On some number densities related to coprimes

Stanislav Sykora, Extra Byte, www.ebyte.it
First published in November 15, 2014

This brief essay explores the limit mean densities of subsets of natural numbers m such that the pair (m, m\b) is either coprime or not coprime. Here m\b = floor(m/b) denotes the operation which can be also characterized as "truncation of the last digit in base b" or "right truncation in base b". The analysis leads to an interesting rational-valued function.

Keywords: math, discrete math, coprime integers, GCD, divisibility, sequence

Introduction

This brief Note explores subsets of natural numbers m which, in a base b > 1, have the property rtc(m;b) such that

$$\gcd(m, \mathrm{floor}(m/b)) = 1 \qquad (1)$$

and those with the complementary property rtnc(m;b) such that

$$\gcd(m, \mathrm{floor}(m/b)) > 1. \qquad (2)$$

Here gcd(n,m) = gcd(m,n) denotes the greatest common divisor of the integer numbers n and m, as defined in [1]. Particularly, for any $n \geq 0$,

$$\gcd(n,1) = 1 \quad \text{and} \quad \gcd(n,0) = n. \qquad (3)$$

To simplify the notation, we will in the following write

$$\mathrm{floor}(m/b) \equiv m\backslash b. \qquad (4)$$

Evidently, the RTC property is equivalent to saying that m and m\b are *coprime*, while the RTNC property is true when m and m\b are *not coprime*[1]. Since the two properties are logically complementary, it follows that:

Lemma 1: Every natural number n satisfies either (1),i.e., has the RTC property, or (2), i.e., has the RTNC property, but not both, because the two properties are mutually exclusive.

For example, in base 10, numbers like 1, 12, 38, 103, 1111, or 8399 belong to the RTC category, while numbers like 2, 3, 4, 26, 39, 147, 729, or 3705 belong to the RTNC category. These two sequences were listed in Online Encyclopedia of Integer Sequences, OEIS [2]; see references [3,4]. Their starting terms are:

RTC: 1, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 23, 25, 27, 29, 31, 32, 34, 35, 37, 38, 41, 43, 45, 47, 49, 51, 52, 53, 54, 56, 57, 58, 59, 61, 65, 67, 71, 72, 73, 74, 75, 76, 78, 79, 81, 83, 85, 87, 89, 91, 92, 94, 95, …

RTNC: 2, 3, 4, 5, 6, 7, 8, 9, 20, 22, 24, 26, 28, 30, 33, 36, 39, 40, 42, 44, 46, 48, 50, 55, 60, 62, 63, 64, 66, 68, 69, 70, 77, 80, 82, 84, 86, 88, 90, 93, 96, 99, 100, 102, 104, 105, 106, 108, 110, 120, 122, 123, 124, …

Similar sequences can be computed for any base b. In base b = 16, for example, the 10000th RTC number is (in hexadecimal notation) 0x42D9, while the 10000th RTNC number is 0x5DFF [5, 6].

From the linearity of the graphs in references [3,4,5,6] it appears that the fraction of each of these subsets of natural numbers in the interval $1 \leq m \leq M$ converges for $M \to \infty$ to a constant mean value.

[1] The acronym RTC stands for "right-truncated is coprime", and RTNC stands for "right-truncated is not coprime".

This Note is concerned with evaluating this limit mean density of RTC and RTNC numbers for every base $b > 1$. To do so, we will concentrate on the limit mean density c(b) of RTC numbers. Given the complementarity of the properties, the limit mean density of RTNC numbers is then $nc(b) = 1 - c(b)$.

Construction of the mean-density formula

First, let us analyze what the RTC and RTNC conditions imply. Any natural number m can be written as

$$m = b^*n + d, \quad 0 \leq d < (b\text{-}1), \tag{5}$$

where d is the last digit in its power expansion in base b. Equation (1) then requires m and n to share no divisor greater than 1, and therefore no prime factor p. In the following, we will find useful

Lemma 2: The RTC condition of equation (1) is equivalent to saying that $n = m\backslash b$ and d, the last digits of m, are coprime. Similarly, the RTNC condition of equation (2) is equivalent to saying that n and d have a common prime divisor.

Proof: whenever n and d have a common prime divisor p, it is also a divisor of m, and therefore a common prime divisor of m and n. Vice versa, when m and n have a common prime divisor p then $(m/p) - b^*(n/p)$ is an integer and, since it equals (d/p), it must be also a divisor of d.

Prime factors p which are not divisors of d are therefore easy to handle because no matter how we choose n, the numbers m and n can not be both divisible by p (note that this covers automatically also all primes $p \geq b$). We will be therefore focus mostly on primes $p < b$ and digits $d < b$ such that p is a divisor of d.

