代数学の世界
―整数論とその応用―

第2回　デジタルの数学、セキュリティーの数学
―符号・暗号理論―

東京大学大学院数理科学研究科
桂　利行

1. 暗号

暗号

秘密で通信するために当事者間のみ了解されるように決めた特殊な記号,

あるいは特殊な方法によって通信しようという文章に変形を加えること.

ここでは後者を扱う.

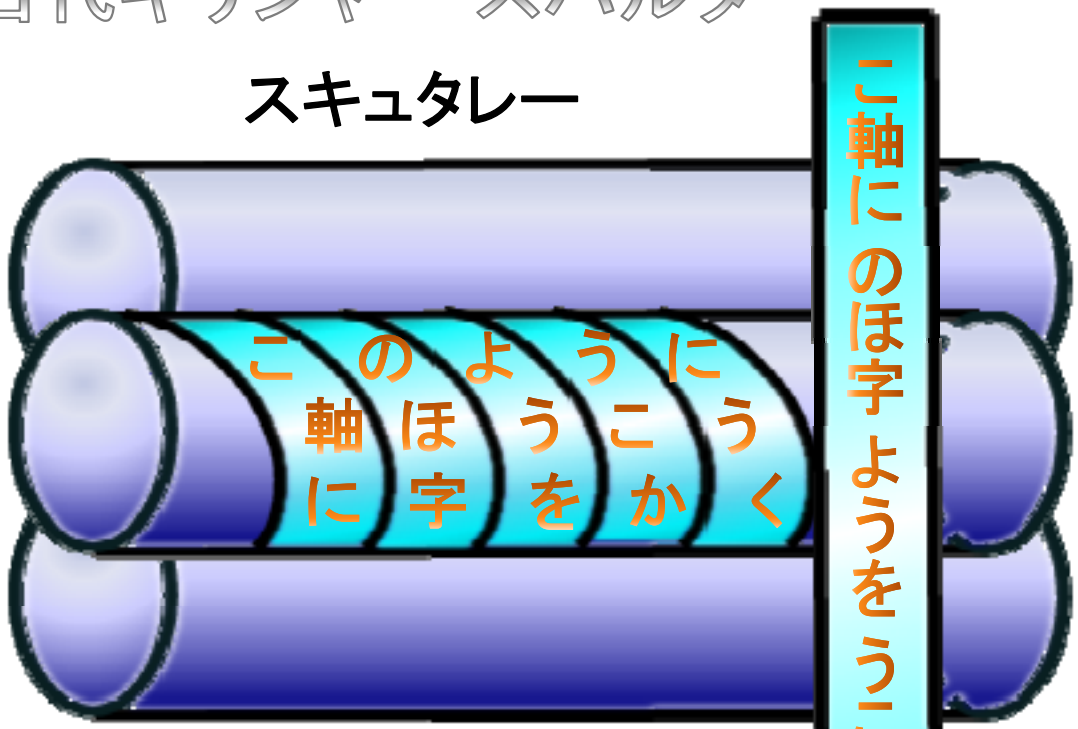
平文　　通信しようとする文章

暗号文　　変形した文章

鍵　　暗号文を平文に戻す道具

古代ギリシャ　スパルタ

スキュタレー



同じ太さの木の棒

細長い紙

細く長い紙を棒にずらしながらまきつけ
棒の軸方向に字を書く.

シーザー暗号

アルファベットを一定の数だけずらす.

