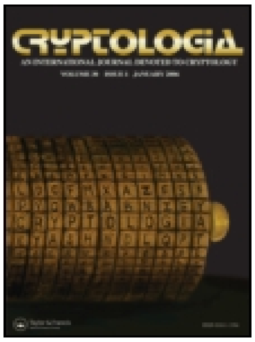

Cryptologia

ISSN: 0161-1194 (Print) 1558-1586 (Online) Journal homepage: https://www.tandfonline.com/loi/ucry20

SECURITY OF NUMBER THEORETIC PUBLIC KEY CRYPTOSYSTEMS AGAINST RANDOM ATTACK, II

Bob Blakley & G. R. Blakley

SECURITY OF NUMBER THEORETIC PUBLIC KEY CRYPTOSYSTEMS AGAINST RANDOM ATTACK, II

Bob Blakley and G. R. Blakley

This paper continues a discussion begun in the last issue of CRYPTOLOGIA. It shows that RSA cryptosystems and more general number theoretic public key cryptosystems appear most resistant to cryptanalysis when the encoding modulus m is the square free product of safe primes, namely primes p of the form $p = 2a + 1$ where a is an odd prime. There is reason to believe that at least 1 in every 100000 one hundred digit numbers is a safe prime. When m is the square free product of safe primes anybody who can find integers x and e larger than 1 such that

$$x^e \equiv x \bmod(m)$$

can break the cryptosystem by factoring the coding modulus m. However, either x or e must in this case exceed $p/3$; where p is the smallest prime factor of m. In a typical RSA cryptosystem the two safe prime factors of m will be chosen so that $p/3$ exceeds $\sqrt{m}/100$.

If every message receiver uses the same coding exponent c there are certain economies possible without any apparent loss of security. The Fermat prime 65537 seems to be a reasonable candidate for universal coding exponent c. If certain routine precautions described below are taken by the message receiver who sets up a number theoretic public key cryptosystem the cryptanalyst's prospects appear dismal, although there is no proof that this is the case. The third paper in this series, which will appear in the next issue of CRYPTOLOGIA, is devoted to pathological examples, and to the proofs of results in I and II.

5. Complementarity properties of safe primes. We continue the section numbering scheme of I. Hence this section is numbered 5. Suppose that p and q are safe primes whose product is m. Any cryptanalyst who can find one nontrivial pair x, e such that $x{\uparrow}e \equiv x \bmod(m)$ can factor m. If x is small then e must be large, and conversely. Results of this latter type are called complementarity principles by analogy with the celebrated notion in physics. The exact statements of these results follow.

Lemma 5.1: Let T be a finite set of safe primes. Suppose that $m = \Pi\{p|\ p \in T\}$. Let $b = \mathrm{MIN}\ \{a(p)\ |\ p \in T\}$. If $2 \leq x \leq 2b-1$ then $\Pi\{a(p)\ |\ p \in T\} \leq \mathrm{per}[x,m]$.

Lemma 5.1 itself is less important than the following immediate corollary.

Theorem 5.1: Let T be a finite set of safe primes and let $m = \Pi\{p|\ p \in T\}$. If $2 \leq |x|$ and $\mathrm{per}[x,m] < \Pi\{a(p)\ |\ p \in T\}$ then $\mathrm{MIN}\ \{p-1|\ p \in T\} \leq |x|$.

Corollary 5.1: Let p and q be distinct safe primes whose product is m. If $2 \leq |x|$ and $\mathrm{per}[x,m] < a(p)a(q)$ then $\mathrm{MIN}\ \{p-1,q-1\} \leq |x|$.

Comment: 7 and 23 are safe primes and $\mathrm{per}[6,161] = \mathrm{ord}[6,161] = 22$. Also $\mathrm{per}[2,161] = \mathrm{ord}[2,161] = 33$. In these two senses Corollary 5.1 is best possible.

Theorem 5.2: Let p and q be safe primes whose product is m. Let x be a positive integer. If $p < q$ and $\mathrm{per}[x,m] = a(p)$ then $q \leq x$. If $\mathrm{per}[x,m] = a(q)$ then $p+1 \leq x$.

Comment: 7 and 23 are safe primes. We know that $\mathrm{per}[23,161] = 3$, and that $\mathrm{per}[8,161] = 11$. Hence Theorem 5.2 is best possible.

Example 5.1: 7 and 23 are safe primes. So are 11 and 59. Clearly $per[2,161] = 33$, and $per[2,649] = 290$. Thus $a(p)a(q)$ and $2a(p)a(q)$ can be multiplicative periods of 2 modulo pq, where p and q are safe primes.

Theorem 5.3: Let p and q be distinct safe primes whose product is m. Let x be a positive integer. If $per[x,m] = p-1$ then $q-1 \leq x$.

Example 5.2: Theorem 5.3 is best possible whether or not $p < q$. On the one hand 7 and 23 are safe primes and $per[6,161] = 22$. On the other hand 7 and 11 are safe primes and $per[10,77] = 6$.

Theorem 5.4: Let T be a set of n odd primes. Let $m = \Pi\{p| \ p \in T\}$. Then m is square free. There are exactly $2\uparrow n$ nonnegative integers x less than m such that $per[x,m] = 1$. If $per[x,m] = 1$ and $p \in T$ then $x \equiv 0 \bmod(p)$ or $x \equiv 1 \bmod(p)$. Evidently 0 and 1 have multiplicative period 1 modulo m. There are exactly $3\uparrow n - 2\uparrow n$ nonnegative integers less than m such that $per[x,m] = 2$. If $per[x,m] = 2$ and $p \in T$ then $x \equiv 0 \bmod(p)$ or $x \equiv 1 \bmod(p)$ or $x \equiv -1 \bmod(p)$. If $per[x,m] = 2$ there is at least one $q \in T$ such that $x \equiv -1 \bmod(q)$. Clearly $m-1$ has multiplicative period 2 modulo m.

