega(n)}}{\varphi(n)} \ll \frac{(\log x)^z}{2-z}. \tag{3}$$

Applying this with $z = \alpha$ as in [11, Corollary 2.5], we have

$$\sum_{\substack{P(n)\le x\\ \Omega(n)\le \alpha \log\log x}} \frac{1}{\varphi(n)} \le \sum_{P(n)\le x} \frac{\alpha^{\Omega(n)-\alpha\log\log x}}{\varphi(n)} \ll (\log x)^{\alpha-\alpha\log\alpha}. \ \blacksquare$$

LEMMA 3.4. *Let α, s, t be real numbers with $0 < \alpha \le 1$ and $3 \le s \le t$. Uniformly in α, s, t and each positive integer a, the number of integers $n \le t$ with $an+1$ prime and $\Omega_s(n) \le \alpha \log\log s$ is*

$$O\bigg(\frac{a}{\varphi(a)} \cdot \frac{t(\log\log t)^2}{(\log t)(\log s)^{1+\alpha\log\alpha-\alpha}}\bigg).$$

Proof. Write $n = pm$ where $p = P(n)$. If $p \le t^{1/\log\log t}$, then (1) shows that these values of n are negligible. So assume that $p > t^{1/\log\log t}$, and thus $m < t^{1-1/\log\log t}$. Further, we may assume that $\Omega_s(m) \le \alpha\log\log s$. For each such m we count the number of primes $p \le t/m$ with $apm+1$ prime. By the second part of Lemma 3.1, this count is

$$O\bigg(\frac{am}{\varphi(am)} \cdot \frac{t/m}{(\log(t/m))^2}\bigg) = O\bigg(\frac{a}{\varphi(a)\varphi(m)} \cdot \frac{t(\log\log t)^2}{(\log t)^2}\bigg).$$

But, writing $m = m_1m_2$, where $P(m_1) \le s$ and $p \mid m_2$ implies $p > s$, we get

$$\sum_{\substack{m\le t\\ \Omega_s(m)\le \alpha\log\log s}} \frac{1}{\varphi(m)} \le \sum_{\substack{P(m_1)\le s\\ \Omega(m_1)\le \alpha\log\log s}} \frac{1}{\varphi(m_1)} \sum_{\substack{m_2\le t\\ (m_2,\lfloor s\rfloor !)=1}} \frac{1}{\varphi(m_2)}$$

$$\ll (\log s)^{\alpha-\alpha\log\alpha} \frac{\log t}{\log s},$$

where we used Lemma 3.3 for the sum over m_1, and the fundamental lemma of the sieve (see [10, Theorem 2.5]) for the sum over m_2. The present lemma now follows by summing the prior estimate over m. $\blacksquare$

LEMMA 3.5. *Let α, s, t be real numbers with $0 < \alpha \le 1$ and $3 \le s \le t$, and let a, b be positive integers with $a < b$. Uniformly, the number of integers $n \le t$ with $an+1$ and $bn+1$ both prime and $\Omega_s(n) \le \alpha \log\log s$ is*

$$O\bigg(\frac{ab(b-a)(a,b)}{\varphi(ab(b-a))\varphi((a,b))} \cdot \frac{t(\log\log t)^3}{(\log t)^2(\log s)^{1+\alpha\log\alpha-\alpha}}\bigg).$$

Proof. This follows in the same way as Lemma 3.4 except that we use the second part of Lemma 3.2 instead of the second part of Lemma 3.1. ∎

4. Preliminaries for the lower bound. We use the following result which follows directly from [1, Theorem 2.1].

PROPOSITION 4.1. *There is an absolute constant M_0 such that, for all sufficiently large values of x, there is a set $\mathcal{M}_0(x)$ of at most M_0 integers all at least $\log x$ such that, if $d \le x^{1/3}$ is a positive integer not divisible by any member of $\mathcal{M}_0(x)$ and z is a number with $dx^{2/3} \le z$, then the number of primes $p \le z$ with $p \equiv 1 \pmod d$ is at least $z/(2\varphi(d)\log z)$.*

Associated with the parameter x we have a secondary parameter y that is chosen optimally at the end of the proof. We shall always have

$$\exp((\log x)^{1/10}) \le y \le x^{1/4}. \tag{4}$$

When we write the expression $o(1)$ it is always under the assumption that $x \to \infty$, which is equivalent to the assumption that $y \to \infty$. It is uniform in other parameters. For two real numbers A, B, when we write

$$A \approx_y B$$

we mean that $|A - B| \le (\log\log y)^{2/3}$.

We let $\mathcal{P}(y)$ denote the set of primes p such that

- $y^{1-1/\log\log y} < p \le y$,
- $\Omega(p-1) \approx_y \log\log y$,
- $P_2(p-1) > y^{1/\log\log y}$,
- for m the largest divisor of $p-1$ with $P(m) \le \log y$, we have $(p-1)/m$ squarefree and $\Omega(m) \le 5\log_3 y$,
- $p-1$ is not divisible by any member of $\mathcal{M}_0(x)$.

LEMMA 4.2. *We have, as $y \to \infty$,*

$$\sum_{p\in\mathcal{P}(y)} \frac{1}{p} = \frac{1+o(1)}{\log\log y}.$$

Proof. This estimate holds for the primes p that satisfy the first item in the definition of $\mathcal{P}(y)$, so it suffices to show that the latter four conditions do not eliminate too many primes. That is, we will show that for t in the interval $[y^{1-1/\log\log y}, y]$, the number of primes $p \le t$ such that any of the remaining conditions in the definition of $\mathcal{P}(y)$ fails is $o(t/\log t)$ as $y \to \infty$.

