金曜高校生講座

# 素敵な数、素数

東京大学数理科学研究科　寺杣　友秀

平成２７年1月２３日

## 1. 数理科学研究科の紹介

数理科学研究科は純粋数学をはじめ、数学と他の分野の融合分野を研究しているところです。大学における数学について、大雑把に紹介したいと思います。数学の分野は大きく、代数、幾何、解析があり、さらにその他の物理学、生物学、経済への数学の応用を見込んだ、数学の応用分野があります。高校での数学との関連でいうと、代数学は多項式や多項式を用いた方程式や、整数などの数の体系を扱う分野です。今日お話する素数とその性質を調べることは、この分野とのかかわりが深いところです。方程式を考えるとき、たとえば２次方程式でもそうですが、グラフなどの幾何との関係を積極的に用いて考える手法も取り入れられています。幾何は図形に関する性質を研究する分野です。高校では直線と円、それから放物線、楕円などを図形として扱いますが、実際に存在する図形はこのように単純なものばかりではありません。もう少し一般の図形の曲がり方などの性質を研究することも幾何学の分野です。考える世界も２次元といわれる平面、３次元といわれる空間からもっと高い次元の「空間」で考えたり、２次元でも平面の上ばかりではなく、球面のように曲がった空間の上で数学を展開したりすることも幾何において重要な課題です。自然科学でさまざまな量を扱うときに必要となってくるのが、関数の考え方ですが、関数において値の変化の様子をしらべたり、極限操作を用いて研究する分野が解析です。高校でも微分積分を習いますが、これは解析分野の代表的な手法です。以上に述べた事でもわかると思いますが、ここまでが代数の範囲とか、ここまでは幾何や解析といったはっきりした境界があるわけではなく、だいたいの方向性を示す分野わけといえます。今日お話する「素数」の話も代数分野に関連が深いものですが、中では解析分野の考え方もたくさん使われます。

また、ほとんどの自然科学において分析の仕方や論理的な考察の基礎には数学が使われていますから、自然科学をより深く研究しようとするならば、基礎的な数学の考え方は必要です。きちんとした分析をして、確かな事実を積み重ねていくことは、新しい発想を実らせていくための一番の近道であることが多いものです。そういった科学の土台となる数学は大学の教養課程で学ぶことができます。

## 2. 素数はたくさんあるか

自然数とは $n = 1, 2, 3, \ldots$ といった数のことですが、$a, b$ を自然数としたとき、加法 $a + b$ と乗法 $ab$ を考えることができます。$ac = b$ となる自然数 $a, c$ があるとき、$b$ は $a$ で割り切れるといいます。このとき $b$ は $a$ の倍数であるといいい、$a$ は $b$ の約数であるといいます。$p$ を 2 以上の自然数とします。$p$ がどんな自然数であっても 1 と $p$ は $p$ の約数ですが、この二つ以外に約数がないとき、$p$ は素数とよばれます。たとえば 7 はそういう性質をもっているので、素数です。28 は 2 が 1 と 28 以外の約数になりますから、素数ではありません。素数でない数を、(1 は除外します。) 合成数といいます。1 とその数以外に約数があったらどんどん積の形に分解して書いてゆくことにより、最終的にどんな 2 以上の自然数も素数の積にかけることがわかります。たとえば

$$28 = 4 \times 7 = 2 \times 2 \times 7$$

という具合です。べつの仕方で

$$28 = 2 \times 14 = 2 \times 2 \times 7$$

とやってもかまいません。実は、どういうやり方でやっても、最終的に素数の積の形に書いたとき、掛け算の順番を変えれば、どれも同じになる事が証明できます。この性質を素因数分解の一意性といわれます。皆さんはこれを当たり前だと思うかも知れませんが、証明しなければいけない事です。今は便利なので、インターネットを探せばその証明は見つかるかもしれませんが、自分で考えてみるのもよいでしょう。証明するときには無条件で使ってよい事実の範囲は定めておかなければなりませんが、その範囲をどう定めるかというのも含めて、考えてみると面白いと思います。

