Publications mathématiques de Besançon
Algèbre et théorie des nombres

Barry S. Fagin

Minimal idempotency, partial idempotency, search heuristics and constructive algorithms for idempotent integers

2024, p. 7-21.
https://doi.org/10.5802/pmb.53

Publication éditée par le laboratoire de mathématiques de Besançon, UMR 6623 CNRS/UFC

Les Publications mathématiques de Besançon *sont membres du Centre Mersenne pour l'édition scientifique ouverte*
http://www.centre-mersenne.org/
e-ISSN : 2592-6616

Publications mathématiques de Besançon – 2024, 7-21

MINIMAL IDEMPOTENCY, PARTIAL IDEMPOTENCY, SEARCH HEURISTICS AND CONSTRUCTIVE ALGORITHMS FOR IDEMPOTENT INTEGERS

by

Barry S. Fagin

Abstract. — Previous work established the set of square-free integers n with at least one factorization $n = \bar{p}\bar{q}$ for which $\bar{p}$ and $\bar{q}$ produce valid RSA keys, whether they are prime or composite. These integers are exactly those with the property $\lambda(n = \bar{p}\bar{q}) \mid (\bar{p}-1)(\bar{q}-1)$, where λ is the Carmichael totient function. We refer to these integers as *idempotent*, because $\forall\ a \in Z_n, a^{k(\bar{p}-1)(\bar{q}-1)+1} \equiv_n a$ for any positive integer k. This set includes the semiprimes and the Carmichael numbers, but is not limited to them. Numbers in this last category have not been previously analyzed in the literature.
We discuss the structure of idempotent integers here, and present heuristics to assist in finding them. We introduce the notions of *partial idempotency* and *minimal idempotency*, give appropriate definitions for them, and present preliminary results.

Résumé. — Un travail antérieur décrit l'ensemble des entiers n sans facteur carré admettant au moins une factorisation $n = \bar{p}\bar{q}$ pour laquelle $\bar{p}$ et $\bar{q}$ produisent des clés RSA valides, qu'ils soient premiers ou non. Ces entiers sont exactement ceux jouissant de la propriété que $\lambda(n = \bar{p}\bar{q}) \mid (\bar{p}-1)(\bar{q}-1)$, où λ est la fonction de Carmichael. Nous appelons ces entiers *idempotents*, parce que $\forall\ a \in Z_n, a^{k(\bar{p}-1)(\bar{q}-1)+1} \equiv_n a$ pour tout entier positif k. Cet ensemble inclut les nombres semi-premiers et les nombres de Carmichael, mais n'est pas seulement composé de ceux-ci. Les nombres de cette dernière catégorie n'ont pas encore été analysés dans la littérature.
Dans cet article nous discutons la structure des entiers idempotents et présentons des heuristiques pour aider à les trouver. Nous introduisons les notions de *partiellement idempotent* et d'*idempotence minimale*, en donnons des définitions appropriées et présentons des résultats préliminaires.

1. Background and definitions

In [3], we introduced the notion of *idempotent* integers, the set of square-free integers n with at least one factorization $n = \bar{p}\bar{q}$ such that $\lambda(n) \mid (\bar{p}-1)(\bar{q}-1)$, where λ is the Carmichael totient function. We refer to these integers as idempotent because $\forall\ a \in Z_n, a^{k(\bar{p}-1)(\bar{q}-1)+1} \equiv_n a$ for any positive integer k. We showed that these integers are exactly those for which $\bar{p}$ and $\bar{q}$ generate valid keys in the 2-prime RSA protocol [13], regardless of whether they are prime or composite. This set was initially known to include the semiprimes [13], and later the

Carmichael numbers [9]. [3] showed it is not limited to them, and gave the definition for the complete set.

While only the semiprimes have useful cryptographic properties [11], idempotent integers deserve study in their own right as they lie at the border of hard problems in number theory and computer science. In particular, we propose that for every composite square-free integer $\bar{p}$ there exists a composite square-free integer $\bar{q_c}$ such that $n = \bar{p}\bar{q_c}$ is idempotent. This suggests there exist infinitely many "ordinary" (p, q) pairs (composite and non-Carmichael) that can be used to generate correct keys using the RSA 2-prime protocol.

Let $n = p_1 p_2 \dots p_m$ be a square-free integer, with $i < j \iff p_i < p_j$. Let $a_i = p_i - 1$. We will call a_i the predecessor of p_i and p_i the successor of a_i. It is a known property of λ that $\lambda(n) = \mathrm{lcm}(a_1, a_2, \dots, a_m)$, where lcm denotes the least common multiple. We will write λ instead of $\lambda(n)$ when the meaning is clear. We write $\bar{p}$ as shorthand for $\prod p_i$.

A factorization of n is a 2-partition of the set of p_i's into products $\bar{p}$ and $\bar{q}$. An idempotent factorization is a 2-partition of $n = \bar{p}\bar{q}$ for which $\lambda \mid (\bar{p}-1)(\bar{q}-1)$. We will refer to an integer n that admits an idempotent factorization as being idempotent itself when the meaning is clear.

All semiprimes (products of two primes) are trivially idempotent. We do not consider them further here.

Any square-free integer n,with m factors, has $\binom{m}{1} = m$ factorizations of the form $\bar{p} = p_i, \bar{q} = p_{j\neq i}$, $\binom{m}{2}$ factorizations of the form $\bar{p} = p_i p_j, \bar{q} = p_{k\neq i,j}$, and so forth. Each partition corresponds to a single equation in $n, \bar{p}$ and $\bar{q}$ that represents a possible idempotent factorization. We refer to these as single-factor partitions/equations/factorizations, double-factor, etc. We call idempotent single-factor partitions semi-composite factorizations of n. All other factorizations are fully composite.

The first eight square-free n with three or more factors and fully composite idempotent factorizations are shown in Table 1 [3].

