On a Deconcatenation Problem

Henry Ibstedt

Abstract: In a recent study of the *Primality of the Smarandache Symmetric Sequences* Sabin and Tatiana Tabirca [1] observed a very high frequency of the prime factor 333667 in the factorization of the terms of the second order sequence. The question if this prime factor occurs peridically was raised. The odd behaviour of this and a few other primefactors of this sequence will be explained and details of the periodic occurence of this and of several other prime factors will be given.

Definition: The n*th* term of the Smarandache symmetric sequence of the second order is defined by S(n)=123...n_n...321 which is to be understood as a concatenation[1] of the first n natural numbers concatenated with a concatenation in reverse order of the n first natural numbers.

Factorization and Patterns of Divisibility

The first five terms of the sequence are: 11, 1221, 123321, 12344321, 1234554321. The number of digits D(n) of S(n) is growing rapidly. It can be found from the formula:

$$D(n) = 2k(n+1) - \frac{2(10^k - 1)}{9} \text{ for n in the interval } 10^{k-1} \le n < 10^k - 1 \qquad (1)$$

In order to study the repeated occurrance of certain prime factors the table of S(n) for n≤100 produced in [1] has been extended to n≤200. Tabirca's aim was to factorize the terms S(n) as far as possible which is more ambitious then the aim of the present calculation which is to find prime factors which are less than 10^8. The result is shown in table 1.

The computer file containing table 1 is analysed in various ways. Of the 664579 primes which are smaller than 10^7 only 192 occur in the prime factoriztions of S(n) for 1≤n≤200. Of these 192 primes 37 occur more than once. The record holder is 333667, the 28693th prime, which occurs 45 times for 1≤n≤200 while its neighbours 333647 and 333673 do not even occur once. Obviously there is something to be explained here. The frequency of the most frequently occurring primes is shown below..

Table 2. Most frequently occurring primes

p	3	33367	37	41	271	9091	11	43	73	53	97	31	47
Freq	132	45	41	41	41	29	25	24	14	8	7	6	6

[1] In this article the concatenation of a and b is written a_b. Multiplication ab is often made explicit by writing a.b. When there is no reason for misunderstanding the signs "_" and "." are omitted. Several tables contain prime factorizations. Prime factors are given in ascending order, multiplication is expressed by "." and the last factor is followed by ".." if the factorization is incomplete or by Fxxx indicating the number of digits of the last factor. To avoid typing errors all tables are electronically transferred from the calculation program, which is DOS-based, to the wordprocessor. All editing has been done either with a spreadsheet program or directly with the text editor. Full page tables have been placed at the end of the article. A non-proportional font has been used to illustrate the placement of digits when this has been found useful.

The distribution of the primes 11, 37, 41, 43, 271, 9091 and 333667 is shown in table 3. It is seen that the occurance patterns are different in the intervals $1\le n\le 9$, $10\le n\le 99$ and $100\le n\le 200$. Indeed the last interval is part of the interval $100\le n\le 999$. It would have been very interesting to include part of the interval $1000\le n\le 9999$ but as we can see from (1) already S(1000) has 5786 digits. Partition lines are drawn in the table to highlight the different intervals. The less frequent primes are listed in table 4 where primes occurring more than once are partitioned.

From the patterns in table 3 we can formulate the occurance of these primes in the intervals $1\le n\le 9$, $10\le n\le 99$ and $100\le n\le 200$, where the formulas for the last interval are indicative. We note, for example, that 11 is not a factor of any term in the interval $100\le n\le 999$. This indicates that the divisibility patterns for the interval $1000\le n\le 9999$ and further intervals is a completely open question.

Table 5 shows an analysis of the patterns of occurance of the primes in table 1 by interval. Note that we only have observations up to n=200. Nevertheless the interval $100\le n\le 999$ is used. This will be justified in the further analysis.