素数は自然数の掛け算を元に考えて、一番基本的な構成要素といえます。素数を小さいほうから書いていくと、

$$2, 3, 5, 7, 11, 13, 17, \ldots$$

となります。そこで問題です。

**問題 2.1.** 素数はどれくらいあるのでしょうか？

もし素数が有限個しかなければ、それらのべきと積であらわせるので、とてもらくちんですが、世の中そんなに甘くありません。実際は次の定理が成り立ちます。

**定理 2.2.** 素数は無限にある。

証明. もし素数が有限個しかなければ、素数のすべてを $p_1, p_2, \ldots, p_n$ とおく。このとき $k = p_1 p_2 p_3 \cdots p_n + 1$ とおくと、$k$ は $p_1$ で割ると 1 余るので $p_1$ では割り切れない。同様に $p_2$ でも、$p_3, \ldots, p_n$ でも割り切れないことがわかる。従って、この数を素因数分解すれば、$p_1, \ldots, p_n$ 以外の素因数が出て来て、素数のすべてを $p_1, \ldots, p_n$ としたことに反する。したがって、素数は無限個ある。 □

上のように、定理を証明したいときにそれがもし成り立っていなければ、不合理が生じる、という仕方で証明する証明方法を背理法といいます。上の証明方法はユークリッド原論にも書かれていて、素数は無限にある事の証明はそのころには知られていたことになります。

## 3. 素数の密度

無限にあることはわかりましたが、始めの問題に戻って、どれくらいあるのでしょうか？無限にあるもの（ここでは自然数の部分集合を考えます）を、何を基準にしてどう図ればよいのでしょうか？たとえば偶数と 2 のべきつまり

$$1, 2, 2^2 = 4, 2^3 = 8, 2^4 = 16$$

はどちらが多いでしょうか？偶数と 2 のべきには $2n$ と $2^{n-1}$ を対応させることにより、1 : 1 の対応があるので、そういう見方をすれば、同じ個数だけあるともいえるのですが、ここではそういう考え方をせず、数の集合の密度で考えてみましょう。たとえば、1 から 100 までで考えると偶数は 50 個、2 のべきは $1, 2, 4, 8, \ldots, 512$ までで、10 個なので 2 のべきのほうが少ないと考えます。これがもっとたくさん、たとえば 10000 まで考えれば、この差はもっと開くはずです。

個数を比較するのに密度を考えるのは自然です。たとえば $N$ を十分に大きな数として $A_N = \{n \mid n \text{ は偶数}, n \leq N\}$, $B_N = \{n \mid n \text{ は } 2 \text{ のべき}, n \leq N\}$ としたとき $\dfrac{A_N}{B_N}$ が $N$ が無限にいったときの極限は無限になることが証明できます。このようなことが成り立つとき、偶数の密度は 2 のべきであらわされる数の密度よりおおきいということにします。もちろん 1 で始まる偶数の集合と 2 で始まる自然数の集合のように比較できないこともあります。

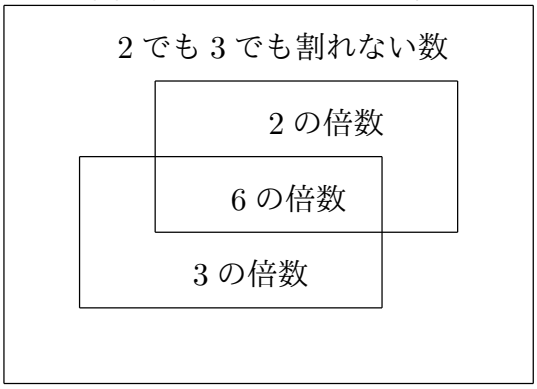
それでは、素数はどれくらいあるかというに問題に戻る前にどれくらい素数があるか実験してみましょう。まず、10 の付近の素数の密度を考えるのに 10 の次にくる素数をみてみます。つぎに $10^2 = 100$ の付近の素数の密度を考えるのに 10 の次にくる素数をみてみます。このようにして $10^3$ の付近、$10^4$ の付近と見てゆき、$10^{20}$ くらいまで見ると下の表のようになります。この表に $10^{20}$ のところは数値をいれていませんが、どれくらいになるか予想してみましょう。（答えは講義の時に発表します。）

| $10^n$ | その次の素数 | $10^n$ | その次の素数 |
|---:|---:|---:|---:|
| 1 | 2 | 10000 | 10007 |
| 10 | 11 | 100000 | 100003 |
| 100 | 101 | 1000000 | 1000003 |
| 1000 | 1009 | $10^{20}$ | ??? |

