チュートリアル

量子力学と情報処理

丸山 耕司

量子力学を適用する必要のないマクロな世界の力学を古典力学とよぶように，量子情報に相対する通常の情報は古典情報とよばれる．われわれが現在持っている実用的な計算機のすべては，マクロな物理的状態によって表現された古典情報を処理している．マクロな物理状態とは，膨大な数の電子による電荷の有無とか，巨大な磁気モーメントを持つ磁石の向きとか，とにかく量子力学が典型的に相手にするような一個の電子・光子や，電子一個がもつスピン磁気モーメントなどと比べてケタ違いに大きい物理的実体のことである．“古典的”計算機は，巨大な物理的違いをもつ“0”と“1”を用い，ANDやORなどの論理演算を行い，計算結果を出力する．これに対し，量子計算機では，“0”と“1”を，例えば電子ひとつの上向きスピン状態$|\uparrow\rangle$と下向きスピン状態$|\downarrow\rangle$などに割り振って，(サイズ的には)チマチマと論理演算を繰り返していく．さらに，0と1だけでなく，それらの量子力学的重ね合わせをも情報として扱って，量子力学の法則に従って系を制御(＝処理)していくわけである．たったこれだけのことなのだが，周知のように量子力学的世界では古典力学での常識・直感が通用しない現象が起こり，これが量子情報の世界を実におもしろくしてくれている．

“量子情報の世界”などと書くと大げさに聞こえるが，情報処理という側面でとらえれば，応用量子力学とでもよぶのがふさわしいような気もする．しかし，一方，量子情報と古典情報の違いを深く考察することは，量子力学の基本原理を深く掘り下げることにもなり，一概に基礎だの応用だのとくくることはできず，非常に広い領域をカバーすることになる．量子力学という巨大な理論体系を情報という視点で見つめる，というのが量子情報の世界だと思っている．

そこで本チュートリアルでは，できるだけ“情報”を意識しながら，量子情報理論の基本的なイメージをお伝えできるように努力したい．量子情報・計算はすでに巨大な研究領域となっており，その中のすべての重要トピックをここで網羅することは到底できない．よって，内容をぐっと絞らざるを得ないが，情報と物理の密接な関連の雰囲気を感じていただくために，まずは情報理論と(古典)物理学の深い縁の話から始め，この縁を量子力学の領域までひきずりこむ．それから，量子計算の理論と実装について，目指すべきものと現状について概観し，さらに情報処理において最大の問題となる“エラー”への対処法(量子誤り訂正符号)などにも触れる予定である．

ところで情報とは何か．ここでは簡単に，情報とは，その受け手にとって知ることが意味や価値のあること，驚きのあること，としておこう．宝くじの当選番号，未来の株価，地球周辺の小惑星の軌道……．どれも確実に知ることが何らかの意味を持っている．また，情報とは最終的には我々人間が認識して，初めて意味があるものだということも，頭の片隅に意識しておいてもらいたい．

まるやま こうじ．(独)理化学研究所基幹研究所．

まずはひとつ，手っ取り早く古典情報と量子情報の大きな違いを示す重要な例を示そう．no-cloning 定理[1]とよばれるものだ．我々が日常経験しているように，古典情報は何度でもコピーが可能だ．もし可能でなければ，コンピュータ内のファイルのバックアップもとれないし，デジタル放送のコピー可能回数でもめることもない．しかし，量子情報ではそうはいかない．今，ある量子状態 $|\phi\rangle$ が与えられたとして，これと同じ状態をもうひとつ作りたい(つまりコピーしたい)としよう．紙情報をコピーするときはコピー先として白紙が必要であるように，この場合もある基準となる量子“白紙”状態 $|w\rangle$ を準備する．“原稿”状態と“白紙”状態を合わせた複合系は，$|\phi\rangle|w\rangle$ で表せる．孤立量子系の時間発展はユニタリー演算子 U で表されるから，コピーが可能であるとするとその効果は，