Corollary 5.2: Let p and q be distinct odd primes whose product is m. Then there are exactly four nonnegative integers x smaller than m such that $per[x,m] = 1$. Two of them are 0 and 1. There are exactly five nonnegative integers x smaller than m such that $per[x,m] = 2$. One of them is $m-1$. More specifically an integer t has multiplicative period 1 modulo pq if and only if there is an ordered pair $(y,z) \in \{0,1\} \times \{0,1\}$ such that $t \equiv y \bmod(p)$ and $t \equiv z \bmod(q)$. An integer t has multiplicative period 2 modulo pq if t does not have period 1 modulo pq but there is an ordered pair $(r,s) \in \{-1,0,1\} \times \{-1,0,1\}$ such that $t \equiv r \bmod(p)$ and $t \equiv s \bmod(q)$.

Let p and q be distinct odd primes. There are integers P and Q such that $Pp + Qq = 1$. Consequently there are nonnegative integers R, S, T and U such that $Rp = Sq + 1$ and $Tp + 1 = Uq$. Let A be the smallest nonnegative integer to which there corresponds a nonnegative integer B such that $Ap = Bq + 1$. Consider the nonnegative integer B above. Clearly B is the smallest nonnegative integer to which there corresponds a nonnegative integer C such that $Cp = Bq + 1$. Evidently $C = A$. Moreover it is clear that A and B are nonzero. Hence they are positive integers. So there is no harm in speaking of the smallest nonnegative integers A and B such that $Ap = Bq + 1$, and in appealing to the fact that A and B are positive.

Lemma 5.2: Let p and q be distinct odd primes. Let A and B be the smallest positive integers such that $Ap \equiv 1 \bmod(q)$ and $Bp + 1 \equiv 0 \bmod(q)$. Then $A + B = q$.

It is a corollary of Lemma 5.2 that

Lemma 5.3: Let p and q be distinct odd primes. Let x and y be the smallest positive integers such that

$$\begin{array}{ll} x \equiv 0 \bmod(p), & x \equiv 1 \bmod(q), \\ y \equiv 1 \bmod(p), \text{ and} & y \equiv 0 \bmod(q). \end{array}$$

Let A, B, C and D be the smallest nonnegative integers such that $x = Ap = Bq + 1$ and $y = Dq = Cp + 1$. Then $A + B + C + D = p + q$.

Lemma 5.4: Let p and q be distinct odd primes. Let A, B, C and D be the smallest nonnegative integers such that $Ap = Bq + 1$ and $Cp + 1 = Dq$. Then either the inequalities $0 < A < q/2$ and $0 < B < p/2$ both hold, or else the inequalities $0 < C < q/2$ and $0 < D < p/2$ both hold.

Theorem 5.5: Let p and q be distinct odd primes. Let $m = pq$. Let x be the smallest nontrivial positive integer which has multiplicative period 1 modulo m. Then $x < m/2$.

Example 5.3: The bound in Theorem 5.5 is a good one. Consider the twin primes $p = 99989$ and $q = 99991$. Let $m = pq$, so that $m = 99980\ 00099$. Let

$$\begin{aligned} x &= 49989\ 50055 = 1 + 49994*99991 = 49995*99989 \\ y &= 49990\ 50045 = 1 + 49996*99989 = 49995*99991. \end{aligned}$$

Then $\mathrm{per}[x,m] = \mathrm{per}[y,m] = 1$ and $0.49999 < x/m < 0.5 < y/m < 0.50001$.

Theorem 5.6: Let p and q be distinct odd primes and let $m = pq$. Let x be a nontrivial positive integer which has multiplicative period 1 modulo m. Then $\mathrm{MAX}\ \{p,q\} \le x$.

23 and 47 are safe primes. Moreover 47 and 1035 have multiplicative period 1 modulo 1081. Therefore Theorem 5.6 is best possible.

Lemma 5.5: Let p and q be distinct odd primes. Let A, C, E and G be the smallest nonnegative integers such that

$$\begin{array}{ll} Ap \equiv -1 \bmod(q), & Cp + 1 \equiv -1 \bmod(q), \\ Ep - 1 \equiv \ \ 0 \bmod(q), \text{ and} & Gp - 1 \equiv \ \ 1 \bmod(q). \end{array}$$

Then they are all positive and $A + E = q$, and $C + G = q$.

Corollary 5.3: Let p and q be distinct odd primes. Let A, B, C, D, E, F, G and H be the smallest nonnegative integers such that

$$\begin{array}{ll} Ap = Bq - 1, & Cp + 1 = Dq - 1, \\ Ep - 1 = Fq, \text{ and} & Gp - 1 = Hq + 1. \end{array}$$

Then these eight integers are all positive and

$$\begin{array}{ll} A + E = q, & B + F = p, \\ C + G = q, \text{ and} & D + H = p. \end{array}$$

Theorem 5.7: Let p and q be distinct odd primes whose product is m. Let x be the smallest positive integer with multiplicative period 2 modulo pq. Then $x < m/3$.