The number of primes $p \le t$ where the second item in the definition of $\mathcal{P}(y)$ fails is $o(t/\log t)$, as can be seen by the method of Erdős [4], or more explicitly by Timofeev [14] (see also [5, Lemmas 2.1 and 2.2]).

We consider primes $p \le t$ for which the third item in the definition of $\mathcal{P}(y)$ fails. For $y^{1-1/\log\log y} \le t \le y$, the number of integers $n \le t$ with $P(n) \le y^{1/\log\log y}$ is $O(t/(\log t)^{10})$ as can be seen by the theorem of de Bruijn used for (1). Write an integer $1 < n \le t$ as mq where $q = P(n)$ and assume that $q > y^{1/\log\log y}$. Suppose that $P(m) \le y^{1/\log\log y}$. If $m > y^{1/2}$, then the number of such integers $mq \le t$ is at most

$$\sum_{q<y^{1/2}} \sum_{m\le t/q} 1 \ll \sum_{q<y^{1/2}} \frac{t}{q(\log t)^{10}} \ll \frac{t}{(\log t)^9}.$$

It thus suffices to count integers $mq \le t$ with $m \le y^{1/2}$, $P(m) \le y^{1/\log\log y}$, and both q and $mq+1$ prime. For a fixed m, the number of primes $q \le t/m$ with $mq+1$ prime is, by Lemma 3.1, at most

$$O\bigg(\frac{t}{\varphi(m)(\log t)^2}\bigg).$$

It remains to note that, with r running over primes,

$$\begin{aligned}\sum_{P(m)\le y^{1/\log\log y}} \frac{1}{\varphi(m)} &= \prod_{r\le y^{1/\log\log y}} \left(1 + \frac{1}{r-1} + \frac{1}{r(r-1)} + \cdots\right) \\ &= \prod_{r\le y^{1/\log\log y}} \left(1 + \frac{r}{(r-1)^2}\right) \ll \frac{\log y}{\log\log y} \ll \frac{\log t}{\log\log t}.\end{aligned}$$

Thus, the number of primes $p \le t$ where the third item in the definition of $\mathcal{P}(y)$ fails is $O(t/(\log t \log\log t)) = o(t/\log t)$.

Now consider primes $p \le t$ where the fourth condition fails. Even without the primality requirement, the number of $n \le t$ such that $n-1$ is divisible by the square of a prime exceeding $\log y$ is $O(t/(\log y \log\log y)) = o(t/\log t)$ as $x \to \infty$. If $\Omega(m) > 5\log_3 y$, let m_0 be the smallest divisor of m with $\Omega(m_0) > 5\log_3 y$, so that $m_0 < (\log y)^{1+5\log_3 y} = y^{o(1)}$ as $y \to \infty$. It then follows from Lemma 3.1 that the number of primes $p \le t$ with $p-1$ divisible by such a number m_0 is bounded by a constant times

$$\frac{t}{\log t} \sum_{\substack{P(m_0)\le \log y \\ \Omega(m_0)>5\log_3 y}} \frac{1}{\varphi(m_0)}.$$

We now use (3) to estimate this sum, with $z = 2 - 1/(5\log_3 y)$, getting

$$\sum_{\substack{P(m_0)\le \log y \\ \Omega(m_0)>5\log_3 y}} \frac{1}{\varphi(m_0)} \le z^{-5\log_3 y} \sum_{P(m_0)\le \log y} \frac{z^{\Omega(m_0)}}{\varphi(m_0)} \ll (\log_2 y)^{2-5\log 2} \log_3 y.$$

Since $2 - 5\log 2 < -1$, it follows that the number of primes $p \le t$ where $\Omega(m) > 5\log_3 y$ is $O(t/(\log t \log_2 y)) = o(t/\log t)$.

For primes $p \le t$ where the fifth item in the definition of $\mathcal{P}(y)$ fails, we see immediately from Proposition 4.1 and the Brun–Titchmarsh inequality that the number of them is $O(t/(\log t \log x)) = o(t/\log t)$. This completes the proof of Lemma 4.2. ∎

We shall consider integers j chosen so that

$$\tfrac{1}{10}\log\log y \le j \le \tfrac{9}{10}\log\log y. \tag{5}$$

For such an integer j and $p \in \mathcal{P}(y)$, we let $\mathcal{D}_{p,j,y}$ denote the set of divisors d of $p-1$ such that

- $P(p-1) \mid d$ and $P_2(p-1) \mid (p-1)/d$,
- all of the prime factors of $(p-1)/d$ exceed $\log y$,
- $\Omega(d) = j$.