どういう感想を持ちましたか？さてそれでは素数の密度について考えてみましょう。$N$ を十分に大きい数とすると、$N$ が素数かどうか確かめるのには $N$ より小さい素数について考えればよいことになります。ここで自然数をとってきた時、それが 2 の倍数にならない確率を考えるとだいたい $\dfrac{1}{2}$ になることがわかります。また 3 の倍数にならないということは 3 で割ったあまりが 1,2 ということなので、その確率は $\dfrac{3-1}{3} = \dfrac{2}{3}$ になります。これらが同次に成り立つ確率、つまり 3 でも 2 でも割り切れない確率は 6 で割ったときのあまりが 1 または 5 となる時なのでその確率は

$$\frac{1}{2} \cdot \frac{2}{3} = \frac{2}{6}$$

となります。これは下のベン図を用いて次のように考えることができます。




この図を用いると

$$1-\frac{1}{2}-\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}=(1-\frac{1}{2})(1-\frac{1}{3})$$

として求めてもよいことになります。$p$ を $N$ 以下の最大の素数として、$N$ までの自然数の中では素数であることは、$2,3,5,\ldots,p$ で割り切れないことと同値になりますから、その確率は

$$C_p=(1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{5})\cdots(1-\frac{1}{p})$$

となると考えましょう。上の積は $p$ までの素数をわたるものです。大体 $p$ と $N$ が等しいとして、これは $N$ が大きくなるとどれくらい小さくなるでしょうか？確率ですからもちろん 1 より小さい値になり、1 より小さい数をかけていくのでどんどん小さくなるはずです。この素数の密度を、もう少しよくわかった $N$ の関数、たとえば $\frac{1}{\sqrt{N}}$ などと比較できるでしょうか？ $N$ 個中（$N\geq 2$）で、素数という条件を満たすものの密度なので $\frac{1}{N}$ よりもちろん大きくなります。

## 4. $\frac{1}{C_p}$ の大きさ

$C_p$ の大きさを測るのに、便利な方法があります。まず等比級数の和の公式

$$1+r+r^2+\cdots+r^{n-1}=\frac{1-r^n}{1-r}$$

を $r=\frac{1}{2}$ として用いて、$n$ が無限になるときを考えると、

$$1+\frac{1}{2}+(\frac{1}{2})^2+(\frac{1}{2})^3+(\frac{1}{2})^4+\cdots=\frac{1}{1-\frac{1}{2}}$$

となります。したがって

$$\begin{aligned}\frac{1}{C_p} =&\left(1+\frac{1}{2}+(\frac{1}{2})^2+(\frac{1}{2})^3+(\frac{1}{2})^4+\cdots\right)\\ &\left(1+\frac{1}{3}+(\frac{1}{3})^2+(\frac{1}{3})^3+(\frac{1}{2})^4+\cdots\right)\\ &\cdots\\ &\left(1+\frac{1}{p}+(\frac{1}{p})^2+(\frac{1}{p})^3+(\frac{1}{p})^4+\cdots\right)\end{aligned}$$

となります。これを展開すると、$m$ を素因数分解したときに $2,3,\ldots,p$ の因数ものをもつ自然数にわたって $\dfrac{1}{m}$ を加えたものに等しくなります。1 から $N$ までの因数はすべて $p$ 以下ですから、

$$1+\frac{1}{2}+\cdots\frac{1}{3}+\cdots+\frac{1}{N}\leq\frac{1}{C_p}$$

がなりたちます。右辺の式はよく出てくる形なので $\sum_{i=1}^{N}\dfrac{1}{i}$ と書きます。この和をあらわす簡単な公式はないのですが、だいたいの大きさは求めることができます。積分を使って求めるほうがより精密にできるのですが、ここでは初等的な方法を用いることのできる、$N=2^n$ の場合を考えます。

$$\begin{aligned}&1+\frac{1}{2}+\underbrace{\frac{1}{3}+\frac{1}{4}}+\underbrace{\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}}+\cdots+\underbrace{\frac{1}{2^{n-1}+1}+\cdots+\frac{1}{2^n}}\\ \geq&1+\frac{1}{2}+\underbrace{\frac{1}{4}+\frac{1}{4}}+\underbrace{\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}}+\cdots+\underbrace{\frac{1}{2^n}+\cdots+\frac{1}{2^n}}\\ \geq&1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\cdots+\frac{1}{2}=1+\frac{n}{2}\end{aligned}$$