Example 5.4: The bound in Theorem 5.7 is a good one. The numbers $p = 49993$ and $q = 99989$ are primes. Let $m = pq$ so that $m = 49987\ 50077$. Let

$$\begin{aligned} x &= 16662\ 16696 = -1 + 33329*49993 = 16664*99989 \\ y &= 16663\ 16684 = 1 + 33331*49993 = -1 + 16665*99989 \\ z &= 33324\ 33393 = -1 + 66658*49993 = 1 + 33328*99989 \\ t &= 33325\ 33380 = 66660*49993 = -1 + 33329*99989. \end{aligned}$$

Then $\mathrm{per}[x,m] = \mathrm{per}[y,m] = \mathrm{per}[z,m] = \mathrm{per}[t,m] = 2$ and

$$0.3333 < x/m < 1/3 < y/m < 0.3334 < 0.6666 < z/m < 2/3 < t/m < 0.6667.$$

Lemma 5.6: Let p and q be safe primes such that $p < q$ and $pq = m$. Then $\mathrm{per}[2,m] \in \{a(p)a(q), 2a(p)a(q)\}$ and $\mathrm{per}[p-1,m] = 2a(q) = q - 1$, and $\mathrm{per}[q,m] \in \{a(p), 2a(p)\}$. If, moreover, $q \not\equiv 2 \bmod(p)$ then $\mathrm{per}[q-1,m] \doteq 2a(p)$.

Theorem 5.8: Let p and q be distinct odd primes and let $m = pq$. Let x be a positive integer which has multiplicative period 2 modulo m. Then $\mathrm{MAX}\ \{p-1, q-1\} \le x$.

7 and 23 are safe primes. The four smallest positive integers with multiplicative period 2 modulo 161 are 22, 69, 91 and 139. Therefore Theorem 5.8 is best possible.

Theorem 5.9: Let T be a finite set of safe primes whose product is m. Let t be a positive integer, and let x be an integer such that $x\uparrow(1+t) \equiv x \bmod(m)$. Suppose that $\mathrm{GCD}\ \{x-1,m\} = \mathrm{GCD}\ \{x,m\} = \mathrm{GCD}\ \{x+1,m\} = 1$. Then $\Pi\{a(p) \mid p \in T\}$ is a divisor of t.

Theorem 5.10: Let p and q be distinct safe primes whose product is m. Let $a = a(p)$ and $b = a(q)$. Let x and t be integers such that

$$2 \le x \le m-2, \qquad 1 \le t \le m, \text{ and } \qquad x\uparrow(1+t) \equiv x \bmod(m).$$

Let

$$F = \mathrm{GCD}\ \{x-1,m\}, \qquad G = \mathrm{GCD}\ \{x,m\}, \text{ and } \qquad H = \mathrm{GCD}\ \{x+1,m\}.$$

If, on the one hand, $\{F,G,H\} \ne \{1\}$ then $\{F,G,H\} \cap \{p,q\} \ne \emptyset$. If, on the other hand, $F = G = H = 1$ then ab is a factor of t. In this latter case let s be the largest factor of t such that $s < 6$ and let $f = t/s$. If p and q are both larger than 13 then it is true that $f = ab$. It therefore follows that $\{i,j\} = \{p,q\}$, where

$$\begin{aligned} 2i &= m - 4f + 1 + (1 + 16f\uparrow2 + m\uparrow2 - 8f - 2m - 8mf)\uparrow(1/2) \\ 2j &= m - 4f + 1 - (1 + 16f\uparrow2 + m\uparrow2 - 8f - 2m - 8mf)\uparrow(1/2). \end{aligned}$$

Comment: The assumption that $2 \le x \le m-2$ is no restriction since there is no knowledge to be gained from considering integers congruent to one of the trivial values 0, 1 or -1 modulo m. The assumption that $1 \le t \le m$ is no restriction since all multiplicative periods m lie in this interval. The assumption that $\{p,q\} \cap \{7,11\} = \emptyset$ is no restriction since the receiver and the cryptanalyst would routinely check m for small factors. From the cryptanalyst's viewpoint this theorem is important. It says that if a search turns up any nontrivial pair x,t of positive integers smaller than m for which $x\uparrow(1+t) \equiv x \bmod(m)$ then the modulus m can be factored immediately and the code broken.

Example 5.5: Consider the safe primes $p = 47$ and $q = 59$ of Example 4.2 again. For the moment let us take for granted that all the congruences in Table 4 below hold modulo 2773, and apply the methods of Theorem 5.10 to them.

We see that $\text{GCD }\{x,2773\} \in \{47,59\}$ whenever $x \in \{235,236,708,1180,2068,2303,2537,2538\}$, that $\text{GCD }\{x-1,2773\} \in \{47,59\}$ whenever $x \in \{236,471,2302,473,945,2304,2538,2539\}$, and that $\text{GCD }\{x+1,2773\} \in \{47,59\}$ whenever $x \in \{235,471,943,1415,1832,2067,2302,2537\}$. Thus 18 of the 30 bases in Table 4 are keys to factoring 2773 and thereby breaking the code. The other 12 of these bases are useless, so we turn to the exponents associated with them. They belong to the set $A = \{668,1335,2002,2669\}$. When the four integers 667, 1334, 2001, and 2668 are examined for factors smaller than 6 we are left with a single value of f in the application of Theorem 5.10, namely $f = 667$. Recall that $m = 2773$. Thus $m - 4f + 1 = 106$ and $1 + 16f\uparrow 2 + m\uparrow 2 - 8f - 2m - 8mf = 144$. Consequently $\{2i,2j\} = \{106 + 12, 106 - 12\}$ so that $\{i,j\} = \{59,47\}$. Example 4.2 contains, or readily yields, the multiplicative periods modulo 2773 of all the bases appearing in the 30 congruences displayed in Table 4. From this information it is clear that all 30 do in fact obtain.