Table 5. Divisibility patterns

Interval	p	n	Range for j
1≤n≤…	3	2+3j	j=0,1,…
1≤n≤…		3j	j=1,2,…
1≤n≤9	11	All values of n	
10≤n≤99		12+11j	j=0,1, … ,7
		20+11j	j=0,1, … ,7
100≤n≤999		None	
1≤n≤9	37	2+3j	j=0,1,2
		3+3j	j=0,1,2
10≤n≤99		12+3j	j=0,1,…,28,29
100≤n≤999		122+37j	j=0,1,…,23
		136+37j	j=0,1,…,23
1≤n≤9	41	4+5j	j=0,1
		5	
10≤n≤999		14+5j	j=0,1,…,197
1≤n≤9	43	None	
10≤n≤99		11+21j	j=0,1,3,4
		24+21j	j=0,1,2,3
100≤n≤999		100	
		107+7j	j=0,1,…,127
1≤n≤9	271	4+5j	j=0,1
		5	
10≤n≤999		14+5j	j=0,1,…,197
1≤n≤999	9091	9+5j	j=0,1,…,98
1≤n≤9	333667	8,9	
10≤n≤99		18+9j	j=0,1,…,9
100≤n≤999		102+3j	j=0,1,…,299

We note that no terms are divisible by 11 for $n>100$ in the interval $100\le n\le 200$ and that no term is divisible by 43 in the interval $1\le n\le 9$. Another remarkable observation is that the sequence shows exactly the same behaviour for the primes 41 and 271 in the intervals included in the study. Will they show the same behaviour when $n\ge 1000$?

Consider

$$S(n)=12...n\_n...21.$$

Let p be a divisor of S(n). We will construct a number

$$N=12...n\_0..0\_n...21 \qquad (2)$$

so that p also divides N. What will be the number of zeros? Before discussing this let's consider the case p=3.

Case 1. p=3.

In the case p=3 we use the familiar rule that a number is divisible by 3 if and only if its digit sum is divisible by 3. In this case we can insert as many zeros as we like in (2) since this does not change the sum of digits. We also note that any integer formed by concatenation of three consecutive integers is divisible by 3, cf a_a+1_a+2, digit sum 3a+3. It follows that also a_a+1_a+2_a+2_a+1_a is divisible by 3. For a=n+1 we insert this instead of the appropriate number of zeros in (2). This means that if $S(n)\equiv 0$ (mod 3) then $S(n+3)\equiv 0$ (mod 3). We have seen that $S(2)\equiv 0$ (mod 3) and $S(3)\equiv 0$ (mod 3). By induction it follows that $S(2+3j)\equiv 0$ (mod 3) for j=1,2,... and $S(3j)\equiv 0$ (mod 3) for j=1,2,... .

We now return to the general case. S(n) is deconcatenated into two numbers 12...n and n... 21 from which we form the numbers

$$A=12...n\cdot 10^{1+[\log_{10}B]} \text{ and } B=n...21$$

We note that this is a different way of writing S(n) since indeed A+B=S(n) and that $A+B\equiv 0$ (mod p). We now form $M=A\cdot 10^s+B$ where we want to determine s so that $M\equiv 0$ (mod p). We write M in the form $M=A(10^s-1)+A+B$ where A+B can be ignored mod p. We exclude the possibility $A\equiv 0$ (mod p) which is not interesting. This leaves us with the congruence

$$M\equiv A(10^s-1)\equiv 0 \pmod p$$

or

$$10^s-1\equiv 0 \pmod p$$

We are particularly interested in solutions for which

$$p\in\{11,37,41,43,271,9091,333667\}$$

By the nature of the problem these solutions are periodic. Only the two first values of s are given for each prime.

Table 6. $10^s-1\equiv 0$ (mod p)

p	3	11	37	41	43	271	9091	33367
s	1,2	2,4	3,6	5,10	21,42	5,10	10,20	9,18

We note that the result is independent of n. This means that we can use n as a parameter when searching for a sequence C=n+1_n+2_...n+k_n+k_...n+2_n+1 such that this is also divisible by p and hence can be inserted in place of the zeros to form S(n+k) which then fills the condition $S(n+k)\equiv 0$ (mod p). Here k is a multiple of s or s/2 in case s is even. This explains the results which we have already obtained in a different way as part of the factorization of S(n) for $n\leq 200$, see tables 3 and 5. It remains to explain the periodicity which as we have seen is different in different intervals $10^u\leq n\leq 10^u-1$.