Table 1. The first 8 integers with fully composite idempotent factorizations

n	factorization	partition $n = \bar{p}\bar{q}$	λ
210	2·3·5·7	10·21	12
462	2·3·7·11	21·22	30
570	2·3·5·19	10·57	36
1155	3·5·7·11	21·55	60
1302	2·3·7·31	6·217	60
1330	2·5·7·19	10·133	36
1365	3·5·7·13	15·91	12
1785	3·5·7·17	21·85	48

The smallest integer with two fully composite idempotent partitions is 2730, when factored into 10·273 and 21·130. The complete list of all $n < 2^{27}$ with fully composite idempotent factorizations is available at [7].

2. Cumulative statistics on factorizations

Cumulative statistics for idempotent factorizations for $n < 2^{30}$ are shown below. R_{sf} indicates the ratio of numbers with idempotent factorizations to the total number of candidates n, those

square-free numbers with strictly more than two factors. R_N indicates the ratio to all n in the indicated interval. The first entry in R_{cpu} is the computation time on the author's computer for the indicated interval. Remaining entries are the ratio of computation time of the current interval to the previous interval. For example, in Table 2, it took 2.7 seconds to check all integers up to 2^{15}, 2.7*11.3 = about 30 seconds to check all integers up to 2^{18}, 30*10.6 = about five minutes to check up to 2^{21}, and so forth. An entry of the form $i : j$ in row with #factors = F indicates there are j integers $< 2^{30}$ with F prime factors and i idempotent factorizations.
All answers are rounded to the indicated number of decimals. We ignore order when counting factorizations. These results were obtained on a HP 640 G4 notebook running Windows 10 Enterprise with a 7th generation Intel ®Core ™i5-7200U CPU@2.5 GHz and 8G of memory. The code was written in Python 3.7 using Robert Campbell's numbthy.py library, augmented with additional code and library routines by the author.

TABLE 2. Proportion of integers with idempotent factorizations

$\max n$	2^{12}	2^{15}	2^{18}	2^{21}	2^{24}	2^{27}	2^{30}
R_{sf}	.61	.37	.28	.21	.17	.13	.11
R_N	.09	.09	.08	.07	.06	.05	.04
R_{cpu}	-	2.7s	11.3	10.6	13.3	9.8	10.4

TABLE 3. Factor distribution of idempotent factorizations $< 2^{30}, < 8$ factorizations

# factors	0	1	2	3	4	5	6	7
3	184510285	34215577	0	15189	0	0	0	0
4	132479584	11347214	4448	15678	28	235	0	315
5	50515758	1733232	6530	13743	93	599	1	441
6	10004651	242377	6143	6906	167	586	12	302
7	931270	35473	2994	1597	124	286	22	102
8	29211	2956	477	158	39	43	5	6
9	99	28	7	2	1	0	1	1

TABLE 4. Factor distribution of idempotent factorizations $< 2^{30}, \geq 8$ factorizations

# factors						
5	8:2	9:6	11:18	15:2		
6	8:3	9:10	11:31	15:20		
7	8:3	9:5	10:1	11:24	15:3	31:1
8	8:1	9:2	11:4			

3. Constructing idempotent products for a given $\bar{p}$

Instead of fixing n and determining whether or not it admits idempotent factorizations, we may instead fix $\bar{p}$ and search for a $\bar{q}$ that forms an idempotent product. The following function

TABLE 5. Proportion of integers with semi-composite idempotent factorizations

$\max n$	2^{12}	2^{15}	2^{18}	2^{21}	2^{24}	2^{27}	2^{30}
R_{sf}	.49	.36	.27	.21	.16	.13	.11
R_N	.010	.09	.08	.07	.06	.05	.04
R_{cpu}	.19s	9.9	11.9	13.3	9	9.6	10.3

TABLE 6. Distribution of semi-composite idempotent factorizations $< 2^{30}$

# factors	0	1	2	3	4	5
3	184510285	34215577	0	15189	0	0
4	132498612	11331335	16992	248	315	0
5	50583104	1676008	10740	550	21	2
6	10083389	175380	2276	149	14	0
7	963273	8502	123	6	0	1
8	32776	126	0	0	0	0
9	139	0	0	0	0	0

TABLE 7. Proportion of integers with fully-composite idempotent factorizations

$\max n$	2^{12}	2^{15}	2^{18}	2^{21}	2^{24}	2^{27}	2^{30}
R_{sf}	.0183	.0149	.0088	.0050	.0025	.0013	.0006
R_N	.0037	.0038	.0026	.0016	.0009	.0005	.0002
R_{cpu}	-	1.5s	11.9	10.0	9.5	10.5	10.7

TABLE 8. Factor distribution of fully composite idempotent factorizations $< 2^{30}, < 8$ factorizations

# factors	0	1	2	3	4	5	6	7
4	143809069	37868	250	315				
5	52184272	79032	5960	645	471	20	5	18
6	10172969	78072	8464	939	432	237	33	12
7	938127	29014	3634	615	272	140	60	9
8	29269	2902	507	126	55	25	11	11
9	99	28	7	2	1	0	1	1

will be useful:

$$D(x) = \frac{\lambda(x)}{\gcd(\lambda(x), x-1)}$$

$D(x)$ is the product of all the factors of $\lambda(x)$ that are not in $x-1$, and is therefore the smallest number containing them all. We have $1 \leq D(x) \leq \lambda(x)$. $D(x) = 1 \iff x$ is prime or a Carmichael number, $D(x) = \lambda(x) \iff \gcd(\lambda(x), x-1)) = 1$.
We need the following theorem:

Theorem 3.1. — *Let $\bar{p}, \bar{q}$ be square-free and coprime. There exist integers $k_p, k_q \geq 1$ s.t. $\bar{p} = k_p D(\bar{q}) + 1, \bar{q} = k_q D(\bar{p}) + 1 \iff n = \bar{p}\bar{q}$ is an idempotent factorization.*