- 1文字ずらす例

  平文　　HUKANKOUGI

  暗号文　　IVLBOLPVHJ

  鍵　　1文字後へずらす

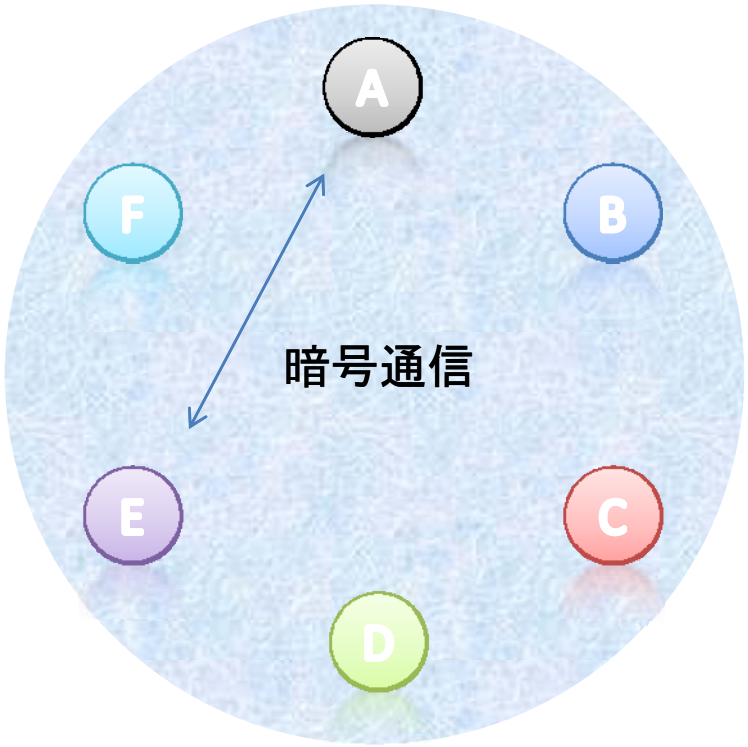
共通鍵暗号

暗号化の鍵と解読するための鍵を，発信者と受信者が共有して用いる方式の暗号.

1976年　W. Diffie，M.Hellman

暗号化の鍵を公開しても，多数の人の中で不特定の2人が暗号通信を行うことが可能であることを発見

公開鍵暗号



鍵を公開

公開鍵暗号の原理

たとえ強力なコンピュータを用いても計算を完了するためには途方もない時間がかかり，事実上解読できない.

構成法

(1)　素因数分解問題の難しさに基づくもの

(2)　離散対数問題の難しさに基づくもの

ここでは (1) を用いたRSA暗号を紹介する.

2. 整数論の定理

定理（Fermatの小定理）

p を素数, a を p で割り切れない整数とすれば,

$$a^{p-1} \equiv 1 (\mathrm{mod}\ p)$$

が成り立つ.

証明

$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ の $\overline{0}$ 以外の元の集合を $\{\overline{1}, \overline{2}, \cdots, \overline{p-1}\}$ とする.

a は p で割り切れないから,

$$\{\overline{1}, \overline{2}, \cdots, \overline{p-1}\} = \{\overline{a} \cdot \overline{1}, \overline{a} \cdot \overline{2}, \cdots, \overline{a} \cdot \overline{p-1}\}$$

故に,

$$\begin{aligned} \overline{1} \cdot \overline{2} \cdot \cdots \cdot \overline{p-1} &= \overline{a}\,\overline{1} \cdot \overline{a}\,\overline{2} \cdot \cdots \cdot \overline{a}\,\overline{p-1} \\ &= \overline{a}^{p-1} \cdot \overline{1} \cdot \overline{2} \cdot \cdots \cdot \overline{p-1} \end{aligned}$$

$\overline{1} \cdot \overline{2} \cdot \cdots \cdot \overline{p-1}$ は $\mathbb{F}_p$ で $\overline{0}$ ではないから

$$\overline{a}^{p-1} = \overline{1}$$

例1

$p = 71$ 素数
$a = 1687$ はpで割り切れない.

$$1687^{70} \equiv 1 \pmod{71}$$

例2

$q = 97$ 素数
$a = 1687$ はqで割り切れない.

$$1687^{96} \equiv 1 \pmod{97}$$

Fermatの小定理を少し一般化して暗号に用いる.

p, q　2つの素数

$\mathbb{Z}/pq\mathbb{Z}$　可換環

定理

a を pq と互いに素な整数とするならば

$$a^{(p-1)(q-1)} \equiv 1 (\mathrm{mod}\ pq)$$

が成り立つ.

