大偏差原理[1]

稲濱　讓[2]・角田　謙吉[3]

[1]2025 年 10 月 3 日・改訂第 2 版
[2]九州大学大学院数理学研究院
[3]九州大学大学院数理学研究院

まえがき

確率論において最も基本的な問題は確率変数の列に対する極限定理であろう．独立確率変数の経験平均に対する大数の法則は日常生活における経験則としてよく知られている．また独立確率変数の列に対する中心極限定理も，学力試験における偏差値を通して我々の生活に深く浸透している．経験則としての大数の法則と中心極限定理だけでなく，昨今ではこれらの極限定理は確率・統計の分野における数学的基礎知識としてほぼ全ての理系分野の人に定着している．しかしながらこれらの極限定理に比べて，本書で扱う大偏差原理は確率論の中心的話題であるにも関わらず，確率論の業界外の人にとっては馴染みのない話題であるように思われる．さらに確率論の専門家に限っても，大偏差原理が幾分独特な問題を扱うためか，この話題に精通していない人も見受けられる．このような現状を僅かでも改善したいという著者の期待を込めて本書を執筆した．大偏差原理の一般論と典型例について解説することが本書の目的であり，本書を通じて大偏差原理に関する知識が普及することを願う．

現代の確率論業界では大偏差原理は極めて盛んに研究されており，その程度はおそらく業界外の人の想像を遥かに超える．解析学を学んだ者なら誰でも知っているように，測度論（Lebesgue 積分論）では「収束定理三羽烏」として単調収束定理・Fatou の補題・Lebesgue の収束定理（優収束定理）が非常に重要視されている．この伝で言えば，現代確率論の「極限定理三羽烏」は大数の法則・中心極限定理・大偏差原理であろう．誇張して言えば，現在の確率論業界の少なからぬ部分においては，多種多様な確率模型に対してこれら三種の極限定理が成立するかどうかを調べている状況である．これだけでも大偏差原理の重要性がわかるが，実は確率論業界に身を置くものの「生活実感」としては，大偏差原理は他の二種の極限定理に比べてより盛んに研究されていると感じる．これは業界外では大数の法則と中心極限定理のほうが，大偏差原理に比べて圧倒的に有名であることと好対照をなす．数学論文の検索エンジンとして世界で最も利用されている MathSciNet の「題名検索」機能を用いて，このことを少し定量的に見てみよう（2025 年 1 月現在）．

- “law of large numbers” 2,039 篇．
- “central limit theorem” 3,589 篇．
- “large deviations” 4,324 篇（“large deviation principle” 514 篇）．

もちろん題名検索だけでそれほど正確なことがわかるはずはないが，それでもこのデータは大偏差原理の重要性を十分証明しているであろう．現在では大偏差原理は確率論の枠を越えて様々な分野で研究されている．例としては，統計物理学・情報理論・数理統計学・力学系・作用素代数などが挙げられる．さらに「現代大偏差原理の父」とでも呼ぶべき S.R.S. Varadhan は 2007 年に数学界最高の賞とされる Abel 賞を受賞したが，そ

の最大の理由として大偏差原理の理論を確立したことが挙げられている．当然ながら大偏差原理に関する英文書は既に多数出版されている.[1]

このような状況にもかかわらず，残念ながら大偏差原理に関する日本語文献は非常に少ない．試しに偶然手近にあった某数学辞典をめくってみると，大数の法則と中心極限定理についての項目は立っているが，大偏差原理に関する項目は見当たらない．また大偏差原理に関する和書は現時点では出版されていない．日本の確率論および周辺分野の将来を考えると，これは憂うべき状況と言ってよいであろう．本書を執筆した動機はまさにこの点を改善するためであり，大偏差原理や関連する話題に興味を持つ読者にとって本書が少しでも役に立つことを期待している．さらに本書の読者の中からそのような分野の研究を志す人が現れれば，それは著者にとって望外の喜びである．

数式を少しだけ用いて，大偏差原理で扱う問題を簡単に紹介する．$X_1, X_2, \ldots$ を独立同分布な実数値確率変数の列として，X_1 は平均 0，分散 1 の正規分布に従うとする．このとき各自然数 n に対して経験平均を

$$Z_n = \frac{X_1 + \cdots + X_n}{n}$$

とおくと，Z_n は平均 0，分散 $1/n$ の正規分布に従う．よって任意の正数 δ に対して

$$\mathbb{P}(Z_n \geq \delta) = \left(\frac{n}{2\pi}\right)^{1/2} \int_{\delta}^{\infty} \exp\left(-\frac{nx^2}{2}\right) dx$$

が成立する．簡単な計算により $n \to \infty$ のとき右辺は 0 に収束することがわかるので，左辺の確率も 0 に収束する.（これは大数の弱法則の特別な場合である.）大偏差原理ではこの確率が 0 に収束する “指数的な速度” を特定する．さらに次が成立することも簡単にわかる．

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq \delta) = -\frac{\delta^2}{2}. \tag{$\spadesuit$}$$

大偏差原理で扱うのはこの種の問題である．

実は ($\spadesuit$) の形で定式化すると，一般化に際して問題が生じる．第一に，一般的な確率変数（確率測度）や事象の確率を扱う場合には極限の存在は保証できない．しかし極限の形で述べられなくとも，上下からの評価のみ得られる場合が多い．第二に，この定式化は１次元（もしくは有限次元）の場合に限定されている．しかし確率論で典型的に扱いたいのは，Wiener 測度のような無限次元空間上の確率測度である．例えばこのような理由により，($\spadesuit$) の左辺にあるような極限を計算したり証明することは可能な場合もあるが，問題がある場合のほうが多い．

---

[1]著者が知っているだけでも [29, 31, 32, 33, 34, 35, 36, 41, 42, 43, 46, 47] などがある．それ以外にも多くあるだろう．

現在では大偏差原理とは Varadhan による定式化を指す．歴史的には，(♠) に現れる極限や確率微分方程式に関する Freidlin-Wentzell の一連の研究などが存在する中で，Varadhan による定式化に最終的に至った．Varadhan は M.D. Donsker との一連の研究の中で，Laplace 原理を無限次元空間上で展開する必要があり，そのため (♠) のような素朴な形ではなく一般性の高い定式化を与えた．その定式化は確率論全般において非常に有効であることが判明し，大数の法則や中心極限定理と同様に多くの確率模型において現れる普遍的な現象であることが次第に認識された．このことは先に見たように “large deviations” と題した論文数の多さからも明らかであろう．

本書が主に想定しているのは測度論的な確率論をある程度知っているが，大偏差原理についてはそれほど詳しくない読者である．大偏差原理には確率論の他の話題には現れない独特な議論が多いので，そのような読者のために本書では冗長に見える危険を冒しても，一般的すぎない設定の下でなるべく丁寧に説明をするように心がけた．また大偏差原理を実感を持って理解するためには，有名な確率模型に対する大偏差原理の具体例をいくつか学ぶことが欠かせないが，その際に難しいのは大偏差原理そのものだけではなく，当該の確率模型に関する基礎知識を学習する点にもある．（例えば第 10 章で論ずる Freidlin-Wentzell の大偏差原理の場合は，後者に相当するのは確率微分方程式に関する基本事項である．）二つ目の負担を減らすために，本書では思いきって非常に有名な具体例のみを紹介することにした．その代わりに各具体例に対して複数の証明を与えることにより，読者が個別模型を学ぶことに煩らわされる時間を減らし，大偏差原理そのものの「感覚」を学ぶことに集中しやすくしたつもりである．

本書を読むために必要な予備知識として，解析系分野を専攻する数学科の学生が学部三年次までに学ぶ内容以外に次を仮定する．

- 測度論的な確率論に関する基礎知識．これに関する教科書は多く出版されているので，忘れた事項などあれば必要に応じて参照せよ．
- 可分完備な距離空間上の確率測度に関する基本事項．例えば，確率測度の列の弱収束・確率測度の族の緊密性・Prokhorov の定理・Skorokhod の表現定理・正則条件付き確率など．[2]
- Schilder の大偏差原理を論じる第 9 章においては，Brown 運動に関する初歩的な知識を仮定する．これに関する教科書は多く出版されている．
- Freidlin-Wentzell の大偏差原理を論じる第 10 章においては，確率微分方程式に関する基本事項を仮定する．係数行列に関する大域的 Lipschitz 条件の下で確率微分方程式に一意的な強解が存在するが，この大定理までの内容を知っていれば充分である．確率微分方程式に関する教科書は現在ではかなり多くあるので，必要に応じて参照せよ．