TABLE 9. Factor distribution of fully composite idempotent factorizations $< 2^{30}, \geq 8$ factorizations

# factors						
5	10:2					
6	8:25	9:5	11:13	12:6	13:1	
7	8:7	9:14	10:9	12:3	26:1	
8	8:1	9:1	10:2	11:2	12:3	26:1

Proof. —

$\Longrightarrow$. — Assume $\bar{p}, \bar{q}$ are square-free and coprime, $\bar{p} = k_p D(\bar{q}) + 1, \bar{q} = k_q D(\bar{p}) + 1$ for some positive integers k_p, k_q. Consider $\lambda(\bar{p}\bar{q})$. Since both $\bar{p}$ and $\bar{q}$ are square-free and coprime, we have

$$\lambda(\bar{p}\bar{q}) = \text{lcm}(\lambda(\bar{p}), \lambda(\bar{q})) = L$$

Without loss of generality, assume L contains all the factors of $\lambda(\bar{p})$ and a subset (possibly improper) of all the factors of $\lambda(\bar{q})$. Consider first the prime factors of $\lambda(\bar{p})$. Every one is either contained in $\bar{p}-1$ or $D(\bar{p})$. If it is contained in $\bar{p}-1$, it will be contained in $(\bar{p}-1)(\bar{q}-1)$. If it is contained in $D(\bar{p})$, it will be contained in $k_q D(\bar{p}) = \bar{q}-1$, and therefore it will be contained in $(\bar{p}-1)(\bar{q}-1)$. Similar reasoning applies to the (possibly improper) subset of the factors of $\lambda(\bar{q})$: They are either contained in $\bar{q}-1$, or in $D(\bar{q})$ and therefore in $k_p D(\bar{q}) = \bar{p}-1$. Thus all the factors of L are contained in $(\bar{p}-1)(\bar{q}-1)$, and therefore $n = \bar{p}\bar{q}$ is an idempotent factorization.

$\Longleftarrow$. — Assume $\bar{p}, \bar{q}$ are square-free and coprime, $n = \bar{p}\bar{q}$ an idempotent factorization. We have

$$\lambda(\bar{p}\bar{q}) | (\bar{p}-1)(\bar{q}-1) \rightarrow \text{lcm}(\lambda(\bar{p}), \lambda(\bar{q})) \mid (\bar{p}-1)(\bar{q}-1)$$

Again without loss of generality, assume L contains all the factors of $\lambda(\bar{p})$ and a subset (possibly improper) of all the factors of $\lambda(\bar{q})$. Consider the set of prime factors of $\lambda(\bar{p}) \notin (\bar{p}-1)$. Because $\bar{p}\bar{q}$ is idempotent, every such factor must be contained in $\bar{q}-1$. So $\bar{q}-1 \equiv_{m_i} 0$ for every prime factor m_i of $\lambda(\bar{p}) \notin (\bar{p}-1)$. By the construction of $D(\bar{p})$ and the Chinese Remainder Theorem, $\bar{q}-1 \equiv_{D(\bar{p})} 0 \rightarrow \bar{q} \equiv_{D(\bar{p})} 1 \rightarrow \bar{q} = k_q D(\bar{p}) + 1$ for some $k_q \geq 1$. Similar results apply for the (possibly improper) subset of the factors of $\lambda(\bar{q}) \notin (\bar{q}-1)$. □

The definition of $D(x)$ and the theorem above yield the following lemma:

Lemma 3.2. — *Let $\bar{p}$ be square-free. Let q, $\bar{q}_c$, $\bar{q}$ be square-free and coprime to $\bar{p}$.*

a. *If $\bar{p}$ is a prime or a Carmichael number, it is idempotent with and only with q prime, or $\bar{q}_c$ a Carmichael number, or $\bar{q}_{\not c}$ a composite non-Carmichael number with $\bar{p} \equiv_{D(\bar{q})} 1$. Since $D(6) = 2$, all primes and Carmichael numbers $\bar{p} \geq 5$ are idempotent with $\bar{q} = 6$.*

b. *Otherwise $\bar{p}$ is idempotent with and only with $q \equiv_{D(\bar{p})} 1$ and either q prime, or $q = \bar{q}_c$ a Carmichael number, or $\bar{q}_{\not c}$ a composite non-Carmichael number such that $\bar{p} \equiv_{D(\bar{q})} 1$.*

This gives a simple algorithm for finding an idempotent $\bar{q}$ for a given a square-free $\bar{p}$:
The above procedure may be iterated to produce sequences of n that admit either semi-composite or fully composite idempotent factorizations, by starting with $\bar{p}$ square-free, finding

Algorithm 1: Finding an idempotent $\bar{q}$, given a square-free integer $\bar{p}$
Input: A square-free integer $\bar{p}$
Output: The first possible idempotent $\bar{q}$

```
Calculate D(p̄)
k ← 1
q̄ ← kD(p̄) + 1
Calculate D(q̄)
while not (q̄ is prime or q̄ is a Carmichael number or p̄ ≡ 1 (mod D(q̄))) do
    k ← k + 1
    q̄ ← kD(p̄) + 1
    Calculate D(q̄)
return q̄
```

a suitable $\bar{q}$, calculating $n = \bar{p}\bar{q}$, setting $\bar{p} = n$ and then repeating. This may be used to construct an idempotent n with a specific number of factors.
For any prime p, there are an infinite number of Carmichael numbers coprime to it, so there are an infinite number of semi-composite factorizations. Similarly, for any Carmichael number, there is at least one Carmichael number coprime to it, so there are an infinite number of fully composite factorizations. Strong impostors (introduced in [2], list available online at [8]) s have $D(s) = 2$ which implies $p \equiv_{D(s)} 1$ for any odd prime p, and there are an infinite number of strong impostors, so there are an infinite number of semi-composite factorizations that do not involve Carmichael numbers.