例

$p = 71, q = 97, pq = 6887$ とする.

このとき

$$(p-1)(q-1) = 6720$$

$a = 1687$ は p, q で割り切れない.

$$1687^{6720} \equiv 1\ (\mathrm{mod}\ 6887)$$

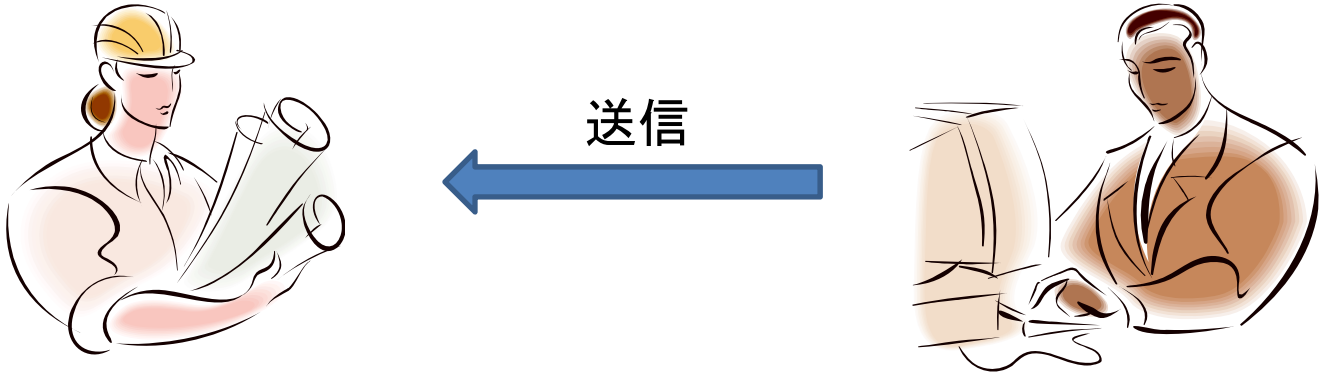
3. RSA 暗号

1978年

R.Rivest, A.Shamir, L.N.Adleman が発表

●2つの大きな素数の積を因数分解することが困難なことに基づく公開鍵暗号.

ユーザーA（受信者）　　ユーザーB（送信者）



ユーザーA（受信者）

大きな素数 p, q を選び

$$n = pq$$

とおく.