したがって

$$C_p\leq\frac{2}{n+2}$$

となります。対数関数を用いれば、$n=\log_2(N)$ と表すことができます。つまりおおよそ、次の不等式が成り立ちます。

$$(1\text{ から N までの数の中の素数の確率})\leq\frac{2}{\log_2 N+2}$$

したがって素数の確率は $N$ に対して非常にゆっくり 0 に近づく関数で上から評価される、ということになります。

## 5. 素数定理

上はずいぶん大雑把な評価ですが、思ったよりはよい評価で、実は次の定理が成り立つことが知られています。

**定理 5.1** (素数定理、アダマール、ド・ラバレ・プサン)**.** $P(n)$ を 1 から $n$ までの間の素数の数とすると、

$$\lim_{n\to\infty} \frac{P(n)\log(n)}{n} = 1$$

となる。つまり、1 から $n$ までにある素数の密度は $\dfrac{1}{\log(x)}$ である。

上の定理において対数関数 $\log(x)$ の底はネーピアの定数（あるいは自然対数の底）と呼ばれる定数で、その近似値は

$$e = 2.71828182845904\cdots$$

となっています。下に $\log(x)$ の表を示します。

| $x$ | $\log(x)$ | $x$ | $\log(x)$ |
|---|---|---|---|
| 10 | 2.30258509 | 10000 | 9.21034037 |
| 100 | 4.60517018 | 100000 | 11.51292546 |
| 1000 | 6.90775527 | 1000000 | 13.81551055 |

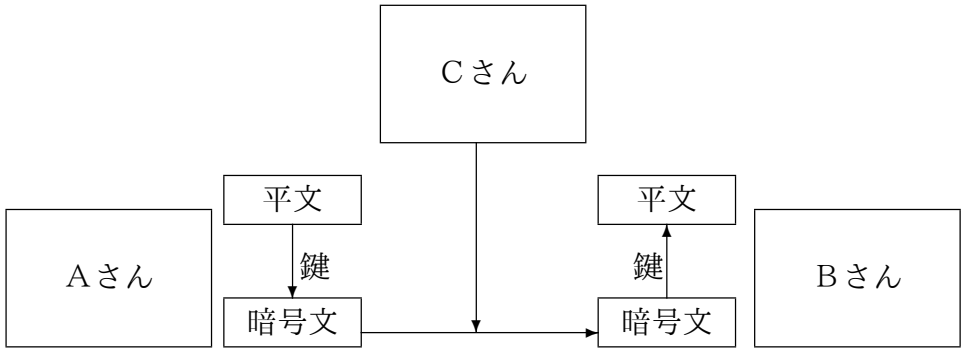
つまりこの表によれば、10 万までの数の範囲で考えれば、11 個か 12 個に一つは素数があることになります。素数定理として、1 から $n$ までの範囲での素数の密度に関する定理としてのべましたが、実は $x$ のまわりの素数の密度というのも定義できます。$\log(x)$ の増大の仕方が非常にゆっくりであることから、結果としてはこれも同じ $\dfrac{1}{\log(x)}$ で与えられることが知られています。だからたとえば 6 桁の数を考えたときに確率的には 10 個から 14 個に一つは素数があることになります。これはたとえば 6 桁の素数をある人が考えたときに別の人が言い当てることはとても難しいことを意味します。たとえば 6 桁の数は全部で 900,000 個あるわけですが、そのうち仮に素数が 14 個に一個であったとしても約 $\dfrac{900,000}{14} = 64286$ 個の素数があることになります。その中の一つを言い当てるのは難しくて、逆によい素数判定さえできれば、素数を見つけることは簡単になります。(実際 6 桁の素数は計算機で計算したところ、68906 個ありました。) 実際、素数を判定する効率的なアルゴリズムはあるのですが、二つの素数を掛けた合成数を素因数分解する効率的なアルゴリズムは現在のところありません。