$$U(|\phi\rangle|w\rangle) = |\phi\rangle|\phi\rangle \qquad (1)$$

となるはずである．$|\phi\rangle$ とは異なる状態 $|\phi\rangle$ に対しても同様の効果が期待できるはずなので，

$$U(|\phi\rangle|w\rangle) = |\phi\rangle|\phi\rangle \qquad (2)$$

である．これら二式の内積をとると

$$\langle\phi|\phi\rangle = (\langle\phi|\phi\rangle)^2 \qquad (3)$$

となり，この式が成り立つには $|\phi\rangle$ と $|\phi\rangle$ が等しい ($\langle\phi|\phi\rangle=1$) か，直交する ($\langle\phi|\phi\rangle=0$) しかなく，任意の状態に対して U がコピーを作るとした仮定に反する．直交する二状態は確率1で区別することができるから，古典的なビットと同じである．したがって，古典情報であれば量子系でも我々の日常と同じように完全なコピーを作ることができるが，それら二状態の任意の重ね合わせ状態のコピーを作ることはできないのである．ちなみに，観測なども含めたさらに一般的なユニタリーでない制御が加えられても，no-cloning 定理が成り立つことが知られている．

古典的世界ではもっとも基本的な情報処理のひとつである，コピーをとるという作業が，量子系ではできない．この違いは大きい．なにか不便な感じもするが，それを補って余りある様々な物理的帰結の元になっている．ここでその方向への深入りはしないが，直感的に分かりやすい例をあげれば，量子暗号の安全性がある．一言で言えば，非直交状態を送ることで暗号通信を行う(秘密鍵を共有する)のが量子暗号プロトコルだが，非直交状態を完全にコピーできるとすると完全な盗聴が可能ということになってしまう．no-cloning 定理のおかげで量子暗号の安全性が保障されているのである†1．

さて，情報に着目して話を進める上では，当然情報を定量的に扱うことが必要である．情報を定量化するにはどうすればよいか．ナイーブに考えて，情報量とは，ある情報を得たときの有益さのようなものとするのが妥当だろう．ほとんど天気が変化しない砂漠のような地域での天気予報より，都市部での局地的な豪雨の予報などの方が，情報として有益だろうから情報量も大きいとするわけである．この有益さの数学的違い(経済的な有益さではない！)を生むのは，各事象の確率である．珍しいことが明日起こるという予報の情報量は大きいし，いつも起こっていることが明日も起こるという予報の情報量は小さい．ある情報の“量”は，その情報がもたらされる確率の単調減少(連続)関数と見てよさそうだ．

さらに，複数の情報が与えられるときの全情報量はどうなるだろうか．A, B というふたつの独立な情報がもつ全情報量は当然，これらふたつの情報が得られる確率 $p(A, B)$ の関数だ．また，全情報量は，(A と B が独立であれば)A がもつ情報量と B がもつ情報量の和であるとするのは自然な要請だろう．つまり，情報量を I で表せば，$I(A, B)=I(A)+I(B)$ となる．一方，確率については，$p(A, B)=p(A)p(B)$ であるので，I の関数形としては積を和に変えるようなものが相

†1 厳密には，量子暗号の無条件での安全性を証明するには no-cloning 定理だけでは不十分である．盗聴者が量子力学の許す範囲内で(量子コンピューターも含めて)無限の情報処理リソースを持っていたとしても，有限の情報量が盗聴者へ漏れることがないことを示さなければならない．