3.1. Non-Carmichael fully composite idempotent factorizations. — We see from the above lemma that Carmichael numbers are the composite numbers that most easily form idempotent factorizations with other composite numbers. A Carmichael number C is idempotent with any coprime prime p, any coprime Carmichael number, and any composite square-free coprime number $\bar{q}$ such that $C \equiv_{D(\bar{q})} 1$.
For those square-free $\bar{p}$ for which the smallest $\bar{q}$ that forms an idempotent product is a Carmichael number $\bar{q}_c$, we might ask if there is a $\bar{p}\bar{q}$-idempotent product with $\bar{q} = \bar{q}_{\complement}$ where $\bar{q}_{\complement}$ is not a Carmichael number. We call an idempotent factorization $n = \bar{p}\bar{q}$ *pure* if $\bar{p}$ and $\bar{q}$ are composite and non-Carmichael numbers. We might ask if for any composite square-free $\bar{p}$ there exists a $\bar{q}_{\complement}$ for which $\bar{p}\bar{q}_{\complement}$ is a pure idempotent factorization.
If $\bar{p}-1$ is prime, the required $\bar{q}_{\complement}$ must be very special indeed. It must be square-free, coprime to $\bar{p}$, and of the form $kD(\bar{p}) + 1$ as previously noted. (Note if $\bar{p} - 1$ is prime, $D(\bar{p}) = \lambda(\bar{p})$, its maximum value). Additionally, $\lambda(\bar{q}_{\complement})$ must have exactly one prime factor not contained in $\bar{q}_{\complement} - 1$, and that factor must be $\bar{p} - 1$.
Despite these constraints, we conjecture that such a $\bar{q}_{\complement}$ can always be found, and we have verified this conjecture for all $\bar{p} < 2^{15}$ [6]. The upper limit is smaller than previous calculations because excluding Carmichael numbers can significantly increase the search time. For $\bar{p} = 7214, D(\bar{p}) = \lambda(\bar{p}) = 3606$, and almost 500 billion candidates of the form $kD(\bar{p}) + 1$ must be examined until the required non-Carmichael $\bar{q}_{\complement} = 1772915178061$ is found ($k = 491657010, \bar{q}_{\complement} - 1 = 2^2 \cdot 3^2 \cdot 5 \cdot 13 \cdot 29^2 \cdot 601 \cdot 1499, \lambda(\bar{q}_{\complement}) = 489474180 = 2^2 \cdot 3^2 \cdot 5 \cdot 13 \cdot 29 \cdot 7213$).

3.2. Cumulative statistics. — The table below shows $D(\bar{p})$, the smallest number q, closest number q, smallest composite $\bar{q}$, smallest Carmichael number $\bar{q}_c$, and the smallest non-Carmichael $\bar{q}_{\complement}$ that form idempotent products for the first 18 square-free $\bar{p}$ (recall that semiprime idempotent factorizations are excluded). We show up to $\bar{p} = 30$ to emphasize how much larger the required $\bar{q}_{\complement}$ can be when $\bar{p} - 1$ is prime. We will say more about this below.

TABLE 10. Idempotent factorizations with $\bar{p}$ fixed

$\bar{p}$	D	q	closest q	$\bar{q}$	$\bar{q}_c$	$\bar{q}_{\complement}$
2	1	561	561	561	561	-
3	1	91	91	91	1105	91
5	1	6	6	6	561	6
6	2	5	5	217	1105	217
7	1	6	6	6	561	6
10	4	13	13	21	561	21
11	1	6	15	6	1105	6
13	1	6	14	6	561	6
14	6	13	13	265	1105	265
15	2	7	13	91	1729	91
17	1	6	15	6	1729	6
19	1	6	21	6	561	6
21	3	10	19	10	1105	10
22	10	21	21	21	2821	21
23	1	6	15	6	561	6
26	12	37	37	217	6601	217
29	1	6	30	6	561	6
30	4	13	29	1729	1729	4537

Cumulative statistics for the smallest number q, composite $\bar{q}$, Carmichael $\bar{q}_c$ and non-Carmichael $\bar{q}_{\complement}$ are shown in the tables below. $\max_k$ indicates the largest k value needed to form the required idempotent product for a composite $\bar{p}$, $@\bar{p}$ the corresponding $\bar{p}$ value where it occurs. $\max(q/\bar{p})$ and similar entries indicate the largest ratio of the appropriate q value to a composite $\bar{p}$ on the indicated interval, and again $@\bar{p}$ the corresponding $\bar{p}$ value where it occurs. R_c is the fraction of $\bar{q}$ values that are Carmichael numbers. R_{c16} is the fraction of $\bar{q}$ values that are the first 16 Carmichael numbers.

TABLE 11. Smallest number q

$\max n$	2^{15}	2^{18}	2^{21}	2^{24}	2^{27}
$\max_k$	60	98	130	196	258
$@\bar{p}$	20337	204297	736069	6083833	28547269
$\max(q/\bar{p})$	15	27.5	37.5	45.0	67.5
$@\bar{p}$	13558	199246	1773094	10934194	66708946
R_c	.0035	.0029	.0022	.0015	.0009
R_{cpu}	3s	13	12.3	12.5	11.1

TABLE 12. Smallest composite number $\bar{q}$

$\max n$	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	2^{15}
$\max_k$	102120	108360	2014320	12149568	55691880	217999188
$@\bar{p}$	998	1362	2270	7122	14754	24474
$\max(\bar{q}/\bar{p})$	50957.7	50957.7	401089.3	2023222.1	9278205.3	36324290.6
$@\bar{p}$	998	998	2270	7122	14754	24474
R_c	.1814	.1841	.1835	.1811	.1803	.1815
R_{cpu}	19.2s	3	12.2	7.3	10	5.6