最近はコンピュータの性能があがっているので、例えば私のＰＣでは 6 桁の素数かける 6 桁の素数を計算してそれを素因数分解させると、一瞬にして答えを返して来ます。しかし、50 桁の素数かける 50 桁の素数としてあらわせる数、つまり 100 桁くらいの数では簡単には答えはでませんでした。実は次の章で述べるように現在の暗号理論で暗号の安全性は素因数分解の困難さを基礎としています。現在主に使われている暗号に使われる二つの素数の積は数は 1024 ビット＝約 300 桁です、これには解読コンテストがあって、そこに掲示されている問題は、未だに因数分解できていないようです。この問題にかかっている賞金は 10 万ドル=1000 万円ですから挑戦してみますか？(インターネットで調べてみたら、1000 億円くらいのスーパーコンピューの京を使えば 1 年もあれば因数分解できるのではないかと評価している人もいるそうです。)

## 6. 公開鍵暗号

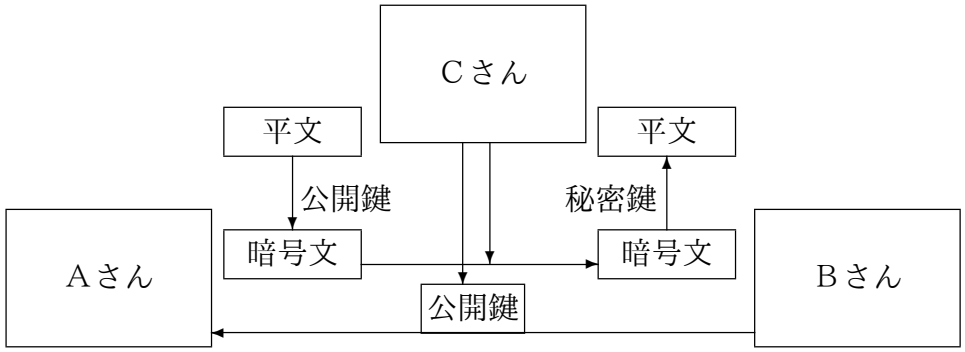
大きい数になると素数の確率は徐々に減ってゆくのですが、それでもかなりたくさんの素数がある、ということを見ました。そしてそのことから、大きな数であると素因子を見破ることは難しく、二つの大きな素数を掛けて得られる数に対して素因数分解する効率のよいアルゴリズムも現在ないのです。実はこの事実は公開鍵暗号の理論に応用されています。

まず公開鍵暗号とはどういうものかをお話しましょう。暗号化して情報を送るというのは、　インターネットなどでクレジットカードの番号のように人に知られては困る情報を送ることを想定しています。（ご存知の様に、クレジットカードの番号を知っていると、それを使って自由に買い物もできますし、お金も引き出せてしまいます。私は隠し事をしない主義ですという人でも、この番号は取引相手以外に知られてはいけないものです。）下の図を見てください。インターネットはさまざまな情報をネットワークを使ってやりとりする仕組みですが、途中で悪意をもっている第三者にも情報を傍受することは可能です。下の図ではＡさんがＢさんに情報を送ろうとしているところで、Ｃさんは途中にいて、傍受することができます。現代のインターネットの仕組みを図式化したものです。




Ｃさんにわかってしまってはまずいですから、Ａさんは情報を暗号化してＢさんに送って、Ｂさんはその情報を解読するという方法をとります。たとえば「あしたのあさこい」といった内容の文章を送りたいとします。このように送りたい文章のことを平文（ひらぶん）といいます。。そこで５０音に従って「３文字後にずらす」という操作をして送るとします。この規則に従って、文章を変換すると、「えそてふえせすお」と変換されることになります。もとの文章になおすには、「３文字前にずらす」操作をすればよいわけなので３文字という解読のための情報をしっていれば、Ｂさんは送られたをもとの文章にもどせるわけです。このように戻すことを復号といい、そのために必要な「後に３文字」という情報を暗号の鍵といいます。