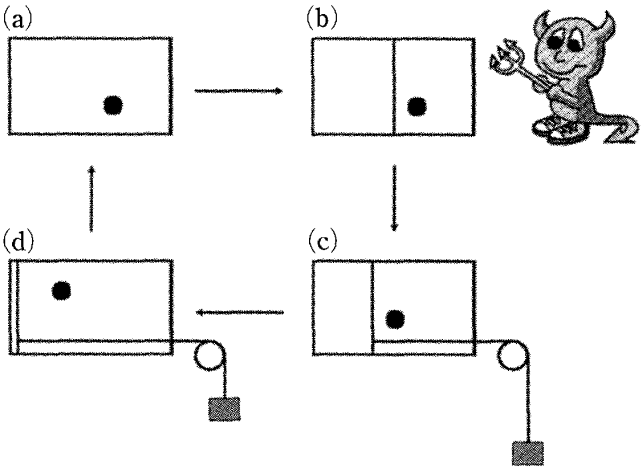

図1 Szilard が考案した一分子気体エンジンによる思考実験
分子の位置情報を得ることにより，熱力学第2法則を破る(ようにみえる)．

応しいことになる．これにぴったりあてはまるものとして，log があることがすぐに分かる(実際，これが上の要請下の唯一の関数形である)．そこで，確率 p で現れる情報のもつ情報量は，$I=-\log_2 p$ と定義しよう．対数の底の2は，最後まで残っていた定数係数分の不定性をなくしつつ，もっとも基本的な二値情報を基準とするためである．I は発案者の名前から“Shannon エントロピー”，あるいは単に“エントロピー”とよばれ，単位はビットである．

ある確率変数 X をひとつずつ独立に生成する情報源 $\mathcal{X}$ を考えよう．確率変数 X が有限個の $x_i(i\in\{1,2,\cdots,N\})$ からなり，各々の x_i が現れる確率が常に p_i で表されるとき，各 x_i がもつエントロピー $-\log p_i$ の平均

$$H(X):=-\sum_{i=1}^{N} p_i \log p_i \qquad (4)$$

を，この情報源のもつ(Shannon)エントロピーとよぶことにする．これを $H(p_i)$ とか $H(\{p_i\})$ などと表記することもある．ここで，ある j について $p_j=0$ となる場合は，$0\log 0$ が現れるが，これは極限の値を採用してゼロと定義する．二値情報の場合，どちらかの確率が1でもう片方が0の場合は $H=0$，どちらの情報(確率変数)が現れる確率も1/2の場合は $H=1$ となって最大値をとる．これは先の“情報の有益さ”の議論とも合致する．エントロピー H が大きいほど，次に現れる情報についての不確定さが大きく，その情報を得たときの価値が高いことになる．

物理の話に戻ろう．物理と情報が明確に結びついたもっとも歴史ある例は，Maxwell の悪魔のパラドックス[2]だろう．19世紀半ばに Maxwell は，もし個々の気体分子の状態を観測できる知的存在(悪魔)がいれば，温度一定の気体を高温部と低温部に分離することができ，第2法則が破られる可能性を提示した．このパラドックスを，特に情報(量)と関連づけてとらえて解決を試みたのは，20世紀初めの Szilard(“シラード”が近い発音らしい)である．彼は，**図1**のような分子がひとつだけからなる“気体”の入った体積 V の箱を考えた．初め，図で(a)の状態では誰も分子が箱中のどこでどう運動しているかは分からない．ここで，薄い壁をさっと挿入し，箱の体積を二等分する．観測者(=悪魔)は，分子の運動状態を変えることなく，分子が壁の右側にあるか左側にあるかだけを測定する(**図1**(b))．ここでは右側にあったとしよう．観測者は挿入した壁の右側におもりを結びつけ，箱を温度 T の熱浴に接触させ，この一分子気体を体積 V まで等温膨張させる．この膨張過程(**図1**(b)→(c)→(d))で，気体は熱浴から熱量 Q を受け取り，おもりに対し，

$$W=kT\int_{\frac{V}{2}}^{V} V^{-1}dV=kT\ln 2 \qquad (5)$$

だけの仕事を行う(kはBoltzmann定数)[†2]．等温過程であるから，$W=Q$である．気体は再び体積Vを占める状態になるわけだから，もとの**図1**(a)に戻った，つまり熱機関としてサイクルが閉じたことになる．ところが，この熱機関は熱浴から熱量Qを力学的仕事Wに完全に変換しており，熱浴のエントロピーは$Q/T=k\ln 2$だけ下がっている．気体の状態に変化はないから，明らかに熱力学第2法則を破っている．これがSzilardエンジンによるMaxwellの悪魔の表現である．