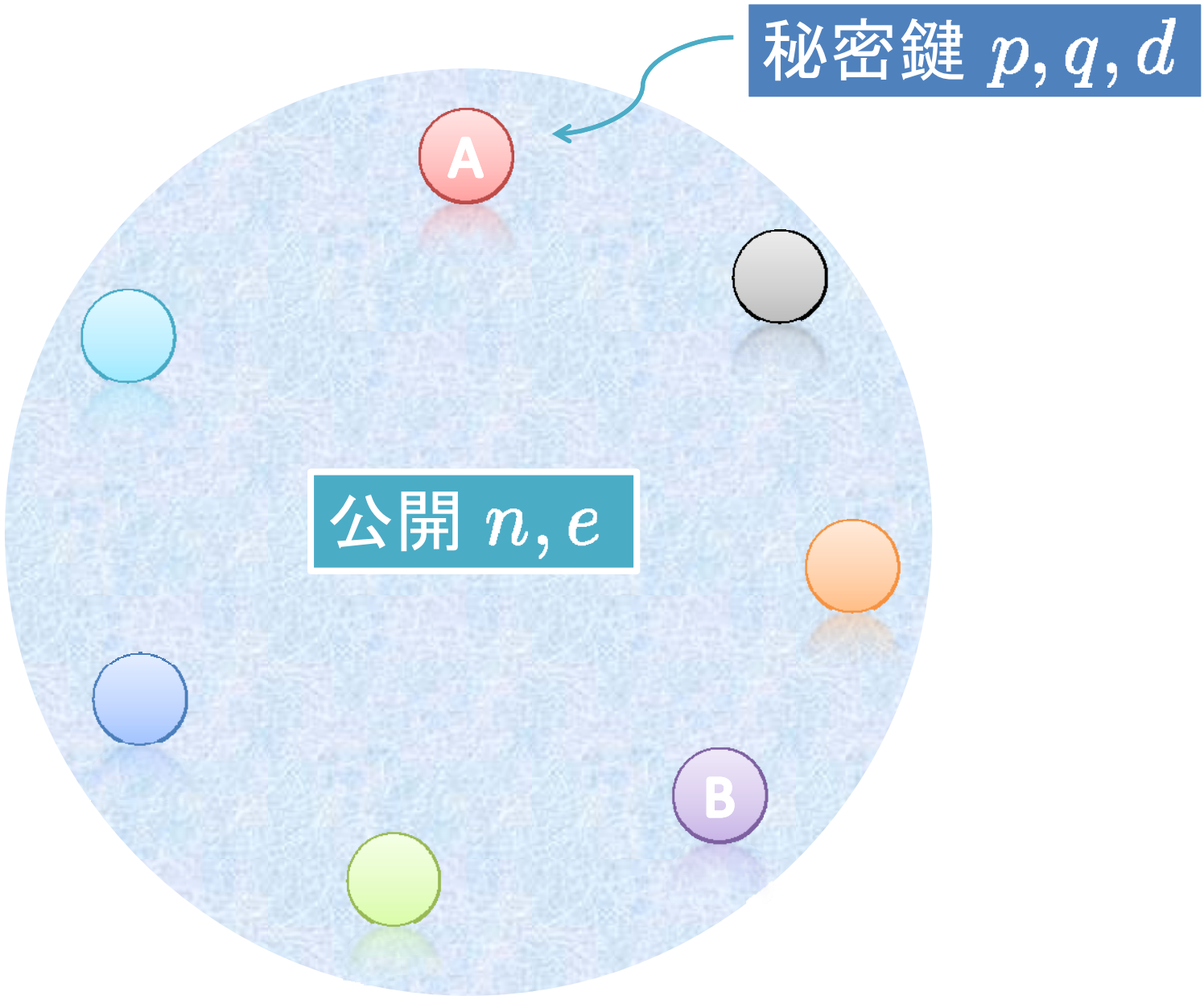
$\mathbb{Z}/(p-1)(q-1)\mathbb{Z}$ の元 $\bar{e}$ で

$$ed \equiv 1 \pmod{(p-1)(q-1)}$$

となる整数 d が存在するものをランダムに選ぶ.

公開　n, e

秘密鍵　p, q, d


秘密鍵 p, q, d
A
公開 n, e
B

ユーザーB（発信者）

n と互いに素な平文Mを発信する.

暗号化　　M^e

原理　nの素因数分解が困難なため第三者は d を計算できない.