暗号を解読するために必要となる鍵は、Ｂさんにだけわかって、Ｃさんにはわからないようにしたいわけです。解読に必要な鍵をＡさんが決めて暗号化しても、解読に必要なその鍵をＢさんに送らなくてはなりません。これをインターネットでおくると、Ｃさんは傍受していますから、Ｃさんにもばれて、その鍵を使って解読できてしまうことになります。そこで考えられたのが公開鍵暗号というしくみです。これは下のような図式であらわされます。




上の図式と違うところは、２点あります。まず鍵には暗号化用鍵＝公開鍵と復号化用鍵＝秘密鍵の二種類があることです。もう一つは、この二つの鍵は情報を受け取るＢさんが作ることです。暗号化は公開鍵を使ってすることにします。そしてそれをもとの文章に復号するには秘密鍵を知らなければ、著しく困難であるようにな仕組みをつくっておきます。公開鍵から原理的には秘密鍵が計算できても、そのために著しい時間がかかるのであれば、秘匿性（秘密を守る性質）が十分高く、実用的であるというわけです。公開鍵暗号において要請される性質を列挙するとつぎのようになります。

(1) 公開鍵、秘密鍵は用意に作り出すことができる。
(2) 公開鍵をもちいて、平文は容易に暗号化できる。
(3) 暗号文は秘密鍵を用いて容易に復号できる。
(4) 暗号文は公開鍵をしっていても、秘密鍵を知らなくては、復号化は著しく困難である。

## 7. RSA 暗号

そのような性質をもつ暗号として考えられたのが RSA 暗号です。これは素数に関する次の性質を使います。

(1) 大きな数においても素数はたくさんあり、かつそれが、素数かどうか判定するのは容易である。
(2) 二つの大きな素数の積が与えられたとき、これを因数分解するのは困難である。

実はある数が素数かどうか判定する（あるいはかなりの精度で確からしい確率で判定する）には効率のよい判定法があります。小さいほうから割り切れないことをチェックするのは大変なのですが、それをせずに判定する方法があるのです。そして、大きな数が素数かどうかは簡単に判定できても、素数でないときに、因数分解をするための効率のよいアルゴリズムが現在、知られておらず、従って因数分解は非常に難しいのです。

素数判定が早くできることは認めて、RSA 暗号の仕組みの概要を説明しましょう。まず二つの異なる素数 $p$ と $q$ をとってきます。そして $e{\cdot}d = c{\cdot}(p-1){\cdot}(q-1)+1$ となるような $e, d, c$ をとってきます。たとえば $p = 137$, $q = 251$ を二つの素数とすれば $13 \times 13077 = 5 \times (137-1) \times (251-1) + 1$ が成り立ちますから、$e = 13, d = 13077, c = 5$ とすれば十分です。このとき次の定理がなりたちます。

**定理 7.1** (一般化されたフェルマーの小定理)**.** $a$ を $p$ でも $q$ でも割り切れない自然数で $pq$ より小さいとする。このとき $a^{c(p-1)(q-1)+1}$ を $pq$ で割ったあまりは $a$ になる。

この定理を上の例で考えてみると、$0 < a < 137 \times 251$ なる $a$ で 137 でも 251 でも割り切れなければ、$a^{5\cdot 136\cdot 250+1}$ を $137 \times 251$ で割った余りは $a$ になります。このようにして求めた数 $p, q, e, d$ を用いて次のように公開鍵、秘密鍵、暗号化の方法、復号化の方法を定めます。以下、自然数 $a$ に対して $a$ を $n$ で割った余りを $a \pmod n$ と書きます。

(1) まず受け手であるBさんは上に挙げた方法で $p, q, e, d$ を定めます。($c$ は使いません。)

(2) Bさんは $p$ と $q$ の積の $n$ を計算して、$n$ と $e$ の組 $(n, e)$ を公開鍵としてAさんに送ります。この鍵はCさんに傍受されてもかまいません。この際Bさんは $p, q, d$ は秘密にしておかなければなりません。あとで復号のときに使うために $n$ と $d$ の組 $(n, d)$ を秘密鍵として記憶しておきます。これは他人には見せないようにします。$p, q$ は忘れてもかまいません。上の例の場合は $137 \times 251 = 34387$ ですから、$(n, e) = (34387, 13), (n, d) = (34387, 13077)$ となります。