TABLE 13. Smallest Carmichael $\bar{q}_c$

smallest $\bar{q}_c$				
$\max\ n$	2^{12}	2^{15}	2^{18}	2^{21}
R_{c16}	.55	.29	.13	.05
$\max C_n$	C_{2253}	C_{22323}	C_{353833}	$C_{1399648}$
$@\bar{p}$	2374	31366	258838	2058046
R_{cpu}	2.2s	2.9	9.2	67

TABLE 14. Smallest non-Carmichael $\bar{q}_{\not c}$

$\max n$	2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}
$\max_k$	114774	655200	1752744	15593760	63901950	648165996
$@\bar{p}$	230	354	642	1790	2082	5870
$\max(\bar{q}_{\not c}/\bar{p})$	21956.8	107349.2	725767.0	725767.0	12500690.8	245760351.8
$@\bar{p}$	230	354	734	734	3734	7214
R_{cpu}	12.6s	4.8	10	12.5	10	22.0

3.3. Using heuristics with relaxed constraints. — All the minimal q and $\bar{q}$ values discussed so far were obtained using brute force: Trying all possible values of $kD+1$ starting with $k=1$ until a $\bar{q}$ is found that meets the requirements. If instead of finding the smallest $\bar{q}$ that meets constraints, we simply wish to find any satisfying $\bar{q}$, other techniques can produce faster results. As mentioned previously, for example, computing D and then searching through a table of Carmichael numbers [12] to find one of the form $kD+1$ can find a composite $\bar{q}$ faster than brute force examination of all composite $\bar{q} = kD+1$ for sufficiently large $\bar{p}$. Similarly, we can limit the search to q or $\bar{q}$ values of the form $kD(\bar{p})+1$, which will by construction be coprime to $\bar{p}$.

Empirically, the most difficult $\bar{p}$ to find corresponding $\bar{q}_{\not c}$ all have $\bar{p}-1$ prime (note the disproportionately larger values for $\bar{p} = 6, 14, 30$ in Table 10 above). Satisfying $\bar{q}_{\not c}$ for larger $\bar{p}$ with prime predecessors can often be found faster by restricting the search space of possible $\bar{q}_{\not c} = kD+1$ to those for which $\lambda(\bar{q}_{\not c})$ will be guaranteed to have the required $\bar{p}-1$ as a factor. This is done by exploiting a known property of the lambda function: If $\bar{q}$ contains a prime factor of the form $p'_j = jp'+1$ with $j \geq 1$, then $\lambda(\bar{q})$ contains p' as a factor.

This suggests the following heuristic. Let $p' = \bar{p}-1$, let $p'_j = jp'+1$. To find the desired $\bar{q}_{\not c}$, loop through successive values of j up to some limit $\max_j$, and for each prime p'_j check only those $\bar{q}_{\not c} = kD+1 \equiv 0 \bmod p'_j$ up to some limit $\max_k$.

Because $kD+1$ is an arithmetic progression, once the first k_0 is found with $kD+1 \equiv 0 \bmod p'_j$, the remaining k are all of the form $k_0, k_0 + p'_j, k_0 + 2p'_j \ldots$ up to $\max_k$. This algorithm can find a satisfying non-Carmichael $\bar{q}$ much faster than brute force, depending on the specific values of j and k for a given $\bar{p}$. Table 15 shows the most troublesome $\bar{p}$ for which $\bar{q}_{\mathfrak{c}}$ was found using this technique.

The above heuristic works well at finding a suitable $\lambda(\bar{q}_{\mathfrak{c}})$. However, we know of no way to determine the p'_j that will produce the k associated with the smallest satisfying $\lambda(\bar{q}_{\mathfrak{c}})$. We can only try the k values associated with a given p'_j until a satisfying $\bar{q}$ is found or until a pre-specified limit is exceeded. Thus this technique does not guarantee that the $\lambda(\bar{q}_{\mathfrak{c}})$ it finds, if it finds one at all, will be the smallest. A brute force examination of all $kD+1$ from $k=1$ remains the only algorithm known to the author that guarantees discovery of the smallest idempotent non-Carmichael $\lambda(\bar{q}_{\mathfrak{c}})$ for a given $\bar{p}$.

Table 14 required over 300 hours of CPU time and is worth a closer look (recall each R_{cpu} entry must be multiplied by all entries to its left to obtain the total CPU time in seconds). Figure 1 shows the amount of time required to compute the smallest idempotent non-Carmichael $\bar{q}_{\mathfrak{c}}$ at intervals of 100. We see that the time for each interval varies widely, due to the existence of certain $\bar{p}$ for which the required k that yields non-Carmichael $\bar{q}_{\mathfrak{c}} = kD+1$ is unusually large.

TABLE 15. $\bar{p} < 2^{15}$ with $\bar{q}_{\mathfrak{c}} = kD+1, k \geq 10^8$, found with heuristic

$\bar{p}$	$\bar{q}_{\mathfrak{c}}$	D=$\lambda(\bar{p})$	k
12378	47,363,331,572,281	2062	22,969,607,940
16190	25,323,514,587,001	3236	7,825,560,750
18230	105,036,969,940,001	3644	28,824,635,000
19478	40,617,829,334,017	9738	4,171,064,832
20442	7,290,062,243,101	3406	2,140,358,850
21122	120,882,182,886,001	5162	23,417,703,000
21738	10,376,471,976,571	3622	2,864,845,935
22398	13,587,398,312,041	3732	3,640,781,970
23082	12,415,474,130,401	3846	3,228,152,400
23174	353,912,754,896,521	11586	30,546,586,820
23538	15,418,820,683,729	3922	3,931,366,824
26022	13,619,438,243,281	4336	3,141,014,355
27282	50,255,219,357,281	4546	11,054,821,680
27998	197,483,339,983,441	13998	14,107,968,280
28790	19,315,677,464,137	5756	3,355,746,606
29130	4,153,082,916,601	1940	2,140,764,390
30518	31,896,923,275,945	15258	2,090,504,868
30594	554,147,100,501,913	5098	108,698,921,244
31034	20,970,363,294,721	7598	2,759,984,640
31190	211,342,409,502,841	6236	33,890,700,690
32030	40,877,987,645,089	6404	6,383,196,072
32234	19,001,029,407,361	7910	2,402,152,896