For $0 < c < 1$, let

$$\Delta(c) = -c\log c - (1-c)\log(1-c).$$

LEMMA 4.3. *For an integer j satisfying* (5) *and a prime $p \in \mathcal{P}(y)$, we have $\#\mathcal{D}_{p,j,y} = (\log y)^{\Delta(c)+o(1)}$, where $c = j/\log\log y$.*

Proof. Let $k = \Omega(p-1)$ and let $i = 2 + \Omega_{\log y}(p-1)$. Then

$$\#\mathcal{D}_{p,j,y} = \binom{k-i}{j-(i-1)}.$$

Indeed, by the various properties of p, the top two prime factors of $p-1$ are distinct as are all of the prime factors of $p-1$ exceeding $\log y$. Thus for $d \in \mathcal{D}_{p,j,y}$, it already has $i-1$ primes (1 from the top two and all $i-2$ of them, with multiplicity, dividing $p-1$ that are at most $\log y$), so there are $j-(i-1)$ left to choose from the remaining $k-i$ primes in $p-1$, and all of these primes appear with exponent 1 in the prime factorization of $p-1$. An elementary estimate using $k \approx_y \log\log y$, $j = c\log\log y$, $i \le 2 + 5\log_3 y$, and Stirling's formula completes the proof. ∎

5. The lower bound. In this section we prove Theorem 1.2. In doing so we shall count only a subset of λ-values; all of the values we count are of the form $\lambda(pq)$. Given x, y satisfying (4) and j satisfying (5), let $r(n) = r_{j,y}(n)$ denote the number of triples a, b, d of positive integers with

- $n = abd$, $(a,b) = 1$,
- $p := ad + 1 \in \mathcal{P}(y)$, $d \in \mathcal{D}_{p,j,y}$,
- $bd + 1$ is prime,
- $\Omega_y(b) \approx_y \log\log y$.

If $n = abd$ as above with $p = ad + 1$ and $q = bd + 1$, then $\lambda(pq) = n$. Thus,

$$V_\lambda(x) \ge \sum_{\substack{n\le x\\ r(n)>0}} 1,$$

and so, from the Cauchy–Schwarz inequality, we have

$$V_\lambda(x) \ge \frac{(\sum_{n\le x} r(n))^2}{\sum_{n\le x} r(n)^2}. \tag{6}$$

Thus, our strategy is to get a lower bound for $\sum_{n\le x} r(n)$, an upper bound for $\sum_{n\le x} r(n)^2$, and then choose our parameters j, y optimally with respect to our bounds.

5.1. The sum $\sum_{n\le x} r(n)$. For $p \in \mathcal{P}(y)$ and $d \in \mathcal{D}_{p,j,y}$, we count choices for b. Thus,

$$\sum_{n\le x} r(n) = \sum_{p\in\mathcal{P}(y)} \sum_{d\in\mathcal{D}_{p,j,y}} \sum_{\substack{b\le x/(p-1)\\ (b,(p-1)/d)=1\\ \Omega_y(b)\approx_y \log\log y\\ bd+1 \text{ prime}}} 1. \tag{7}$$

For $p \in \mathcal{P}(y)$ and $d \in \mathcal{D}_{p,j,y}$, write $p = ad + 1$. From the definition of $\mathcal{P}(y)$ it follows that d is not divisible by any member of $\mathcal{M}_0(x)$, defined in Proposition 4.1. It thus follows from the size restrictions for p and y that the number of choices for an integer $b \le x/(p-1)$ with $bd+1$ prime is at least

$$\frac{x/a}{2\varphi(d)\log(x/a)} \gg \frac{x}{a\varphi(d)\log x}.$$

Further, from Lemma 3.1, this lower bound also stands as an upper estimate; that is,

$$\sum_{\substack{b\le x/(p-1)\\ bd+1 \text{ prime}}} 1 \asymp \frac{x}{a\varphi(d)\log x}.$$

This ignores the other two conditions on b, namely $\Omega_y(b) \approx_y \log\log y$ and $(b, a) = 1$. From the definitions of $\mathcal{P}(y)$ and $\mathcal{D}_{p,j,y}$ it follows that a has all of its $O(\log\log y)$ prime factors greater than $\log y$, so that by Lemma 3.1, the number of choices for $b \le x/ad$ with $bd + 1$ prime and $(b, a) > 1$ is

$$O\Big(\frac{x\log\log y}{a\varphi(d)\log x\log y}\Big),$$

which is negligible compared with the prior estimate.

Showing that the restriction $\Omega_y(b) \approx_y \log\log y$ does not significantly affect the count takes a little more work. The argument is similar to the proof of Lemma 3.4; we give the details. The number of integers $b \le x/(ad)$ with $P(b) \le x^{1/\log\log x}/(ad)$ is $O(x/(ad(\log x)^{10}))$, by the result (1) of de Bruijn,

so these values of b are negligible. For the remaining values of b, write them as $b = b'r$ where $r = P(b)$, and so $\Omega_y(b) = \Omega_y(b')$ or $\Omega_y(b) = 1 + \Omega_y(b')$. For each choice of $b' \le x^{1-1/\log\log x}$, we count primes $r \le x/adb'$ with $b'rd + 1$ prime. By Lemma 3.1, this is

$$O\Big(\frac{x(\log\log x)^2}{a\varphi(db')(\log x)^2}\Big) = O\Big(\frac{x(\log\log x)^3}{a\varphi(d)b'(\log x)^2}\Big). \tag{8}$$

We have, with $w := \log_2 y - (\log_2 y)^{2/3}$,

$$\sum_{\substack{b'\le x\\ \Omega_y(b')\le w}} \frac{1}{b'} \le \sum_{\substack{P(b_1)\le y\\ \Omega(b_1)\le w}} \frac{1}{b_1} \sum_{\substack{b_2\le x\\ (b_2,\lfloor y\rfloor!)=1}} \frac{1}{b_2}.$$

For the sum over b_1 we use Lemma 3.3 with $\alpha = w/\log_2 y$, and for the sum over b_2 we use the fundamental lemma of the sieve (in [10]), so that

$$\sum_{\substack{b'\le x\\ \Omega_y(b')\le w}} \frac{1}{b'} \ll (\log y)^{1-1/(2(\log\log y)^{2/3})} \cdot \frac{\log x}{\log y} = \frac{\log x}{\exp\big(\frac12(\log\log y)^{1/3}\big)}.$$