Admittedly Examples 4.2 and 5.5 beat the RSA cryptosystem based on the safe primes 47 and 59 to death. But they exemplify well what the foregoing theoretical development says about the effect of an attack of the type Simmons and Norris [7] proposed on an RSA public key cryptosystem based on safe primes p and q. Their attack will inevitably succeed *because* p and q were chosen to be safe primes. The foregoing theoretical development makes heavy use of that assumption. But its expected cost will be prohibitive, as Rivest noted [6], *because* p and q are safe primes. This is because a cryptanalyst who finds numbers x and f such that $x\uparrow f \equiv x \bmod(pq)$ knows that either $2f \geq \text{MIN }\{p-1,q-1\}$ or $x \geq \text{MAX }\{p-1,q-1\}$. Rivest, Shamir and Adleman advocate [5, p. 125] a gauge g well in excess of 300. Therefore either $\log(f) \geq 300$ or $\log(x) \geq 300$. Although numbers f of the form $f = c\uparrow n$ in the size range $n \log(c) \geq 300$ can be obtained for small n (n less than 100 if c exceeds 10) there is no reason to assume that their distribution is related to p or q. Thus Rivest's comments in [6] seem persuasive, even though they are not actually proven.

Example 5.6: Let

$$p = 97, \qquad q = 109, \qquad m = 10573,$$
$$\lambda(p) = p-1 = 96, \qquad \lambda(q) = q-1 = 108, \text{ and } \qquad \lambda(m) = \text{LCM }\{96,108\} = 864.$$

The primes 97 and 109 are unsafe. So it is instructive to see how badly the conclusions of Corollary 3.1 and many results in Sections 4 and 5 fail under these unsafe circumstances. There are, of course, still only nine solutions to the congruence $x\uparrow 3 \equiv x \bmod(10573)$. These are the numbers 0, 1, 872, 873, 1745, 8828, 9700, 9701, 10572. To verify this note that the equations

$$\begin{aligned} 872 &= 8 * 109 &&= 9 * 97 - 1 \\ 873 &= 8 * 109 + 1 &&= 9 * 97 \\ 1745 &= 16 * 109 + 1 &&= 18 * 97 - 1 \end{aligned}$$

$$\begin{aligned} 8828 &= 81 * 109 - 1 = 91 * 97 + 1 \\ 9700 &= 89 * 109 - 1 = 100 * 97 \\ 9701 &= 89 * 109 \qquad = 100 * 97 + 1 \end{aligned}$$

have obvious interpretations as congruences modulo 97 and modulo 109. It is easy to find many residue classes modulo 10573 which have small multiplicative periods modulo 10573. In fact it is well known [4, p. 49] that if s is a prime then to each divisor t of $s-1$ there correspond $\phi(t)$ pairwise incongruent numbers whose multiplicative order modulo s is equal to t. When we apply this result to the primes 97 and 109 we obtain Table 6. If we apply Theorem 3.1 to Table 6 and do the necessary double entry bookkeeping on the 12 by 12 array which results, we come up with the information contained in Table 7. One look at Table 7 shows that any number theoretic public key cryptosystem (c, d, 10573) based on $10573 = 97 * 109$ is more vulnerable to an attack of the type G. J. Simmons and J. N. Norris [7] describe than a number theoretic public key cryptosystem based on safe primes. Evidently 325 residue classes x modulo 10573 satisfy the congruence $x\uparrow 25 \equiv x \bmod(10573)$ and 1813 residue classes y modulo 10573 satisfy the congruence $y\uparrow 145 \equiv y \bmod(10573)$. If p and q are safe primes, and $p = 2a + 1$, and $q = 2b + 1$ then $a + b < ab$. Therefore $6ab > (2a+1)(2b+1) = pq = m$. It follows from Corollary 3.1 that $|d| > m/3c - 1$ for every number theoretic public key cryptosystem (c, d, m) in which the modulus m is the product of two safe primes. But that need not be the case in this unsafe example. In fact consider (c, d, m) = (5, 173, 10573). Evidently $5 * 173 \equiv 1 \bmod(864)$ so that 5 can be taken as a public coding exponent and 173 as a secret decoding exponent. But the conclusion of Corollary 3.1 fails, since

$$d = 173 < 1035 < 5184/5 - 1 = 2*48*54/5 - 1 = 2*[(p-1)/2]*[(q-1)/2]/c - 1.$$

The conclusion of Lemma 4.2 fails. The congruence $x\uparrow f \equiv x \bmod(m)$ holds identically in x for any of the thirteen members f of the set

$$\{1, 865, 1729, 2593, \ldots, 9505, 10369\} = \{1 + 864t | 0 \le t \le 12\}$$

The conclusion of Corollary 4.1 fails. Table 7 exhibits 24 multiplicative periods modulo m. Similarly the conclusion of Theorem 4.3 fails. Now we note in passing that

$$\begin{aligned} 1199 &= 11 * 109 = 12 * 97 + 35 \\ 1962 &= 18 * 109 = 20 * 97 + 22 \\ 327 &= 3 * 109 = 3 * 97 + 36 \\ 2943 &= 27 * 109 = 30 * 97 + 33. \end{aligned}$$

