情報理論とその応用シリーズ

3

符号理論とその応用

情報理論とその応用学会

［編］

培風館

第３分冊　編集担当者

辻井茂男（委員長）
今井秀樹（副委員長）
平田康夫（主査）

執 筆 者

1章
- 1.1　平澤茂一（早稲田大学）・有本卓（立命館大学）
- 1.2　平田康夫・大橋正良（KDDI（株））
- 1.3　今井秀樹（東京大学）

2章
- 2.1　笠原正雄（大阪学院大学）・森井昌克（徳島大学）
- 2.2　笠原正雄・森井昌克
- 2.3　坂庭好一（東京工業大学）

3章
- 3.1　大橋正良・笹野博（近畿大学）
- 3.2　田島正登（富山大学）・山口和彦（電気通信大学）

4章
- 4.1　西島利尚（法政大学）
- 4.2　山口和彦

5章　笠原正雄

6章
- 6.1　山崎彰一郎（（株）東芝）
- 6.2　安田豊（KDDI（株））・大橋正良
- 6.3　宮垣嘉也（岡山理科大学）
- 6.4　斉藤洋一（神奈川大学）
- 6.5　田中良紀（富士通（株））
- 6.6　植松友彦（東京工業大学）

7章　藤原英二（東京工業大学）

8章　井上徹（広島修道大学）

刊行に際して

情報化社会という言葉が聞かれて久しいが，最近のLSI技術を基礎とする広義のディジタル技術は，従来技術に基づく制度的・経営的枠組みを越えて業界再編成を促すまでに情報の処理加工・生成性と疎通性を高め，いよいよ本格的な情報ネットワーク社会が始まろうとしている。

広くて新しい意味の情報理論は，こうした情報ネットワーク社会の基盤技術の主要な理論的バックボーンとなるものである。

本レクチャーノートシリーズは，情報理論とその応用学会 (SITA) が10周年を迎えたのを機に，1990年頃企画され，編集委員会において議論を重ねた結果，大学院学生や第一線の研究者・技術者の研究開発意欲を鼓舞すべく，SITAの対象分野の最前線を生き生きとわかりやすく解説することを目的として，次の5分冊の形で出版することとなった。

分冊1　「情報符号化とデータ圧縮技術」
分冊2　「確率過程――応用と話題」
分冊3　「符号理論とその応用」
分冊4　「暗号と認証」
分冊5　「情報伝送の理論と方式」

分冊1は，半世紀に及ぶ伝統的な理論と方式を踏まえつつ，常に新しい展開をみせ，最近では，知的符号化のような知識処理まで包含するに至った情報源符号とデータ圧縮について述べたものである。いかにメモリーの低価格化が進んでも，大量のマルチメディア情報を蓄積して，効率よく検索するためには，高度の情報圧縮が不可欠である。また，通信分野においても，携帯情報端末の発展にともなう限られた電波資源の有効利用の重要性を考えれば，この分野への期待がきわめて大きいことが理解されよう。

分冊2も，歴史と伝統を誇るとともに，情報通信技術の基礎理論として発展を続ける確率過程論について，10章にわたって解説したものである。その内

容は，点過程の概論と最近応用の進んでいる点時系列解析，確率場の表現と予測理論，および確率場としての画像の統計的処理に関する話題や確率微分方程式，時系列データ間の因果性，Wiener 過程の非線形汎関数理論等，他の既刊書にみられない新しく，応用上重要性の高い内容が選ばれて解説されている。

分冊 3 は，誤り訂正符号とその通信系・蓄積系への多岐にわたる応用について説明したものである。この分野もシャノンの通信路符号化定理とハミング符号に端を発する数十年の歴史を有するが，代数幾何学的符号にみられるように現代数学の諸成果をも貪欲に吸収しつつ，とどまるところなく発展を遂げつつある熱気を帯びた分野である。また，情報源符号化・データ圧縮と同様に，LSI 技術の限りない発展は，高度な誤り訂正符号理論の実用化を可能にしており，その応用分野も大きく広がっている。例えば，移動通信の分野もディジタル化の時代を迎えたが，そこでは，音声信号を大きく圧縮するとともに，無線伝送路の劣悪さを誤り訂正技術によって克服している。

分冊 4 は，情報ネットワーク社会の技術的保障要件ともいえる情報セキュリティ技術の中核をなす現代暗号理論について述べたものである。暗号は人類の歴史とともに古いが，現代暗号は 1976 年の公開鍵暗号の概念提案に始まるといえよう。これは，情報の秘匿と並んで，現代暗号の重要な役割となる認証を強く意識したものであったが，1980 年代中端に至って，零知識相互証明という魅力ある認証理論が誕生した。本分冊はこのような暗号の認証機能を中心にまとめている。

分冊 5 は，信号理論変復調理論等，通信の基礎理論に始まって，21 世紀の実用化が期待される量子状態制御光通信理論まで，通信の清新な理論的側面を解説するとともに，最近とみに重要性を増しているスペクトル拡散通信方式や，通信主体としての人間にまで踏み込んだヒューマンコミュニケーションシステムについて展望している。

本シリーズは，このような分野を対象として，先人達の偉業を伝承しつつ，常に新しい分野を創造する SITA の人々の総力を結集して数十名からなる執筆人によってまとめられたものであり，研究・勉学意欲のある人々に魅力ある内容を盛り込めたのではないかと考えている。

中央大学　理工学部　情報工学科
編集委員長　辻井重男

目　次

1
概　論

1.1　通信路符号化の歴史と展望

1.1.1　はじめに

通信路符号化は情報源符号化と並んで情報理論の双璧をなしている．本稿ではこの通信路符号化の問題について，今日までに得られている成果をたどりながら展望を述べる．

C.E.Shannon による通信路符号化定理[38]は，その後 A.Feinstein [16]，R.M.Fano [15]，R.G.Gallager [19]，A.J.Viterbi [45]，有本 卓[3]，I.Csiszár と J.Körner [10]，韓 太舜[22]らにより精密化・定量化・一般化されたが現在でもなお，最も重要な結論の一つであることに変わりはない．ここでは，通信路符号化定理について Shannon の証明から出発し，次の三つの流れについてふりかえる．

(1) 信頼度関数の精密化
(2) 構成的（代数的）符号化による定理の証明
(3) 定理の一般化

通信路符号化は雑音のある通信路において，符号化により高信頼化を図る問題を取り扱う．**通信路容量** C を超えない正の**情報伝送速度** R で**符号長** N を大とすることにより，**復号誤り確率** $\Pr(\mathscr{E})$ を指数的に 0 に収束できる符号化が存在するという工学的に極めて重要な結論が導かれている．(1) は $\Pr(\mathscr{E})$ の 0 への収束速度を与える信頼度関数に関する研究であり，この分野の中心的話題

の一つである．現在でも低情報伝送速度では信頼度関数の上界と下界との間に差があり，真にきつい限界式は求まっていない．これは情報理論の最も基本的問題であるが，未だ解決されていない．**漸近的距離比** $\delta(R)$ が **Gilbert 下界式**[21]を等式で満たすという Shannon ら[37]の推測によれば，削除誤り指数 $E_{\rm ex}(R)$ が真の信頼度関数を与えるが，この推測に反例はないが証明もされていない．

一方，代数学を背景にもつ符号理論は数多くの実用的な誤り訂正符号を与えてきた．(2) は通信路符号化のもう一つの重要な問題として，ランダム符号化により証明された通信路符号化定理を非ランダムな構成的符号化により証明することである．**連接符号**を用いた試み[13, 40]があるが未だ完全ではない．通信路符号化定理で存在することが保証された最良符号を求めることも困難とされている[9]．

1970 年後半から，Csiszár と Körner [10]，R.E.Blahut [9]，韓[22]らにより情報理論の再構築が図られた．(3) は通信路符号化に対し新しい視点や新しい解釈と一般化に関する話題である．「**タイプの概念**」を用いて，通信路の特性には依存しない符号化・復号化による**ユニバーサル通信路符号化定理**が導出された[10]．さらに「**情報スペクトルの概念**」を用いて，自然な一般化がされている[22]．後者では，通信路に対する制約はなく，非定常でも非エルゴード的でもよい．また，通信路の入出力アルファベットは可算無限個が許される[22]．

以下では，特にことわらない限り離散的無記憶通信路を仮定し，ブロック符号についてのみとり上げる．たたみ込み符号や連接符号の詳しい議論については，筆者らの解説論文[1, 26]を参照されたい．なお，対数と指数の底は 2 とし，情報量の単位は [bits] とする．

1.1.2 Shannon の通信路符号化定理

離散的無記憶通信路の入力アルファベットを $\mathscr{X}$，出力アルファベットを $\mathscr{Y}$ とし，通信路行列を $P=[\Pr(y|x)]$ $(x\in\mathscr{X}, y\in\mathscr{Y})$ とする．ここで，行列 P は既知とする．Shannon はこの通信路の**通信路容量** C は $\mathscr{X}$ と $\mathscr{Y}$ の間の相互情報量 $I(\mathscr{X};\mathscr{Y})$ の入力アルファベット $\mathscr{X}$ 上の確率分布 $\Pr(x)$ に関する最大値として，次式のように与えられることを示した．

$$C=\max_{\Pr(x)} I(\mathscr{X};\mathscr{Y}) \tag{1.1}$$

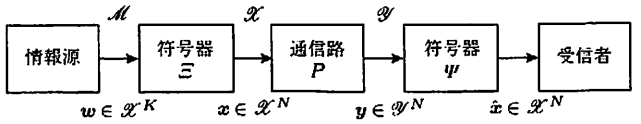

図 1.1 通信路符号化システム

ここで

$$I(\mathscr{X};\mathscr{Y})=\sum_{y\in\mathscr{Y}}\sum_{x\in\mathscr{X}}\Pr(x)\Pr(y|x)\log_2\frac{\Pr(y|x)}{\sum_{x'\in\mathscr{X}}\Pr(x')\Pr(y|x')} \tag{1.2}$$

である．Shannon [38]によって示された通り，通信路容量 C は復号誤り確率 $\Pr(\mathscr{E})$ を小さくできるよい符号が存在するか否かの境い目を与えている．ここで，式 (1.2) からその最大値を求める計算は単純な通信路でも必ずしも容易ではない．これを求めるアルゴリズムは二重最大化を交互に繰返す **Arimoto の算法**[2]として知られ，また Blahut [6]により独立に求められている．

ブロック符号による符号化 Ξ は通報 m の集合 $\mathscr{M}=\{1,2,\cdots,M\}$ を長さ K の**情報系列** $\boldsymbol{w}\in\mathscr{X}^K$ に 1 対 1 に対応させ，さらにこれを長さ N の**符号系列**（**符号語**）$\boldsymbol{x}$ に写像することにより行われる（図 1.1 参照）．すなわち，$\Xi: m\to(\boldsymbol{w}_m\to)\boldsymbol{x}_m$ である．ここで，$\|\mathscr{M}\|=M$ とするとき，$M=\|\mathscr{X}\|^K$ である[1]．また，情報伝送速度 R は

$$R=\frac{1}{N}\log_2 M \qquad [\text{bits/symbol}] \tag{1.3}$$

で与えられる．

通信路入力系列（符号系列）を $\boldsymbol{x}\in\mathscr{X}^N$，通信路出力系列（**受信系列**）を $\boldsymbol{y}\in\mathscr{Y}^N$ とするとき，復号化 Ψ は $\boldsymbol{y}$ から $\boldsymbol{x}$ の推定符号語 $\hat{\boldsymbol{x}}\in\mathscr{X}^N$ を求める操作である．すなわち，$\Psi:\boldsymbol{y}\to\hat{\boldsymbol{x}}$ である．図 1.1 に通信路符号化システムを示す．