For the purposes of this section, we assume that $2b \leq m < M = b^*N+2b$, where N is a large number which will be eventually allowed to grow to infinity[2]. It is also convenient to concentrate on the RTNC condition which is slightly simpler to handle.

For every one of the b possible values of d, there are N possible choices of n and therefore N pairs (m,n). If these were not at all limited, we would end up with a density of the subset equal to $b^*N/(M\text{-}2b) = 1$. Vice versa, if none of these passed a condition, the density would be 0. Of course, the RTNC condition imposes more specific selections which we will now analyze.

Consider first the case $d = 0$. For all such numbers (N values), equation (5) generates (m, n) non-coprime pairs that need to be included. When b=10 and N=8, for example, these are the numbers 20, 30, ..., 90.

Another special case is $d = 1$ for which equation (5) generates N pairs (m,n) which are all coprime and therefore must be excluded. In our example of b=10 and N=8, these are the numbers 21, 31, ..., 91.

Note that when $b = 2$, the two above cases exhaust all possibilities, giving $c(2) = 1/2$.

Now consider a base $b > 2$, and all the d values, $1 < d < b$, which contain a prime factor $p_1 < b$. There are $(b\text{-}1)\backslash p_1$ such cases and for each of them we must include the $N\backslash p_1$ values of n which are divisible by p_1. The total is $((b\text{-}1)\backslash p_1)(N\backslash p_1)$ included values for every prime $p_1 < b$.

If all digits $d > 1$ were divisible by just one prime, we would have:

$$t(b,N) = N + \sum_{\text{prime } p1<b} [(b\text{-}1)\backslash p_1](N\backslash p_1),$$

$$nc(b) = \lim_{N\to\infty} t(b,N)/(bN) = [\lim_{N\to\infty} t(b,N)/N]/b = [1 + \sum_{\text{prime } p1} [(b\text{-}1)\backslash p_1]/p_1]/b,$$

[2] The "offset" of 2b is used just to avoid the limited number of cases which might require the use of special conventions for the greatest common divisor function, namely gcd(1,n) = 1, and gcd(0,n) = n for $n \neq 1$. With b = 10, for example, these arise whenever m < 20. Fortunately, it is clear that the limit density of any infinite subset of natural numbers does not change when a finite number of starting elements is skipped.

where t(b,N) denotes the total number of included values of n. In the last passage, we have exploited the obvious fact that $\lim_{N\to\infty} (N\backslash p)/N = 1/p$, and we have dropped the relation $p_1 < b$ which is superfluous because when $p_1 \geq b$ the corresponding term evaluates to zero. This completes the cases b = 3, 4, 5, and 6, for which nc(b) evaluates to 1/2, 13/24, 8/15, and 26/45, respectively.

However, for bases b > 6, some of the digits d have two distinct prime divisors, say $p_1 < p_2$. Consider, for example, the case of d = 6, divisible by both 2 and 3. In such cases the last formula is in error because we have counted twice all those n which are divisible by both p_1 and p_2. To correct the situation, we must add to the count e(b,N) a negative correction term of the type

$$-\sum_{\text{prime } p1<p2} [(b-1)\backslash(p_1p_2)](N\backslash(p_1p_2))$$

This resolves the problem for d values containing two distinct prime divisors, but not for those containing three of them (the first such case occurs for d = 2*3*5 = 30, which can occur in bases b > 30). These were "over-corrected" and need another correction term of the form

$$+\sum_{\text{prime } p1<p2<p2} [(b-1)\backslash(p_1p_2p_3)](N\backslash(p_1p_2p_3)),$$

and so on, according to the iterative inclusion-exclusion principle [7].

Collecting all the terms one finally obtains the correct count of all the included pairs:

$$t(b,N) = N + \sum_{k>0} (-)^{k-1} \sum_{\text{prime } p1<p2<\ldots<pk} [(b-1)\backslash(p_1p_2\ldots p_k)](N\backslash(p_1p_2\ldots p_k))$$

and $$nc(b) = [\lim_{N\to\infty} t(b,N)/N] / b.$$

It is convenient to define a new function s(b) of b, such that:

$$s(b) = \lim_{N\to\infty} t(b+1,N)/N =$$

$$= 1 + \sum_{k>0} (-)^{k-1} \sum_{\text{prime } p1<p2<\ldots<pk} [b\backslash(p_1p_2\ldots p_k)]/(p_1p_2\ldots p_k) \qquad (6)$$

$$nc(b) = s(b-1)/b \quad \text{and} \quad c(b) = 1 - nc(b) = 1 - s(b-1)/b. \qquad (7)$$

Notice once again that the summation for s(b) is always finite and the resulting values are all rational. Moreover, the formula for s(b) admits the extension to s(0), namely s(0) = 1.