Making straightforward use of these equations, which have obvious interpretations as congruences, and simply calculating the multiplicative periods modulo 97 of several residue classes modulo 97 we see that no analog of Lemma 5.1 holds, since

$$\begin{aligned} \mathrm{per}[35,97] &= \mathrm{per}[1199, 10573] = 3 \\ \mathrm{per}[22,97] &= \mathrm{per}[1962, 10573] = 4 \\ \mathrm{per}[36,97] &= \mathrm{per}[\ 327, 10573] = 6 \\ \mathrm{per}[33,97] &= \mathrm{per}[2943, 10573] = 8. \end{aligned}$$

$235\uparrow 5 \equiv 235$	$1415\uparrow 231 \equiv 1415$
$236\uparrow 5 \equiv 236$	$4\uparrow 668 \equiv 4$
$471\uparrow 9 \equiv 471$	$9\uparrow 668 \equiv 9$
$2538\uparrow 30 \equiv 2538$	$16\uparrow 1335 \equiv 16$
$2068\uparrow 30 \equiv 2068$	$2537\uparrow 1335 \equiv 2537$
$2304\uparrow 30 \equiv 2304$	$2\uparrow 1335 \equiv 2$
$473\uparrow 70 \equiv 473$	$3\uparrow 1335 \equiv 3$
$708\uparrow 70 \equiv 708$	$5\uparrow 1335 \equiv 5$
$1832\uparrow 117 \equiv 1832$	$11\uparrow 1335 \equiv 11$
$2067\uparrow 117 \equiv 2067$	$25\uparrow 2002 \equiv 25$
$2303\uparrow 117 \equiv 2303$	$2302\uparrow 2669 \equiv 2302$
$2539\uparrow 117 \equiv 2539$	$6\uparrow 2669 \equiv 6$
$943\uparrow 231 \equiv 943$	$7\uparrow 2669 \equiv 7$
$945\uparrow 231 \equiv 945$	$10\uparrow 2669 \equiv 10$
$1180\uparrow 231 \equiv 1180$	$19\uparrow 2669 \equiv 19$

Table 4

Some congruences modulo 2773

Let p and q be distinct safe primes. Let $(p-1)/2 = a < b = (q-1)/2$.

If the multiplicative *period* v *modulo* pq belongs to the set	then a sharp lower bound for nontrivial *positive integers* with multiplicative period v is	and a sharp upper bound for the smallest positive integers with multiplicative period v is
$\{1\}$	MAX $\{p,q\}$	$pq/2$
$\{2\}$	MAX $\{p-1,q-1\}$	$pq/3$
$\{a\}$	q	
$\{b\}$	p	
$\{2a\}$	$q-1$	
$\{2b\}$	$p-1$	
$\{ab\}$	2	
$\{2ab\}$	2	
$\{a,2a\}$	$q-1$	q
$\{b,2b\}$	$p-1$	p
$\{ab,2ab\}$	2	2

Table 5

Multiplicative period u modulo 109	Number of residue classes modulo 109 having multiplicative period u modulo 109	Multiplicative period v modulo 97	Number of residue classes modulo 97 having multiplicative period v modulo 97
1	2	1	2
2	1	2	1
3	2	3	2
4	2	4	2
6	2	6	2
9	6	8	4
12	4	12	4
18	6	16	8
27	18	24	8
36	12	32	16
54	18	48	16
108	36	96	32
	total 109		total 97

Table 6

Multiplicative period w modulo 10573	Number of residue classes modulo 10573 having multiplicative period w modulo 10573
1	4
2	5
3	12
4	16
6	28
8	20
9	24
12	104
16	40
18	60
24	136
27	72
32	80
36	228
48	272
54	180
72	288
96	544
108	684
144	576
216	864
288	1152
432	1728
864	3456
	total 10573

Table 7

This, incidentally, shows that the conclusion of Theorem 5.1 fails. Nothing like Theorem 5.10 applies either. To summarize the example we may say that the guaranteed search difficulty which safe primes assure is missing. There are small secret decoding exponents and many messages with small multiplicative period modulo 10573. Thus a Simmons-Norris cryptanalytic attack which concentrates on decoding individual messages is often cheap. But the cryptanalyst is perhaps denied the triumph that comes with decoding even one nontrivial message in a cryptosystem based on safe primes, the pleasure of thereby breaking the entire code by quickly inferring the factorization of the coding modulus m.

Hopefully the significance of safe primes is clear from Example 5.6. Number theoretic public key cryptosystems based on them have a certain brittle resistance to cryptanalysis. It is hard to make a dent in such a cryptosystem. But if you do, you have broken it completely. Cryptosystems based on unsafe primes are easier to break partially. But work still remains after a partial break.

How common, then, are safe primes? The famous approximation $\pi(x) \sim \mathrm{li}(x)$ yields, when "differentiated", the rule of thumb that a positive integer x is prime with probability not far from $P(x) = 1/\ln(x)$. *If the events* $\{w$ *is a prime which is congruent to* 3 modulo $4\}$ and $\{(w-1)/2$ is a prime$\}$ were stochastically independent then a positive integer v would be a *safe prime with probability not far from*

$$Q(v) = 1/[2\ \ln(v)\ln((v-1)/2)] = 1/[2\ \ln(v)[\ln(v-1) - \ln(2)]].$$

Among positive integers of 100 digits we know from the value of $P(x)$ that primes occur more often than thrice in 1000. In the same size range safe primes would be expected, from the value of $Q(x)$, to occur at least once in 1 00000. This density would make their use in constructing public key cryptosystems feasible. But is the stochastic independence assumption behind $Q(x)$ a correct one? It may be cheaper to search for safe primes in any specified size range than to try to prove or disprove stochastic independence. We will content ourselves with the following observation. There are 205 primes between 6 and 1284. We have previously exhibited all safe primes between 6 and 1284. There are 27 of them. Integrating $Q(x)$, which is derived on the basis of the stochastic independence assumption, and $P(x)$ we find that