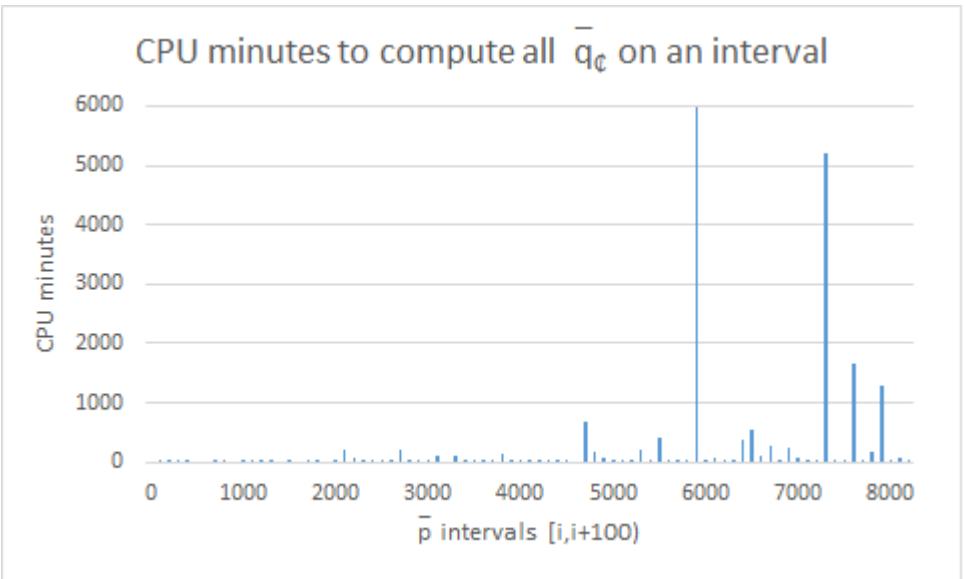

FIGURE 1. CPU minutes to compute all $\bar{q}_{\text{¢}}$ on an interval

4. Partial idempotency

Idempotent factorizations $n = \bar{p}\bar{q}$ have the property that $(\bar{p}, \bar{q})$ behave like primes in the RSA protocol, in that messages will be encrypted and decrypted correctly for any (e, d) s.t. $ed \equiv_{(\bar{p}-1)(\bar{q}-1)} 1$. Factorizations that are not idempotent, however, may still exhibit idempotency for certain (e, d). In particular, if in addition to $ed \equiv_{(\bar{p}-1)(\bar{q}-1)} 1$ we have $ed \equiv_{\lambda(n)} 1$, then by similar reasoning in the proof of Section 2, we have $a^{ed} \equiv_n a \; \forall \; a \in Z_n$.

Rather than consider idempotency an all-or-nothing quantity, we may regard the idempotency of a factorization $n = \bar{p}\bar{q}$ as a number between 0 and 1. Let $\phi'(\bar{p}, \bar{q}) = (\bar{p} - 1)(\bar{q} - 1)$ denote the *pseudototient* function, used when p and/or q are not necessarily prime. We define a factorization's *idempotency ratio* as the number of valid (e, d) pairs that are inverses mod $\phi'(\bar{p}, \bar{q})$ and mod $\lambda(\bar{p}\bar{q})$, divided by the number of valid (e, d) pairs that are inverses mod $\phi'(\bar{p}, \bar{q})$.

Let $\phi'(\bar{p}, \bar{q}) = (\bar{p} - 1)(\bar{q} - 1)$. Using Iverson brackets [10], we define the idempotency ratio $R(\bar{p}, \bar{q})$ of a factorization $n = \bar{p}\bar{q}$ as

$$R(\bar{p}, \bar{q}) = \frac{\sum_{e=0}^{n-1} [\gcd(e, \phi'(\bar{p}, \bar{q})) = 1][ed \equiv_{\lambda(\bar{p},\bar{q})} 1] \text{where } d = e^{-1} \bmod \phi'(\bar{p}, \bar{q})}{\sum_{e=0}^{n-1} [\gcd(e, \phi'(\bar{p}, \bar{q})) = 1]}$$

Idempotent factorizations $n = \bar{p}\bar{q}$ as defined previously have $R(\bar{p}, \bar{q}) = 1$, since $\lambda(\bar{p}\bar{q}) \mid \phi'(\bar{p}, \bar{q}) \rightarrow (ed \bmod \phi'(\bar{p}, \bar{q}) \equiv 1 \iff ed \equiv_{\lambda(\bar{p}\bar{q})} 1)$, making the numerator and denominator equal. We define the idempotency ratio of an integer n as the sum of the idempotency ratios of all its factorizations divided by the number of its factorizations.

The idempotency ratios for the factorizations of the first four square-free n with > 2 factors are shown in Table 16.

Good approximations exist for the denominator in the definition of R above. The formula for the number of integers on an interval coprime to some n is known [1], but it is cumbersome

Table 16. Idempotency ratios

n	$\bar{p}$	$\bar{q}$	R
30	2	15	0.538
30	3	10	0.5
30	5	6	1
42	2	21	0.235
42	6	7	1
42	3	14	0.263
66	3	22	0.211
66	6	11	1
66	2	33	0.212
70	5	14	0.094
70	2	35	0.121
70	7	10	0.522

and needlessly complex for our purposes. Let $\Phi' = \phi'(\bar{p}, \bar{q})$. There are $\phi(\Phi')$ integers between 0 and Φ' coprime to Φ', where $\phi(x)$ is Euler's totient function. Call this quantity Φ. We may write n as $Q\Phi' + r$, where $Q = \lfloor n/\Phi' \rfloor$ and r is the remainder of integers in the sum when divided by Φ'. We have $n \equiv r \bmod \Phi'$, so on the assumption the totatives are proportionally distributed throughout the integers in r, the denominator may be approximated by

$$\text{denom} \;=\; \Phi \cdot Q + (\Phi/\Phi') \cdot (n \bmod \Phi')$$

Computer calculations indicate this assumption to be greater than 99% accurate even for small $n = \bar{p}\bar{q}$, with accuracy increasing with n.
Simplifying the numerator is more difficult, as we will see shortly.