定理 1.1 (通信路符号化定理（Shannon の第 2 符号化定理）) [38]
雑音のある離散的通信路の通信路容量を C[bits/symbol]，情報伝達速度を R[bits/symbol] とする．もし，$R<C$ ならば符号長 N を十分大とすることにより，誤り確率 $\Pr(\mathscr{E})$ をいくらでも 0 に近づけることのできる符号化が存在する[2]． □

[1] $\|\mathscr{A}\|$ は集合 $\mathscr{A}$ の要素の数を示す．
[2] $R>C$ のときの結果についても示されているが，ここでは省略する．

Shannon は $\|\mathscr{X}\| = \|\mathscr{Y}\| = 2$（すなわち，2元通信路）を仮定し，符号語 $\boldsymbol{x}$ をランダムに割当てるランダム符号化により，**大数の法則**を用いてこの証明を与えた．すなわち，通報 $m \in \mathscr{M}$ はランダムに $M_{\mathscr{X}} \doteq 2^{NH(\mathscr{X})}$ 個 $(M < M_{\mathscr{X}})$ の中から選ばれた符号系列 $\boldsymbol{x}_m \in \mathscr{X}^N$ に割当てる[3]．一方，受信系列 $\boldsymbol{y} \in \mathscr{Y}^N$ を観測する原因となった通信路入力系列 $\boldsymbol{x} \in \mathscr{X}^N$ の数 $M_{\mathscr{X}|\mathscr{Y}} \doteq 2N^{H(\mathscr{X}|\mathscr{Y})}$ である．この中に符号系列 $\boldsymbol{x}_m$ はもちろん含まれる．しかし，符号系列でない系列，あるいは他の符号系列 $\boldsymbol{x}_{m'}(m' \neq m)$ も含まれる可能性がある．もし，他の符号系列 $\boldsymbol{x}_{m'}$ が含まれていなければ $\boldsymbol{y} \to \boldsymbol{x}_m$ と正しく復号できる．ところで，$M_{\mathscr{X}}$ 個の任意の1つの系列に通報が割当てられている確率は

$$\frac{M}{2^{NH(\mathscr{X})}} = 2^{-N[H(\mathscr{X})-R]} \qquad (M = 2^{NR}) \tag{1.4}$$

であるから，逆に割当てられていない確率は $1 - 2^{-N[H(\mathscr{X})-R]}$ である．したがって $M_{\mathscr{X}|\mathscr{Y}}$ 個の中の通信路入力系列が真の符号系列 $\boldsymbol{x}_m$ を除いて他の通報に1つも割当てられない確率 $\Pr(\mathscr{C})$ は

$$\begin{aligned}\Pr(\mathscr{C}) &\doteq \{1 - 2^{-N[H(\mathscr{X})-R]}\}^{2^{NH(\mathscr{X}|\mathscr{Y})}} \\ &\doteq 1 - 2^{-N[H(\mathscr{X})-R-H(\mathscr{X}|\mathscr{Y})]}\end{aligned} \tag{1.5}$$

である．これは，$\boldsymbol{y}$ を受信する原因となった $M_{\mathscr{X}|\mathscr{Y}}$ 個の $\boldsymbol{x}_m$ の復号領域が互いに交わらない確率，すなわち正しく復号できる確率であり，したがって復号誤り確率 $\Pr(\mathscr{E}) = 1 - \Pr(\mathscr{C})$ は簡単な計算の結果

$$\Pr(\mathscr{E}) \doteq 2^{-N\delta} \quad (\delta \doteq H(\mathscr{X}) - H(\mathscr{X}|\mathscr{Y}) - R \leq C - R) \tag{1.6}$$

となる．よって，$R < C$ ならば $\delta > 0$，したがって $\Pr(\mathscr{E}) \to 0$ $(N \to \infty)$ とできる．以上において，$H(\mathscr{X}), H(\mathscr{X}|\mathscr{Y})$ は完全事象系 $\mathscr{X}, \mathscr{Y}$ のそれぞれエントロピーおよび条件付情報量である．

Shannon のこの証明は情報源 $\mathscr{X}$ から生成される長さ N の系列の数 $M_{\mathscr{X}}$ は平均 $2^{NH(\mathscr{X})}$ 個であり，$N \to \infty$ と共に，これら**典型的系列**（標準系列）以外の系列の生起確率は0に近づくという大数の法則に基づいた平均的な値しか示していない．しかし，$I(\mathscr{X};\mathscr{Y}) = H(\mathscr{X}) - H(\mathscr{X}|\mathscr{Y})$ であるから，式 (1.1) の通信路容量 C が $\Pr(\mathscr{E}) \to 0$ $(N \to \infty)$ とできる限界として重要な意味をもつことがわかる．

1.1.3 信頼度関数の精密化

通信路符号化定理は，その後 Feinstein [16]により精密な証明が与えられ，さらに Fano [15]は復号誤り確率 $\Pr(\mathscr{E})$ が符号長 N と共に指数的に0に収束す

[3] $\doteq$ は近似式であることを示す．

ることを示した．すなわち

$$\Pr(\mathscr{E}) \leq \exp_2[-N \cdot \alpha] \quad (\alpha > 0, \quad 0 \leq R < C) \tag{1.7}$$

であり，α は R の減少関数で後に信頼度関数 $E(R)$（reliability function）[4]と呼ばれている．**信頼度関数** $E(R)$ は次式で定義される[20]．

$$E(R) = \lim_{N\to\infty} \sup \frac{-\log_2 \Pr(\mathscr{E})}{N} \tag{1.8}$$

supremum は符号長 N，情報伝送速度 R のすべてのブロック符号に対してとる．関数 $E(R)$ を定めるために

$$E_L(R) \leq E(R) \leq E_U(R) \quad (0 \leq R < C) \tag{1.9}$$

となる $E(R)$ 関数の**上界** $E_U(R)$ と**下界** $E_L(R)$ を

$$\Pr(\mathscr{E}) \leq \exp_2[-NE_L(R)] \quad (0 \leq R < C) \tag{1.10}$$

$$\Pr(\mathscr{E}) \geq \exp_2[-NE_U(R)] \quad (0 \leq R < C) \tag{1.11}$$

を評価して求めていく．$E_L(R) = E_U(R)$ となる $E(R)$ を定めることが目的なので，式 (1.10)，(1.11) は数学的にきつい（tight）評価式である程望ましい．様々な符号化，復号法，通信路が与えられた条件の下に $0 \leq R < C$ 全領域にわたって厳密な $E(R)$ 関数を求める研究がなされているが，一般には $R = 0$，$R_{\text{crit}} \leq R < C$ の領域を除いて未だ定まっていない[5]．ここで，R_{crit} は**臨界情報伝送速度**（critical rate）と呼ばれ後に定義する．

また，ここまでは $0 \leq R < C$ としたが，$R > C$ においても通信路符号化の逆定理として，どのような符号を用いても復号誤り確率 $\Pr(\mathscr{E})$ は一定値以下にはできないことが示される[16, 3]．

(1) 信頼度関数の上界と下界

通信路容量 C 以下の情報伝送速度 R における符号化定理について明らかにする．証明の詳細は成書にゆずり，ここでは工学的視点から重要な Gallager [19]による復号誤り確率の上界を与えるランダム符号化指数（randam coding exponent）$E_r(R)$ の導出の流れを中心に示す．

[4] 式 (4.7) の α は正確にはその下界である．また，信頼度関数は簡単に誤り指数（error exponent）とも呼ばれる．

[5] 高雑音通信路では近似的に誤り指数の上界と下界が一致する．

(a) ランダム符号化

Gallager は Shannon よりもう一歩進めて，符号長 N，情報伝送速度 $R(M=2^{NR})$ のすべての符号（符号化法）をランダムに選ぶ方法を用いた．$\mathscr{X}^N$ の集合から重複を許して M 個の符号語を選択する写像 Ξ は $(\|\mathscr{X}\|^N)^M$ 個存在する．R を一定とするとき，N が大，したがって $M=\lceil 2^{NR}\rceil$ が[6]指数的に増大すると符号 $\mathcal{C}_i$ の復号誤り確率 $\Pr(\mathscr{E}|\mathcal{C}_i)$ を実際に計算することは困難となる．さらに，符号化の方法は $i=1,2,\cdots,\|\mathscr{X}\|^{NM}$ という膨大なものとなるから，その中から最小の復号誤り確率のものを選びこれを評価することはとうてい不可能となる．ここで，符号語 $\boldsymbol{x}_m\in\mathcal{C}_i$ $(m=1,2,\cdots,M)$ は等確率で生起するとすると，$\Pr(\mathscr{E}|\mathcal{C}_i)$ は

$$\Pr(\mathscr{E}|\mathcal{C}_i)=\sum_{m=1}^{M}\frac{1}{M}\Pr(\mathscr{E}|\boldsymbol{x}_m) \tag{1.12}$$

で与えられる．$\Pr(\mathscr{E}|\boldsymbol{x}_m)$ は $\boldsymbol{x}_m$ を送信したときの復号誤り確率で，次式で与えられる．

$$\Pr(\mathscr{E}|\boldsymbol{x}_m)=\sum_{\boldsymbol{y}\in\mathscr{Y}^N}\Pr(\boldsymbol{y}|\boldsymbol{x}_m)I_m(\boldsymbol{y}) \tag{1.13}$$

$$I_m(\boldsymbol{y})=\begin{cases}0, & \boldsymbol{x}_{\hat{m}}=\boldsymbol{x}_m;\\ 1, & \boldsymbol{x}_{\hat{m}}\neq\boldsymbol{x}_m\end{cases} \tag{1.14}$$

ここで，$\boldsymbol{y}$ は受信系列，$\boldsymbol{x}_{\hat{m}}\in\mathcal{C}_i$ は $\boldsymbol{y}$ より推定，すなわち復号して得られた符号語である．また，$I_m(\boldsymbol{y})$ は正しく復号されたとき 0，誤って復号されたとき 1 となるインディケータである．