$$22 \le \int dx/[2\ \ln(x)\ln((x-1)/2)] \le 23, \text{ and} \qquad\qquad 213 \le \int dx/[\ln(x)] \le 214,$$

where the integrals above are both from 6 to 1284. The latter integral exemplifies the well known fact that $\pi(x) < \mathrm{li}(x)$ for small x. *The former integral does not weaken the stochastic* independence assumption. In fact it shows that there are more small safe primes than that assumption suggests. It is nevertheless possible that there are *only finitely many safe primes*.

Example 5.7: It is evident that $13 * 19 = 247$ and that $61 * 85 = 1 + 24 * 216$. Therefore $(61, 85, 247)$ is an RSA number theoretic public key cryptosystem but $61 = 12 * 5 + 1$. Consequently $x\uparrow 61 \equiv x \bmod(13)$ for every integer x. Also $x\uparrow 61 \equiv x\uparrow 7 \equiv x \bmod(19)$ for x belonging to one of the equivalence classes 0, 1, 7, 8, 11, 12, 18 modulo 19. Therefore 91 of the 247 residue classes modulo 247 consist of integers x such that $x\uparrow 61 \equiv x \bmod(247)$. This number theoretic public key cryptosystem therefore conceals fewer than 64% of the messages which can be sent in it. More than 36% of possible messages, in other words, have the property that their cleartext is the same as their ciphertext.

6. The directorate and the message receiver in an RSA public key cryptosystem. If the directorate chooses a width $w > 2$ then it allows every message receiver N to pick primes $p(N)$ and $q(N)$ at random such that

$$g < \log(p(N)) < g + 1 < g + w/2 < g + w < \log(q(N)) < g + 3w/2 .$$

This guarantees that $2g < 2g + w < \log(p(N)q(N)) < 2g + 2w$ and that $2p(N) < q(N)$, whence random search for factors of $m = p(N)q(N)$ near $\sqrt{m}$ becomes expensive for a cryptanalyst if the gauge g is large.

Every message receiver N takes some pains to provide a coding exponent $c(N)$ which is relatively prime to $[p(N)-1][q(N)-1]$. Why? Nothing in [5] or in this paper justifies such an effort. The directorate can choose a positive integer u as a universal coding exponent. In other words the directorate can require that $c(N) = u$ for every receiver N until such time as a reason be adduced for having every receiver N provide a distinctive coding exponent $c(N)$. Such a decision will cut the size of the directory by almost half, and will standardize the coding process. Thus the first page of the directory will contain g, w and u. If the universal coding exponent u is a large prime there will be a high likelihood that $\text{GCD}\,\{u, [(p-1)(q-1)]\} = 1$ for randomly chosen large positive integers p and q which are prime to all intents and purposes.

Rivest, Shamir and Adleman note [5, p. 125] the desirability of having a coding exponent $c(N)$ satisfy the inequality $c(N) > 2g + 2w$. Since $\log(m) < 2g + 2w$ it is clear that $x\uparrow(c(N)) > m$ for every positive integer $x \geq 2$. Thus every nontrivial message must undergo reduction modulo $m(N)$, or *wraparound,* in their terminolgy. Under these circumstances, assuming that $300 < g < 400$ and that $3 \leq w \leq 10$ we can see that the universal coding exponent u (which is equal to $c(N)$ for every receiver N) might be chosen so that $u > 820$. On the other hand we know from Theorem 3.3 that the receiver N has a decoding exponent $d(N)$ which satisfies the inequality

$$|d(N)| > \lambda(m(N))/c(N) - 1 = \lambda(m(N))/u - 1.$$

The directorate has the best interests of the message receiver at heart. So it wants $d(N)$ to be large for every receiver N. This makes a cryptanalyst's search for $d(N)$ expensive.

If u is a Fermat [3, pp. 14-16] prime then $u-1$ is a power of 2. Therefore $\text{GCD}\,\{u-1, \lambda(m)\} = 2$ for every square free integer m which is a finite nonvoid product of primes which are congruent to 3 modulo 4. Hence u is a deranging exponent for such m if u is not a factor of $\lambda(m)$. Consequently $u = 65537$ is an attractive candidate for a universal coding exponent. It is so large that:

1. Every nontrivial message x undergoes wraparound if $g \leq 30000$ and $w \leq 2000$;
2. 99.99% of positive integers are relatively prime to u. Hence it is easy to search for primes p such that u is not a factor of $p-1$.

And it is so small that:

3. It takes fewer than 17 multiplications to form a uth power;
4. The inequality $2g - 20 < \log(d)$ holds when u is the coding exponent and d is the decoding exponent in an RSA number theoretic public key cryptosystem based on safe primes p and q.

If p and q are safe primes and $m = pq$ the receiver knows that the cryptanalyst has a guaranteed way to decode messages. But the expected cost is prohibitively high.

The receiver N will not act exactly as described in the paper [5] of Rivest, Shamir and Adleman. For one thing he is not free to choose his coding exponent $c(N)$. The directorate provides him with the prime u which he, like all receivers, will use as coding exponent. For another thing, he wants his secret primes p and q to be safe primes. Also the receiver N will take the precautions suggested by [1] to guarantee that the cryptosystem conceals as many messages as possible. Therefore he will do the following, assuming that $w > 2$.