ユーザーA（受信者）

$(M^e)^d \equiv M \pmod{n}$ よりMを復元

検証

Fermatの小定理の一般化から

$$M^{(p-1)(q-1)} \equiv 1 \pmod{n}$$

$ed \equiv 1 \pmod{(p-1)(q-1)}$ だから，ある整数 s があって

$$ed = (p-1)(q-1)s + 1$$

故に，

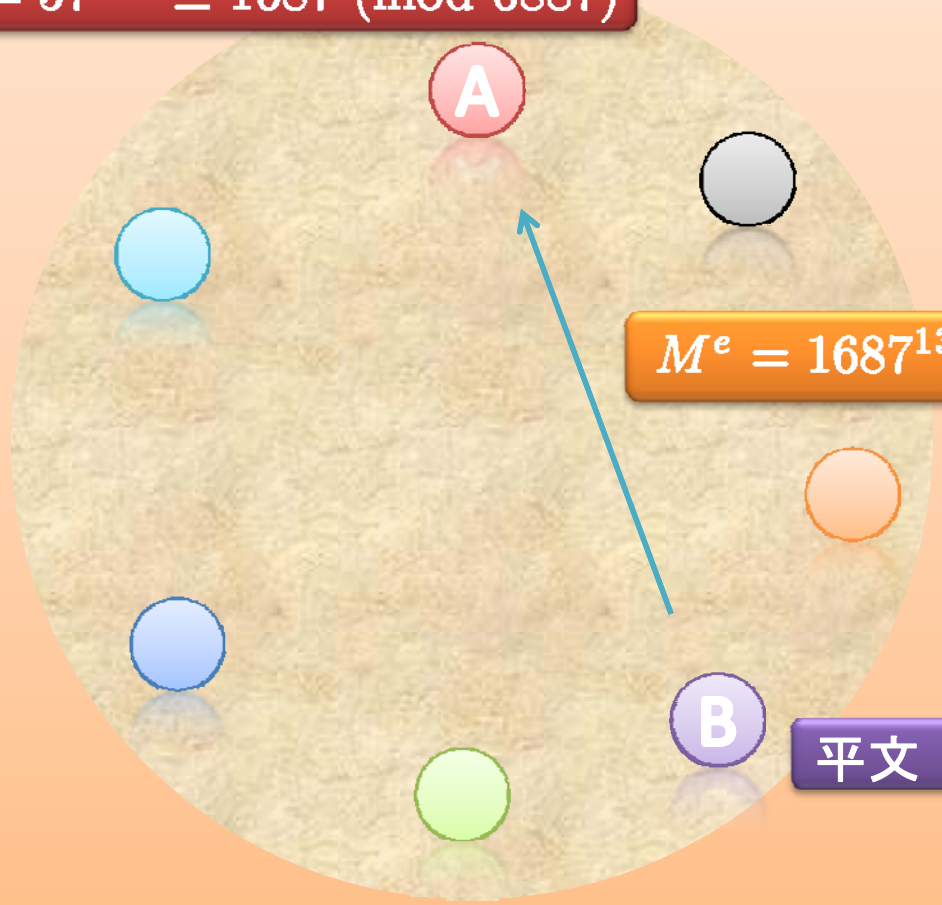
$$\begin{aligned} M^{ed} \pmod{n} &\equiv M^{(p-1)(q-1)s+1} \pmod{n} \\ &\equiv (M^{(p-1)(q-1)})^s M \pmod{n} \\ &\equiv M \pmod{n} \end{aligned}$$