個々の符号 $\mathcal{C}_i$ に対し $\Pr(\mathscr{E}|\boldsymbol{x}_m)$ を求めることをあきらめ，$\mathscr{X}^N$ 上に任意の確率測度 $\Pr(\boldsymbol{x})$ $(\boldsymbol{x}\in\mathscr{X}^N)$ を導入し，$\Pr(\mathcal{C}_i)=\prod_{m=1}^{M}\Pr(\boldsymbol{x}_m)$ と符号化 $\mathcal{C}_i$ をランダムに生成する．このとき，平均復号誤り確率 $\overline{\Pr(\mathscr{E})}=\sum_{\mathcal{C}_i}\Pr(\mathcal{C}_i)\Pr(\mathscr{E}|\mathcal{C}_i)$ であるから，もし平均値が計算できればこれより大きいものと小さいものがあり，したがって $\Pr(\mathscr{E}|\mathcal{C})\le\overline{\Pr(\mathscr{E})}$ となる符号 $\mathcal{C}$ が存在することになる．実際，平均値を求めることは比較的容易な場合があり，特に $N\to\infty$ では，Chernoff 限界式[11]を用いた確率密度関数のすその評価により復号誤り確率が与えられる．

以上の結果，最小の復号誤り確率の上界を平均値 $\overline{\Pr(\mathscr{E})}$ で与える．このように符号化定理は符号の存在のみが示された存在定理であり，どの符号が最良であるか定めていない．また，これを定めるのは困難であるとされている[9]．

(b) 最尤復号法

復号化の写像 Ψ は $\boldsymbol{y}$ より $\boldsymbol{x}_{\hat{m}}\in\mathcal{C}$ となる $\boldsymbol{x}_{\hat{m}}$ を推定する．符号語 $\boldsymbol{x}$ の事前確率 $\Pr(\boldsymbol{x})$ が等しく $1/M$ のとき，次式で示される**最尤復号法**が復号誤り確率を最小にすることが知られている[15]．

$$\boldsymbol{y}\to\boldsymbol{x}_{\hat{m}}:\hat{m}=\arg\max_{\boldsymbol{x}_i\in\mathcal{C}}\Pr(\boldsymbol{y}|\boldsymbol{x}_i), \tag{1.15}$$

[6] $[x]$ は x より大きいか等しい最小の整数を示す．

$\boldsymbol{x}_m$ を送信したときの復号誤り事象 $\mathscr{E}=\bigcup_{\hat{m}}\{\boldsymbol{x}_{\hat{m}}\neq\boldsymbol{x}_m\}$ であるから，式 (1.15) より式 (1.14) は直ちに，次式で与えられる．

$$I_m(\boldsymbol{y})=\begin{cases}0, & {}^{\forall}i\neq m,\quad \Pr(\boldsymbol{y}|\boldsymbol{x}_m)>\Pr(\boldsymbol{y}|\boldsymbol{x}_i);\\ 1, & {}^{\exists}i\neq m,\quad \Pr(\boldsymbol{y}|\boldsymbol{x}_m)\le\Pr(\boldsymbol{y}|\boldsymbol{x}_i).\end{cases} \tag{1.16}$$

(c) ランダム符号化指数の導出

式 (1.16) のインディケータ $I_m(\boldsymbol{y})$ は個々の符号の距離構造と通信路が与えられれば計算できるが，一般には困難である．そこで Gallager は式 (1.16) において，$I_m(\boldsymbol{y})$ が $0<\rho\le 1$ に対し

$$I_m(\boldsymbol{y})\le\left[\frac{\sum_{i\neq m}\Pr(\boldsymbol{y}|\boldsymbol{x}_i)^{\frac{1}{1+\rho}}}{\Pr(\boldsymbol{y}|\boldsymbol{x}_m)^{\frac{1}{1+\rho}}}\right]^{\rho} \tag{1.17}$$

となることに注目した．その後の式の誘導はさほど困難ではない[19]．式 (1.17) を式 (1.13) に代入し $\Pr(\mathscr{E}|\boldsymbol{x}_m)$ を求め，これを式 (1.12), (1.13) に代入してランダム符号化を行う．期待値を計算して結局，次式を得る．

$$\overline{\Pr(\mathscr{E})}\le(M-1)^{\rho}\sum_{\boldsymbol{y}\in\mathscr{Y}^N}\left[\sum_{\boldsymbol{x}\in\mathscr{X}^N}\Pr(\boldsymbol{x})\Pr(\boldsymbol{y}|\boldsymbol{x})^{\frac{1}{1+\rho}}\right]^{1+\rho} \tag{1.18}$$

確率測度 $\Pr(\boldsymbol{x})$ $(\boldsymbol{x}\in\mathscr{X}^N)$ は任意でよいから，$\Pr(\boldsymbol{x})=\prod_{i=1}^{N}\Pr(x_i)$ $(x_i\in\mathscr{X})$ と考える．通信路は無記憶であるから，$\Pr(\boldsymbol{y}|\boldsymbol{x})=\prod_{i=1}^{N}\Pr(y_i|x_i)$ $(x_i\in\mathscr{X},y_i\in\mathscr{Y})$ である．これらを式 (1.18) に代入して次式を得る．

$$\overline{\Pr(\mathscr{E})}\le\exp_2[-NE_r(R)] \tag{1.19}$$

ただし

$$E_r(R)=\max_{\boldsymbol{q},0\le\rho\le 1}[E_0(\rho,\boldsymbol{q})-\rho R] \tag{1.20}$$

$$E_0(\rho,\boldsymbol{q})=-\log_2\sum_{y\in\mathscr{Y}}\left[\sum_{x\in\mathscr{X}}\Pr(x)\Pr(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho} \tag{1.21}$$

$$\boldsymbol{q}=(q_1,q_2,\cdots,q_{\|\mathscr{X}\|})\qquad(q_i=\Pr(\xi_i),\ \xi_i\in\mathscr{X}) \tag{1.22}$$

である．ここで，式 (1.21) は **Gallager 関数**とよばれている[7]．以上より，復号誤り確率 $\Pr(\mathscr{E})$ が

$$\Pr(\mathscr{E})\le\exp_2[-NE_r(R)]\quad(0\le R<C) \tag{1.23}$$

となる符号長 N，情報伝送速度 R のブロック符号の存在することが示された．式 (1.23) は Gallager [19]により全く新しい方法により導出された．これを **Gallager の上界**とよぶ．

[7] Gallager 関数は，G.D.Forney, Jr. [18]により，さらに一般化されている．

(d) その他の信頼度関数の上界と下界

(c) で求めたランダム符号化指数 $E_r(R)$ は，信頼度関数の下界式の 1 つである．現在知られているその他の限界式は次の通りである．

$$E_{sp}(R) = \max_{\boldsymbol{q},\rho\geq 0}[E_0(\rho,\boldsymbol{q}) - \rho R] \tag{1.24}$$

$$E_{\mathrm{ex}}(R) = \max_{\boldsymbol{q},\rho\geq 1}[E_x(\rho,\boldsymbol{q}) - \rho R] \tag{1.25}$$

ただし

$$E_x(\rho,\boldsymbol{q}) = -\rho\log_2 \sum_{x\in\mathscr{X}}\sum_{x'\in\mathscr{X}} \Pr(x)\Pr(x')\left[\sum_{y\in\mathscr{Y}}\sqrt{\Pr(y|x)\Pr(y|x')}\right]^{\frac{1}{\rho}} \tag{1.26}$$

とする．このとき，$E_r(R) = E_{sp}(R)$ $(R_{\mathrm{crit}} \leq R < C)$ が成り立つ．$R_x \leq R \leq R_{\mathrm{crit}}$ で，$E_x(1,\boldsymbol{q}) = E_0(1,\boldsymbol{q})$ であるから $E_r(R)$ と $E_{\mathrm{ex}}(R)$ は共に $\rho = 1$ で最大値をとり，$E_r(R) = E_{\mathrm{ex}}(R) = R_{\mathrm{comp}} - R$ $(R_x \leq R \leq R_{\mathrm{crit}})$．ただし $R_{\mathrm{comp}} = \max_{\boldsymbol{q}} E_0(1,\boldsymbol{q})$ である．また

$$E_{sl}(R) = \lambda E_{\mathrm{ex}}(0) + (1-\lambda)E_{sp}(R_{sl}) \tag{1.27}$$

$$0 < R = (1-\lambda)R_{sl} \leq R_{sl} \quad (0 \leq \lambda \leq 1)$$

ただし，$R = R_{sl}$ で $E_{sl}(R) = E_{sp}(R)$ である．

以上の結果，信頼度関数の上界 $E_U(R)$ と下界 $E_L(R)$ をまとめると次の通りである．

$$E_L(R) = \max[E_r(R), E_{\mathrm{ex}}(R)] \tag{1.28}$$

$$= \begin{cases} E_{\mathrm{ex}}(R), & 0 \leq R \leq R_x; \quad \text{(low-rate segment)} \\ R_{\mathrm{comp}} - R, & R_x \leq R \leq R_{\mathrm{crit}}; \quad \text{(straight-line segment)} \\ E_r(R), & R_{\mathrm{crit}} \leq R < C \quad \text{(high-rate segment)} \end{cases}$$

$$E_U(R) = \min[E_{sp}(R), E_{sl}(R)] \tag{1.29}$$

$$= \begin{cases} E_{\mathrm{ex}}(0), & R = 0; \quad \text{(zero-rate exponent)} \\ E_{sl}(R), & 0 \leq R \leq R_{sl}; \quad \text{(tangent exponent)} \\ E_{sp}(R), & R_{sl} \leq R \leq R_{\mathrm{crit}}; \quad \text{(high-rate segment extension)} \\ E_r(R), & R_{\mathrm{crit}} \leq R < C \quad \text{(high-rate segment)} \end{cases}$$

工学的な視点から次の定理を与える．

定理 1.2 (通信路符号化定理)

通信路容量 C の離散的無記憶通信路において，復号誤り確率 $\Pr(\mathscr{E})$ が次式を満足する符号長 N，情報伝送速度 R のブロック符号が存在する．

$$\Pr(\mathscr{E}) \leq \exp_2[-NE_L(R)] \quad (0 \leq R < C) \tag{1.30}$$

しかし，次式を満足する符号は存在しない．

$$\Pr(\mathscr{E}) \leq \exp_2[-N\{E_U(R) + o(1)\}] \quad (0 \leq R < C) \tag{1.31}$$

ここで，$o(1) \to 0$ $(N \to \infty)$ である． □

定理 1.3

$R < C$ ならば，$E_L(R) > 0$ である． □

図 1.2 に誤り確率 $p = 0.01$ の 2 元対称通信路の信頼度関数の上界と下界を示す．先に示した $E_{\mathrm{ex}}(R), E_{sp}(R), E_{sl}(R)$ について述べておこう．