[2] これら全てを（証明を含めて）手際よくまとめている和書は残念ながら思いあたらない．洋書であるが [39, Chapter I] や [40] などによくまとまっている．

本書の構成を説明する．本書は大偏差原理の一般論を紹介する第 1 章と大偏差原理の典型例を解説する第 2–10 章の二つの部分からなる．一般論の部は大偏差原理に関する必須の教養であり，この話題に興味を持つ全ての読者に関係する．その一方で，個別具体例の間には論理的な関係がほとんどないので，読者は自分が知りたい具体例に関する部分だけを読むことが可能である.（例えば確率微分方程式だけに興味を持つ読者は，第 1 章を読み終えた後に Schilder の大偏差原理と Freidlin-Wentzell の大偏差原理を論ずる第 9–10 章をいきなり読んでも基本的に問題は生じない.）

本書で扱う大偏差原理の具体例とその証明方法を箇条書きする．

- Cramér の大偏差原理.
  - 1 次元の場合（第 2.3 節）．確率測度の変更に基づく方法.
  - 多次元の場合（第 2.4 節）．指数型 Chebyshev の不等式と確率測度の変更に基づく方法.
  - Gärtner-Ellis の大偏差原理の系（注意 3.0.2）.
  - 弱収束法（第 5.5 節）．相対エントロピーの連鎖律と指数型の積率公式に基づく方法.
  - 応用：Curie-Weiss 模型の大偏差原理（第 2.6 節）.
- Gärtner-Ellis の大偏差原理（第 3 章）.
  - 応用：独立同分布な確率変数列に対する中偏差原理（第 3.3 節）.
  - 応用：有限状態 Markov 連鎖に対する大偏差原理（第 3.4 節）.
- Sanov の大偏差原理.
  - Cramér の大偏差原理の一般化に基づく方法（第 4 章）.
  - 弱収束法（第 5.4 節）．相対エントロピーの連鎖律と指数型の積率公式に基づく方法.
  - 情報理論的証明（第 6 章）.
- 対数気体に対する大偏差原理（第 7 章）.
- Mogulskii の大偏差原理（第 8 章）．射影極限法を用いる方法.
- Schilder の大偏差原理（第 9 章）.
  - Brown 運動の見本路の基本性質（特に原点の近傍からの脱出確率の評価）を用いる方法（第 9.4 節）.
  - Fernique の定理を用いる方法（第 9.5 節）.

- Mogulskii の大偏差原理を用いる方法（第 9.6 節）.

- Freidlin-Wentzell の大偏差原理（第 10 章）.
  - 確率微分方程式に対する Euler・丸山近似と大偏差原理に関する指数的に良い近似を組み合わせる方法（第 10.2 節）.
  - 大偏差評価を用いる方法（第 10.3 節）.
  - 弱収束法（第 10.4 節）. 指数型の Wiener 汎函数に対する変分表示公式に基づく方法（第 C 章）.

大偏差原理の具体例は無数にあるので，非常に重要であるが本書では扱わなかった具体例も当然多くある．例えば Donsker-Varadhan の大偏差原理が挙げられる.[3] また簡単なものを除いて，本書では大偏差原理の応用についても扱わなかった．大偏差原理を示すこと自体が重要であることと同様に，個々の確率模型の解析に応用することも当然ながら重要である．代表的な応用例として，Freidlin-Wentzell の大偏差原理の脱出問題への応用や Donsker-Varadhan の大偏差原理を用いた Wiener sausage の研究が挙げられる．しかしその種の応用を紹介するためには各模型の詳細に立ち入る必要があり，それは大偏差原理の基礎事項の解説を目指すという本書の趣旨から外れるため，本書では扱わなかったことをご了承頂きたい.

謝辞

本書の内容について，濱口雄史氏は早い段階で原稿全体を通読し，非常に多くの点について有益な指摘をくださった．針谷祐氏には Boué-Dupuis の変分表示公式について詳しくご教示していただいた．付録第 C 章はそれに基づいている．江崎翔太氏，須田颯氏，高野凌史氏，田中亮吉氏，千野由喜氏，永沼伸顕氏，難波隆弥氏，森隆大氏，渡辺樹氏の各氏には本書の原稿の一部を精読していただき，数々の貴重なご意見をお寄せいただいた．また村山拓也氏，蛯名真久氏，柴田修平氏，崎原雄大氏に九州大学において，桑江一洋氏，藪奥哲史氏，徐梓健氏，谷本浩彩氏，工藤綺星氏に福岡大学においてそれぞれ本書に関するセミナーを行なっていただき，有益なご意見を多くいただいた．これらは本書の改善に大変役立った．貴重なご意見をいただいたことに，著者として改めて深く感謝いたします．一部の内容については九州大学大学院数理学府における著者の講義でも扱った．受講生の皆様にも大変感謝いたします.

なお著者は本書の執筆に際して，JSPS 科研費 JP23K20216 と JP22K13929 の助成を受けた.

[3] Donsker-Varadhan の大偏差原理については例えば [32, 17] を見よ.

目 次

記号表

- $\mathbb{N}=\{1,2,\ldots\}$：自然数の全体
- $\mathbb{N}_0=\{0\}\cup\mathbb{N}$
- $\mathbb{Z}$：整数の全体
- $\mathbb{Q}$：有理数の全体
- $\mathbb{R}$：実数の全体
- $\overline{\mathbb{R}}=\mathbb{R}\cup\{\pm\infty\}$
- $[0,\infty]=[0,\infty)\cup\{\infty\}$
- $\mathbb{C}$：複素数の全体
- $\sup A$：部分集合 $A\subset\overline{\mathbb{R}}$ の上限（ただし $\sup\emptyset=-\infty$）
- $\inf A$：部分集合 $A\subset\overline{\mathbb{R}}$ の下限（ただし $\inf\emptyset=\infty$）
- $\max A$：部分集合 $A\subset\mathbb{R}$ の最大
- $\min A$：部分集合 $A\subset\mathbb{R}$ の最小
- $a\wedge b=\min\{a,b\}$：$a,b\in\overline{\mathbb{R}}$ の最小
- $a\vee b=\max(a,b)$：$a,b\in\overline{\mathbb{R}}$ の最大
- $\lfloor x\rfloor$：$x\in\mathbb{R}$ 以下の最大の整数
- $\lceil x\rceil$：$x\in\mathbb{R}$ 以上の最小の整数
- $\mathbf{1}_A$：部分集合 A の定義関数
- $|A|$：有限集合 A の元の個数
- $\|f\|_\infty$：関数 $f:A\to\mathbb{R}$ の上限ノルム（第 8 章では本質的上限ノルム）
- $\mathcal{B}(S)$：位相空間 S の Borel 加法族
- $\mathcal{A}\vee\mathcal{B}$：集合族 $\mathcal{A}$ と $\mathcal{B}$ を含む最小の σ 加法族
- A°：位相空間における部分集合 A の内部
- $\overline{A}$：位相空間における部分集合 A の閉包