Szilardはパラドックスをここまで単純化し，第2法則が破られないために必要であろう仕事Wの消費理由(エントロピー$k\ln 2$の増大理由)を考察した．彼は，**図1**(b)で悪魔が分子の位置を測定することによる情報取得が，エントロピー増大を伴うと考え，$k\ln 2$というエントロピー変化量をこの情報取得に関連した基本的量であるとした[3]．定数係数は別として，これはまさに先に導入した1ビットの情報量に対応する量である．Shannonによる情報理論確立の20余年も前の，物理の中での情報の役割がはっきりと認識された最初の例である．

さて，悪魔のパラドックスについてはこの後，Brillouin[†3]がSzilardのアイデアをさらに推し進め，分子運動の観測に光を用いた場合の，光の散乱による系のエントロピー変化を計算し，これがパラドックス中のエントロピー減少を補うのに十分な量であることを示した[4]．これで悪魔は退治されたと信じられた期間が10年ほど続いたのだが，実はまだ健在であったことがIBMのLandauerが提示したアイデアによって明らかになった．

†2 分子ひとつだけの気体にPやTなどを含む通常の状態方程式が適用できるのか，という疑問がわくかもしれない．しかし，Szilardエンジンの議論については，これを正当化できる．それを説明するには，残念ながら紙面が足りない……．

†3 固体物理でお馴染みのブリルアン・ゾーンのブリルアンである．

Landauerは，情報処理という作業を真っ向から物理的視点でとらえた．キーとなる事実は，すべての情報は何らかの物理的実体を媒体として保持・伝達され，すべての情報処理はそれらに直接(物理的に)働きかけることで実行される，ということである．媒体として使われるのは，紙，フィルム，電荷，等々，実に様々だが，どれも立派な物理的なモノである．ということは，互いに異なる情報は，常に互いに異なる物理状態に対応していなければならない．そうでなければ，異なる情報を表すはずの複数の物理的状態が，受信者や計算機にとっては区別できず，情報としての意味をなさないからだ．さらに，情報が常に媒体の物理的状態に対応している以上，情報処理はその媒体に物理的に働きかけて状態を変化させる，物理的な過程であるはずである．

ところで，情報処理とは，与えられた入力に対して，ある情報を出力する写像(の連続)である．写像fとしてのある情報処理が，単射である場合(例えば，NOTゲート＝ビット反転処理)，つまり入力xと出力$y=f(x)$が1対1に対応する場合は，逆写像f^{-1}によって出力yから元の入力$x=f^{-1}(y)$を一意に得ることができる．言い換えれば，ある情報処理が単射写像であるときは，その処理は“可逆”である．上で述べたように，情報処理を物理的過程ととらえれば，可逆な情報処理に対応する物理的過程はやはり可逆であることになる．もちろん，我々が日常目にする現実世界では，可逆な物理過程などというものはほとんどあり得ない．しかし，ここで我々が扱っているのは，情報と物理の本質的・原理的な関連であるので，理想化された場合のみを考え，可逆な情報処理は可逆な物理的過程によって実行される，という言い方をする．

次に，入力と出力が1対1に対応しない情報処理を考えよう．入力としてありうる値の個数よりも出力の値の個数の方が小さい場合である．2ビットの入力から1ビットの出力を得る処理はすべてこれに相当するが，1ビット処理でも，0か1