Putting this into our prior calculation (8), we find that the number of choices for $b \le x/ad$ with $bd + 1$ prime and $\Omega_y(b) \le w$ is of order at most

$$\frac{x(\log\log x)^3}{a\varphi(d)(\log x)\exp\big(\frac12(\log\log y)^{1/3}\big)} \ll \frac{x}{a\phi(d)(\log x)\exp\big(\frac15(\log\log x)^{1/3}\big)}$$

(using (4)), which is negligible. A similar calculation shows the same inequality holds for the number of b's with $\Omega_y(b) \ge z := \log_2 y + (\log_2 y)^{2/3}$. Hence, the inner sum in (7) satisfies

$$\frac{x}{a\varphi(d)\log x} \ll \sum_{\substack{b\le x/(p-1)\\ (b,(p-1)/d)=1\\ \Omega_y(b)\approx_y \log\log y\\ bd+1 \text{ prime}}} 1 \ll \frac{x}{a\varphi(d)\log x} \ll \frac{x\log\log y}{ad\log x} \ll \frac{x\log\log y}{p\log x}. \tag{9}$$

Using (7), we conclude that, with $c = j/\log\log y$,

$$\sum_{n\le x} r(n) = \frac{x(\log\log y)^{O(1)}}{\log x} \sum_{p\in\mathcal{P}(y)} \frac1p \sum_{d\in\mathcal{D}_{p,j,y}} 1 = \frac{x}{\log x}(\log y)^{\Delta(c)+o(1)}, \tag{10}$$

where for the last estimate we used Lemmas 4.2 and 4.3.

5.2. The sum $\sum_{n\le x} r(n)^2$. In the argument above, we counted triples a, b, d of positive integers with $abd \le x$ as in (7), with $a = (p-1)/d$. Note that $\sum_{n\le x} r(n)^2$ counts the number of 6-tuples a, b, d, a', b', d' with

$abd = a'b'd' \le x$, with both a, b, d and a', b', d' as in $\sum_{n\le x} r(n)$. There are four possibilities:

1. $a = a'$ and $b = b'$,
2. $a \ne a'$ and $b = b'$,
3. $a = a'$ and $b \ne b'$,
4. $a \ne a'$ and $b \ne b'$.

Let $R_\nu(x, y)$ denote the number of 6-tuples in cases $\nu = 1, 2, 3, 4$. Thus,

$$\sum_{n\le x} r(n)^2 = R_1(x, y) + R_2(x, y) + R_3(x, y) + R_4(x, y). \tag{11}$$

Our task now is to find upper estimates for each $R_\nu(x, y)$.

The case of $\nu = 1$ is simple, since $R_1(x, y)$ is equal to $\sum_{n\le x} r(n)$, which we have already estimated in (10). So,

$$R_1(x, y) = S_1(x, y)(\log y)^{o(1)}, \text{ where } S_1(x, y) := \frac{x}{\log x}(\log y)^{\Delta(c)}. \tag{12}$$

The estimation for $R_2(x, y)$ is also easy. Here we have $\lambda(pq) = \lambda(pq')$ with $q \ne q'$. Looking at the triple summation in (7), we amend this by choosing two unequal divisors d, d' of $p - 1$ in $\mathcal{D}_{p,j,y}$ in the middle summation, which can be done in $(\log y)^{2\Delta(c)+o(1)}$ ways (Lemma 4.3). Then we choose an integer $b \le x/(p-1)$ coprime to both $(p-1)/d$ and $(p-1)/d'$, and with both $bd+1$ and $bd' + 1$ prime. By Lemma 3.2, the number of choices for b is at most $x(\log\log x)^{O(1)}/p(\log x)^2$. Thus, using Lemma 4.2, and an argument similar to the one we used to estimate (9) and (10), we get

$$R_2(x, y) \le S_2(x, y)(\log y)^{o(1)}, \text{ where } S_2(x, y) := \frac{x}{(\log x)^2}(\log y)^{2\Delta(c)}. \tag{13}$$

We will work harder in estimating $R_3(x, y)$. Here we are counting the number of quadruples a, d, d', q with $p = ad + 1, p' = ad' + 1 \le y$ unequal primes, $\Omega(d) = \Omega(d') = j = c\log\log y$, $\Omega(a) \approx_y (1 - c)\log\log y$, q prime, $q \le x/a + 1$, $q \equiv 1 \pmod{[d, d']}$ and $\Omega_y((q - 1)/d) \approx_y \log\log y$. Further, we have $a > y^{1/\log\log y}$. (There are also the conditions that a is coprime to both $(q-1)/d$ and $(q-1)/d'$, but we shall ignore these.) Let $u = (d, d')$ and write $d = uv, d' = uv'$. Let i, θ be such that $\Omega(v) = i = \theta c\log\log y$, so that

$$\Omega(v) = \Omega(v') = i = \theta c\log\log y, \quad \theta \in [0, 1]. \tag{14}$$

Note that $[d, d'] = uvv'$, so that $\Omega_y((q - 1)/[d, d']) \approx_y (1 - \theta c)\log\log y$. Lemma 3.4 implies that the number of such primes $q \le x/a + 1$ is at most

$$\frac{x}{auvv'\log x}(\log y)^{-1-(1-\theta c)\log(1-\theta c)+1-\theta c+o(1)}.$$