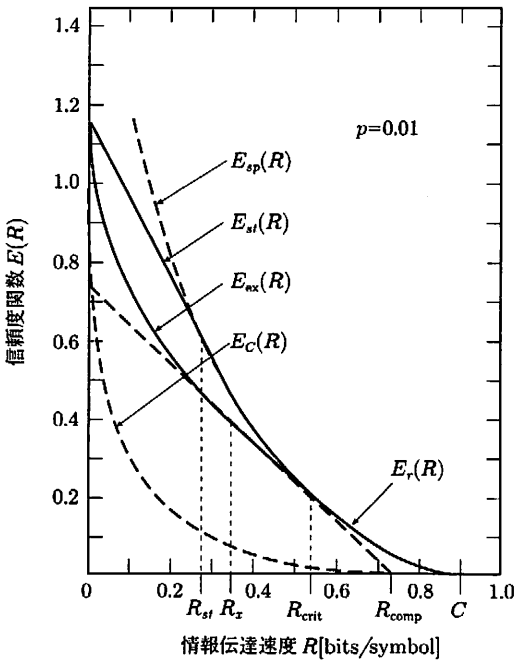

図 1.2 $p = 0.01$ の 2 元対称通信路における信頼度関数の限界

i) **削除誤り指数** (expurgated error exponent) [8] $E_{ex}(R)$ は Gallager [19]によって導かれた．$2M$ 個の符号語を用意し，これより M 個の不良な符号語を削除して得られる[9]．ただし，削除する技法は本質的に必要なものではない[43]．また，2 元入力出力対称通信路[10]では削除する技法を用いず **Bhattacharyya 限界式**[5] を用いて直接導出することも可能である[32]．一方，復号誤り確率の下界は $M = O(N)$，すなわち，$R \doteq 0$ とし符号語間の最小 Hamming 距離を与え，Chernoff 限界式を用いて，$\Pr(\mathscr{E}) \geq \exp\{-N[E_{ex}(0) + O(\frac{1}{\sqrt{N}})]\}$ が得られ結局，漸近的に $E_{ex}(0)$ は真の誤り指数を与えることがわかる[37]．

ii) **球充填誤り指数** (sphere-packing exponent) $E_{sp}(R)$ は最初 Fano [15]により論じられ，続いて Shannon ら[37]により示された信頼度関数の上界である．$E_{sp}(R)$ は，超球上に配置された M 個の符号語間の最大の最小 Hamming 距離と Chernoff 限界式を用いて導出される．信頼度関数の下界の 1 つである式 (1.20) のランダム符号化指数と対比すれば明らかな通り，ρ のとる領域が異なるだけのエレガントな形で与えられている．ただちに $0 \leq \rho \leq 1$ の範囲では両者は一致し，真の誤り指数を与えていることがわかる．

iii) **直線誤り指数** (straight-line exponent) $E_{sl}(R)$ も Shannon ら[37]により求められ，任意の $E_U(R)$ と $E_{sp}(R)$ を用いて，式 (1.27) のように与えられる．式 (1.27) は点 $(0, E_{ex}(0))$ と $E_{sp}(R)$ 上の任意の点を結ぶ直線である[11]．

以上の議論により，$E_{ex}(0)$ と $E_r(R) = E_{sp}(R)$ $(R_{crit} \leq R < C)$ が真の誤り指数，すなわち，信頼度関数 $E(R)$ として確立した．しかし低情報伝送速度では $E(R)$ は決定されていない．なお，誤り指数の計算法は Arimoto [4]により，また同時期に J.R.Lesh [28]により見出されている．

(2) 通信路符号化の逆定理

情報伝送速度 R が通信路容量 C に近づくときは信頼度関数の上界と下界は共に 0 に近づく．すなわち，$R \to C - 0$ のとき $E(R) \to 0$ である．したがって，$R > C$ のとき余り良い結果は望めそうにない．実際，どんな符号化復号化によらずある $\gamma > 0$ が存在し，シンボル当たりの平均復号誤り確率は γ 以上になる．これは Fano の不等式を用いて証明される[16]．J.Wolfowiz [46]，Gallager [20] に続いて，ブロック符号に対して Arimoto [3]により，次の結果

[8] 復号誤り確率の上界を与えるので**削除上界**ともよばれる．

[9] 低情報伝送速度でのランダム符号化では，同一の符号語を選ぶ等の多くの不良な符号語を含み，復号誤り確率の上界の評価がゆるくなってしまうため削除する技法を用いる．

[10] 代表的通信路の 1 つである 2 元対称通信路や 2 元入力加法的白色ガウス雑音通信路も含まれる．

[11] もし $E(R)$ が下に凸の関数であれば，$E(R)$ の上界 $E_U(R)$ の任意の 2 点を結ぶ直線はまた，$E_U(R)$ を与える．しかし $E(R)$ が下に凸の関数であることは証明されていないので，式 (1.27) は $L \geq 2$ 個のリスト復号器と符号長 N_1, N_2 $(N_1 + N_2 = N)$ の 2 つの符号を連接した符号化を用いて証明される[37]．

が示されている．

定理 1.4 (通信路符号化の逆定理) [3]
通信路容量 C の離散的無記憶通信路において，符号長 N，情報伝送速度 R のブロック符号の復号誤り確率 $\Pr(\mathscr{E})$ は次式で与えられる．

$$\Pr(\mathscr{E}) = 1 - \Pr(\mathscr{C}) \tag{1.32}$$

$$\Pr(\mathscr{C}) \leq \exp_2[-NE_{sc}(R)] \quad (R > C) \tag{1.33}$$

$$E_{sc}(R) = \max_{-1 \leq \rho < 0, \boldsymbol{q}} [E_0(\rho, \boldsymbol{q}) - \rho R] \tag{1.34}$$

ただし，符号語の生起は等確率とし，$E_0(\rho, \boldsymbol{p})$ は式 (1.21) で与えられる Gallager 関数である． □

$E_{sc}(R)$ は式 (1.20) の $E_r(R)$ と双対の形になる．その違いは ρ が $[-1, 0)$ であり，$E_r(R) > 0$ $(R < C)$ の双対として $E_{sc}(R) > 0$ $(R > C)$ が示される．このような符号の存在することは G.Dueck と J.Körner [12]により示されている．

(3) 漸近的距離比の限界式から信頼度関数の導出

符号理論の分野では q 元 (N, K, D) 線形符号[12]$(q \geq 2$ は素数のべき$)$ が存在するための限界式 (通常は D の上界と下界) が議論される．ここで，情報伝送速度 $R = (K/N)\log_2 q$ [bits/symbol] である．

Shannon ら[37]は**漸近的距離比** $\delta(R)$ を

$$\delta(R) = \lim_{N \to \infty} \sup D/N \tag{1.35}$$

と定義し，誤り確率 p の 2 元対称通信路で

$$E(R) \leq \delta(R) \log_2 \frac{1}{\sqrt{4p(1-p)}} + o(1) \tag{1.36}$$
$$o(1) \to 0 \quad (N \to \infty)$$

であることを示した[37]．一方，式 (1.25), (1.26) より誤り確率 p の 2 元対称通信路における削除誤り指数 $E_{ex}(R)$ は

$$E_{ex}(R) = -z_p \log_2 \frac{1}{\sqrt{4p(1-p)}} \leq E(R), \quad R = 1 - H(z_p) \tag{1.37}$$

であることが導かれる[37, 43]．Shannon らは，式 (1.36) の $\delta(R)$ が Gilbert 下界式 (1.42) を等式で満たすと推測した．このとき，式 (1.37) の後半の R は

[12] 符号長 N，情報記号数 K，最小距離 D の符号を (N, K, D) 符号と記す．

$z_p = \delta(R)$ とおけば **Gilbert 下界式**，式 (1.42) を等式で満たした式であり，したがって，式 (1.36), (1.37) より $E(R) = E_{ex}(R)$ となり真の誤り指数が与えられることになる．しかし，式 (1.42) を等式で満たすことは推測であり [35]，反例はないが証明されていない．

なお，Chernoff 限界式を用いて

$$E(R) \geq -\frac{\delta(R)}{2}\log_2 p - \left(1-\frac{\delta(R)}{2}\right)\log_2(1-p) - H\left(\frac{\delta(R)}{2}\right) \qquad (1.38)$$

であることも示されている [37]．その結果，式 (1.36) と組み合わせて

$$\frac{\delta(R)}{2} = \lim_{p\to 0}\frac{E(R)}{-\log_2 p} \qquad (1.39)$$

が導かれる [37]．

一方，符号理論の分野では符号の限界式としていくつか求められており，主な漸近的限界式は

$$\delta(R) \leq 2H^{-1}(1-R) \qquad \text{Hamming bound [25]} \qquad (1.40)$$

$$\delta(R) \leq H^{-1}\left(\frac{1-R}{2}\right) \qquad \text{Plotkin bound [36]} \qquad (1.41)$$

$$\delta(R) \geq H^{-1}(1-R) \qquad \text{Gilbert bound [21]} \qquad (1.42)$$

ただし

$$H(x) = -x\log_2 x - (1-x)\log_2(1-x) \qquad (0 \leq x \leq 1)$$

である．符号理論のテキスト [33, 34] でも，これらと 2 項分布を用いて誤り確率 p の 2 元対称通信路における誤り指数の上界・下界が詳細に議論されている．その結果，Hamming 上界式を用いて誤り指数の上界 $E_{sp}(R)$ が，Gilbert 下界式を用いて誤り指数の下界 $E_{ex}(R), E_r(R)$ が導かれる．

1.1.4 構成的符号化による定理の証明

1.1.3.(3) で符号の最小距離の限界式を用いて誤り指数を導出でき，したがって符号化定理を証明できることを示した．しかし，漸近的距離比の限界式はそれを満足する符号が存在することを示したにすぎない．したがって，その結果の符号化定理もまた存在定理である．しかも，漸近的距離比の上界式と下界式の間には依然として隙があり真の限界式は定まっていない．これは，低情報伝送速度で信頼度関数の上界と下界がうめられていないことに相当する．ここでは，構成的 (constructive) な符号化を用いて符号化定理を証明することを考える．

代数的符号はいずれも符号語の生成方法，すなわち符号化方法を符号長 N の高々多項式のオーダで決定でき，また復号法も代数的演算により多くは $O(N^2)$ 程度で実現できる．このような符号を構成的とよび，通信路符号化定理の証明に用いたランダム符号化と比べ，実現性において著しい違いがあり，**非ランダム符号**ともよばれる．しかし残念ながら，実用化されている BCH (Bose–Chaudhuri–Hocquenghem) 符号や RM (Reed-Muller) 符号などの優れた符号の多くは，符号長 N が 10^3 程度以下の有限長では高い訂正能力をもつが，$N \to \infty$ と共に急速に能力が低下する．逆に $N \to \infty$ と共に能力が低下しない符号を漸近的に良い (asymptotically good) 符号という[13]．もしこの様な符号が発見されれば，通信路符号化定理をランダム符号化によらず証明することが可能である．

いま，$\Pr(\mathscr{E}) \to 0$ $(N \to \infty)$ を実現する構成的符号を通信路符号化定理を証明する非ランダム符号とすれば，このような符号として P.Elias [14] の**繰返し符号** (iterated codes) と G.D.Forney,Jr. [17] の**連接符号** (concatenated codes) のただ 2 つが知られているのみである [33]．さらに構成的で漸近的に良い符号としては，連接符号の 1 つのクラスである Justesen [27] による **Justesen 符号**のみが知られている．