- ∂A：位相空間における部分集合 A の境界
- $B(x,r) = B_S(x,r)$：距離空間 (S,d) における中心 $x \in S$，半径 $r > 0$ の開球

$$B(x,r) = \{y \in S \mid d(x,y) < r\}$$

- $\overline{B}(x,r) = \overline{B}_S(x,r)$：距離空間 (S,d) における中心 $x \in S$，半径 $r > 0$ の閉球

$$\overline{B}(x,r) = \{y \in S \mid d(x,y) \leq r\}$$

- $|\cdot|$：$\mathbb{R}^d$ のノルム
- $\|\cdot\| = \|\cdot\|_X$：無限次元ベクトル空間 X のノルム
- $x \cdot y = \langle x, y \rangle = \langle x, y \rangle_X$：内積空間 X における x と y の内積
- $\{e_1, \ldots, e_d\}$：$\mathbb{R}^d$ の標準基底
- δ_x：点 x における Dirac 測度（点測度）
- $\mathcal{M}_1(S)$：可測空間 S 上の確率測度全体
- $C_{\mathrm{b}}(S)$：位相空間 S 上の有界連続関数全体
- $M_{\mathrm{b}}(S)$：可測空間 S 上の有界可測関数全体
- ∇：$\mathbb{R}^d$ 上の勾配作用素
- δ_{ij}：Kronecker のデルタ記号
- $\mathbb{P}(A)$：事象 A の確率
- $\mathbb{E}[X]$：実数値（ベクトル値）確率変数 X の期待値

一般的な規約表

- 「$A := B$」は A を既知の量 B で定義することを意味する．例えば, $f(x) := x^2$ $(x \in \mathbb{R})$ は新しい関数 f を右辺で定義したことを意味する．

- $\overline{\mathbb{R}}$ には $\mathbb{R}$ の２点コンパクト化としての位相を導入する（$[0,1]$ と同相になる）．$[0,\infty]$ には $\overline{\mathbb{R}}$ からの相対位相を入れる（これも $[0,1]$ と同相になる）．これは $[0,\infty)$ の１点コンパクト化としての位相と同じである．

- ベクトル空間の係数体は明記しない限り実数体 $\mathbb{R}$ とする．

- ベクトル空間の部分集合 A, B と $c \in \mathbb{R}$ に対して，次のように定める．

$$A + B := \{x + y \mid x \in A, y \in B\}, \quad cA := \{cx \mid x \in A\}.$$

- 単に確率変数と言うときは，ある確率空間上で定義されている確率変数を指し，その引数（見本）もしばしば省略する．またその確率空間を明示することが重要でない場合は特に言及しない．

- 確率変数に関する事象を表すときはその引数（見本）を明示しないことが多い．例えば，$\{\omega \in \Omega \mid X(\omega) \geq 0\}$ は単に $\{X \geq 0\}$ と略記する．

- ２つの確率変数 X と Y の法則が等しいことを $X \stackrel{\mathrm{Law}}{=} Y$ と表す．また確率変数 X の法則と確率測度 μ が等しいことを $X \stackrel{\mathrm{Law}}{=} \mu$ と表す．

- $0 \log 0 := 0$ と定める．また $\log 0 := -\infty$，$\log \infty := \infty$ と拡張することにより，対数関数を $[0,\infty]$ 上の関数 $\log : [0,\infty] \to \overline{\mathbb{R}}$ と見なす．

- $\overline{\mathbb{R}}$ 値の数列 $\{a_n\}_{n \in \mathbb{N}}$ と $\alpha \in \overline{\mathbb{R}}$ に対して，極限 $\lim_{n\to\infty} a_n = \alpha$ に対して，$\overline{\lim}_{n\to\infty} a_n \leq \alpha$ を上からの評価，$\underline{\lim}_{n\to\infty} a_n \geq \alpha$ を下からの評価と呼ぶ．正数 ε で添字づけられた族 $\{a_\varepsilon\}_{\varepsilon > 0}$ に対して，極限 $\lim_{\varepsilon \searrow 0} a_\varepsilon$ を考える場合も同様である．

第1章　大偏差原理の一般論

大偏差原理の学習を本格的に始める際に，最初に知っておくべき一般的な基礎知識を本章でまとめておく．個別具体例を論ずる次章以降とは異なり，本章の内容は大偏差原理に興味を持つ全ての読者に関係する．

1.1　大偏差原理の定義

本節の目的は大偏差原理の定義を与えることである．それにあわせて弱大偏差原理と指数的緊密性も導入する．

大偏差原理は位相空間上で定義された確率測度の族に対する極限定理である．本書を通じて，大偏差原理を論じる際に確率測度の族が定義されている位相空間 $\mathcal{X}$ は空でない Hausdorff 空間であると仮定する.[1] 特に断らない限り，$\mathcal{X}$ 上の σ 加法族としては Borel 加法族 $\mathcal{B}(\mathcal{X})$（すなわち $\mathcal{X}$ の開部分集合全体から生成される σ 加法族）を考え，$\mathcal{X}$ 上の確率測度としては $\mathcal{B}(\mathcal{X})$ 上で定義されているもの（すなわち Borel 確率測度）のみを扱う．

まず速度関数の定義をするが，その前に下半連続性の定義を復習する．拡大実数値関数 $f\colon \mathcal{X} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ が下半連続であるとは，任意の $r \in \mathbb{R}$ に対して $f^{-1}([-\infty, r]) = \{x \in \mathcal{X} \mid f(x) \le r\}$ が閉集合になることをいう．また f が上半連続であるとは，$-f$ が下半連続であることをいう．

定義 1.1.1. Hausdorff 空間 $\mathcal{X}$ 上の速度関数 I とは，下半連続な拡大非負値関数 $I\colon \mathcal{X} \to [0, \infty]$ のことである．速度関数 I が有限の値を取る集合，すなわち $I^{-1}([0, \infty)) = \{x \in \mathcal{X} \mid 0 \le I(x) < \infty\}$ を I の実効定義域と呼ぶ．任意の $r \in [0, \infty)$ に対して $I^{-1}([0, r]) = \{x \in \mathcal{X} \mid 0 \le I(x) \le r\}$ がコンパクト集合になるとき，I を良い速度関数と呼ぶ.[2]

$\mathcal{X}$ 上の確率測度の族に対する大偏差原理を考えるときは，大雑把に分類するとその族が小さな正数 $0 < \varepsilon \ll 1$ で添字付けられている場合と，自然数 $n \in \mathbb{N}$ で添字付けられている場合の 2 種類がある．ここでは主に前者についてまず解説し，後者については後で軽く触れる．

---

[1] 本章では $\mathcal{X}$ が空ではないことを暗黙のうちに仮定する．

[2] 文献によっては, 速度関数が良いことを速度関数の定義や大偏差原理の定義に含めてあるので，少し注意する必要がある．[34, 29] などがそういった例である．

$\mathcal{X}$ 上の確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ に対する大偏差原理を考えるとき，速度とは広義単調減少する正値関数 $a\colon(0,1]\to(0,\infty)$ で $\lim_{\varepsilon\searrow 0} a_\varepsilon=\infty$ を満たすもののことである.[3] 速度の最も典型的な例は $a_\varepsilon=1/\varepsilon$ または $a_\varepsilon=1/\varepsilon^2$ である.[4]

定義 1.1.2. 以下の 2 条件が満たされているとき，Hausdorff 空間 $\mathcal{X}$ 上の確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のときに速度関数 I に対して速度 a_ε で大偏差原理を満たすという.

$$\varliminf_{\varepsilon\searrow 0}\frac{1}{a_\varepsilon}\log\mu_\varepsilon(O)\ge -\inf_{x\in O}I(x),\qquad O \text{ は } \mathcal{X} \text{ の任意の開集合.} \tag{1.1.1}$$

$$\varlimsup_{\varepsilon\searrow 0}\frac{1}{a_\varepsilon}\log\mu_\varepsilon(F)\le -\inf_{x\in F}I(x),\qquad F \text{ は } \mathcal{X} \text{ の任意の閉集合.} \tag{1.1.2}$$

(1.1.1) を大偏差原理の下からの評価と呼び，(1.1.2) を大偏差原理の上からの評価と呼ぶ.