例

秘密鍵: $p = 71, q = 97, d = 517$

公開: $n = pq = 6887, e = 13$

$57^d = 57^{517} \equiv 1687 \pmod{6887}$



$M^e = 1687^{13} \equiv 57 \pmod{6887}$

平文 $M = 1687$

4. 符号理論

アナログからデジタルへ

Compact Disc

著作権処理の都合で、

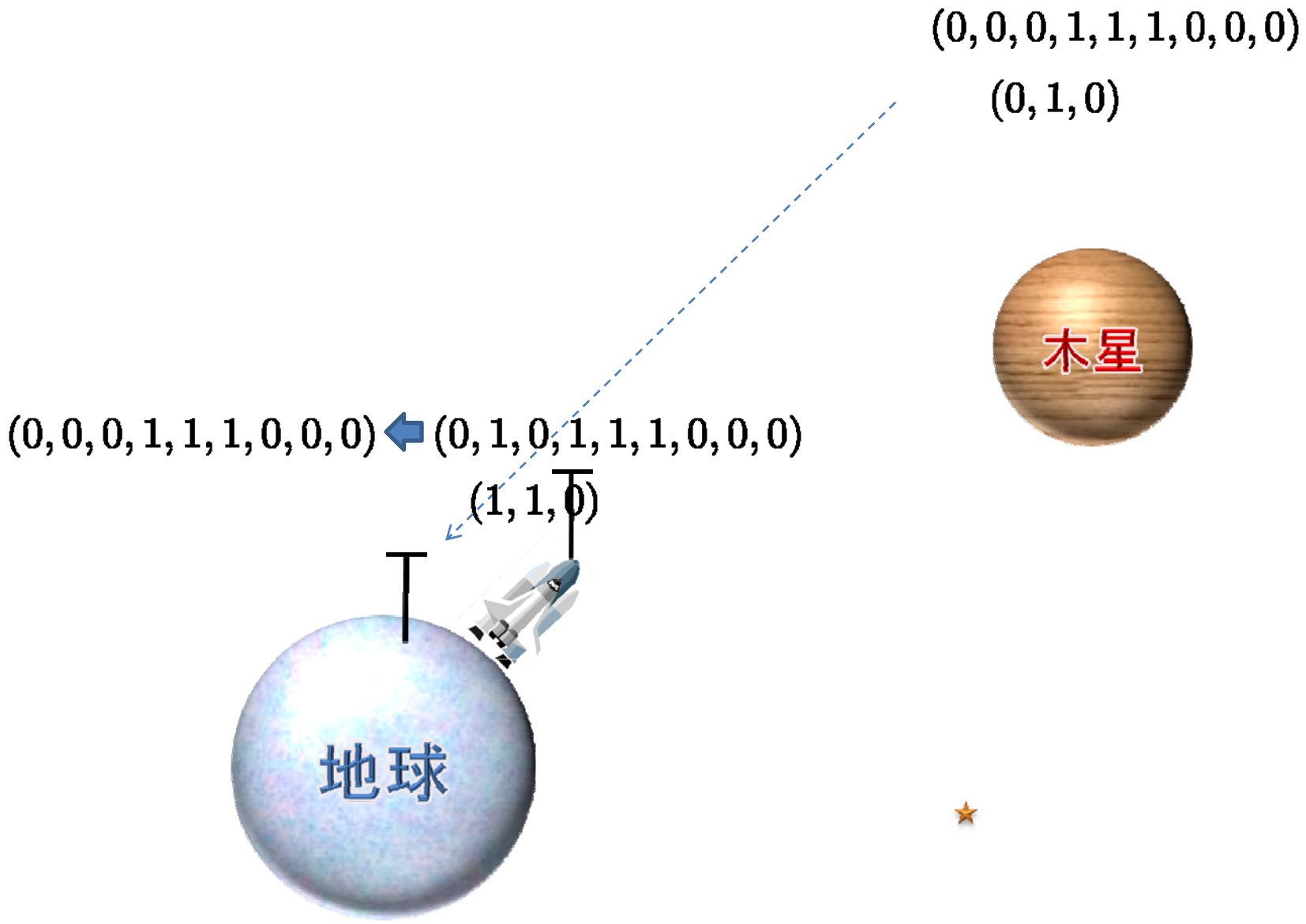
この場所に挿入されていた”電子機器の図版”を

省略させていただきます。

デジタルのあるところ誤りあり

誤り訂正符号

誤り検出符号


(0, 0, 0, 1, 1, 1, 0, 0, 0)
(0, 1, 0)
木星
(0, 0, 0, 1, 1, 1, 0, 0, 0)
(0, 1, 0, 1, 1, 1, 0, 0, 0)
(1, 1, 0)
地球


情報源
符号器
通信路
復号器
宛先

$\mathbb{F}_q^n$:$\mathbb{F}_q$ 上の n 次元横数ベクトル空間
$(x_1, x_2, \cdots, x_n)$ を語(alphabet)とよぶ.

$\mathbb{F}_q^n \supset C$
C を符号 (code)という.
Cの元を情報のalphabetとしてもちいる.
n : C の符号長 (word length)という.

$\mathbb{F}_q^n \setminus C$: 冗長部分.
この部分を誤り訂正に用いる.

Cが大きい方が多くの情報を伝えることができる

$\mathbb{F}_q^n \setminus C$が大きいほうが一般に誤り訂正能力が高い.

↳ 相反するこの両方の条件を満たす
できるだけ効率のよい符号をつくることを
めざすのである.