This may be best done by using concrete examples. Let us use the sequences starting with n=12 for p=37, n=12 and n=20 for p=11 and n=102 for p=333667. At the same time we will illustrate what we have done above.

Case 2: n=12, p=37. Period=3. Interval: $10 \leq n \leq 99$.

```
S(n)=  123456789101112_____________121110987654321
N=     123456789101112000000000000121110987654321
C=                    131415151413
S(n+k)=123456789101112131415151413121110987654321
```

Let's look at C which carries the explanation to the periodicity. We write C in the form

C=101010101010+30405050403

We know that $C \equiv 0 \pmod{37}$. What about 101010101010? Let's write

$101010101010=10+10^3+10^5+...+10^{11}=(10^{12}-1)/9 \equiv 0 \pmod{37}$

This congruence mod 37 has already been established in table 6. It follows that also

$30405050403 \equiv 0 \pmod{37}$

and that

$x \cdot (101010101010) \equiv 0 \pmod{37}$ for x = any integer

Combining these observations we se that

$232425252423,\ 333435353433,\ ...\ 939495959493 \equiv 0 \pmod{37}$

Hence the periodicity is explained.

Case 3a: n=12, p=11. Period=11. Interval: $10 \leq n \leq 99$.

```
S(12)=12_.._12________________________________________________12_.._21
S(23)=12_.._1213141516171819202122232322212019181716151413 12_.._21
C=          131415161718192021222323222120191817161514 13=
C1=         10101010101010101010101010101010101010101010+
C2=          3040506070809101112131312111009080706050403
```

From this we form

```
2·C1+C2=      2324252627282930313233333231302928272625242 3
```

which is NOT what we wanted, but $C1 \equiv 0 \pmod{11}$ and also $C1/10 \equiv 0 \pmod{11}$. Hence we form

```
2·C1+C1/10+C2=24252627282930313233343433323130292827262524
```

which is exactly the C-term required to form the next term S(34) of the sequence. For the next term S(45) the C-term is formed by 3·C1+2·C1/10+C2 The process is repeated adding C1+C1/10 to proceed from a C-term to the next until the last term <100, i.e. S(89) is reached.

Case 3b: n=20, p=11. Period=11. Interval: $10 \leq n \leq 99$.

This case does not differ much from the case n=12. We have

```
S(20)=12_.._20________________________________________________20_.._21
S(31)=12_.._20212223242526272829303131302928272625242322212 0_.._21
C=          2122232425262728293031313029282726252423222 1=
C1=         10101010101010101010101010101010101010101010+
C2=          1020304050607080910111111009080706050403020 1
```

The C-term for S(42) is

```
3·C1+C1/10+C2=32333435363738394041424241403938373635343332
```

In general C=x·C1+(x-1)·C1/10+C2 for x=3,4,5, ..,8. For x=8 the last term S(97) of this sequence is reached.

Case 4: n=102, p=333667. Period=3. Interval: 100≤n≤999.

```
S(102)=12_.._101102__________________102101_.._21
S(105)=12_.._101102103104105105104103102101_.._21
C=                 103104105105104103              ≡0 (mod 333667)
C1=                100100100100100100              ≡0 (mod 333667)
C2=                  3004005005004003              ≡0 (mod 333667)
```

Removing 1 or 2 zeros at the end of C1 does not affect the congruence modulus 333667, we have:

```
C1'=                10010010010010010              ≡0 (mod 333667)
C1''=                1001001001001001              ≡0 (mod 333667)
```

We now form the combinations:

x·C1+y·C1'+z·C1''+C2≡0 (mod 333667)

This, in my mind, is quite remarkable: All 18-digit integers formed by the concatenation of three consecutive 3-digit integers followed by a concatenation of the same integers in descending order are divisible by 333667, example 376377378378377376≡0 (mod 333667). As far as the C-terms are concerned all S(n) in the range 100≤n≤999 could be divisible by 333667, but they are not. Why? It is because S(100) and S(101) are not divisible by 333667. Consequently n=100+3k and 101+3k can not be used for insertion of an appropriate C-value as we did in the case of S(102). This completes the explanation of the remarkable fact that every third term S(102+3j) in the range 100≤n≤999 is divisible by 333667.