Thus, the contribution to $R_3(x,y)$ corresponding to the parameter θ is at most

$$\frac{x}{\log x}(\log y)^{-\theta c-(1-\theta c)\log(1-\theta c)+o(1)}\sum_{a,u,v,v'}\frac{1}{auvv'}. \tag{15}$$

For the summation, we first sum over a given u,v,v'. The important conditions for us are that $a > y^{1/\log\log y}$, $\Omega(a) \approx_y (1-c)\log\log y$, and both $auv+1$ and $auv'+1$ are prime. Since $v \neq v'$ in this case, Lemma 3.5 and partial summation imply that

$$\sum_{a,u,v,v'}\frac{1}{auvv'} \le (\log y)^{-2-(1-c)\log(1-c)+1-c+o(1)}\sum_{u,v,v'}\frac{1}{uvv'}.$$

To complete the estimate we use $\Omega(u) = (c-\theta c)\log\log y$ and (14). Thus, using Lemma 3.3,

$$\sum_{u,v,v'}\frac{1}{uvv'} \le (\log y)^{-(c-\theta c)\log(c-\theta c)+c-\theta c-2\theta c\log(\theta c)+2\theta c+o(1)}.$$

Putting this estimate together with the previous one, we have

$$\sum_{a,u,v,v'}\frac{1}{auvv'} \le (\log y)^{-1-(1-c)\log(1-c)-(c-\theta c)\log(c-\theta c)+\theta c-2\theta c\log(\theta c)+o(1)},$$

and together with (15) we have the contribution to $R_3(x,y)$ corresponding to the parameter θ being

$$\frac{x}{\log x}(\log y)^{-1-(1-\theta c)\log(1-\theta c)-(1-c)\log(1-c)-(c-\theta c)\log(c-\theta c)-2\theta c\log(\theta c)+o(1)}.$$

We consider all integers $i \in [0,j]$ which correspond to evenly spaced values of $\theta \in [0,1]$, the spacing being $1/j = 1/(c\log\log y)$. However, it can only give a larger estimate if we consider θ as a continuous variable in $[0,1]$, and so we seek that value of θ (as a function of c) which maximizes the above expression. A short calculation shows this occurs when $\theta = 1/(1+c)$, and so using this one value of θ in the above expression and multiplying by $\log\log y$ gives us our estimate for $R_3(x,y)$. After a little algebraic simplification this gives us

$$R_3(x,y) \le S_3(x,y)(\log y)^{o(1)}, \tag{16}$$
$$\text{where } S_3(x,y) := \frac{x}{\log x}(\log y)^{-1+(1+c)\log(1+c)-c\log c+\Delta(c)}.$$

Now we estimate $R_4(x,y)$. This quantity is the number of solutions to $adb = a'd'b' \le x$ where $p = ad+1$, $p' = a'd'+1 \in \mathcal{P}(y)$ as in the definition of $r(n)$ and $bd+1 = q \le x/a+1$, $b'd'+1 = q' \le x/a'+1$ again as in the definition of $r(n)$, where p,p',q,q' are primes with $p \neq p'$, $q \neq q'$. Further, as before, we assume that $\Omega(d) = \Omega(d') = j = c\log\log y$. Dividing the

equation $adb = a'd'b'$ by $u = (ad, a'd')$, we see that there is an integer h with

$$b = \frac{a'd'}{u}h, \qquad b' = \frac{ad}{u}h.$$

Assume that $\Omega(u) = i$. We have $1 \le i \le z = \log_2 y + (\log_2 y)^{2/3}$. We again use an auxiliary variable $\theta \in [0,1]$, this time defined by the equation $i = \theta z$. Since

$$\Omega(ad), \Omega(a'd'), \Omega_y(b), \Omega_y(b') \approx_y \log\log y$$

and $\Omega(u) = \theta z$, it follows that

$$|\Omega_y(h) - \theta z| \le 2(\log\log y)^{2/3}.$$

Since $b \le x/(ad)$, the variable h ranges up to $x/[ad, a'd'] = x/[p-1, p'-1]$. Further,

$$\frac{a'd'd}{u}h + 1 \quad \text{and} \quad \frac{add'}{u}h + 1$$

are the different primes q and q', respectively. Thus, the number of integers h with these conditions is, by Lemma 3.5, at most

$$\frac{x}{[ad, a'd'](\log x)^2}(\log y)^{-1+\theta-\theta\log\theta+o(1)}.$$

We now sum on a, d, a', d'. By Lemma 4.3, this gives

$$\frac{x}{(\log x)^2}(\log y)^{2\Delta(c)-1+\theta-\theta\log\theta+o(1)} \sum_{p,p'} \frac{1}{[p-1, p'-1]}. \tag{17}$$

We compute this sum over primes $p, p' \in \mathcal{P}(y)$. Recall that $u = (p-1, p'-1)$ and $\Omega(u) = \theta z$. Write $p-1 = uv, p'-1 = uv'$, so that $\Omega(v), \Omega(v') \approx_y (1-\theta)z$. We are to consider $\sum 1/uvv'$ under these conditions. Either $u > y^{1/3}$ or both $v, v' > y^{1/3}$. Assume first that $u > y^{1/3}$. Given $v \ne v'$, we note that u has the properties that $u > y^{1/3}$, $uv+1$ is prime, $uv'+1$ is prime, and $\Omega(u) = \theta z$. Thus, by Lemma 3.5 and partial summation,

$$\sum_{\substack{u,v,v' \\ u > y^{1/3}}} \frac{1}{uvv'} \le (\log y)^{-2+\theta-\theta\log\theta+o(1)} \sum_{v,v'} \frac{1}{vv'}.$$