(1) 繰返し符号

Elias は数少ない完全符号の 1 つとして知られている Hamming 符号を拡大し J 段 $(J > 1)$ に多次元に用いた繰返し符号[14]を提案し，代数的復号法により，$\Pr(\mathscr{E}) \to 0$ $(N \to \infty)$ とできることを示した [14]．2 元 $(N_j, K_j, 4)$ 拡大 Hamming 符号を第 j 段の符号とするとき，繰返し符号の生成行列は $j = 1, 2, \cdots, J$ の各段の生成行列のテンソル積となる．ただし $N_j = 2^{m+j-1}, m \geq 4$ である[15]．$j = 1, 2, \cdots, J$ として符号長を 2 倍ずつ伸長することにより，式 (1.44) の符号化比率 R_j^* を正に保ち，復号誤り確率 $\Pr(\mathscr{E}) \to 0$ $(N \to \infty)$ とすることが可能となる．ここで，q 元 (N, K, D) 符号の $R^* = K/N$ を**符号化比率** (code rate) という[16]．

[13] 符号理論において，符号は誤り訂正能力により評価される．通常，漸近的距離比 $\delta(R)$ を用い，$\delta(R) > 0$ $(R \neq 0)$ を満たす符号を漸近的に良い符号という．

[14] Elias は最初，誤りなしの符号化 (error free coding) とよんでいる．

[15] 特別な場合として，$m = 2, 3$ も考えることができる．

最小距離 4 の拡大 Hamming 符号を代数的に復号する．そのとき，1 [bit] の誤りは正しく訂正でき，2 [bits] の誤りはそのまま誤り検出にとどめるものとする．もし，3 [bits] 以上の誤りが生じると 1 [bit] 誤りを追加する可能性がある．これを評価すると

$$\Pr(\mathscr{E}) \le (2^{m+1}p)^{2^J}, \tag{1.43}$$

$$R_J^* = \prod_{j=1}^{J}\left[1 - \frac{m+j}{2^{m+j-1}}\right] \tag{1.44}$$

となる．ただし，2 元対称通信路の誤り確率を p とする．$J \to \infty$ とすると，$R_J^* \to R_0^*$ で

$$R_0^* > 1 - \frac{m+2}{2^{m-1}} \quad (m \ge 4) \tag{1.45}$$

すなわち，$R_0^* > 0$ が得られる．

$N_0 = \prod_{j=1}^{J} N_j$ とし，式 (1.43)，(1.45) を用い式 (1.10) と似た形に変換してみると，幸い指数部が N_0 の関数と R_0^* の関数に分離できて，次の定理を得る [30]．

定理 1.5

誤り確率 p の 2 元対称通信路において，符号長 N_0，符号化比率 R_0^* の繰返し符号の誤り確率 $\Pr(\mathscr{E})$ は次式を満足する．

$$\Pr(\mathscr{E}) \le \exp_2\left[-f(N_0)\cdot E_I(R_0^*)\right] \tag{1.46}$$

ここで

$$f(N_0) = \exp_2\left[\sqrt{2\log_2 N_0}(1 - o(1))\right] \tag{1.47}$$

$$E_I(R_0^*) = -(m+1) - \log_2 p \tag{1.48}$$

ただし，$o(1) \to 0$ $(N_0 \to \infty)$ である． □

式 (1.47) の $f(N_0)$ は N_0 が大となると共に，N_0 よりゆるやかに増加する関数である．注意すべきことは，式 (1.48) から明らかな通り $E_I(R_0^*) > 0$ とできるのは通信路容量よりもかなり小さな R_0^* である．しかも，誤り確率 p がある値以下の 2 元対称通信路に限られる（例えば，$m = 2$ としても $p < \frac{1}{8}$ でなければならない）．なお，明らかに繰返し符号の漸近的距離比 $\delta_E(R_0^*)$ は

[16] 符号化比率 R^* は $q = 2$ のとき，[bit] を単位とする情報伝送速度 R に等しい．

$$\begin{aligned}\delta_E(R_0^*) &= \lim_{N_0\to\infty} D_0/N_0 \\ &= \lim_{J\to\infty}\prod_{j=1}^{J}\frac{4}{2^{m+j-1}} \\ &\to 0\end{aligned} \tag{1.49}$$

である[17]．ただし，$D_0 = \prod_{j=1}^{J} D_j$ である．また，復号化は代数的復号法によるから，その計算量 $\chi_I(N_0)$ は高々 $O(N_0)$ である [30]．なお，繰返し符号を重畳符号化あるいは部分符号化を用いて修正し，同じ復号誤り確率の上界を達成するのに従来より符号化比率を大きくできることが示されている [29]．

(2) 連接符号

Forney [17] は GF(q^K) 上の (n,k,d)RS (Reed–Solomon) 符号を外部符号とし，GF(q) 上の (N,K,D) 符号を内部符号とする (N_0,K_0,D_0) 連接符号を提案した．q は任意の素数のべきとする．$N_0 = nN, K_0 = kK, D_0 \ge dD$ である．$q = 2$ のとき，$n = 2^K$ であるから，$N_0 = O(2^{N(1+o(1))})$ である．それゆえ，BCH 符号など優れた内部符号と RS 符号を組み合せて符号長 N_0 が十分大となる強力な符号を作ることができる．その結果，任意の通信路に適用可能とするために内部符号は最尤復号化し，その復号誤り確率をある一定値以下におさえれば，外部符号の代数的復号化 (GMD 復号化) により $\Pr(\mathscr{E}) \to 0$ $(N_0 \to \infty)$ とすることが可能となる．このとき，次の定理が成り立つ．

定理 1.6

通信路容量 C の離散的無記憶通信路において，符号長 N_0，情報伝送速度 R_0 の連接符号の復号誤り確率 $\Pr(\mathscr{E})$ は次式を満足する．

$$\Pr(\mathscr{E}) \le \exp_2[-N_0 E_C(R_0)] \quad (0 \le R_0 < C) \tag{1.50}$$

$$E_C(R_0) = \max_{rR=R_0}(1-r)E_L(R) \quad (0 \le r \le 1, \quad 0 \le R < C) \tag{1.51}$$

□

ここで，内部符号を最尤復号とするために $O(2^{NR})$ の比較を n 回必要とするが，外部符号が代数的に復号可能であるため $O(n^2\log^4 n)$ 程度で実現でき，全

[17] このことは，$\Pr(\mathscr{E}) \to 0$ $(N \to \infty)$ となる符号は，必ずしも $\delta(R_0^*) > 0$ を要しないことを示す例である．

体として復号化には $O(N_0^2 \log^2 N_0)$ の計算量でよい[23]．このことは，実用性が極めて高いことを示している．

連接符号化誤り指数 $E_C(R_0)$ を図 1.2 に点数で示す．なお，部分符号をもつ内部符号と符号化比率の異なる外部符号を複数個用いて連接符号を構成し，多段復号化することによって得られる一般化された連接符号の誤り指数は従来の連接符号のそれより大となることが示されている[23, 24]．

(3) Justesen 符号

Justesen は連接符号を用いて漸近的距離比 $\delta(R_0^*) > 0 \quad (0 < R_0^* < 1)$ となる符号を見出した[27]．外部符号は GF(2^K) 上の RS 符号とする．GF(2) 上の (N, K) 内部符号に集合として Gilbert 下界式を満たす Wozencraft の集合を用いる．すなわち長さ K の情報記号と，これに β^i を乗じたものを検査記号として得られる符号である．ただし，β は GF(2^K) 上の原始元である．RS 符号の符号長は $n = 2^K - 1$ であるから，$i = 1, 2, \cdots, n$ とすべてを使い切ると n 個の内部符号の生成行列はすべて異なるが，集合として Gilbert 下界式を満たすという性質を用いることができる．その結果，次の定理が与えられる．

定理 1.7 [27]
符号化比率 R_0^* の Justesen 符号の漸近的距離比 $\delta_J(R_0^*)$ は，次式で与えられる．

$$\delta_J(R_0^*) \geq \max_{r^* R^* = R_0^*, R^* \geq 1/2} (1 - r^*) H^{-1}(1 - R^*) \tag{1.52}$$

□

Justesen は同時に復号法も与え，2 元対称通信路においてその計算量は $O(N_0^2 \log N_0)$ であることを示している．$E(R)$ に対する $E_C(R_0)$ と同様，Gilbert 下界式に対し $\delta_J(R_0^*)$ ははるかおよばない．しかし，$R > 0$ で $\delta(R^*) \neq 0$ の符号の存在価値は大きい．なお，Justesen 符号を複数個重畳することにより，高符号化比率で Zyablov 下界式[18][47]を超える漸近的に良い符号も示されている[39]．

また，Justesen 符号の内部符号は $R \geq \dfrac{1}{2}$ でなければならない．この条件は外部符号に代数幾何符号を用いることにより除去でき，その結果低符号化比率において改良され，漸近的距離比は Zyablov 下界式と一致する[42]．

[18] Zyablov 下界式は $\delta_Z(R_0^*) \geq \max_{r^* R^* = R_0^*} (1 - r^*) H^{-1}(1 - R^*)$ で与えられる．

以上の結果，2 元対称通信路において，$\delta_J(R_0^*) > 0 \ (0 < R_0^* < 1 \ (R_0^* \neq 0))$ の構成的符号化である Justesen 符号により，式 (1.38) を用いて通信路符号化定理が証明される．なお，連接符号，Justesen 符号の考え方を用い代数的構成方法で直接誤り指数を導出しようとする研究もある[13]．

1.1.5 通信路符号化定理の一般化

(1) タイプの概念によるユニバーサル符号化[10]

情報源の特性が未知のとき，その特性を全く意識しないで符号化を行うのがユニバーサル（情報源）符号化である．Ziv–Lempel 符号や Bayes 符号などがよく知られており，実際にデータ圧縮のアルゴリズムとして用いられている．通信路符号化においても，通信路の特性が未知のときや変動するとき，通信路の統計的性質に依存しない符号化と復号法を考える．これを用いて拡張したものがユニバーサル（通信路）符号化である．

ユニバーサル符号化定理を導くために，まずタイプの考え方を示す．通信路の入力アルファベットの集合を $\mathscr{X} = \{\xi_1, \xi_2, \cdots, \xi_{\|\mathscr{X}\|}\}$，出力アルファベットを $\mathscr{Y} = \{\psi_1, \psi_2, \cdots, \psi_{\|\mathscr{Y}\|}\}$ とする．符号語 $\boldsymbol{x}_m \in \mathscr{X}^N$ において，$\xi_i \in \mathscr{X}$ の出現する相対頻度を q_i とするとき，確率分布 $\boldsymbol{q} = (q_1, q_2, \ldots, q_{\|\mathscr{X}\|})$ を系列 $\boldsymbol{x}_m$ のタイプといい，$q_i = P_{\boldsymbol{x}_m}(\xi_i)$ と表す．$\boldsymbol{q}$ を用いてエントロピー $H(\boldsymbol{q})$ が定義できる．同様に符号語 $\boldsymbol{x}_m \in \mathscr{X}^N$ と受信系列 $\boldsymbol{y} \in \mathscr{Y}^N$ の対 $(\boldsymbol{x}_m, \boldsymbol{y})$ に対し，$\xi_i \in \mathscr{X}$ と $\psi_j \in \mathscr{Y}$ の出現する同時（結合）相対頻度を Q_{ij} とするとき，これを対 $(\boldsymbol{x}_m, \boldsymbol{y})$ の同時タイプといい，$Q_{ij} = P_{\boldsymbol{x}_m, \boldsymbol{y}}(\xi_i, \psi_j)$ と表す．したがって，条件付確率 $P_{\boldsymbol{x}_m, \boldsymbol{y}}(\psi_j | \xi_i) = P_{\boldsymbol{x}_m, \boldsymbol{y}}(\xi_i) / P_{\boldsymbol{x}_m}(\xi_i)$ が定義でき，$\boldsymbol{x}_m$ と $\boldsymbol{y}$ の間の相互情報量 $I(\boldsymbol{x}_m; \boldsymbol{y})$ は次式で与えられる．

$$I(\boldsymbol{x}_m; \boldsymbol{y}) = \sum_{\xi_i, \psi_j} P_{\boldsymbol{x}_m, \boldsymbol{y}}(\xi_i, \psi_j) \log_2 \frac{P_{\boldsymbol{x}_m, \boldsymbol{y}}(\xi_i, \psi_j)}{P_{\boldsymbol{x}_m}(\xi_i) P_{\boldsymbol{y}}(\psi_j)} \tag{1.53}$$