These three cases have shown what causes the periodicity of the divisibility of the Smarandache symmetric sequence of the second order by primes. The mechanism is the same for the other periodic sequences.

Beyond 1000

We have seen that numbers of the type:

10101010…10, 100100100…100, 10001000…1000, etc

play an important role. Such numbers have been factorized and the occurrence of our favorite primes 11, 37, …, 333667 have been listed in table 7. In this table a number like 100100100100 has been abbreviated 4(100) or q(E), where q and E are listed in separate columns.

Question 1. Does the sequence of terms S(n) divisible by 333667 continue beyond 1000?

Although S(n) was partially factorized only up n=200 we have been able to draw conclusions on divisibility up n=1000. The last term that we have found divisible by 333667 is S(999). Two conditions must be met for there to be a sequence of terms divisible by p=333667 in the interval 1000≤n≤9999.

<u>Condition 1.</u> There must exist a number 10001000...1000 divisible by 333667 to ensure the periodicity as we have seen in our case studies.
In table 7 we find q=9, E=1000. This means that the periodicity will be 9 – if it exists, i.e. condition 1 is met.

Condition 2. There must exist a term S(n) with $n \geq 1000$ divisible by 333667 which will constitute the first term of the sequence.
The last term for $n<1000$ which is divisible by 333667 is S(999) from which we build

S(108)=12...999_1000_..._1008_1008_...1000_999-...21

where we deconcatenate 100010011002...10081008...10011000 which is divisible by 333667 and provides the C-term (as introduced in the case studies) needed to generate the sequence, i.e. condition 2 is met.

We conclude that $S(1008+9j) \equiv 0 \pmod{333667}$ for j=0,1,2, ... 999. The last term in this sequence is S(9999). From table 7 we see that there could be a sequence with the period 9 in the interval $10000 \leq n \leq 99999$ and a sequence with period 3 in the interval $100000 \leq n \leq 999999$. It is not difficult to verify that the above conditions are filled also in these intervals. This means that we have:

$S(1008+9j) \equiv 0 \pmod{333667}$	for j=01,2,...,999, i.e. $10^3 \leq n \leq 10^4-1$
$S(10008+9j) \equiv 0 \pmod{333667}$	for j=01,2,...,9999, i.e. $10^4 \leq n \leq 10^5-1$
$S(100002+3j) \equiv 0 \pmod{333667}$	for j=01,2,...,99999, i.e. $10^5 \leq n \leq 10^6-1$

It is one of the fascinations with large numbers to find such properties. This extraordinary property of the prime 333667 in relation to the Smarandache symmetric sequence probably holds for $n>10^6$. It easy to loose contact with reality when plying with numbers like this. We have $S(999999) \equiv 0 \pmod{333667}$. What does this number S(999999) look like? Applying (1) we find that the number of digits D(999999) of S(999999) is

$$D(999999)=2\cdot 6\cdot 10^6-2\cdot(10^6-)/9=11777778$$

Let's write this number with 80 digits per line, 60 lines per page, using both sides of the paper. We will need 1226 sheets of paper – more that 2 reams!

Question 2. Why is there no sequence of S(n) divisible by 11 in the interval $100 \leq n \leq 999$?

Condition1. We must have a sequence of the form 100100.. divisible by 11 to ensure the periodicity. As we can see from table 7 the sequence 100100 fills the condition and we would have a periodicity equal to 2 if the next condition is met.