定義から直ちにわかることを 2 つ注意する．上記の大偏差原理 (1.1.1) かつ (1.1.2) は次の条件と同値である．任意の $A\in\mathcal{B}(\mathcal{X})$ に対して

$$-\inf_{x\in A^\circ}I(x)\le\varliminf_{\varepsilon\searrow 0}\frac{1}{a_\varepsilon}\log\mu_\varepsilon(A)\le\varlimsup_{\varepsilon\searrow 0}\frac{1}{a_\varepsilon}\log\mu_\varepsilon(A)\le-\inf_{x\in\overline{A}}I(x).$$

ここで A° と $\overline{A}$ はそれぞれ A の内部と閉包である．大偏差原理が成り立っている場合は $A=\mathcal{X}$ とすればすぐわかるように $\inf_{x\in\mathcal{X}}I(x)=0$ である．後出の補題 1.2.6 で説明するが，I が良いときは I は $\mathcal{X}$ 上で最小値を持つので，実は $\min_{x\in\mathcal{X}}I(x)=0$ である．

直感的にはこの定義は $A\in\mathcal{B}(\mathcal{X})$ に対して，$\varepsilon>0$ が十分小さいときに

$$\mu_\varepsilon(A)\fallingdotseq\exp\Bigl(-a_\varepsilon\inf_{x\in A}I(x)\Bigr)$$

ということを言っている.[5] A が速度関数の零点から離れている（すなわち $\inf_{x\in A}I(x)>0$ の）ときは，$\mu_\varepsilon(A)$ が「指数的な速さで減少する」という現象を速度関数 I と部分集合 A を用いてうまく表している．さらに直感的な表現として，物理学の文脈では

$$\mu_\varepsilon(x)\fallingdotseq\exp\Bigl(-a_\varepsilon I(x)\Bigr),\quad x\in\mathcal{X}$$

として表現されることもある．式が表すように,「値 $x\in\mathcal{X}$ が現れる（非常に小さい）確率は $e^{-a_\varepsilon I(x)}$ である」と解釈される.（本段落の説明は厳密ではない.）

次は確率変数の族に対する大偏差原理を導入する．確率論では $\mathcal{X}$ 上の確率測度がなんらかの確率空間上で定義された $\mathcal{X}$ 値確率変数の法則（すなわち像測度）として得られている場合が非常に多い．そこで簡単のために $\mathcal{X}$ 値確率変数の族の法則が $\mathcal{X}$ 上で大偏差原理を満たすことをもって，その確率変数の族は大偏差原理を満たすということにする.

---

[3] $\varepsilon\searrow 0$ のときの極限を議論するので，添字 ε が属する集合を $(0,1]$ から $(0,c]$ や $(0,\infty)$ に代えても結局同値である．ここで c は任意の正数である．簡単のため本書では $c=1$ とする.

[4] よく似た用語であるが,「速度関数」と「速度」ははっきり違う．前者は “rate function” の訳語であり，後者は “speed” の訳語である.

[5] 大偏差原理の初心者で log の中に確率が入っているのを気持ち悪く感じる人は，このように指数関数の形で書いてみるとよい.

定義 1.1.3. $\mathcal{X}$, I および a は定義 1.1.2 と同じとする．各 $\varepsilon \in (0,1]$ に対して，X_ε はある確率空間 $(\Omega_\varepsilon, \mathcal{F}_\varepsilon, \mathbb{P}_\varepsilon)$ 上で定義された $\mathcal{X}$ 値確率変数だとし，その法則を $\mathbb{P}_\varepsilon \circ (X_\varepsilon)^{-1}$ と書く.[6] $\mathcal{X}$ 上の確率測度の族 $\{\mathbb{P}_\varepsilon \circ (X_\varepsilon)^{-1}\}_{\varepsilon\in(0,1]}$ が定義 1.1.2 の条件 (1.1.1) かつ (1.1.2) を満たすときに，$\{X_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに速度関数 I に対して速度 a_ε で大偏差原理を満たすという．

次は弱大偏差原理を定義する．これは上からの評価を閉集合に対して要請する代わりに，コンパクト集合に対してのみ要請するものである．下からの評価はまったく同じである．大偏差原理から弱大偏差原理が導けることは自明だが，実は前者は後者より真に強い命題であることが知られている.[7]（煩雑になるので書かないが，$\mathcal{X}$ 値確率変数の族 $\{X_\varepsilon\}_{\varepsilon\in(0,1]}$ の弱大偏差原理も定義 1.1.3 と同じ要領で定義する.）

定義 1.1.4. 以下の２条件が満たされているとき，Hausdorff 空間 $\mathcal{X}$ 上の確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに速度関数 I に対して速度 a_ε で弱大偏差原理を満たすという．

$$\varliminf_{\varepsilon\searrow 0} \frac{1}{a_\varepsilon} \log \mu_\varepsilon(O) \geq -\inf_{x\in O} I(x), \qquad O \text{ は } \mathcal{X} \text{ の任意の開集合.} \tag{1.1.3}$$

$$\varlimsup_{\varepsilon\searrow 0} \frac{1}{a_\varepsilon} \log \mu_\varepsilon(K) \leq -\inf_{x\in K} I(x), \qquad K \text{ は } \mathcal{X} \text{ の任意のコンパクト集合.} \tag{1.1.4}$$

(1.1.3) を弱大偏差原理の下からの評価と呼び，(1.1.4) を弱大偏差原理の上からの評価と呼ぶ.

３つ目に指数的緊密性を定義する.（煩雑になるので書かないが，$\mathcal{X}$ 値確率変数の族 $\{X_\varepsilon\}_{\varepsilon\in(0,1]}$ の指数的緊密性も定義 1.1.3 と同じ要領で定義する.）

定義 1.1.5. Hausdorff 空間 $\mathcal{X}$ 上の確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに（速度 a_ε で）指数的に緊密であるとは，任意の $r \in [0,\infty)$ に対してある $\mathcal{X}$ のコンパクト集合 K_r が存在して

$$\varlimsup_{\varepsilon\searrow 0} \frac{1}{a_\varepsilon} \log \mu_\varepsilon(\mathcal{X} \setminus K_r) \leq -r$$

が成立することをいう.

大雑把に言えば「指数的な尺度で測ったとしても，μ_ε のほとんどの重みはある共通のコンパクト集合に乗っている」という条件である．なお上記の定義の中で「コンパクト」を「相対コンパクト」に変えても同値な定義であることは明らかであろう.

---

[6] すなわち，任意の $A \in \mathcal{B}(\mathcal{X})$ に対して $\mathbb{P}_\varepsilon \circ (X_\varepsilon)^{-1}(A) := \mathbb{P}_\varepsilon((X_\varepsilon)^{-1}(A))$ とおくことにより定まる $\mathcal{X}$ 上の確率測度のことである.

[7] 例えば $\mathcal{X} = (0,\infty)$ の場合に，$I \equiv \infty$ と点測度の族 $\{\delta_\varepsilon\}_{\varepsilon\in(0,1]}$ を考えてみよ.

注意 1.1.6. 測度の族（または確率変数の族）の添字集合が $\mathbb{N}$ の場合（つまり「列」の場合）でも，定義 1.1.2, 定義 1.1.4, 定義 1.1.5 が自然に修正できることは明らかであろうから省略する．この場合，速度とは広義単調増大する非負値数列 $a = \{a_n\}_{n\in\mathbb{N}}$ で，$\lim_{n\to\infty} a_n = \infty$ を満たすもののことである．自然な例は $a_n = n^d$ $(d \in \mathbb{N})$ であり，最重要例は当然 $a_n = n$ である．(例えば大偏差原理の定義を例に取ると，定義 1.1.2 に登場する添字 $\varepsilon \in (0, 1]$, $\varepsilon \searrow 0$ をただ単に $n \in \mathbb{N}$, $n \to \infty$ に書き直せばよい.)