1.1.3 の証明で用いた最尤復号法は明らかに通信路行列 $P = [\Pr(y|x)] \ (x \in \mathscr{X}, y \in \mathscr{Y})$ に依存している．そこで，これを

$$\boldsymbol{y} \to \boldsymbol{x}_{\hat{m}} : \hat{m} = \arg \max_{\boldsymbol{x}_m \in C} I(\boldsymbol{x}_m; \boldsymbol{y}) \tag{1.54}$$

とする．右辺の $\boldsymbol{x}_m \in C$ は $m = 1, 2, \ldots, M$ のすべての対 $(\boldsymbol{x}_m, \boldsymbol{y})$ に対し同時タイプ $P_{\boldsymbol{x}_m, \boldsymbol{y}}(\xi_i, \psi_j)$ を計算する．この値は $(\boldsymbol{x}_m, \boldsymbol{y})$ が与えられたとき，$x_l = \xi_i$，かつ $y_l = \psi_j$ となる両系列の要素の位置 l の相対頻度である．したがって，$\boldsymbol{x}_m$ の復号領域は通信路行列 P にはよらず一意に定まる．符号語 $\boldsymbol{x}_m$ はもちろん通信路に依存しないから，式 (5.2) はユニバーサルな復号法となっている．これは V.D.Goppa により提案された**最大相互情報量**（maximum mutual information : MMI）**復号法**である．

MMI 復号法によれば復号領域は通信路行列 P によらない．ここで，すべての符号語

$\boldsymbol{x}_m$ $(m=1,2,\ldots,M)$ が唯一のタイプ $\boldsymbol{q}$ をもつ符号化[19]を考える．さらに，通信路行列 P を考え，その集合を $\mathscr{P}$ とする．$\mathscr{P}$ が与えられたとき，$\boldsymbol{x}\in\mathscr{X}^N$ に対し条件付タイプ P^* をもつ $\boldsymbol{y}\in\mathscr{Y}^N$ の集合 $\mathscr{T}_{\boldsymbol{x}}^N(P^*)$ を

$$\mathscr{T}_{\boldsymbol{x}}^N(P^*)=\{\boldsymbol{y}\in\mathscr{Y}^N:\boldsymbol{y}=\boldsymbol{x}P^*,P^*\in\mathscr{P}\} \tag{1.55}$$

と表す．式 (1.55) は送信符号語を $\boldsymbol{x}$ とするとき，通信路行列 P^* による受信系列 $\boldsymbol{y}$ の集合である．$\boldsymbol{x}$ がタイプ $\boldsymbol{q}$ をもつとき，$\mathscr{T}_{\boldsymbol{x}}^N(P^*)\neq\phi$ となるような条件付タイプ P^* の集合を $\mathscr{P}_N^*(\boldsymbol{q})$ と表すと

$$\|\mathscr{P}_N^*(\boldsymbol{q})\|\le(N+1)^{\|\mathscr{X}\|\|\mathscr{Y}\|} \tag{1.56}$$

を満足する (type counting lemma)．なぜならば $\Pr(y|x)$ のとり得る値は，高々 $0/N, 1/N, \ldots, N/N$ の $N+1$ 通りであるから，P の次元 $\|\mathscr{X}\|\|\mathscr{Y}\|$ から異なるタイプの数（通信路行列の個数）は，式 (1.56) で与えられる．

同様に，$\boldsymbol{x}$ に対する $P\in\mathscr{P}_N^*(\boldsymbol{q})$ による集合 $\boldsymbol{y}$ の個数は

$$\frac{1}{(N+1)^{\|\mathscr{X}\|\|\mathscr{Y}\|}}2^{NH(P|\boldsymbol{q})}\le\|\mathscr{T}_{\boldsymbol{x}}^N(P)\|\le 2^{NH(P|\boldsymbol{q})} \tag{1.57}$$

$$H(P|\boldsymbol{q})=-\sum_i q_i\sum_j P_{ij}\log_2 P_{ij} \tag{1.58}$$

を満足する．また，条件付タイプ P により $\boldsymbol{x}$ を送信したとき，P とは異なる通信路行列 $Q\in\mathscr{X}\mathscr{Y}$ によって得られる受信系列 $\boldsymbol{y}$ の生起確率を P と Q のダイバージェンスで評価すると

$$Q(\boldsymbol{y}|\boldsymbol{x})=2^{-N[D(P\|Q|\boldsymbol{q})+H(P|\boldsymbol{q})]} \tag{1.59}$$

を満足する．ここで，$D(\cdot\|\cdot|\cdot)$ は条件付ダイバージェンスである．

復号誤り確率を小さくするためには，式 (1.55) で示した $\boldsymbol{x}_m$ $(m=1,2,\ldots,M)$ の M 個の空間をなるたけ互いに交わらないようにすればよい．符号語 $\boldsymbol{x}_m$ がタイプ $\boldsymbol{q}$ をもつならば，どのような通信路行列 P,P' についても次式を満たす M 個の系列が存在する (packing lemma)．

$$\frac{\|\mathscr{T}_{\boldsymbol{x}_m}^N(P)\cap[\bigcup_{m'\neq m}\mathscr{T}_{\boldsymbol{x}_{m'}}^N(P')]\|}{\|\mathscr{T}_{\boldsymbol{x}_m}^N(P)\|}\le 2^{-N|I(\boldsymbol{q},P')-R|^+},\quad |\alpha|^+=\max\{\alpha,0\} \tag{1.60}$$

ただし

$$I(\boldsymbol{q},P')=H(\boldsymbol{q}P')-H(P'|\boldsymbol{q}) \tag{1.61}$$

$$H(\boldsymbol{q}P')=-\sum_j r_j\log_2 r_j\qquad\left(r_j=\sum_i q_i P'_{ij}\right)$$

である．すなわち，$\boldsymbol{x}_m$ による $\boldsymbol{y}$ の集合で，他の $\boldsymbol{x}_{m'}$ による $\boldsymbol{y}$ と交わる割合は式 (1.60) の右辺で上界される．この補題はランダム符号化を用い，式 (1.56) より証明される．これらの結果はすべて通信路によらないユニバーサルなものである．

[19] すなわち固定タイプ符号．

ところで，送信符号語を $\boldsymbol{x}_m$，受信系列を $\boldsymbol{y}$ としたときの復号誤り $\mathscr{E}_m$ は式 (1.54) より $\hat{m}\neq m$，すなわち $I(\boldsymbol{q},P')>I(\boldsymbol{q},P)$ で $\boldsymbol{y}\in\mathscr{T}_{\boldsymbol{x}_m}^N(P)\cap\mathscr{T}_{\boldsymbol{x}_{m'}}^N(P')$ のとき生起する．$\boldsymbol{y}\to\boldsymbol{x}_{m'}$ と復号するときその和集合をとり，式 (1.56)-(1.61) を用いて $\mathscr{E}_m$ の生起確率 $\Pr(\mathscr{E}_m)$ を求めると，$|I(\boldsymbol{q},P')-R|^+\ge|I(\boldsymbol{q},P)-R|^+$ であるから結局

$$\Pr(\mathscr{E}_m)\le\sum_{P,P'}\exp_2\{-N[D(P\|P'|\boldsymbol{q})+|I(\boldsymbol{q},P)-R|^+]\}$$

となり，次の定理が得られる．

定理 1.8 (ユニバーサル通信路符号化定理) [10]
任意の離散的無記憶通信路 Q において，固定タイプ[20]の符号長 N，情報伝送速度 $R_a>0$ の符号化を行う．このとき，$\delta>0$ に対し 復号誤り確率 $\Pr(\mathscr{E})$ は次式を満足する．

$$\Pr(\mathscr{E})\le\exp_2[-N[E_r(R,\boldsymbol{q},Q)-\delta]\quad(R<I(\boldsymbol{q},Q)\le H(\boldsymbol{q}))$$

$$R_a>R-\delta \tag{1.62}$$

$$E_r(R,\boldsymbol{q},Q)=\min_P[D(P\|Q|\boldsymbol{q})+|I(\boldsymbol{q},P)-R|^+] \tag{1.63}$$

である． □

$D(\cdot\|\cdot|\cdot)\ge 0$ であるから，次の系が成り立つ．

定理 1.9
$R<I(\boldsymbol{q},Q)$ ならば，$E_r(R,\boldsymbol{q},Q)>0$ である． □

なお，式 (1.62) を導く途中，$(N+1)^{2\|\mathscr{X}\|\|\mathscr{Y}\|}\le 2^{N\delta}(N\to\infty)$ としている．また，$E_r(R,\boldsymbol{q},Q)=E_r(R)$ である．したがって，式 (1.62) は

$$\Pr(\mathscr{E})\le(N+1)^{2\|\mathscr{X}\|\|\mathscr{Y}\|}\exp_2[-NE_r(R)] \tag{1.64}$$

と表現できる．式 (1.23) と式 (1.64) より，ユニバーサル符号化による劣化は符号長 N の多項式オーダ $(N+1)^{2\|\mathscr{X}\|\|\mathscr{Y}\|}$ である．式 (1.62) では信頼度関数 $E_r(\cdot)$ に δ の劣化分となって現れる．これはユニバーサル情報源符号化でも平均符号長の上界が大きくなってしまうのと同様，通信路特性が未知であることによるペナルティである．削除誤り指数 $E_{\mathrm{ex}}(R)$ についても同様の代償を払ってユニバーサル化される．なお，本節の議論は[41]に詳しいので参照されたい．

[20] 固定コンポジットともいう．

(2) 情報スペクトル的方法による一般化[22]

従来の通信路符号化定理は離散的無記憶通信路を主体に議論され精密化されてきた．加法的白色ガウス雑音通信路，高雑音通信路など個々の通信路や記憶のある通信路などへの拡張もあるが，それでも基本的には定常性，エルゴード性を仮定している．そのために厳密な議論が可能となる．しかし，実際の通信路はそのような理想的・典型的な場合ばかりではないから，必要に応じ厳密な結果を多少緩和して，適当な類推により理論と実際の狭間をうめる方法がとられる[15]．