Step 1. Choose at random an odd positive integer a such that $g - 1 < \log(a) < g + w/2 - 1$.

Step 2. Form GCD $\{r, a\}$ and GCD $\{r, 2a+1\}$ for every prime $r \leq u$. Form GCD $\{a, (u-1)/2\}$. If any of these numbers is unequal to 1, forget a and return to Step 1.

Step 3. Test whether a and $2a+1$ are both *prime to all intents and purposes*. If either is demonstrably composite, forget a and return to Step 1.

Step 4. Choose at random an odd positive integer b such that

$$g + w - 1 < \log(b) < g + 3w/2 - 1.$$

Step 5. Form GCD $\{r, b\}$ and GCD $\{r, 2b+1\}$ for every prime $r \leq u$. Form GCD $\{b, (u-1)/2\}$. If any of these numbers is unequal to 1, forget b and return to Step 4.

Step 6. Test whether b and $2b+1$ are both *prime to all intents and purposes*. If either is demonstrably composite, forget b and return to Step 4.

Step 7. Form GCD $\{a,b\}$, GCD $\{a, 2b+1\}$, GCD $\{2a+1,b\}$ and GCD $\{2a+1, 2b+1\}$. If any of these numbers is unequal to 1, forget a and b and return to Step 1.

Step 8. Solve the six pairs of simultaneous linear congruences:

$$\begin{array}{ll}
A \equiv 0 \bmod(2a+1), \text{ and} & A \equiv 1 \bmod(2b+1); \\
B \equiv 1 \bmod(2a+1), \text{ and} & B \equiv 0 \bmod(2b+1); \\
C \equiv 0 \bmod(2a+1), \text{ and} & C \equiv -1 \bmod(2b+1); \\
D \equiv -1 \bmod(2a+1), \text{ and} & D \equiv 0 \bmod(2b+1); \\
E \equiv 1 \bmod(2a+1), \text{ and} & E \equiv -1 \bmod(2b+1); \\
F \equiv -1 \bmod(2a+1), \text{ and} & F \equiv 1 \bmod(2b+1).
\end{array}$$

Examine the Hollerith character typescripts which correspond to A, B, C, D, E and F. If all six of these typescripts are hopeless gibberish, go on to Step 9. Otherwise forget a and b and go back to Step 1.

Step 9. Let

$p(N) = 2a + 1,$ $\qquad a(p(N)) = a,$

$q(N) = 2b + 1,$ $\qquad a(q(N)) = b,$

$m(N) = p(N)q(N) = (2a+1)(2b+1),$ and $\qquad v = 2ab.$

Comment: The receiver N now knows that $m(N)$ is square free if both $p(N)$ and $q(N)$ are square free. He believes, of course, that the four integers $a(p(N))$, $a(q(N))$, $p(N)$ and $q(N)$ are all primes. His belief need not be correct.

Step 10. Solve the linear congruence $ud \equiv 1 \bmod(v)$ for d. Call its smallest positive integer solution $d(N)$.

Step 11. Send the list $(N, m(N))$ to the directorate for inclusion as a listing in the directory. It satisfies the condition $2g < \log(m(N)) < 2g + 2w$.

Step 12. Keep $p(N)$, $q(N)$ and $d(N)$ secret.

7. The cryptanalyst and the sender in an RSA public key cryptosystem. The cryptanalyst has her work cut out for her. She can try the best known factoring algorithms [5, p. 125] on a given coding modulus m. But m is a 200 digit number. And none of these algorithms is cheap. She can try to decode a message x with respect to m.

Definition 7.1: Let m be a positive integer. Suppose that $2 \le x \le m - 1$. To *decode the message* x with respect to m is to find a positive integer $f < m$ such that $x\uparrow(1+f) \equiv x \bmod(m)$.

If she can decode even one single message, and if m is the product of two distinct safe primes, she can use Theorem 5.10 to factor m and, thus, break the code.

Definition 7.2: Let m be a square free positive integer. To *break the code* of which m is the coding modulus and c is the coding exponent is to find a positive integer $d < m$ such that $x\uparrow cd \equiv x \bmod(m)$ for every integer x.

In particular the cryptanalyst can decode a message x if she can take its coded form $x\uparrow c \bmod(m)$ and find f such that $x\uparrow cf \equiv x \bmod(m)$. But the complementarity principles militate against a small expectation of decoding cost.

There are two integers x such that $2 \le x \le m - 1$ and $x\uparrow 2 \equiv x \bmod(m)$. The smaller of them, call it s, satisfies the inequalities $\mathrm{MAX}\ \{p, q\} \le s < m/2$ because of Theorems 5.5 and 5.6. There are four integers x such that

$$2 \le x \le m - 1, \qquad x\uparrow 3 \equiv x \bmod(m), \text{ and } \qquad x\uparrow 2 \not\equiv x \bmod(m).$$

The smallest of them, call it t, satisfies the inequalities $\mathrm{MAX}\ \{p-1,q-1\} \le t < m/3$ because of Theorems 5.7 and 5.8. Random searches in these ranges are expensive, even though the search for t is only about $1/3$ as expensive as the search for s. The cryptanalyst should test every intercepted message y to see whether $y\uparrow 3 \equiv y \bmod(m(N))$ just on the off chance. It is unlikely, but the code is broken if it occurs.