注意 1.1.7. 確率論では確率測度が定義された空間はほぼ常に可分完備距離空間である．したがって，読者があまりこの分野に詳しくない場合は，$\mathcal{X}$ が可分完備距離空間であるとして本章を読むことを勧める．実際，本書において $\mathcal{X}$ が可分完備距離空間でない場合を本当に扱っているのは

- 射影極限法に関する第 1.9 節
- 射影極限法を用いて Mogulskii の定理を証明する第 8 章
- Mogulskii の定理から Schilder の定理を導く第 9.6 節

の 3 箇所である．これら以外に一般化された Sanov の定理を扱う第 6 章もあるが，これは定義 1.1.2 の意味での大偏差原理ではない．もし読者がこれらの話題に興味がなければ，$\mathcal{X}$ が可分完備距離空間である場合のみ考えても特に実害はない．

1.2　基礎中の基礎

大偏差原理に関する学習を本格的に始める前の準備として，いくつかの簡単な事実を本節でまとめて紹介する．大偏差原理業界ではこれらは常識だと見なされており，多くの文献では特に説明もせずに使うが，本書の趣旨にのっとり本節では証明を与える．

注意 1.2.1. 本章の残り（第 1.2 節から 1.9 節まで）においては，確率測度の族が $\varepsilon \in (0, 1]$ で添字付けられており，大偏差原理（弱大偏差原理，指数的緊密性）の速度が $a_\varepsilon = 1/\varepsilon$ である場合のみを議論する．(一般の速度を持つ場合や自然数で添字付けられている場合などでも，本章中の証明に自明な修正をすることにより事実上同じ命題を証明できる.)

1.2.1　非確率論的な事項

事象の確率の対数を取った量は大偏差原理以外の確率論の話題ではそれほどは現れないので，すぐには馴染めない読者もいるかもしれない．その種の量の極限を計算するときに頻繁に用いる補題を 2 つ紹介する．

補題 1.2.2. $k \in \mathbb{N}$ とし，$A^i\colon (0,1] \to [0,\infty]$ とする（$1 \le i \le k$）．このとき次が成り立つ．

$$\varlimsup_{\varepsilon \searrow 0} \varepsilon \log \left(\sum_{i=1}^{k} A_\varepsilon^i \right) = \max_{1 \le i \le k} \varlimsup_{\varepsilon \searrow 0} \varepsilon \log A_\varepsilon^i. \tag{1.2.1}$$

証明. $k=2$ の場合に (1.2.1) を示せば十分である．そのためには各 $\varepsilon \in (0,1]$ に対して，次の不等式を示せばよい．

$$(\varepsilon \log s) \vee (\varepsilon \log t) \le \varepsilon \log(s+t) \le (\varepsilon \log s) \vee (\varepsilon \log t) + \varepsilon \log 2, \qquad s,t \in [0,\infty]. \tag{1.2.2}$$

実際 $(s,t) = (A_\varepsilon^1, A_\varepsilon^2)$ を代入して上極限を取れば，$\varlimsup_{\varepsilon \searrow 0}$ と $\vee$ は明らかに交換するので，(1.2.1) の $k=2$ の場合が得られる．ちなみに $(s,t) = (0,0)$, $(0,\infty)$, $(\infty,0)$, (∞,∞) のときは (1.2.2) は自明である．それ以外の場合は次の不等式に帰着される．

$$0 \le \log(s+t) - (\log s) \vee (\log t) \le \log 2, \qquad s,t \in (0,\infty). \tag{1.2.3}$$

$s \le t$ のとき明らかに $\log(s+t) - \log t \ge 0$ であり，

$$\log(s+t) - \log t \le \log(2t) - \log t = \log 2$$

となっている．よって対称性により任意の $s,t \in (0,\infty)$ に対して (1.2.3) が成立することがわかり，証明が終わる．□

補題 1.2.3. 任意の $A\colon (0,1] \to [0,\infty]$ と $B\colon (0,1] \to [0,\infty]$ に対して，次が成り立つ．

$$\varliminf_{\varepsilon \searrow 0} \varepsilon \log(A_\varepsilon + B_\varepsilon) \le (\varliminf_{\varepsilon \searrow 0} \varepsilon \log A_\varepsilon) \vee (\varlimsup_{\varepsilon \searrow 0} \varepsilon \log B_\varepsilon) \tag{1.2.4}$$

証明. 簡単のために $\alpha := \varliminf_{\varepsilon \searrow 0} \varepsilon \log A_\varepsilon$ および $\beta := \varlimsup_{\varepsilon \searrow 0} \varepsilon \log B_\varepsilon$ とおく．$\beta < \infty$ の場合のみ示せば十分である．任意の $\beta' \in (\beta, \infty)$ に対して，ある $\varepsilon_0 \in (0,1]$ が存在して，$\varepsilon \in (0,\varepsilon_0]$ であれば $\varepsilon \log B_\varepsilon \le \beta'$ となる．また単調減少して 0 に収束する部分列 $\{\varepsilon_n\}_{n \in \mathbb{N}}$ が存在して，$\alpha = \lim_{n \to \infty} \varepsilon_n \log A_{\varepsilon_n}$ となる．部分列を取ると下極限の値は広義増大するので，

$$\begin{aligned}
\varliminf_{\varepsilon \searrow 0} \varepsilon \log(A_\varepsilon + B_\varepsilon) &\le \varliminf_{n \to \infty} \varepsilon_n \log(A_{\varepsilon_n} + B_{\varepsilon_n}) \\
&\le \varliminf_{n \to \infty} \varepsilon_n \log(A_{\varepsilon_n} + e^{\beta'/\varepsilon_n}) \\
&\le \varlimsup_{n \to \infty} \varepsilon_n \log(A_{\varepsilon_n} + e^{\beta'/\varepsilon_n}) \\
&\le (\varlimsup_{n \to \infty} \varepsilon_n \log A_{\varepsilon_n}) \vee (\varlimsup_{n \to \infty} \varepsilon_n \log e^{\beta'/\varepsilon_n}) = \alpha \vee \beta'
\end{aligned}$$

となる．ここで補題 1.2.2 を使った．最後に $\beta' \searrow \beta$ として，(1.2.4) を得る．□

下半連続関数に関する補題をいくつか与える．f を Hausdorff 空間 $\mathcal{X}$ 上の $(-\infty,\infty]$ 値関数とする．各 $\alpha\in\mathbb{R}$ に対して $f^{-1}((-\infty,\alpha])=\{x\in\mathcal{X}\mid f(x)\le\alpha\}$ が $\mathcal{X}$ の閉集合であるとき，f を下半連続関数と呼ぶ．下半連続関数は Borel 可測であることに注意せよ．また各 $\alpha\in\mathbb{R}$ に対して $f^{-1}((-\infty,\alpha])$ が $\mathcal{X}$ のコンパクト集合であるとき，f を良い下半連続関数と呼ぶ．

$f\colon\mathcal{X}\to(-\infty,\infty]$ が下半連続であれば，定義により

$$f(x)=\sup\left\{\inf_{y\in U}f(y)\,\middle|\,U\subset\mathcal{X}\colon U\text{ は }x\text{ の開近傍}\right\},\qquad x\in\mathcal{X}\tag{1.2.5}$$

となる．特に任意の $x\in\mathcal{X}$ と $-\infty<\kappa<f(x)$ を満たす任意の κ に対して, x の開近傍 U で $\inf_{y\in U}f(y)\ge\kappa$ を満たすものが存在する．実際，(1.2.5) の左辺が右辺以上であることは自明である．$\kappa\in(-\infty,f(x))$ を任意に取ると，下半連続性により $f^{-1}((\kappa,\infty])$ は x の開近傍である．従って右辺の値は κ 以上である．$\kappa\nearrow f(x)$ として (1.2.5) を得る．なお $\mathcal{X}$ が距離空間の場合は，(1.2.5) の右辺は

$$\sup_{\delta>0}\inf\{f(y)\mid y\in B_{\mathcal{X}}(x,\delta)\}=\lim_{\delta\searrow0}\inf\{f(y)\mid y\in B_{\mathcal{X}}(x,\delta)\}$$

に等しいことに注意せよ．

次の命題は関数の下半連続性についての基本的な命題である．

命題 1.2.4. $f:\mathcal{X}\to(-\infty,\infty]$ を Hausdorff 空間 $\mathcal{X}$ 上の下に有界な関数とする．このとき次が成立する．

(1) 次で定義される関数 $\underline{f}:\mathcal{X}\to(-\infty,\infty]$ は下半連続である.[8]

$$\underline{f}(x):=\sup\left\{\inf_{y\in U}f(y)\,\middle|\,U\subset\mathcal{X}:\ U\text{ は }x\text{ の開近傍}\right\},\quad x\in\mathcal{X}.$$