かにかかわらず出力を0にしてしまうような処理は2対1の対応であり，単射ではない．これを情報の消去とよぶ．

ここでひとつ注釈しておく．情報を消去するのに，わざわざ0にせずとも，そのビットを放っておいて二度と使わなければいいではないか，と思われるかもしれない．しかし，熱力学第2法則を論じる以上，熱力学的サイクルを一度閉じて，その上で得られる正味の仕事量や全系のエントロピーの増減を評価しなければならない．サイクルを閉じる，つまり初期状態に戻さなければならないものの中には，当然情報を記憶する領域(メモリー，あるいはレジスター)も含まれる．したがって，再利用するにあたっては，以前に0であったか1であったかにかかわらず，一度どちらか(ここでは0とする)にリセットしておく必要がある．これが情報の消去であり，まさにMaxwellの悪魔のパラドックスを解決するポイントとなるものである．

上述したように，消去は単射写像ではなく，1ビット情報の場合は2対1の写像である．よって，情報処理を請け負う物理系について考えると，消去によって許される自由度が減り，エントロピーが減少，その結果，自由エネルギーが増大する．これは，正のエネルギーの消費が必要であること，その分は環境に散逸してしまうことを意味する．1ビットの情報消去の場合は，エネルギー散逸量は $kT\ln 2$(エントロピー $k\ln 2$)となる．Landauerはこれを一般化し，不可逆な情報処理はすべて物理的には散逸を伴うと主張した[5]．この主張は，Landauerの消去原理とよばれ，今日では広く受け入れられている．逆に，可逆な処理だけで全体の情報処理を構成することができれば，散逸ゼロで計算が実行できることになる．時間発展がユニタリー演算子で与えられる量子計算(の観測を含まない部分)はこれに相当する．

さて，同じくIBMのBennettはLandauerの考察からもう一歩踏み込んで，観測による情報取得は原理的には無限小の散逸を伴うだけで可能であることを示した[6]．これにより，Szilardエンジンで記述したMaxwellの悪魔のパラドックスにおいて，散逸を伴うのは悪魔による観測行為ではなく，悪魔の頭の中の記憶を消去するときであることが確定することになる．消去に必要なエネルギー消費がエンジンから得られるエネルギーを相殺するので，パラドックスは解決する．

ところで，ここに述べた論理展開だけだと，単射でない情報処理に物理的散逸が伴う理由は熱力学第2法則そのものであり，その帰結として第2法則が破られないのは当然である．しかし，1ビットの情報消去に少なくとも $k\ln 2$ のエントロピー増加が必要である[†4]ことは，第2法則を用いなくても導くことができる．その詳細はチュートリアルとしての本稿の範囲を超えるので，興味ある読者は文献[7][8]やレビュー論文[9]等を参照されたい．

では，今度は量子系を(古典)情報の記録・処理媒体として考えよう．量子系に情報を持たせる(エンコードする)とは，情報 $\{i_1, i_2, \cdots, i_N\}$ を，量子状態 $\{\rho_1, \rho_2, \cdots, \rho_N\}$ にそれぞれ対応させることである．各 ρ_n は量子状態を記述する密度演算子である．各々の情報 i_n が現れる確率を p_n とし，量子状態 ρ_n が確率 p_n で現れる系を，$\{p_n, \rho_n\}$ と表記しよう．受け手側からすれば，観測する前には確率分布 $\{p_n, \rho_n\}$ しか分からないから，どのメッセージも ρ_n を“平均”した量子状態

$$\rho = \sum_{i=1}^{N} p_i \rho_i \qquad (6)$$

にあると考えることができる．状態 ρ に保持された情報を消去するのには，どれだけのエントロピー増加が必要であろうか．各 ρ_i がすべて純粋状態 $\rho_i = |\phi_i\rangle\langle\phi_i|$ の場合は簡単で，$k\ln 2S(\rho)$ となる(証明は次回！)．ここで $S(\rho)$ は状態 ρ のvon Neumannエントロピーで，次式で定義される：