Since a manual evaluation of equation (6) for large values of b can be a bit cumbersome, the Appendix offers a PARI/GP script to handle this task.

Numeric values and tests

The rational-valued functions nc(b) and c(b) were evaluated using a PARI/GP [8] script [9], and the separate sequences of their normalized numerators and shared denominators were listed in OEIS [9, 10, 11]. The results for a few values of the base b are shown in the Table below.

As a check, the densities were estimated, for all the selected bases, using a brute force count of numbers up to 10^9 satisfying the RTNC condition (for the used code, see the Appendix). These results are listed in the Table in the column "Estimated nc(b), were compared with the rational values computed from equations (6) and (7). In all cases, a perfect agreement was found up to at least 6 significant digits.

Notice that the functions nc(b) and c(b), while converging to a limit for large b (to be discussed below), are not monotonous.

The numerators of the rational-valued function s(n) were also registered on OEIS [12]. It is interesting to note that the corresponding denominators are already listed because they match those of the partial sums of the phi(n)/n series [13], where phi(n) is the Euler totient function. The corresponding numerators [14], however, are different.

Table 1. Column 1 defines the base b. Column 2 gives the exact rational value of nc(b) computed from equations (6, 7) and column 3 displays the same value in floating point notation. Column 4 shows the "brute-force" value of nc(b) estimated by counting all numbers having the RTNC property in the interval from 1 to 10^9. Finally, for completeness sake, the last column shows the exact (rational) value of the complementary c(b) = 1 – nc(b).

Base b	Exact nc(b)	Exact nc(b)	Estimated nc(b)	Exact c(b)
2	1/2	0.5<u>0</u>	0.500000 ...	1/2
3	1/2	0.5<u>0</u>	0.500000 ...	1/2
4	11/24	0.45833<u>3</u>	0.458333 ...	13/24
5	7/15	0.46666<u>6</u>	0.466666 ...	8/15
6	19/45	0.42222<u>2</u>	0.422222 ...	26/45
7	16/35	0.457142 ...	0.457142 ...	19/35
8	117/280	0.417857 ...	0.417857 ...	163/280
9	269/630	0.426984 ...	0.426984 ...	361/630
10	877/2100	0.417619 ...	0.417619 ...	1223/2100
16	199663/480480	0.415549 ...	0.415549 ...	280817/480480
256	See ref. [8,9]	0.393612 ...	0.393612 ...	See ref. [8,10]
∞	1 - 1/ζ(2)	0.392072 ...	not applicable	1/ζ(2)

The limit value and an alternative interpretation

What if we let b grow to infinity in equation (7)? Clearly, an expression like $[(b-1)\backslash(p_1p_2...p_k)]/b$ has a limit which equals $1/(p_1p_2...p_k)$. Consequently,

$$\begin{aligned}\lim_{b\to\infty} nc(b) &= \Sigma_{k>0}\,(-)^{k-1}\,\Sigma_{\text{prime } p1<p2<...<pk}\, 1/(p_1p_2...p_k)^2 \\ &= 1 - [1+\Sigma_{k>0}\,(-)^{k}\,\Sigma_{\text{prime } p1<p2<...<pk}\, 1/(p_1p_2...p_k)^2] \\ &= 1 - \Pi_{\text{prime } p}\,(1 - 1/p^2) \\ &= 1 - 1/\zeta(2) = 1 - 6/\pi^2 = 0.392072\ ...\ (\text{OEIS A229099}), \qquad (8)\end{aligned}$$

where ζ(n) is the Riemann zeta function [15]. The above passages are relatively simple, provided one uses properly inclusion-exclusion principle and the Euler product theorem. In a somewhat different context, the procedure occurs also in reference [16].

Since c(b) = 1 – nc(b), equation (8) implies that

$$\lim_{b\to\infty} c(b) = 1/\zeta(2) = 6/\pi^2 = 0.607927\ ...\ (\text{OEIS A059956}), \qquad (9)$$

a result that is not unexpected. We have seen, in fact, that the RTC condition of equation (1) coincides with that of m\b being coprime to the last digit of m (Lemma 2). Since the value of n = m\b is unlimited, it is evident that for large values of b the density of numbers satisfying RTC must approach that of relatively coprime pairs among all pairs of nonnegative integers which, by Cesaro's theorem [1, 17], is known to be 1/ζ(2). This reflection leads to the following interpretation:

Lemma 3: *The value of c(b) is also the mean density of coprime pairs among all pairs of nonnegative integers such that one is unconstrained while the other is drawn randomly from the set {0, 1, 2, ..., b-1}.*

The convergence to the limit, however, is very slow and, as already pointed out, somewhat erratic.