最近，韓は情報スペクトル的方法によってこの狭間をうめる必要のない極めて一般的な結論を導き出した[22]．それによれば，通信路は非定常，非エルゴードでもよく，通信路の入出力アルファベットは可算無限個でもよい．

情報スペクトル的方法では，従来のように定常でエルゴード的という1つの通信路モデル[21]だけで考えるのではない．通信路のモデル空間には，非定常，非エルゴードで記憶をもつものが含まれている．これらを相互情報量，あるいはその拡張したもので分類し広がりのあるスペクトルを考えるのである．そのために，与えられた通信路に対し個々の（孤立した）入出力対を区別して扱う．

まず，一般の通信路を次のように考える．通信路の入力系列を $\boldsymbol{x}\in\mathcal{X}^N$，出力系列を $\boldsymbol{y}\in\mathcal{Y}^N$ とするとき，$Q^N(\boldsymbol{y}|\boldsymbol{x})$（ただし，$\forall\boldsymbol{x}\in\mathcal{X}^N:\sum_{\boldsymbol{y}\in\mathcal{Y}^N}Q^N(\boldsymbol{y}|\boldsymbol{x})=\cdot 1$）を一般の通信路と定義する．すなわち，行列 $Q^N(\cdot|\cdot)$ はその要素が $[0,1]$ であり，各行の要素の和が1であればどのような行列でもよい．$\boldsymbol{Q}=\{Q^N\}_{N=1}^{\infty}$ が通信路であるから非定常，非エルゴード，記憶のあるものもすべて含まれる．X^N を $\mathcal{X}^N$ の要素をとる確率変数とすると，一般の通信路の入力過程 $\boldsymbol{X}=\{X^N\}_{N=1}^{\infty}$ と，その出力過程 $\boldsymbol{Y}=\{Y^N\}_{N=1}^{\infty}$ に対し

$$P_{X^N,Y^N}(\boldsymbol{x},\boldsymbol{y})=P_{X^N}(\boldsymbol{x})\,Q^N(\boldsymbol{y}|\boldsymbol{x}) \tag{1.65}$$

が成り立つ．そこで，入出力対 $(\boldsymbol{X},\boldsymbol{Y})$ に対し $\lim_{N\to\infty}P\{z_N<\alpha\}=0$，ここで

$$z_N=\frac{1}{N}\log\frac{Q^N(Y^N|X^N)}{P_{Y^N}(Y^N)} \tag{1.66}$$

となる α の下限 $\inf\alpha$ を $\underline{I}(\boldsymbol{X};\boldsymbol{Y})$ と定義する．ただし，z_N は個々の入出力対に対し定まる通常の相互情報量に相当するもので，$(\boldsymbol{X},\boldsymbol{Y})$ の**相互情報量密度レート**とよばれる．すべての入出力対 $(\boldsymbol{X},\boldsymbol{Y})$ の相互情報量密度レートを求め，同一の情報量密度レートを持つ入出力対をひとまとめにしその（頻度）分布を求めると，相互情報量密度レートに対する確率分布，すなわち相互情報量スペクトル[22]が得られる．通信路符号化で重要な役割を持つのはその下限であり，これが相互情報量スペクトル下限 $\underline{I}(\boldsymbol{X};\boldsymbol{Y})$ である[23]．

[21] 確かにこのような制限下では大数の法則を用いて厳密な議論ができる．しかし，この制限を緩めるとたちまち何も言えなくなってしまう．

[22] このような確率分布を総称して情報スペクトルという．

なお，通常の定常エルゴード的通信路に対する式(1.66)の z_N の値は個々の通信路を区別して考えるまでもなく $N\to\infty$ ですべて同じ値をとり，$\lim_{N\to\infty}z_N=I(\boldsymbol{X};\boldsymbol{Y})$ であるから，相互情報量スペクトルは輝線（単一）スペクトルとなる．これが従来の議論に対応する．

定理 1.10 (一般化された通信路符号化定理) [30]
一般の通信路 $\boldsymbol{Q}$ の通信路容量 C は

$$C=\sup_{\boldsymbol{X}}\underline{I}(\boldsymbol{X};\boldsymbol{Y}) \tag{1.67}$$

で与えられる． □

式(1.67)はもちろん，定常エルゴード的通信路でも成り立ち，$C=\sup_X I(X;Y)$ に帰着する．これは，定常エルゴード性を用いて z_N を展開し積の形を導き，その各項の分散が有界であるから Chebyshev の不等式を適用して導出されるが，その道のりは必ずしも楽ではない．

一方，一般の通信路においては $I(\cdot;\cdot)$ の定義から $z_N>\underline{I}(\boldsymbol{X};\boldsymbol{Y})-\gamma$ $(\gamma>0)$ となる入出力対 $(\boldsymbol{x},\boldsymbol{y})\in\mathcal{X}^N\mathcal{Y}^N$ の集合（典型的系列）を考えれば，ランダム符号化を用いて $R<C$ で $\Pr(\mathscr{E})\to 0$ $(N\to\infty)$ となることが示される．逆に，$R>C$ ならば $\Pr(\mathscr{E})\to 0$ $(N\to\infty)$ とできないことも証明できる．ここで注意すべきことは，確率的下限値の意味は符号長 N，情報記号数 K，復号誤り確率 $\Pr(\mathscr{E})$ の符号 C の情報スペクトルは，$\Pr(\mathscr{E})\to 0$ $(N\to\infty)$ のとき R に確率収束することである．すなわち，$\Pr(\boldsymbol{x})=\dfrac{1}{M}$ $(M=2^{NR})$ のとき $^\forall\gamma>0$ に対し

$$\lim_{N\to\infty}\Pr\{z_N\le R-2\gamma\}=0 \tag{1.68}$$

$$^\forall N>N_0,\Pr\{z_N>R+\gamma\}=0 \tag{1.69}$$

である．

このように，情報スペクトル的方法に基づき，数学的に緻密な演繹によって一般の通信路で通信路符号化定理を導くことができる．しかし，信頼度関数を用いた復号誤り確率などの詳細や通信路を一般化したためにどのような代償を払わねばならないかを明らかにすることは今後の課題である．

[23] これは z_N の確率的下極限とよばれる．確率的下極限（または上極限）の考え方は，筆者らの知る限り従来の確率論にはない．

1.1.6 むすび

通信路符号化定理は1948年Shannon [38]により示された．その証明方法は大数の弱法則にしたがう典型的系列の出現確率に基づくものである．その後，1960年代Gallager [19, 20]らにより大数の法則によらない方法，すなわちChernoff限界式により信頼度関数の上界・下界として復号誤り確率の評価が精密化された．しかし現在，特別な通信路を除いて，低情報伝送速度の信頼度関数はその上界と下界が一致しないという意味で正確に求まっていない．これは通信路符号化定理の完成のための重要な課題である．

一方，ランダム符号化によらない構成的符号化を用いて通信路符号化定理を証明する研究が続けられている．1950年代，当時知られていた数少ない符号の一つであった拡大Hamming符号を多段に組み合わせ，誤り確率$p\left(0<p<\frac{1}{8}\right)$の2元対称通信路において，復号誤り確率を0に収束させることができるはじめての非ランダム符号として，Elias [14]により**繰返し符号**が示された．ただし，(N_0, K_0, D_0)繰り返し符号は$R_0^* = K_0/N_0 > 0$で$D_0/N_0 \to 0$ $(N_0 \to \infty)$となる．符号理論の立場から言えば，拡大Hamming符号は単一段のBCH符号などと同程度漸近的に良い符号ではないにもかかわらず，段数を重ねるにしたがい要素符号の符号長を2倍ずつ伸長することにより，R_0^*を正に保ちながら復号誤り確率$\Pr(\mathscr{E})$を0に収束することが証明される．

また1960年代，BCH符号など漸近的には余り良いとは言えない符号とRS符号を組み合わせて符号長を十分大とできる第2の非ランダム符号として，**連接符号**[17]が登場した．これは1970年代に入って，外部符号の符号長に等しいn個の内部符号を次々相異なる生成行列で符号化する可変内部符号化により構成した**Justesen符号**[27]として脚光を浴びた．その特徴は次のように述べることができる．いま，2元(N, K, D)符号化法がn通りあるとする．その個々の符号では$K/N > 0$で$D/N \not\to 0 (N \to \infty)$を保証できないが，その集合ならば平均的に$D/N > 0$であることが保証できるのがWozencraftの集合である．これを内部符号とし，すべて1回ずつ合計n回$(n = 2^K - 1)$用いて符号長nのRS符号を外部符号とする連接符号がJustesen符号である．その結果，漸近的に良い符号が構成できる．この内部符号の構成法はGallagerによるランダム符号化に対応している．構成的符号化による通信路符号化定理の証明には，$0 < R_0^* = K_0/N_0 < 1$で$D_0/N_0 > 0$ $(N_0 \to \infty)$の(N_0, K_0, D_0)符号が得られれば，復号誤り確率$\Pr(\mathscr{E}) \to 0$示すことができる．Justesen符号の漸近的距離比D_0/N_0 $(N_0 \to \infty)$の値は，$0 < R_0^* < 1$において存在することが保証されているGilbert下界式の値にはるかにおよばないが，$\Pr(\mathscr{E}) \to 0$ $(N_0 \to \infty)$とできるのである．このような議論は，通信路符号化定理がランダム符号化による存在定理であったものを，構成的符号化により具体的符号を用いて証明することを可能にするばかりではない．代数的手順により構成された符号は復号の手順も代数的に実行でき，したがって多くは復号の計算量が符号長の多項式オーダに収まるのである．ここでは内部符号の復号誤り確率をある値以下（たとえば，10^{-3}以下）にすれば，外部符号は代数的復号法により最尤復号法と同じ復号誤り確率の上界を達成する**GMD復号法**[17]が活躍する．Justesen符号は現在のところ理論的興味にとどまるが，実用的符号を探求する符号理論と理論的限界を求める情報理論の橋渡しとして重要である．

通信路符号化定理の拡張・一般化は，1980年代Csiszár と Körner [10]の**タイプの理論**による情報理論の再構築により，通信路の統計的特性によらないユニバーサル符号化定理として注目された．その後，しばらくこの流れは途絶えたかにみえたが，1990年代後半に入りShannonの論文から50年の後に，韓[22]による**情報スペクトル的方法**により通信路の種類・性質によらない一般的な定理として新しい体系化がなされた．今後さらに詳細化・精密化されるであろう．

以上述べたように，通信路符号化定理をめぐる話題は情報理論・符号理論の多くの基本問題を含んでいる．ここでは述べなかったが，ゼロ誤り通信路容量の計算法，多端子通信路の符号化定理などや，信頼度関数の上界・下界，ブロック符号とたたみ込み符号の構造解析，ターボ符号の理論的な性能の解析など未解決の問題が多い．また，漸近的に良い非ランダム符号の改良は代数幾何符号の発展に大きな期待が寄せられている．

参考文献

[1] 有本卓, 平澤茂一：“通信路符号化と信頼度関数”, 信学論 (A),J73-A,2, pp.188-195, 平2.