(2) f が下半連続であることと，$f=\underline{f}$ が成立することは同値である．

証明. まず f は下に有界なので $\underline{f}$ は $(-\infty,\infty]$ 値であることに注意する．$\underline{f}$ が下半連続であることを示すためには，任意の $\alpha\in\mathbb{R}$ に対して $\underline{f}^{-1}((\alpha,\infty])$ が開集合であることを示せばよい．$\underline{f}(x)>\alpha$ を満たす $x\in\mathcal{X}$ を任意に取る．$\underline{f}(x)$ の定義により，x の開近傍 U で

$$\underline{f}(y)\ge\inf_{y'\in U}f(y')>\alpha,\qquad y\in U$$

を満たすものが存在する．よって $U\subset\underline{f}^{-1}((\alpha,\infty])$ である．これで $\underline{f}^{-1}((\alpha,\infty])$ は開集合であることがわかったので，$\underline{f}$ の下半連続性が示された．

(1) により $\underline{f}$ は下半連続なので，$f=\underline{f}$ ならば f は下半連続である．逆に f が下半連続ならば，(1.2.5) の直後で見たように $f=\underline{f}$ である．　□

[8] $\underline{f}$ は f の下半連続化と呼ばれる．

補題 1.2.5. Λ を空でない添字集合とし，$\{f_\lambda\colon \mathcal{X} \to (-\infty,\infty]\}_{\lambda\in\Lambda}$ を Hausdorff 空間 $\mathcal{X}$ 上の下半連続関数の族とする．このとき $\sup_{\lambda\in\Lambda} f_\lambda$ も $\mathcal{X}$ 上の $(-\infty,\infty]$ 値の下半連続関数である．さらにある λ に対して f_λ が良いことも仮定すると，$\sup_{\lambda\in\Lambda} f_\lambda$ も良い．

証明. 任意の $\alpha \in (-\infty,\infty]$ に対して，明らかに

$$\left\{x \in \mathcal{X} \,\middle|\, \sup_{\lambda\in\Lambda} f_\lambda(x) \leq \alpha\right\} = \bigcap_{\lambda\in\Lambda} \{x \in \mathcal{X} \mid f_\lambda(x) \leq \alpha\}$$

となるが，仮定により右辺は閉集合族の共通部分なので閉集合である．ある λ に対して f_λ が良い場合は，上式の右辺にある集合は閉集合とコンパクト集合の共通部分なのでコンパクトである． □

補題 1.2.6. $\mathcal{X}$ を Hausdorff 空間，$f\colon \mathcal{X} \to (-\infty,\infty]$ を良い下半連続関数とする．f が $\mathcal{X}$ の空でない閉集合 F 上において下から有界であれば，f は F 上で最小値を持つ．

証明. 簡単のために $\beta := \inf_{x\in F} f(x) > -\infty$ とおく．$f(z) = \beta$ となる $z \in F$ が存在することを見ればよい．$\beta = \infty$ のときは，f は F 上恒等的に ∞ なので証明することはない．よって $\beta \in \mathbb{R}$ の場合に，有限交叉性を持つ閉集合族を用いたコンパクト性の特徴づけを用いて補題を証明する．補題の主張は $f^{-1}(\{\beta\}) \cap F$ が空でないことなので，これを示す．

f が良いために，各 $n \in \mathbb{N}$ に対して $f^{-1}((-\infty,\beta+1/n]) \cap F$ は空でないコンパクト集合であり，n に関して広義単調減少である．Hausdorff 空間ではコンパクト集合は閉集合なので，有限交叉性により

$$\bigcap_{n\in\mathbb{N}} \{f^{-1}((-\infty,\beta+1/n]) \cap F\} \neq \emptyset.$$

となる．一方，左辺の集合は

$$\bigcap_{n\in\mathbb{N}} \{f^{-1}((-\infty,\beta+1/n]) \cap F\} = f^{-1}((-\infty,\beta]) \cap F = f^{-1}(\{\beta\}) \cap F$$

となるので，$f^{-1}(\{\beta\}) \cap F$ は空でないことがわかった． □

補題 1.2.7. $\mathcal{X}$ を距離空間とする．このとき $f\colon \mathcal{X} \to (-\infty,\infty]$ に対して次の 2 条件は同値である．

(1) f は下半連続である．すなわち任意の $\alpha \in (-\infty,\infty]$ に対して，$f^{-1}((-\infty,\alpha])$ は閉集合である．

(2) 任意の $\mathcal{X}$ の収束列 $\{z_n\}_{n\in\mathbb{N}}$ に対して，$f(\lim_{n\to\infty} z_n) \leq \varliminf_{n\to\infty} f(z_n)$ が成り立つ．

証明. ここでは $z_\infty := \lim_{n\to\infty} z_n$ と書く．(1) を仮定すると，任意の $\delta \in (0,1)$ に対して $\{x \in \mathcal{X} \mid f(x) > f(z_\infty) - \delta\}$ は z_∞ の開近傍である．よって，有限個の例外を除いた全ての z_n はこの開近傍に属する（すなわち $f(z_n) > f(z_\infty) - \delta$ を満たす）．これから $\underline{\lim}_{n\to\infty} f(z_n) \geq f(z_\infty) - \delta$ を得るが，δ の任意性により (2) を得る．

逆を示そう．$\alpha \in (-\infty, \infty]$ を任意とする．$\{z_n\}_{n\in\mathbb{N}}$ を部分集合 $f^{-1}((-\infty, \alpha])$ の点列で，ある $z_\infty \in \mathcal{X}$ に収束しているとする．(2) を用いると $f(z_\infty) \leq \underline{\lim}_{n\to\infty} f(z_n) \leq \alpha$ なので，$z_\infty \in f^{-1}((-\infty, \alpha])$ とわかる．これは $f^{-1}((-\infty, \alpha])$ が閉であることを意味するので，(1) が得られた． □

補題 1.2.8. $I\colon \mathcal{X} \to [0,\infty]$ を Hausdorff 空間 $\mathcal{X}$ 上で定義された良い速度関数だとする．

(1) $\{F_\delta\}_{\delta\in(0,1]}$ を $\mathcal{X}$ の閉集合の族で，δ に関して広義単調増大だとする．[9]このとき

$$\inf_{z\in F_0} I(z) = \lim_{\delta\searrow 0} \inf_{z\in F_\delta} I(z)$$

が成立する．ただし $F_0 := \cap_{\delta\in(0,1]} F_\delta$ とおいた．

(2) さらに $(\mathcal{X}, d_\mathcal{X})$ を距離空間だと仮定する．部分集合 $A \subset \mathcal{X}$ に対して，

$$d_\mathcal{X}(z, A) := \inf_{x\in A} d_\mathcal{X}(z, x) \quad \text{および} \quad A^\delta := \{z \in \mathcal{X} \mid d_\mathcal{X}(z, A) \leq \delta\}$$

と定める．このとき任意の $A \subset \mathcal{X}$ に対して，次の等式が成立する．

$$\inf_{z\in \overline{A}} I(z) = \lim_{\delta\searrow 0} \inf_{z\in A^\delta} I(z).$$

証明. まず (1) を示す．簡単のため示すべき式の右辺の値を β と書く．この式の左辺が β 以上であることは明らかなので，β 以下であることを示せば十分である．$\beta < \infty$ と仮定してよい．$\beta' \in (\beta, \infty)$ を任意に取る．自明な評価 $\beta' > \beta \geq \inf_{z\in F_\delta} I(z)$ と I が良いことにより，各 δ に対して $F_\delta \cap I^{-1}([0,\beta'])$ は空でないコンパクト集合で，さらに δ に関して広義単調である．有限交叉性により

$$F_0 \cap I^{-1}([0,\beta']) = \Big(\bigcap_{\delta\in(0,1]} F_\delta\Big) \cap I^{-1}([0,\beta']) = \bigcap_{\delta\in(0,1]} (F_\delta \cap I^{-1}([0,\beta'])) \neq \emptyset$$