(3) 次にAさんは送りたい情報 $a$ を暗号化します。$a$ は $pq$ より小さく、$p$ でも $q$ でも割り切れない数でなくてはならないのですが、通常は $p$ も $q$ も十分におおきな素数を選びますから、$p$ と $q$ の小さいほうよりも小さい数をおくることにします。はじめから $p$ や $q$ は十分に大きくとるという約束事はすべてのひとに共通の約束事として定めておかなくてはなりません。

（暗号化の方法）$a^e$ を $n$ で割ったあまり、つまり $b = a^e \pmod n$ を求め、それを暗号としてBさんに送ります。上の例を用いて、たとえば送りたい情報 $a$ を 57 としてみましょう。この場合Bさんに送るものは

$$b = a^d \pmod n = 57^{13} \pmod{34487} = 32286$$

となります。

(4) Bさんは受け取った暗号 $b$ に対して $b^d \pmod n$ を計算します。こうすると $a$ が復元できます。なぜなら,

$$b^d \pmod n = (a^e)^d \pmod n = a^{ed} \pmod n$$
$$= a^{c(p-1)(q-1)+1} \pmod n = a \pmod n$$

となるからです。ここで最後の等式にはフェルマーの定理を用いました。今の例で本当にもとに戻るかやってみると

$$b^d \pmod n = 32286^{13077} \pmod{34487} = 57$$

となり確かに 57 となり、もとの情報が回復されました。Bさんは $d = 13077$ と知っているので $b$ から $a$ が回復できるのですが、傍受者Cさんは $d$ を知らないため、このようなやり方では $b$ から $a$ は求められません。

Cさんは　$b$ と $(n, e)$ 知っているのですが、この状態でもとの情報 $a$ を求めようとすると、つぎの二つの方法があります。まず一つ目の方法ですが、$a'$ を 1 から順番に動かして、それぞれに対して $a'^e \pmod n$ を計算して $b$ が得られる

まで計算する。$p$ のほうが $q$ より小さければ、$a'$ は $p-1$ まで動かさなくてはなりませんから、大変な計算量で、とても $a$ を推測することはできません。ただここで、$a$ のとりうる範囲を大きくして、推測しにくくする必要があります。私の手元にあるクレジットカードは 16 桁ですから、これなら $n, e, b$ から $a$ は容易に推測はできないでしょう。

もうひとつの方法は、$n = pq$ と因数分解して $p, q$ を求め $ed = c(p-1)(q-1)+1$ となる $d$ を求める。そしてBと同じ方法で $a$ を計算します。もし因数分解が簡単にできれば、$d$ はユークリッドの互除法で容易に求められます。しかし大きな数 $n$ の素因数分解について、現在は有効なアルゴリズムはないので、これも大変な計算量となります。

こういう理由から暗号の安全性が保障されているのです。

## 8. まとめ

ここで使ったのは整数論の定理のひとつであるフェルマーの定理です。この定理は様々な自然科学において状態を図る数値を扱う数学とは少し赴きの変わった数学です。デジタル化された情報を暗号化するときにはこういった自然数の代数的な性質が有効に使われます。とくに素数に特有の性質を用いて暗号をもちいれば、他人に情報がもれることなく、目的の人にだけわかるような情報のやり取りが可能となります。暗号の安全性には素因数分解の困難さは用いられましたが、本当にそれは困難なのか、ということは大問題です。しかしいまのところは時間のかかるアルゴリズム以外のアルゴリズムが発見されていないというだけで、将来簡単なアルゴリズムができない保障はどこにもありません。しかしそのようなアルゴリズムがないのであれば、素数はほとんど無尽蔵にあるので計算機の性能があがっても、それに応じて堅牢な暗号化ができることになります。

## 9. 付録：フェルマーの定理

フェルマーの定理を証明しましょう。$p, q$ を異なる素数として $n = pq$ とおきます。集合 $A$ を

$$A = \{a \mid a \text{ は } p, q \text{ で割れない自然数で } 0 \leq a \leq n-1 \text{ をみたすもの } \}$$