Condition 2. There must exist a term S(n) with $n \geq 100$ divisible by 11 which would constitute the first term of the sequence. This time let's use a nice property of the prime 11:

$$10^s \equiv (-1)^s \pmod{11}$$

Let's deconcatenate the number a_b corresponding to the concatenation of the numbers a and b: We have:

$$a\_b = a \cdot 10^{1+[\log_{10} b]} + b = \begin{cases} -a+b \text{ if } 1+[\log_{10} b] \text{ is odd} \\ a+b \text{ if } 1+[\log_{10} b] \text{ is even} \end{cases}$$

Let's first consider a deconcatenated middle part of S(n) where the concatenation is done with three-digit integers. For convienience I have chosen a concrete example – the generalization should pose no problem

```
273274275275274273≡2-7+3-2+7-4+2-7+5-2+7-5+2-7+4-2+7-3≡0 (mod 11)
+-+-+-+-+-+-+-+-+-
```

It is easy to see that this property holds independent of the length of the sequence above and whether it start on + or -. It is also easy to understand that equivalent results are obtained for other primes although factors other than +1 and −1 will enter into the picture.

We now return to the question of finding the first term of the sequence. We must start from n=97 since S(97) it the last term for which we know that S(n)≡0 (mod 11). We form:

```
9899100101..n_n..1011009998≡2 (mod 11) independent of n<1000.
+-+-+-+-+-.. _ ..-+-+-+-+-+-
```

This means that S(n)≡2 (mod 11) for 100≤n≤999 and explains why there is no sequence divisible by 11 in this interval.

Question 3. Will there be a sequence divisible by 11 in the interval 1000≤n≤9999?

Condition 1. A sequence 10001000...1000 divisible by 11 exists and would provide a period of 11, se table 7.

Condition 2. We need to find one value n ≥1000 for which S(n)≡0 (mod 11). We have seen that S(999)≡2 (mod 11). We now look at the sequences following S(999). Since S(999)≡2 (mod 9) we need to insert a sequence 10001001..m_m...10011000≡9 (mod 11) so that S(m)≡0 (mod 11). Unfortunately m does not exist as we will see below

```
10001000≡2 (mod 11)
+-+-+-+-
1   1
1000100110011000≡2. (mod 11)
+-+-+-+-+-+-+-+-
1   1   1   1
      1   1
100010011002100210011000≡0 (mod 11)
+-+-+-+-+-+-+-+-+-+-+-+-
1   1   1   1   1   1
      1   2   2   1
10001001100210031003100210011000≡-4≡7 (mod 11)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
1   1   1   1   1   1   1   1
      1   2   3   3   2   1
```

Continuing this way we find that the residues form the period 2,2,0,7,1,4,5,4,1,7,0. We needed a residue to be 9 in order to build sequences divisible by 9. We conclude that S(n) is not divisible by 11 in the interval 1000≤n≤9999.

Trying to do the above analysis with the computer programs used in the early part of this study causes overflow because the large integers involved. However, changing the approach and performing calculations modulus 11 posed no problems. The above method was preferred for clarity of presentation.

Epilog

There are many other questions that may be interesting to look into. This is left to the reader. The author's main interest in this has been to develop means by which it is possible to identify some properties of large numbers other than the so frequently asked question as to whether a big number is a prime or not. There are two important ways to generate large numbers that I found particularly interesting – iteration and concatenation. In this article the author has drawn on work done previously, references below. In both these areas very large numbers may be generated for which it may be impossible to find any practical use – the methods are often more important than the results.

References:

1. Tabirca, S. and T., *On Primality of the Smarandache Symmetic Sequences*, Smarandache Notions Journal, Vol. 12, No 1-3 Spring 2001, 114-121.
2. Smarandache F., *Only Problems, Not Solutions*, Xiquan Publ., Pheonix-Chicago, 1993.
3. Ibstedt H. *Surfing on the Ocean of Numbers*, Erhus University Press, Vail, 1997.
4. Ibstedt H, *Some Sequences of Large Integers*, Fibonacci Quarterly, 28(1990), 200-203.