エラーを調べるために，$\mathbb{F}_q^n$ に距離を定義する．

距離とは

3次元ユークリッド空間 $\mathbb{R}^3$ の2点 $P=(x_1,x_2,x_3)$，$Q=(y_1,y_2,y_3)$ の距離

$$d(P,Q)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2}$$

次の3つの性質を満たす．

(i) $d(x,y)\geq 0$．また，$d(x,y)=0 \Leftrightarrow x=y$．

(ii) $d(x,y)=d(y,x)$

(iii) [三角不等式] $d(x,y)+d(y,z)\geq d(x,z)$

距離にとって大事なのはこの3つの性質であり，したがって逆に，この3つの性質をもつものは距離と呼ぶにふさわしいものである．

定義 $\mathbb{F}_q^n \ni x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n)$ の距離
(ハミング距離)

$$d(x, y) = \sharp\{1 \leq i \leq n \mid x_i \neq y_i\}$$

ここに，集合 S に対して $\sharp S$ は S の元の数を表す.

- これが先のような距離空間の公理を満たすことは容易にたしかめられる.

- たとえば $\mathbb{F}_2^5$ において，$x = (1, 0, 1, 0, 1)$ と $y = (1, 1, 0, 1, 0)$ では4つの成分が異なるから，$d(x, y) = 4$ となる.

言い換えると，$x = (1, 0, 1, 0, 1)$ を送信し $y = (1, 1, 0, 1, 0)$ を受信したとすればハミング距離 $d(x, y) = 4$ だけの誤りが生じたことになる.

定義 $z \in \mathbb{F}_q^n$ と自然数 r に対して

$$B_r(z) = \{x \in \mathbb{F}_q^n \mid d(z,x) \leq r\}$$

を z を中心とする半径 r の球と呼ぶ.

定義 最小距離 d

$\mathbb{F}_q^n$ の部分集合 C に対し, その最小距離 d を次のように定義:

$$d = \min\{d(x,y) \mid x,y \in C,\ x \neq y\}$$

ここに, min は最小値を表す.

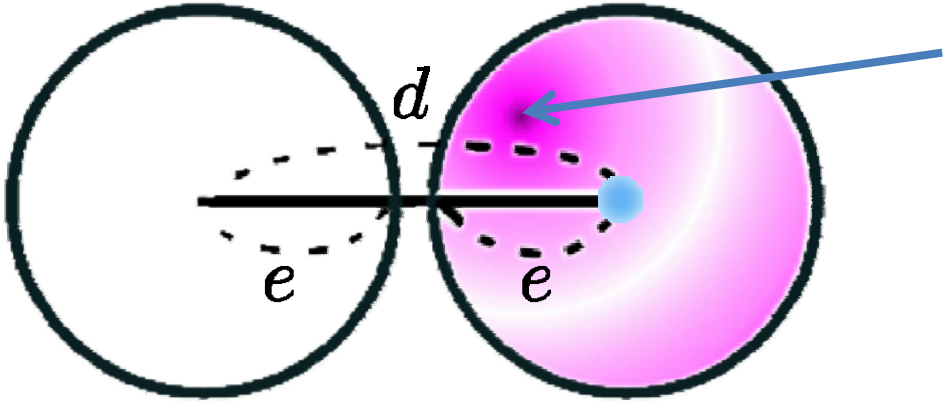
例 $\mathbb{F}_2^2 \supset C = \{(0,0),(1,1)\}$ のとき

最小距離 $d = 2$

エラー訂正の原理

状況を分かりやすくするため，

C の元を中心とする半径 e の球を考えると，それらは互いに共通部分がない．



受信した符号がそれらの球のいずれかに入ったとき，距離的に一番近い**その球の中心である C の元**が送信された符号として復号する．

したがって

最尤復号

このときには，確率的にいって，e 個までのエラーを訂正できることになる．

★ d が偶数のときには $(d-2)/2$ 個のエラーが訂正できる．

問題　$\mathbb{F}_q^n \supset C$

ハミング距離に関して，C の任意の2点ができるだけ離れているような部分集合 C を構成せよ．

ハミング符号，BCH符号，　RS符号，ゴーレイ符号，・・・・