とします。これを $n$ と互いに素となる剰余の集合と呼びます。また自然数 $a$ に対して $a$ を $n$ で割ったあまりを $a \pmod n$ と書きます。$a, b$ を自然数とするとき、$ab \pmod n$ は $a \pmod n$ と $b \pmod n$ をかけたものの $n$ で割ったあまりと等しいことがわかります。式でかけば

$$ab \pmod n = (a \pmod n b \pmod n) \pmod n$$

となります。次の定理が成立します。

**定理 9.1.** (1) $a, b \in A$ ならば $ab \pmod n$ も $A$ の元である。
(2) $a_1, a_2, b \in A$ として $a_1 b \pmod n = a_2 b \pmod n$ とすると、$a_1 = a_2$ となる。

証明. (1) $a, b$ ともに $p$ で割り切れなければ、$ab$ も $p$ で割り切れない。$n$ は $p$ の倍数なので、$ab$ を $n$ で割った余りも $p$ で割り切れない。$q$ についても同様。
(2) $a_1 b \pmod n = a_2 b \pmod n$ とすると、$a_1 b - a_2 b = (a_1 - a_2)b$ は $n$ で割り切れる。$b$ は $p$ を素因数に持たないので $a_1 - a_2$ が $p$ で割り切れなくてはなら

ない。同様にして $a_1 - a_2$ は $q$ でも割り切れなくてならず、その結果 $a_1 - a_2$ は $pq = n$ で割り切れなくてはならない。$a_1, a_2 \in A$ なので $a_1 = a_2$ となる。 □

集合 $A$ の元の個数を $m = \varphi(n)$ と書きます。これは $n$ のオイラー関数と呼ばれます。さらに

$$A = \{b_1, b_2, \ldots, b_m\}$$

とおきます。このとき $m = (p-1)(q-1)$ となります。

**定理 9.2.** $a \in A$ とすると、$a^m (\mathrm{mod}\ n) = 1$ となる。

証明. $c_1, \ldots, c_m$ を

$$c_1 = ab_1 (\mathrm{mod}\ n), \quad c_2 = ab_2 (\mathrm{mod}\ n), \ldots, \quad c_m = ab_m (\mathrm{mod}\ n)$$

とすると定理 9.1 の (1) より $c_1, \ldots, c_m \in A$ であり、さらに (2) より $i$ と $j$ が異なれば、$c_i$ と $c_j$ は異なる。従って、$A$ の元の個数を考えることにより

$$A = \{c_1, \ldots, c_m\}$$

となることがわかる。従って $A$ の元をすべて掛け合わせることにより、

$$c_1 \cdots c_m = b_1 \cdots b_m$$

となる。これらを使うと次の等式が成り立つ。

$$\begin{aligned}
&(a^m (\mathrm{mod}\ n) b_1 b_2 \cdots b_m (\mathrm{mod}\ n))(\mathrm{mod}\ n) = a^m b_1 b_2 \cdots b_m (\mathrm{mod}\ n) \\
=&(ab_1 (\mathrm{mod}\ n) \cdot ab_2 (\mathrm{mod}\ n) \cdots ab_m (\mathrm{mod}\ n))(\mathrm{mod}\ n) \\
=&c_1 c_2 \cdots c_m (\mathrm{mod}\ n) \\
=&b_1 b_2 \cdots b_m (\mathrm{mod}\ n)
\end{aligned}$$

さて定理 9.1 の (1) を繰り返して用いて $b_1 b_2 \cdots b_m (\mathrm{mod}\ n)$ は $A$ の元となることがわかるので (2) より $a^m (\mathrm{mod}\ n) = 1$ となる。 □

**定理 7.1 の証明** 上の定理 9.2 を用いて

$$\begin{aligned}
a^{cm+1} (\mathrm{mod}\ n) &= ((a^m (\mathrm{mod}\ n))^c \cdot a (\mathrm{mod}\ n))(\mathrm{mod}\ n) \\
&= a (\mathrm{mod}\ n) = a
\end{aligned}$$

となる。