4.1. D_I and the idempotency ratio. — The more factors $\lambda(\bar{p}\bar{q})$ and $\phi'(\bar{p}, \bar{q})$ have in common, the higher $R(\bar{p}, \bar{q})$ should be. We define $D_I(\bar{p}, \bar{q})$ as

$$D_I(\bar{p}, \bar{q}) = \lambda(\bar{p}\bar{q}) / \gcd(\phi'(\bar{p}, \bar{q}), \lambda(\bar{p}\bar{q}))$$

$D_I(\bar{p}, \bar{q}) = 1$, its minimal value, when and only when $\bar{p}\bar{q}$ is a fully idempotent factorization. $\bar{p}$ and $\bar{q}$ cannot both be even, otherwise n would not be square free. Thus $\phi'(\bar{p}, \bar{q})$ is always even. Since $\lambda(\bar{p}\bar{q})$ is always even, their smallest gcd is 2, and therefore the largest value of D_I is $\lambda(\bar{p}\bar{q})/2$.

4.2. Finding (e, d) pairs that lend idempotency. — Calculating idempotency ratios involves finding all (e, d) from the "bottom up": Examining all e relatively prime to $\phi'(\bar{p}, \bar{q})$ from 0 to $n = \bar{p}\bar{q} - 1$, finding the corresponding inverse d, and determining if $ed \equiv_{\lambda(\bar{p}\bar{q})} 1$. This method can be used to construct an (e, d) pair that lends idempotency to a given factorization $n = \bar{p}\bar{q}$: iterate through all valid (e, d) pairs until one with the desired property is found.
Alternatively, (e, d) pairs can be discovered from the "top down". From number theory, we have

$$x \equiv_{m_1} a, x \equiv_{m_2} a \iff x \equiv_{\text{lcm}(m_1, m_2)} a.$$

Letting $a = 1, m_1 = \phi'(\bar{p}, \bar{q}), m_2 = \lambda(\bar{p}\bar{q})$, we see the desired (e, d) pair has the property $ed \equiv 1 \bmod L$ where $L = \text{lcm}(\phi'(\bar{p}, \bar{q}), \lambda(\bar{p}\bar{q}))$. Thus given $n = \bar{p}\bar{q}$, we may find a desired (e, d) pair in the following way:

For a given $(\bar{p}, \bar{q})$, calculate $L = \mathrm{lcm}(\phi'(\bar{p}, \bar{q}), \lambda(\bar{p}\bar{q}))$. For $k = 1, 2, \ldots$, calculate $kL + 1$ until a value is found that can be factored into two numbers coprime to $\phi'(\bar{p}, \bar{q})$. These two numbers by construction will be an (e, d) pair that lends idempotency to $n = \bar{p}\bar{q}$. However, as we will see in the next section, while the trivial case of $(e, d) = (1, 1)$ will always lend idempotency, there are cases where a nontrivial solution does not exist.

4.3. Existence conditions and minimally idempotent factorizations of n. — As shown above, the (e, d) pairs that lend idempotency to a factorization of $n = \bar{p}\bar{q}$ are exactly those for which $ed \equiv_L 1$, where $L = \mathrm{lcm}(\phi'(\bar{p}, \bar{q}), \lambda(\bar{p}\bar{q}))$. The desired (e, d) are then exactly those solutions to the 2-variable system of nonlinear modular equations $ed \equiv_{m_1} 1, ed \equiv_{m_2} 1 \ldots ed \equiv_{m_j} 1$, where $m_1, m_2 \ldots m_j$ are the prime power factors of L.
Determining whether or not such systems have solutions and calculating their exact number are known NP-hard problems. Thus simple, efficient calculations of idempotency ratios are likely to prove elusive.
The trivial pair $(e, d) = (1, 1)$ lends idempotency to any factorization of n, so by the previous definition the idempotency ratio of any factorization is never zero. We might inquire, however, if factorizations exist for which the resulting nonlinear system of modular equations has no non-trivial solution. In fact, the answer is yes. We refer to these factorizations as *minimally idempotent*.
The first sixteen minimally idempotent factorizations are shown in Table 17.

Table 17. Minimally idempotent factorizations

n	$\bar{p}$	$\bar{q}$
154	2	77
470	2	235
658	2	329
710	2	355
782	2	391
994	2	497
1034	2	517
1222	2	611
1310	2	655
1474	2	737
1798	2	899
1833	3	611
1886	2	943
1974	14	141
2134	2	1067
2338	2	1169

Most minimally idempotent factorizations have maximal D_I, but not all. Of the sixteen above, only 1833=3·611 does not. Factorizations also exist with maximal D_I that are not minimally idempotent.
As shown by the entry for 1974 in the table above, fully composite minimally idempotent factorizations also exist, although they are considerably rarer. The first sixteen are shown below:

TABLE 18. Fully composite minimally idempotent factorizations

n	$\bar{p}$	$\bar{q}$
1974	14	141
3390	10	339
5170	55	94
5170	22	235
6834	6	1139
8130	15	542
8178	6	1363
10542	14	753
13746	6	2291
14514	118	123
16626	6	2771
16638	118	141
16638	6	2773
17358	22	789
18894	134	141
19722	38	519

5170 is the smallest integer with more than one minimally idempotent factorization, both of which are fully composite.
All entries in this table have maximal D_I. It is an open question whether this holds for all fully composite minimally idempotent factorizations. To date, the author has found no counterexamples.