To complete the estimate, we use $\Omega(v), \Omega(v') \approx_y (1-\theta)z$, giving us the upper bound

$$(\log y)^{-2+\theta-\theta\log\theta+2(1-\theta)-2(1-\theta)\log(1-\theta)+o(1)} = (\log y)^{-\theta-\theta\log\theta-2(1-\theta)\log(1-\theta)+o(1)},$$

where we have used Lemma 3.3. Now we assume that $u \le y^{1/3}$ so that both $v, v' > y^{1/3}$. We have, by Lemma 3.4 and partial summation,

$$\sum_{\substack{u,v,v' \\ v,v'>y^{1/3}}} \frac{1}{uvv'} \le ((\log y)^{-1+(1-\theta)-(1-\theta)\log(1-\theta)+o(1)})^2 \sum_u \frac{1}{u}.$$

Using Lemma 3.3, this final sum is at most $(\log y)^{\theta-\theta\log\theta+o(1)}$, so that we end up with the same estimate for $\sum 1/(uvv') = \sum 1/[p-1, p'-1]$ as in the case $u > y^{1/3}$, namely

$$(\log y)^{-\theta-\theta\log\theta-2(1-\theta)\log(1-\theta)+o(1)}.$$

Thus, by (17), we have

$$R_4(x,y) \le \frac{x}{(\log x)^2}(\log y)^{2\Delta(c)-1+2\Delta(\theta)+o(1)}.$$

We now choose θ in $[0,1]$ so as to maximize this exponent; this is when $\theta = 1/2$. Hence,

$$R_4(x,y) \le S_4(x,y)(\log y)^{o(1)}, \tag{18}$$

where

$$S_4(x,y) := \frac{x}{(\log x)^2}(\log y)^{2\Delta(c)-1+\log 4}.$$

Using (11), (12), (13), (16), and (18), we have

$$\sum_{n\le x} r(n)^2 \le (S_1(x,y) + S_2(x,y) + S_3(x,y) + S_4(x,y))(\log y)^{o(1)}.$$

Since $S_2(x,y) \le S_1(x,y)$, we may simplify this a little to

$$\sum_{n\le x} r(n)^2 \le (S_1(x,y) + S_3(x,y) + S_4(x,y))(\log y)^{o(1)}. \tag{19}$$

5.3. Choosing parameters. Using

$$\sum_{n\le x} r(n) = S_1(x,y)(\log y)^{o(1)}$$

(which follows from (10) and the definition of $S_1(x,y)$), (6), and (19), we have

$$V_\lambda(x) \ge \frac{S_1(x,y)^2}{S_1(x,y) + S_3(x,y) + S_4(x,y)}(\log y)^{o(1)}.$$

Dividing the numerator and denominator by $x(\log y)^{2\Delta(c)}/(\log x)^2$, we get

$$V_\lambda(x) \ge \frac{x}{S'(x,y)}(\log y)^{o(1)},$$

where

$$S'(x,y) := (\log x)(\log y)^{-\Delta(c)} + (\log x)(\log y)^{-1+(1+c)\log(1+c)-c\log c-\Delta(c)} + (\log y)^{-1+\log 4}.$$

We choose the parameter c so that the first two terms in $S'(x,y)$ are equal. This gives c as a solution of the equation

$$1 = (1+c)\log(1-c) - c\log c,$$

so that $c = .5422114$ is a reasonable choice. (Namely, we choose the integer j as a function of y so that $c \to .5422114$ as $y \to \infty$.) Next we choose y to equate the first and third terms in $S'(x,y)$, so that $\log y$ is $(\log x)^{1/(\Delta(c)-1+\log 4)}$, which is approximately $(\log x)^{.929477}$. Thus, (4) holds. We thus get

$$V_\lambda(x) \ge x/(\log x)^{(-1+\log 4)/(\Delta(c)-1+\log 4)+o(1)},$$

so that

$$V_\lambda(x) \ge \frac{x}{(\log x)^{.359052}}$$

for all sufficiently large values of x. This completes our proof of Theorem 1.2.

6. Proof of Theorem 1.3. Since $\lambda(p) = p-1$, we trivially have

$$\#\{\lambda(n) : n \le x\} \ge \pi(x),$$

so that the lower bound implicit in Theorem 1.3 follows immediately from the prime number theorem.

It thus remains to prove that

$$\#\{\lambda(n) : n \le x\} \le \frac{x}{(\log x)^{1+o(1)}} \quad \text{as } x \to \infty. \tag{20}$$

Let $M = \sqrt{\log\log x}$ and, for a positive integer n, let

$$F_M(n) = \prod_{p \le M} p^{v_p(\varphi(n))}, \qquad L_M(n) = \prod_{p \le M} p^{v_p(\lambda(n))}.$$

Here, $v_p(m)$ denotes the integer ν with $p^\nu \,\|\, m$. Suppose that we have both $F_M(n) > (\log x)^2$ and $L_M(n) \le \log x$. Then

$$\lambda(n) = \varphi(n) \cdot \frac{\lambda(n)}{\varphi(n)} \le x \cdot \frac{L_M(n)}{F_M(n)} < \frac{x}{\log x}.$$

Thus, to prove (20) it suffices to show that both

$$\#\{n \le x : F_M(n) \le (\log x)^2\} \le \frac{x}{(\log x)^{1+o(1)}},$$

$$\#\{\lambda(n) : n \le x,\ L_M(n) > \log x\} \le \frac{x}{(\log x)^{1+o(1)}}$$

as $x \to \infty$.