となるが，これは $\inf_{z\in F_0} I(z) \leq \beta'$ を意味する．最後に $\beta' \searrow \beta$ として (1) の証明が終わる．

次に (2) を示す．3 角不等式により $z \mapsto d_\mathcal{X}(z, A)$ の連続性が簡単に確認できるので，A^δ は閉集合であり，明らかに δ に関して広義単調である．よって (1) を適用できる．最後に $\cap_{\delta\in(0,1]} A^\delta = \{z \in \mathcal{X} \mid d_\mathcal{X}(z, A) = 0\} = \overline{A}$ に注意すると証明が終わる． □

[9] $\delta \leq \delta'$ のとき $F_\delta \subset F_{\delta'}$ が成り立つという意味である．

1.2.2　位相空間の分離公理

少し本節の主旨から外れるが，ちょうどよい機会なのでここで少し寄り道して，位相空間の分離公理について復習する．なお距離空間上の大偏差原理のみに興味を持つ読者は読み飛ばしてよい．

以下，位相空間 $\mathcal{Y}$ に対する条件を導入する．

($\mathbf{T}_1$) $\mathcal{Y}$ の任意の相異なる２点 x, y に対して，x の開近傍で y を含まないものが存在する．

($\mathbf{T}_2$) $\mathcal{Y}$ の任意の相異なる２点 x, y に対して，x の開近傍 U と y の開近傍 V で $U \cap V = \emptyset$ となるものが存在する．

($\mathbf{T}_3$) 任意の閉部分集合 $F \subset \mathcal{Y}$ と任意の $x \in F^c$ に対して，$x \in U, F \subset V$ かつ $U \cap V = \emptyset$ となる開部分集合 U と V が存在する．

($\mathbf{T}_2$) を満たす位相空間を Hausdorff 空間という．**($\mathbf{T}_1$)** かつ **($\mathbf{T}_3$)** を満たす位相空間を正則空間という．[10]距離空間は正則である．明らかに **($\mathbf{T}_2$)** は **($\mathbf{T}_1$)** を導く．

定義から簡単にわかる事実を述べる．

- **($\mathbf{T}_1$)** は「任意の１点集合が閉である」という条件と同値である．
- **($\mathbf{T}_2$)** は「$\mathcal{Y} \times \mathcal{Y}$ の対角線集合が $\mathcal{Y} \times \mathcal{Y}$ の閉部分集合である」という条件と同値である．ここで $\mathcal{Y} \times \mathcal{Y}$ には直積位相が備わっている．
- Hausdorff 空間においてはコンパクト集合は閉集合である．
- $\mathcal{Y}$ が正則であれば Hausdorff である．

本書では射影極限法に関係して，距離付け可能ではない位相ベクトル空間を扱うので，位相ベクトル空間と分離公理の関係に関する事実を１つ紹介しよう．ベクトル空間の係数体は $\mathbb{R}$ とする．

$\mathcal{Y}$ が位相ベクトル空間であるとは，$\mathcal{Y}$ が位相空間かつベクトル空間であり，さらにベクトル空間としての演算

$$\mathcal{Y} \times \mathcal{Y} \ni (x, y) \mapsto x + y \in \mathcal{Y}, \qquad \mathbb{R} \times \mathcal{Y} \ni (a, x) \mapsto ax \in \mathcal{Y}$$

がどちらも連続写像であることをいう．ここで $\mathcal{Y} \times \mathcal{Y}$ と $\mathbb{R} \times \mathcal{Y}$ には直積位相が備わっている．位相ベクトル空間 $\mathcal{Y}$ の部分集合 A, B に対して，$A + B := \{x + y \mid x \in A, y \in B\}$ および $-A := \{-x \mid x \in A\}$ と定める．$A = -A$ となるとき，A は対称的であるという．（なお $B = \{y\}$ のときは $A + y := A + \{y\}$ と略記する．）

---

[10]分離公理 **($\mathbf{T}_3$)** や正則空間の定義については２つの流儀があるようである．本書の定義は [19, 21] などに従った．しかし例えば [6] は異なる定義を採用している．

命題 1.2.9. $\mathcal{Y}$ を位相ベクトル空間とする．このとき $\mathcal{Y}$ は $(\mathbf{T}_3)$ を満たす．したがって，１点集合 $\{0\}$ が $\mathcal{Y}$ で閉であれば，$\mathcal{Y}$ は位相空間として正則である．

証明. まず 0 の任意の開近傍 W に対して，0 の対称的な開近傍 O が存在して $O+O \subset W$ となることを示そう．加法の連続性により，$V+U \subset W$ を満たす 0 の開近傍 V と U が存在する．そこで $O := V \cap (-V) \cap U \cap (-U)$ とおくと，作り方により $O = -O$ かつ $O+O \subset W$ となる．

$F \subset \mathcal{Y}$ を $x \in \mathcal{Y}$ を含まない閉部分集合とする．平行移動は同相写像なので，初めから $x = 0$ と仮定してよい．このとき F^c は 0 の開近傍であるので，0 のある対称的な開近傍 O を取れば $O+O \subset F^c$，すなわち $(O+O) \cap F = \emptyset$ を満たす．このとき $O \cap (O+F) = \emptyset$ が従うことに注意せよ．実際，これが空でないとすれば，$x, y \in O$ と $f \in F$ で $x = y+f$ を満たすものが存在するが，対称性により $-y \in O$ なので $x - y = f \in (O+O) \cap F$ となり矛盾する．$O+F = \cup_{x \in F}(O+x)$ と書けるので，これは F を含む開集合である．以上により $\mathcal{Y}$ が $(\mathbf{T}_3)$ を満たすことがわかった．

平行移動は同相写像なので，$\{0\}$ が $\mathcal{Y}$ で閉のときは任意の１点集合が閉である．これは $(\mathbf{T}_1)$ と同値なので，$\mathcal{Y}$ は正則である． □

注意 1.2.10. 命題 1.2.9 は「任意の位相群は $(\mathbf{T}_3)$ を満たす」という有名な事実の特殊な場合である．実際，上記の証明中で -1 倍以外のスカラー倍は全く使っていない.（-1 倍する写像はむしろ逆元を取る操作だと見なすべきであろう.）

1.2.3　確率論的な事項

大偏差原理の学習を始める前に確認しておくべき事項のうち，確率論に関係するものをまとめておく．まず大偏差原理の速度を固定した場合，速度関数は一意的に決まることを見る．

命題 1.2.11. $\{\mu_\varepsilon\}_{\varepsilon \in (0,1]}$ を正則空間 $\mathcal{X}$ 上の確率測度の族で，$\varepsilon \searrow 0$ のときに速度 $1/\varepsilon$ で大偏差原理を満たすとする．このとき，この大偏差原理（と速度 $1/\varepsilon$）に付随する速度関数は一意的である．

証明. ２つの異なる速度関数が I と J がこの大偏差原理に付随すると仮定する．このとき $I(z) \neq J(z)$ となる $z \in \mathcal{X}$ が存在する．対称性により $I(z) > J(z)$ としてよい．これから矛盾を導く．まず $I(z) > \kappa > J(z)$ となる $\kappa \in \mathbb{R}$ を任意に取る．(1.2.5) により z の開近傍 U で $\inf_{x \in U} I(x) \geq \kappa$ を満たすものが取れる．分離公理 $(\mathbf{T}_3)$ を z と閉集合 U^c に適用すると，z の開近傍 O と U^c の開近傍 V で $O \cap V = \emptyset$ となるものが存在する．V^c は閉かつ $O \subset V^c$ なので $\overline{O} \subset V^c$ であるが，これから $\overline{O} \cap U^c \subset \overline{O} \cap V = \emptyset$ となり，結局 $\overline{O} \subset U$ とわかる．したがって特に $\inf_{x \in \overline{O}} I(x) \geq \kappa$ を得る．大偏差原理により

$$-\inf_{x \in O} J(x) \leq \varliminf_{\varepsilon \searrow 0} \varepsilon \log \mu_\varepsilon(O) \leq \varlimsup_{\varepsilon \searrow 0} \varepsilon \log \mu_\varepsilon(\overline{O}) \leq -\inf_{x \in \overline{O}} I(x) \leq -\kappa$$