4.4. Cumulative statistics. — Cumulative statistics for idempotency ratios and minimally idempotent factorizations are shown below. IR_{avg} is the average integer idempotency ratio, IR_{max} the number of maximally idempotent integers, $IR_{max<1}$ the maximum idempotency ratio for an integer that is not maximally idempotent, @n is the n where it occurs. $IR_{min>0}$ is the minimum idempotency ratio for an integer that is not minimally idempotent, @n is the n where it occurs. MIP_j is the number of integers with j minimally idempotent factorizations. No integers where all factorizations are minimally idempotent are known.

TABLE 19. Idempotency ratio analysis

max n	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}
IR_{avg}	.3085	.2697	.2398	.2116	.1868	.1651	.1466	.1307	.1171
$IR_{min>0}$	.0642	.0240	.0187	.0091	.0064	.0032	.0018	.0010	.0006
@n	782	1771	3619	4807	8463	16653	62968	119239	223579
R_{cpu}	1s	3	5.3	3.8	6	5.3	5.1	5.1	4.7

5. Conclusions and future work

We define the class of idempotent integers as those n which can be factored into $\bar{p}\bar{q}$ such that $\lambda(n) \mid (\bar{p}-1)(\bar{q}-1)$. This set includes the primes, semiprimes, and Carmichael numbers,

TABLE 20. Minimally idempotent factorization analysis

$\max n$	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}
R_{sf}	.0427	.0452	.0462	.0440	.0443	.0447	.0442
R_N	.0085	.0099	.0110	.0112	.0120	.0127	.0131
MIP_2	0	2	2	3	3	7	12
R_{cpu}	28s	4.29	5	4.9	4.8	4.8	4.8

TABLE 21. Fully composite minimally idempotent factorization analysis

$\max n$	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}
n with $FCMIP$	2	6	9	24	56	137	308
R_{sf}	.0150	.0169	.0105	.0116	.0116	.0125	.0125
R_N	.0005	.0007	.0005	.0007	.0009	.0010	.0012
$FCMIP_2$	0	1	1	2	2	6	9
R_{cpu}	9s	5.1	3.9	6.3	5.6	5.4	4.8

but is not unique to them. Those members that are not included in the first three classes, while lacking cryptographically useful properties, are worthy of study in their own right as they lie at the boundaries of hard problems in computer science and number theory. We have presented some examples above.

A combination of brute force and heuristics suggests there are infinitely many (p, q) pairs of composite square-free non-Carmichael numbers which will produce correct keys in the 2-prime RSA protocol. While such numbers have no cryptographic utility, the empirical results were a surprise to the author.

Some integers have the property that all their factorizations are idempotent; all their factorizations have an idempotency ratio of 1. We refer to these numbers as *maximally idempotent*. This is also class of numbers worth studying, one whose members have a unique structure that also suggests challenges and open problems. This is work in progress [5], [4].

6. Acknowledgements

The author thanks the anonymous referee, whose exceptionally careful review significantly improved the paper. He also thanks Dr Kurt Herzinger,and Dr Ian Pierce of the USAFA Department of Mathematical Sciences for their assistance. Finally, he is especially grateful to his students and his former Department of Computer and Cyber Sciences colleague, Dr Carlos Salazar, for asking interesting questions.

References

[1] P. ERDŐS, F. LUCA & C. POMERANCE, "On the Proportion of Numbers Coprime to a Given Integer", in *Anatomy of integers*, CRM Proceedings & Lecture Notes, vol. 46, American Mathematical Society, 2008, p. 47-64.

[2] B. FAGIN, "Composite Numbers That Give Valid RSA Key Pairs For Any Coprime p", *Information* **9** (2018), no. 9, article no. 216.

[3] ———, "Idempotent Factorizations of Square-Free Integers", *Information* **10** (2019), no. 7, article no. 232.

[4] ———, "Idempotent Integers: The complete class of numbers that work correctly in RSA", in *Géométrie algébrique, Théorie des nombres et Applications 2021*, 2021, p. 16-20.

[5] ———, "Search Heuristics and Constructive Algorithms for Maximally Idempotent Integers", *Information* **12** (2021), no. 8, article no. 305.

[6] B. Fagin & OEIS Foundation, "Smallest number for the nth square-free number that forms a pure idempotent product", 2018, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A325945`.

[7] ———, "Squarefree n with fully composite idempotent factorizations", 2018, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A306508`.

[8] ———, "Strong impostors $\neq 0 \pmod 4$", 2018, The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A318555`.

[9] E. D. Huthnance & J. Warndof, "On Using Primes for Public Key Encryption Systems", *Appl. Math. Lett.* **1** (1988), no. 3, p. 225-227.

[10] D. E. Knuth, "Two Notes on Notation", *Am. Math. Mon.* **99** (1992), no. 5, p. 403-422.

[11] R. Pinch, "On Using Carmichael Numbers for Public Key Encryption Systems", in *Cryptography and coding. 6th IMA international conference (Cirencester, 1997)*, Lecture Notes in Computer Science, vol. 1355, Springer, 1997, p. 265-269.

[12] ———, "The Carmichael Numbers up to 10^{21}", in *Proceedings of the Conference on Algorithmic Number Theory 2007*, TUCS General Publication, Turku Centre for Computer Science, 2007, `http://www.s369624816.websitehome.co.uk/rgep/cartable.html`, p. 129-131.

[13] R. Rivest, A. Shamir & L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystem", *Commun. ACM* **21** (1978), no. 2, p. 120-126.

Barry S. Fagin, Dept of Computer Science, 2354 Fairchild Drive, US Air Force Academy, Colorado Springs CO 80840 • *E-mail :* `barry.fagin@afacademy.af.edu`