For the first inequality, we use [12, Proposition 2] with the parameter λ set at 2 and with $\mathcal{Q}$ as the set of primes up to M. Then the quantity R is $o(1)$ and it follows that the number of $n \le x$ with $F_M(n) \le (\log x)^2$ is at most $x/(\log x)^{1+o(1)}$.

If $L_M(n) > \log x$, then $\lambda(n)$ is divisible by a number $m > \log x$ with $P(m) \le M$. The number of such integers in $[1, x]$ is at most

$$\sum_{\substack{P(m)\le M\\ m>\log x}} \frac{x}{m} \le \frac{x}{(\log x)^{1-1/M}} \sum_{P(m)\le M} \frac{1}{m^{1/M}} = \frac{x}{(\log x)^{1-1/M}} \prod_{p\le M} \left(1 - \frac{1}{p^{1/M}}\right)^{-1}$$

$$\le \frac{x}{(\log x)^{1-1/M}} (1 - 2^{-1/M})^{-\pi(M)} \le \frac{x}{\log x} e^{O(M)}.$$

Thus, the number of integers $n \le x$ with $L_M(x) > \log x$ is at most

$$\frac{x}{\log x} \exp\left(O(\sqrt{\log\log x})\right) = \frac{x}{(\log x)^{1+o(1)}} \quad \text{as } x \to \infty.$$

This completes our proof of Theorem 1.3.

7. Appendix: Algorithms. Here we present three algorithms for computing $V_\lambda(x)$, all with complexity $x(\log x)^{O(1)}$. The first two, which we have implemented, require a similar amount of space, while the third requires $x^{1/2}(\log x)^{O(1)}$ space.

The first two algorithms are based on the following easy observation. If $v = \lambda(n)$ and n is minimal with this property, then for each prime $p \,\|\, n$, we have $\lambda(n/p)$ a proper divisor of v.

Let $\mathcal{V}_\lambda = \lambda(\mathbb{N})$, the set of all values of λ. The first algorithm assumes that $v > 1$ and one knows all members of $\mathcal{V}_\lambda$ smaller than v, and the issue is whether $v \in \mathcal{V}_\lambda$. Let $p = P(v)$. If $p - 1 \mid v$ and $v/p \in \mathcal{V}_\lambda$, then $v \in \mathcal{V}_\lambda$. If $p^\alpha \,\|\, v$ and there is a divisor up^α of v with $q := up^\alpha + 1$ prime and a divisor $d > 1$ of u with $(v/(q-1), d) = 1$ and $v/d \in \mathcal{V}_\lambda$, then $v \in \mathcal{V}_\lambda$. If neither of these cases occur, then $v \notin \mathcal{V}_\lambda$.

To see the complexity of a procedure based on this plan, one can use a version of the sieve of Eratosthenes to find the prime factorizations of all numbers up to x. For each candidate v there are at most $\tau(v)$ checks of primality, where $\tau(v)$ is the number of divisors of v. These checks of primality are done by looking up the candidate prime in our sieved interval, and the pairs of divisors d, u that one searches over are found using the prime factorization at hand for the candidate v. The number of pairs d, u is at most $\tau(v)^2$, and since the sum of $\tau(v)^2$ for v up to x is of magnitude $x(\log x)^3$, our complexity estimate follows.

The second algorithm recursively builds up the set

$$\mathcal{V}_\lambda(x) := \mathcal{V}_\lambda \cap [1, x].$$

Let $\mathcal{V}_1$ denote the set of numbers $\lambda(q) \le x$ where q runs over primes and prime powers. This is easily accomplished via the sieve of Eratosthenes. Our initial choice for $\mathcal{V}_\lambda(x)$ is $\mathcal{V}_1$. We then run through the members of the current choice for $\mathcal{V}_\lambda(x)$ looking to see if we can append new members. When we get to member v, for each $d \,|\, v$ with $d > 1$ we look for members $\lambda(q) \in \mathcal{V}_1$ with $\lambda(q) \le dx/v$ and $\lambda(q) = kd$ for some integer k coprime to v/d. For each such k found, we append vk to $\mathcal{V}_\lambda(x)$. When we reach x (actually $x/2$), the current set $\mathcal{V}_\lambda(x)$ is the correct set.

Since the number of members of $\mathcal{V}_1$ up to dx/v divisible by d is at most x/v, summing this over $d \,|\, v$ and $v \le x$ gets us to $x(\log x)^2$. In fact, a factor of $\log x$ can be saved since there are usually about $x/(v \log x)$ members of $\mathcal{V}_1$ up to dx/v divisible by d. However, it seems to be more advantageous to consider the $x(\log x)^2$ version by having an array of the even numbers up to x, each with the label "0". We then change some of these labels to "1", and at the end we have the characteristic function for $\mathcal{V}_\lambda(x) \setminus \{1\}$. We begin by labeling the members of $\mathcal{V}_1$ with "1". Then one cycles through even integers d. For each d, one visits sequentially the multiples of d in the table. If the entry there is 0, it is passed over. If it is 1, say the location is ad, then one has a (we are at the ath multiple of d in our tour). For this value of a we visit the succeeding multiples of d, starting at $(a+1)d$ and stopping before surpassing x/a. For each bd found with label 1, one computes (a, b), and if $(a, b) = 1$ with the label in location adb being 0, the label there is changed to 1. (One might look first at location adb and if the label is 1, no gcd calculation need be done and no changes need be made.) As above, the complexity is the number of triples d, a, b with $adb \le x$, which is of magnitude $x(\log x)^2$, with all of the arithmetic (after the initial construction of $\mathcal{V}_1$) very simple. In this variant, for a given d, one need take a only up to $(x/d)^{1/2}$.