†4 ちなみに，$H(p)$ ビットの情報を消去する場合のエントロピー増加は $k\ln 2H(p)$ 以上である．なぜだか考えてみよう．

$$S(\rho) := -\mathrm{Tr}\,(\rho \log_2 \rho). \qquad (7)$$

演算子の log がとっつきにくい印象を与えるかもしれないが，これは ρ を対角化して，その対角要素が $\{\lambda_1, \lambda_2, \cdots, \lambda_d\}$ だったとすると，$S(\rho) = -\sum \lambda_i \log_2 \lambda_i$ で計算される．一般に，$\rho = \sum p_i \rho_i$ が与えられたとき，$S(\rho) \leq H(p_i)$ が成り立ち[†5]，量子状態 $\{p_i, \rho_i\}$ にエンコードされた情報を消去する方が，同じ確率分布を持つ古典的状態にエンコードされた情報を消去するよりも，エントロピー増加が小さい．つまり，ρ には $H(p_i)$ ビットよりも小さい情報量しか持たせられない，ということになる．これは，異なる i と j に対して ρ_i と ρ_j が直交しない場合，それらを観測によって完全には区別できないという重要な事実に由来する．区別できない→観測しても得られる情報が少ない→エントロピーが小さい，となるわけである．量子系を観測することで得られる情報量については，次回で少し詳しく触れる．

ところで，von Neumann エントロピーの式(7)は，Shannon による情報理論的エントロピーの式(4)によく似ているが，(7)の定義の動機は熱力学であって，情報理論ではない．von Neumann は，異なる量子状態をもつ気体分子を異なる種類の分子とみなしたときの混合気体のもつエントロピーが，熱力学的エントロピーと等しくなるように量子系のエントロピー(7)を定義した．実際，von Neumann エントロピーは Shannon エントロピーよりも 10 年以上早く発表されている．

今回は，情報と物理を本質的に結びつける Maxwell の悪魔のパラドックスから，情報の定量化，古典および量子系中の情報消去について述べた．ちなみに，エントロピー概念の由来の違いなどのため，消去原理に関しては未だに議論の的となることがある．しかし，反論が出る以上に正当性を裏付ける(理論的)結果が多く発表されており，これを否定する合理的理由は今のところない．とは言え，健全な批判的精神をもって既存の理論体系を精査することは，(哲学的)理解を深めるのに大いに役に立つので，「あれ？」と思う部分はとことん調べたり考えたりしていただきたい．さて，結果だけを見るととても単純な“消去原理”だが，これだけ(あるいは類似の論理)で量子情報理論におけるいくつかの非常に興味深い事実を導くことができる．次回は，そのうちの量子データ圧縮と量子系から得られる情報量を中心に話を進める．

参考文献

[1] Wootters, W. K. and Zurek, W. H., Nature, 299 (1982), 802.
[2] Leff, H. S. and Rex, A. F., Mexwell's Demon 2, Institute of Physics Publishing, Bristol and Philadelphia, 2003.
[3] Szilard, L., Z. Phys., 53(1929), 840. 文献[2]に英訳が収録されている．
[4] Brillouin, L., J. Appl. Phys., 22(1951), 334. 文献[2]に収録．
[5] Landauer, R., IBM J. Res. Dev., 5(1961), 183. 文献[2]に収録．
[6] Bennett, C. H., Int. J. Theor. Phys., 21(1982), 905. 文献[2]に収録．
[7] Shizume, K., Phys. Rev. E, 52(1995), 3495.
[8] Piechocinska, B., Phys. Rev. A, 61(2000), 062314.
[9] Maruyama, K., Nori, F. and Vedral, V., Rev. Mod. Phys., 81(2009), 1.

†5 等号が成り立つのは，ρ_i が互いに直交する，つまり $i \neq j$ に対して $\mathrm{Tr}[\rho_i \rho_j] = 0$ となるときだけである．