Appendix

The numeric values reported here, as well as those registered in OEIS were computed using a PARI [8] with the following GP scripts. Note that the names of the functions were slightly adapted to match the author's library conventions:

A) The function s(b) of equation (6), here named S(b):

```
S(b) = 1 + S_aux(b,1,1);
```

This calls the following auxiliary function S_aux which encodes the inclusion-exclusion process:

```
S_aux(n,p0,inp) =
/* ----------------------------------------------------------
This is an iterative loop used to evaluate the function
s(b) of the referencehttp://dx.doi.org/0.3247/SL5Math14.005.
p0 is the starting product of increasing prime numbers, which
is to be extended by the next prime factor with index ipn.
The actual evaluation of s(b) is ahieved by calling
S(b) = 1+S_aux(b,1,1)
---------------------------------------------------------- */
{
  my (t=0/1,tt=0/1,in=inp,pp);
  while (1,pp = p0*prime(in);tt = n\pp;
    if (tt==0,break,t += tt/pp - S_aux(n,pp,in++)));
  return (t);
}
```

B) The functions nc(b) and c(b) of equations (7), here named RTnc(b) and RTc(b), respectively:

```
RTnc(b) = S(b-1)/b;

RTc(b) = 1 - RTnc(b);
```

C) The mean densities of RTNC numbers for various b values were computed by calling:

```
DensityNCond1(10^9,IsNotCoprimeToRTrunc,16)
```

The DensityNCond1 function script is:

```
DensityNCond1(nmax,condition,m,monit=0) =
/* ----------------------------------------------------------
Estimates the density of naturals satisfying a test condition.
The condition function must be of the form
  BOOL cond(t_INT n,t_INT m), where m is a parameter.
The argument nmax is the desired maximum value of n to test.
Returns the ratio k/nmax, where k is the number n-values not
exceeding nmax which satisfy the condition.
Example: density of numbers n such that n is coprime to n\16:
  > DensityNCond1(10^7,IsCoprimeRTrunc,16).
---------------------------------------------------------- */
{
  my (k=0,n=0);
  while(1,n++;if(condition(n,m),k++;
    if(monit && (k%1000==0),print(n," ",k," ",0.0+k/n));
    if(n==nmax,break)));
  return(k/n);
}
```

It can be used to count all numbers in the interval 1 to nmax which satisfy a generic condition(n,m), with m being a parameter. In the present case, the condition was defined as

```
IsNotCoprimeToRTrunc(n,b) = gcd(n,n\b)!=1;
```

References

[1] Wikipedia, Coprime integers.
[2] The On-Line Encyclopedia of Integer Sequences, OEIS, published electronically at http://oeis.org.
[3] S. Sykora, *Numbers m that are coprime to floor(m/10)*, OEIS A248499.
[4] S. Sykora, *Numbers m that are not coprime to floor(m/10)*, OEIS A248500.
[5] S. Sykora, *Numbers m that are coprime to floor(m/16)*, OEIS A248501.
[6] S. Sykora, *Numbers m that are not coprime to floor(m/16)*, OEIS A248502.
[7] Wikipedia, Inclusion-exclusion principle.
[8] See the PARI/GP home page.
[9] S. Sykora, *Numerators of the limit mean densities of numbers m such that gcd(m,floor(m/n))=1*, OEIS A250031
[10] S. Sykora, *Numerators of the limit mean densities of numbers m such that gcd(m,floor(m/n))>1*, OEIS A250032.
[11] S. Sykora, *Denominators for sequences A250031 and A250032*, OEIS A250033.
[12] S. Sykora, *Numerators of a coprimes counting function s(n)*, OEIS A250034.
[13] N. J. A. Sloane, *Denominator of Sum_{k=1..n} phi(k)/k*, OEIS A072155 .
[14] N. J. A. Sloane, *Numerator of Sum_{k=1..n} phi(k)/k*, OEIS A071708.
[15] Wikipedia, Riemann zeta function.
[16] Eric Weisstein's World of Mathematics, Euler Product, Equations. 5 and 8.
[17] G.H. Hardy, E. M. Wright (2008), *An Introduction to the Theory of Numbers* (6th edition), Oxford University Press. ISBN 978-0199219865, theorem 332.

History of this document

15 Nov 2014: Assigned a DOI (10.3247/SL5Math14.005) and uploaded online.
Note: At this moment, not all the OEIS entries are yet registered (refs. 9-11).