Here is a third algorithm that requires less space. It can compute the number of λ-values in an interval of length $x^{1/2}$ contained in $[1, x]$ in time and space $x^{1/2}(\log x)^{O(1)}$. With a modified sieve of Eratosthenes, one finds the complete prime factorizations of the numbers in the interval, and using this, the list of divisors of the numbers in the interval. Then, via some fast primality test, one finds a list of divisors of the form $\lambda(q)$, where q is a prime or prime power, for each number in the interval. One then recognizes if a candidate v is a λ-value since it is one if and only if

$$v = \operatorname{lcm}\{\lambda(q) : \lambda(q) \,|\, v,\ q \text{ a prime or prime power}\}.$$

We have run the first algorithm with Mathematica up to 10^7 and the second algorithm with Sage up to $5 \cdot 10^8$; our counts follow. Define $c(x)$ by the equation $V_\lambda(x) = x/(\log x)^{c(x)}$.

x	$V_\lambda(x)$	$c(x)$	x	$V_\lambda(x)$	$c(x)$
10^4	2933	.5524	$5 \cdot 10^6$	1238634	.5100
$2 \cdot 10^4$	5696	.5478	10^7	2445343	.5066
$5 \cdot 10^4$	13836	.5395	$2 \cdot 10^7$	4830396	.5035
10^5	27155	.5335	$5 \cdot 10^7$	11891820	.4995
$2 \cdot 10^5$	53242	.5290	10^8	23523516	.4967
$5 \cdot 10^5$	130116	.5229	$2 \cdot 10^8$	46558154	.4940
10^6	256158	.5187	$5 \cdot 10^8$	114882775	.4907
$2 \cdot 10^6$	504850	.5147			

Acknowledgments. Work by the first author was partially done in Spring of 2007 while he visited Williams College. He would like to thank the college for its hospitality. The second author was supported in part by NSF grant DMS-1001180. The two authors thank Kevin Ford, Andrew Granville, Paul Pollack, and Igor Shparlinski for some enlightening conversations. They also gratefully acknowledge the help of Zeb Engberg in improving and programming the sieve algorithm described in the Appendix.

References

[1] W. R. Alford, A. Granville, and C. Pomerance, *There are infinitely many Carmichael numbers*, Ann. of Math. 140 (1994), 703–722.

[2] W. D. Banks, J. B. Friedlander, F. Luca, F. Pappalardi, and I. E. Shparlinski, *Coincidences in the values of the Euler and Carmichael functions*, Acta Arith. 122 (2006), 207–234.

[3] E. R. Canfield, P. Erdős, and C. Pomerance, *On a problem of Oppenheim concerning "factorisatio numerorum"*, J. Number Theory 17 (1983), 1–28.

[4] P. Erdős, *On the normal number of prime factors of $p-1$ and some related problems concerning Euler's ϕ-function*, Quart. J. Math. Oxford Ser. 6 (1935), 205–213.

[5] P. Erdős and C. Pomerance, *On the normal number of prime factors of $\phi(n)$*, Rocky Mountain J. Math. 15 (1985), 343–352.

[6] P. Erdős, C. Pomerance, and E. Schmutz, *Carmichael's lambda function*, Acta Arith. 58 (1991), 363–385.

[7] P. Erdős and S. S. Wagstaff Jr., *The fractional parts of the Bernoulli numbers*, Illinois J. Math. 24 (1980), 104–112.

[8] K. Ford, *The distribution of totients*, Ramanujan J. 2 (1998), 67–151 (updated version at arXiv:1104.3264).

[9] J. B. Friedlander and F. Luca, *On the value set of the Carmichael λ-function*, J. Austral. Math. Soc. 82 (2007), 123–131.

[10] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.

[11] P. Kurlberg, J. C. Lagarias, and C. Pomerance, *Product-free sets with high density*, Acta Arith. 155 (2012), 163–173.

[12] F. Luca and C. Pomerance, *On composite integers n for which $\varphi(n) \mid n-1$*, Bol. Soc. Mat. Mexicana 17 (2011), 13–21.

[13] H. Maier and C. Pomerance, *On the number of distinct values of Euler's ϕ-function*, Acta Arith. 49 (1988), 263–275.
[14] N. M. Timofeev, *The Hardy–Ramanujan and Halász inequalities for shifted primes*, Math. Notes 57 (1995), 522–535; transl. of: Mat. Zametki 57 (1995), 747–764.

Florian Luca
Mathematical Institute, UNAM Juriquilla
Juriquilla, 76230 Santiago de Querétaro
Querétaro de Arteaga, México
and
School of Mathematics
University of the Witwatersrand
P.O. Box Wits 2050
Johannesburg, South Africa
E-mail: fluca@matmor.unam.mx

Carl Pomerance
Department of Mathematics
Dartmouth College
Hanover, NH 03755–3551, U.S.A.
E-mail: carl.pomerance@dartmouth.edu

Received on 23.1.2013 (7430)