[2] S.Arimoto, “An algorithm for computing the capacity of arbitrary discrete memoryless channel”, *IEEE Trans. Inform. Theory*, IT-18, pp.14-20, 1972.

[3] S.Arimoto, “On the converse to the coding theorem for discrete memolyless channels”, *IEEE Trans. Inform. Theory*, IT-19,3, pp.357-359, 1973.

[4] S.Arimoto, “Computation of random coding exponent functions”, *IEEE Trans. Inform. Theory*, IT-22, pp.665-671, 1976.

[5] A.Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions", *Bull. Calcutta Math. Soc.* 35, pp.99-110, 1943.
[6] R.E.Blahut, "Composition of channel capacity and rate distortion functions", *IEEE Trans. Inform. Theory*, IT-18, pp.460-473, 1972.
[7] R.E.Blahut, "Hypothesis testing and information theory", *IEEE Trans. Inform. Theory*, IT-20, pp.405-417, 1974.
[8] R.E.Blahut, "Composition bounds for channel block codes", *IEEE Trans. Inform. Theory*, IT-23, pp.656-674, 1977.
[9] R.E.Blahut, *Principles and Practice of Information Theory*, Addison Wesley, Reading Mass. or Tokyo, 1987.
[10] I.Csiszár and J.Körner, *Information Theory*, Akademiai Kiado, Budapest, and Academic Press, New York, 1981.
[11] H.Chernoff, "A Measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations", *Ann. Math. Stat*, 23, pp.493-507, 1952.
[12] G.Dueck and J.Körner, "Reliability function of a discrete memoryless channel at rates above capacity", *IEEE Trans. Inform. Theory*, IT-25,1, pp.82-85, 1979.
[13] P.Delsarte and P.Piret, "Algebraic constructions of Shannon codes for regular channels", *IEEE Trans. Inform. Theory*, IT-28, pp.593-599, 1982.
[14] P.Elias, "Error-free coding", *IRE Trans. Inform. Theory*, PGIT-4, pp.29-37, 1954.
[15] R.M.Fano, *Transmission of Information*, MIT Press, Cambridge, Mass., 1961, (宇田川訳, 「情報理論」, 紀伊国屋書店, 昭 40).
[16] A.Feinstein, "A new basic theorem of information theory", *IRE Trans. Inform. Theory*, PGIT-4, pp.2-22, 1954.
[17] G.D.Forney,Jr., *Concatenated Codes*, MIT Press, Cambridge, MA., 1966.
[18] G.D.Forney,Jr., "Exponential error bounds for erasure, list, and decision feedback schemes", *IEEE Trans. Inform. Theory*, IT-14, pp.206-220, 1968.
[19] R.G.Gallager, "A simple derivation of the coding theorem and some applications", *IEEE Trans. Inform. Theory*, IT-11,1, pp.3-18, 1965.
[20] R.G.Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.
[21] E.N.Gilbert, "A comparison of signalling alphabets," *Bell Syst. Tech. J.* 31, pp.504-522, 1952.
[22] 韓 太舜, 情報理論における情報スペクトル的方法, 培風館, 1998.
[23] S.Hirasawa, M.Kasahara, Y.Sugiyama and T.Namekawa, "Certain generalizations of concatenated codes—Exponential error bounds and decoding complexity", *IEEE Trans. Inform. Theory*, IT-26, pp.527-534, 1980.
[24] S.Hirasawa, M.Kasahara, Y.Sugiyama and T.Namekawa, "An improvement of error exponents at low rates for the generalized version of concatenated codes",

IEEE Trans. Inform. Theory, IT-27, pp.350-352, 1981.
[25] R.W.Hamming, "Error detecting and error correcting codes", *Bell Syst. Tech. J.* 29, pp.147-160, 1950.
[26] 平澤茂一, "積符号と連接符号", 信学誌, 69,12, pp.1231-1239, 昭 61.
[27] J.Justesen, "A class of constructive asymptotically good algebraic code", *IEEE Trans. Inform. Theory*, IT-18, pp.652-656, 1972.
[28] J.R.Lesh, *Computational algorithms for coding bound exponents*, Ph. D. dissertation, Univ. California, Los Angeles, 1976.
[29] T.Nishijima, H.Inazumi and S.Hirasawa, "A further improvement of the performance for the original iterated codes", *Trans. IEICE*, vol.E72, no.2, pp.104-110, 1989.
[30] 西島利尚, 平澤茂一, "積符号と繰返し符号の漸近的な能力の比較", 信学論 A, vol.J77-A, no.5, pp.786-793, 1994.
[31] J.P.Odenwalder, "*Optinal Decoding of Convolutional Codes*", Ph. D. Dissertation,University of Califorinia, Los Angeles, 1970.
[32] J.K.Omura : "Expurgated bounds,Bhattacharyya distance and rate distortion functions", *Inform. Contr.* 24, pp.358-383, 1974.
[33] W.W.Peterson and E.J.Weldon,Jr., *Error-Correcting Codes*, 2nd ed., MIT Press, Cambridge, MA., 1972.
[34] W.W.Peterson, *Error-Correcting Codes*, MIT Press, Cambridge, MA., 1961.
[35] J.N.Pierce, "Limit distributions of the minimum distance of random linear codes", *IEEE Trans. Inform. Theory*, IT-13, pp.595-600, 1967.
[36] M.Plotkin, "Binary codes with specified minimum distance", *IRE Trans*, IT-6, pp445-450, 1960. Also Research Divesion Report 51-20,University of Pennsylvania, 1951.
[37] C.E.Shannon, R.G.Gallager and E.R.Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels," *Inform. Contr.*10,pp.65-108(Part I) and pp.522-552(Part II), 1967.
[38] C.E.Shannon, "A mathematical theory of communication", *Bell Syst. Tech. J.* 27, pp.379-423 (Part I) and pp.623-656 (Part II), 1948.
[39] Y.Sugiyama, M.Kasahara, S.Hirasawa and T.Namekawa, "A new class of asymptotically good codes beyond the Zyablov bound," *IEEE Trans. Inform. Theory,* vol IT-24, pp.198-204, 1978.
[40] C.Thomessen, "Error-correcting capabilities of concatenated codes with MDS outer codes on memoryless channels with maximum-likehood decoding", *IEEE Trans. Inform. Theory*, IT-33, pp.632-640, 1987.
[41] 植松友彦, 現代シャノン理論, 培風館, 1998.
[42] 植松友彦, 水野亮, S.Noppanakeepong, "優れた信頼度関数を有する符号の代数的構成法と光通信への応用", 信学論 (A), J77-A, 11, pp.1537–1545, 1994.
[43] A.J.Viterbi and J.K.Omura, *Principles of Digital Communication and Coding*,

McGraw-Hill, New York, 1979.
[44] S.Verdu and T.Han, "A general formura for channel capacity," *IEEE Trans. Inform. Theory*, vol.IT-40,pp.1147-1157, 1994.
[45] A.J.Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm", *IEEE Trans. Inform. Theory*, IT-13, pp.260-269, 1967.
[46] J.Wolforwitz, "The coding of messages subject to chance errors", *Ill. J. Math*, 1, pp.591-606, 1957.
[47] V.V.Zyablov, "On estimation of complexity of construction of binary linear concatenated codes", *Probl. Peredach. Inform*, 7, pp.5-13, 1971.

1.2 誤り訂正符号化技術の基礎

1.2.1 はじめに

誤り訂正符号化とは，伝送しようとする情報に対して，元来表現可能なメッセージ最よりも大きなサイズの符号語への変換を行い，冗長度を付加する操作をさす．この冗長性をうまく利用することにより，伝送路上や記録媒体上などで生じた誤りを自動的に検出し，訂正を行うことが可能となる．ここではごく簡単に誤り訂正符号を理解する上での基礎事項を述べる．なお本稿ではブロック符号に対象を限定する．

ブロック符号は符号長を n とすると $n-$ タプルのベクトルで表すことができる．あるアルファベット $\mathcal{A}$ 上の符号長 n を有する符号語数 M のブロック符号 C を $[n, M]$ 符号とよぶ．伝送するメッセージは $\mathcal{A}$ 上の k 次元ベクトルで表される．$R = k/n$ を符号化率とよぶ．符号の特性を理解するための基本は，符号語間のハミング距離である．

定義 1.1
二つの符号語 $\boldsymbol{x}, \boldsymbol{y}$ 間のハミング（Hamming）距離 $d(\boldsymbol{x}, \boldsymbol{y})$ とは，ベクトル $\boldsymbol{x}, \boldsymbol{y}$ の互いに対応する要素すべてを比較したとき，一致しない要素の数である．ハミング距離は，次の距離の公理を満たす．

(1) $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$, 等号は $\boldsymbol{x} = \boldsymbol{y}$ のときのみ成り立つ．
(2) $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$.
(3) $d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z}) \geq d(\boldsymbol{x}, \boldsymbol{z})$.

符号語間のハミング距離が決まれば，その符号に含まれる全ての符号語間の距離の最小値を求めることにより，符号 C の持つハミング距離 d が定まる．すなわち，

$$d \triangleq \min_{\substack{x,y \in C \\ x \neq y}} \{d(x, y)\} \tag{1.70}$$

誤り訂正符号の訂正限界は，次のようになる．

定理 1.11
C を距離 $d = 2e + 1$ を有する $[n, M]$ 符号とする．このとき C は e 個までのどのような誤りも訂正することができる．誤り検出のみに用いるならば，C は $2e$ 個までのどのような誤りも検出できる．

1.2.2 線形符号

次に線形符号について述べる．メッセージが，q 個の要素からなるある体 F 上の $k-$ タプルのベクトル空間 $V_k(F)$ を構成するものとする．これをメッセージ空間 $\mathcal{M}$ とよぶ．メッセージは q^k 個存在し，その要素は，$\boldsymbol{m} = (m_1, m_2, \ldots, m_k)$, $m_i \in F$ なる k 次元ベクトルで表わされる．

誤り訂正能力を有するためには，符号語には何らかの冗長性が含まれている必要がある．したがって符号長 n を有する誤り訂正符号化は，メッセージ空間 $V_k(F)$ を $n > k$ なる n 次元ベクトル空間 $V_n(F)$ の部分集合に写像する必要がある．そこで符号長 n，メッセージ長 k を持つ符号を $(n, k)-$ 符号と記すとき，線形符号 C を次のように定義する．

定義 1.2
体 F 上の $(n, k)-$ 線形符号 C とは，n 次元ベクトル空間 $V_n(F)$ の k 次元部分ベクトル空間である．

体 F 上の線形符号 C は，次の性質を持つ．

(1) $\mathbf{0} \in C$,
(2) C の任意の符号語 $\boldsymbol{c}_1, \boldsymbol{c}_2 \in C$ に対し，

$$a_1\boldsymbol{c}_1 + a_2\boldsymbol{c}_2 \in C, \text{ ただし } a_1, a_2 \in F. \tag{1.71}$$

ここでメッセージのベクトル表現と符号語のベクトル表現の間には，まだ対応が定められていない．そこで，符号化の方法としては，n 次元ベクトル空間