The cryptanalyst can try to find a positive integer y such that $2\uparrow(1+y) \equiv 2 \bmod (m)$. But $\log(y)$ is necessarily larger than $2g$. Hence $\log(\log(2\uparrow(1+y))) > 2g$ and $2g$ is presumably larger than 600. How can she reduce so large a number as $2\uparrow(1+y)$ modulo m in a reasonable period of time? Every way she turns complementarity says the expected expense is huge. But, still, there are winning tickets in lotteries with small expectations. Why shouldn't the cryptanalyst get lucky, and quickly? No lower bounds exist to answer the question.

Number theoretic public key cryptosystems are analogous to the Gordian knot. Those who sought to undo it fell into three categories. Cryptanalysts take note. The *fumblers* tried to pull every which way on it. This approach corresponds to reliance on a powerful computer without a very deep analysis. A computer the size of the sun running for a billion years has a tiny probability of breaking an RSA cryptosystem if it works at random. No fumbler undid the Gordian knot. The *thinkers* tried to understand the knot. The modern analog of this approach calls for successively deeper and deeper analysis of a mathematical nature. No thinkers undid the knot. Alexander of Macedon sundered the Gordian knot with his sword after trying the other approaches. The *alexandrine* attitude to any cryptosystem is that reading the cleartext message - not necessarily decoding the coded message - is the objective. The sender and the receiver both have a copy of the cleartext message. Moreover there are instances [1] of RSA public key cryptosystems in which many or all messages are unchanged by the coding process.

At present the cryptanalyst's most cost effective strategy may be disruption or corruption of a given receiver's organization. Failing this she may want to try the same approach toward several of the most important senders. However, alexandrine cryptanalysis is both clever and surprising. If a clever stroke, based on a seemingly trivial observation, breaks the code it will be alexandrine by definition.

A sender who wants to transmit a message to receiver N without disclosing its content to anybody else must assume that the cryptanalyst is very rich and very able. He would like the receiver to be at least as motivated to guarantee secrecy, and at least as able to work to that end as himself. It is relatively easy for a receiver to obtain considerable security cheaply. Moreover he will be motivated to think of the most important possible incoming message in formulating his public key cryptosystem. It is quite likely that the hypothetical most important incoming message is more important to him than a sender's current outbound message is to that sender. A sender will take a routine precaution with each message x, and verify that the congruence $x\uparrow 3 \equiv x \bmod (m)$ does not obtain. If it does, he will either inform the receiver that the code has been broken or else go into the cryptanalysis business against that receiver.

8. Summary. Suppose that m is an odd, square free, composite integer. Then there are integers c, d larger than 2 such that the congruence $x\uparrow cd \equiv x \bmod (m)$ holds for every integer x. In fact this occurs when $cd \equiv 1 \bmod (\lambda(m))$.

RSA public key cryptosystems are based on the assumption that m has exactly two prime factors. When the logarithm (to base two) of each factor exceeds 300 the expected cost of breaking the code, *i.e.* factoring m, appears prohibitive. Rivest, Shamir and Adleman [5] considered direct factorization of m at length. We have considered search for f such that the congruence $x\uparrow f \equiv x \bmod(m)$ holds for every integer x, under the assumption that the two prime factors of m are safe primes. These papers concur in recommending the use of safe primes [5, p. 124] to those who seek security in RSA cryptosystems. We have given, in Theorem 5.10, an algorithm for breaking a cryptosystem based on safe primes p and q. But the complementarity results in Section 5 show that it does not have acceptable expected cost.

I and II have concentrated on number theoretic public key cryptosystems in which a receiver is correct in assuming that his coding modulus m is the product of two large safe primes. But the results herein have wider applicability. On the one hand they are the starting point of an analysis of number theoretic public key cryptosystems based on general square free positive integer coding moduli m. On the other hand they make it possible to examine in detail what happens when $m = pq$ is square free but p and q are not both prime, even though a receiver finds c and d such that $cd \equiv 1 \bmod(\mathrm{LCM}\{p-1,q-1\})$, and uses c for a coding exponent and d for a decoding exponent. In this case, $\lambda(m) \neq \mathrm{LCM}\{(p-1),(q-1)\}$. Finally, it is instructive to consider what happens when m is not even square free. The next paper in this series, III, will appear in the next issue of CRYPTOLOGIA. It will consider some examples of such behavior and will provide proofs of the results stated heretofore.

REFERENCES

1. Blakley, G. R. and Borosh, I., Rivest-Shamir-Adleman public key cryptosystems do not always conceal messages, Computers and Mathematics with Applications, 5 (1979) to appear.

2. Diffie, W. and Hellman, M. E., New directions in cryptography, IEEE Trans. on Infor. Th., IT-22, 6 (1976), 644-654. Also private communication regarding safe primes in public key distribution systems, April 1978.

3. Hardy, G. H. and Wright, E. M., An Introduction to the Theory of Numbers, Fourth Edition. (London: Oxford University Press, 1965).

4. LeVeque, W. J., Topics in Number Theory, Volume 1, First Edition. (Reading, Massachusetts: Addison-Wesley Publishing Company, 1958).

5. Rivest, R. L., Shamir, A. and Adleman, L., A method for obtaining digital signatures and public key cryptosystems. Comm. ACM, 21 (1978), 120-126.

6. Rivest, R. L., Remarks on a proposed cryptanalytic attack of the "M.I.T. public key cryptosystem", CRYPTOLOGIA, 2 (1978), 62-65.

7. Simmons, G. J. and Norris, J. N., Preliminary comments on the M.I.T. public key cryptosystem, CRYPTOLOGIA, 1 (1977), 406-414.