が成り立つ．よって $\inf_{x\in O} J(x) \geq \kappa$ を得る．一方で $\inf_{x\in O} J(x) \leq J(z) < \kappa$ であるため矛盾が生じた． □

確率測度の族の重みがある部分集合にのみ乗っている場合，大偏差原理が全体集合上で満たされることと，その部分集合上で満たされることの関係はどうであろうか．次の命題で議論するが，その前にまず相対位相の定義を思い出そう．

この段落では $\mathcal{X}$ を位相空間とし，$\mathcal{Z}$ をその部分集合とする．$O' \subset \mathcal{Z}$ が $\mathcal{Z}$ 位相に関する開集合であることは，$\mathcal{X}$ 位相に関する開集合 $O \subset \mathcal{X}$ が存在して $O' = O \cap \mathcal{Z}$ と書けることと同値である．同様に $F' \subset \mathcal{Z}$ が $\mathcal{Z}$ 位相に関する閉集合であることは，$\mathcal{X}$ 位相に関する閉集合 $F \subset \mathcal{X}$ が存在して $F' = F \cap \mathcal{Z}$ と書けることと同値である．なお $\mathcal{Z}$ が $\mathcal{X}$ 位相に関する閉集合である場合は，$F' \subset \mathcal{Z}$ が $\mathcal{Z}$ 位相に関する閉集合であることと $\mathcal{X}$ 位相に関する閉集合であることは同値である．

命題 1.2.12. $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ を Hausdorff 空間 $\mathcal{X}$ 上の確率測度の族，$\mathcal{Z}$ を $\mathcal{X}$ の Borel 可測集合とし，任意の $\varepsilon \in (0,1]$ に対して $\mu_\varepsilon(\mathcal{Z}) = 1$ が成り立つとする．$\mathcal{Z}$ には $\mathcal{X}$ からの相対位相（部分集合としての位相）が備わっている．

(1) $\mathcal{Z}$ は閉集合で，$\varepsilon \searrow 0$ のとき $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{Z}$ 上で速度関数 $I'\colon \mathcal{Z} \to [0,\infty]$ に対して大偏差原理を満たすとする．$\mathcal{Z}^c$ 上では値を ∞ とおいて I' の拡張 $I\colon \mathcal{X} \to [0,\infty]$ を定める．このとき，I は速度関数であり，$\varepsilon \searrow 0$ のとき $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{X}$ 上で I に対して大偏差原理を満たす．

(2) $\varepsilon \searrow 0$ のとき $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{X}$ 上で速度関数 $I\colon \mathcal{X} \to [0,\infty]$ に対して大偏差原理を満たすとし，さらに $I^{-1}([0,\infty)) \subset \mathcal{Z}$ とする．このとき $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{Z}$ 上で速度関数 $I' := I \restriction_{\mathcal{Z}}$ [11]に対して大偏差原理を満たす．

証明. (1) を示す．拡張の仕方により，$r \in [0,\infty)$ のとき $I^{-1}([0,r]) = (I')^{-1}([0,r])$ となるが，I' が $\mathcal{Z}$ 上の速度関数であるために右辺は $\mathcal{Z}$ で閉であり，その結果 $\mathcal{X}$ でも閉である．これで I が速度関数であることがわかった．$\mathcal{X}$ 位相に関する開集合 $O \subset \mathcal{X}$ を任意に取る．このとき $O \cap \mathcal{Z}$ は $\mathcal{Z}$ 位相に関して開である．また明らかに $\inf_{x\in O} I(x) = \inf_{x\in O\cap\mathcal{Z}} I'(x)$ となる．$\mathcal{Z}$ 上の大偏差原理の下からの評価により

$$-\inf_{x\in O} I(x) = -\inf_{x\in O\cap\mathcal{Z}} I'(x) \leq \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(O \cap \mathcal{Z}) = \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(O)$$

となり，$\mathcal{X}$ 上の大偏差原理の下からの評価を得る．上からの評価も全く同じ論法で証明できるので省略する．

(2) を示す．$\mathcal{Z}$ 位相に関する開集合 $O' \subset \mathcal{Z}$ を任意に取る．このとき，ある $\mathcal{X}$ 位相に関する開集合 O に対して $O' = O \cap \mathcal{Z}$ となる．仮定 $I^{-1}([0,\infty)) \subset \mathcal{Z}$ により I は $\mathcal{Z}^c$ 上で

[11] $I \restriction_{\mathcal{Z}}$ は I を $\mathcal{Z}$ に制限した関数を表す．

は恒等的に ∞ であるため $\inf_{x\in O} I(x) = \inf_{x\in O'} I'(x)$ となる．$\mathcal{X}$ 上の大偏差原理の下からの評価により

$$-\inf_{x\in O'} I'(x) = -\inf_{x\in O} I(x) \le \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(O) = \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(O')$$

となり，$\mathcal{Z}$ 上の大偏差原理の下からの評価を得る．上からの評価も全く同じ論法で証明できるので省略する．　□

注意 1.2.13. $\mathcal{Z}$ が閉集合のときは，命題 1.2.12 (2) の仮定 $I^{-1}([0,\infty)) \subset \mathcal{Z}$ は常に満たされる．実際，開集合 $\mathcal{Z}^c$ は重さが常に 0 なので，$\mathcal{X}$ 上の大偏差原理の下からの評価により

$$-\inf_{x\in\mathcal{Z}^c} I(x) \le \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(\mathcal{Z}^c) = -\infty$$

であるが，これは I は $\mathcal{Z}^c$ 上で恒等的に ∞ であること，すなわち $I^{-1}([0,\infty)) \subset \mathcal{Z}$ と同値である．

大偏差原理の定義に現れる関数は速度関数（下半連続）であることが要請される．実は下半連続とは限らない場合に大偏差原理の上下評価が成立すれば，その関数が下半連続になるように取り替えることができる．この主張を以下の命題で正確に述べる．

命題 1.2.14. $\mathcal{X}$ を Hausdorff 空間とし，$I : \mathcal{X} \to [0,\infty]$ を下半連続とは限らない関数とする．$\mathcal{X}$ 上の確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は，任意の $A \in \mathcal{B}(\mathcal{X})$ に対して

$$-\inf_{x\in A^\circ} I(x) \le \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(A) \le \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(A) \le -\inf_{x\in\overline{A}} I(x) \tag{1.2.6}$$

を満たすと仮定する．このとき，$\mathcal{X}$ 上の確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに速度関数 $\underline{I}$ に対して大偏差原理を満たす.

証明. $\underline{I}$ が速度関数であることは命題 1.2.4(1) による．ほぼ明らかに $I \ge \underline{I}$ が成立するので，(1.2.6) により速度関数 $\underline{I}$ に対して大偏差原理の上からの評価が成立する．大偏差原理の下からの評価を示すために，開集合 $O \subset \mathcal{X}$ と $x \in O$ を任意に取る．x の開近傍 $U \subset O$ に対して，(1.2.6) により

$$\varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(O) \ge \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(U) \ge -\inf_{y\in U} I(y)$$

が成立する．このとき x の開近傍 U について下限を取ると，右辺は $-\underline{I}(x)$ になる．さらに $x \in O$ について上限を取ることにより

$$\varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(O) \ge -\inf_{x\in O} \underline{I}(x)$$

を得る．　□

定義から容易に予想できることだが，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ が $\mathcal{X}$ 上で速度関数 I に対して大偏差原理を満たしているとき，重みは I の零点の周りに集中しそうである．少なくとも $\mathcal{X}$ が距離空間で I が良いときは，以下のようにこのことを厳密に示せる．U が $I^{-1}(0)$ の開近傍であれば，補題 1.2.6 により $\min_{x\in U^c} I(x)$ $(=: a)$ は存在し，作り方から $a>0$ であるため，大偏差原理の上からの評価により $\overline{\lim}_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(U^c)\le -a$，特に $\lim_{\varepsilon\searrow 0}\mu_\varepsilon(U^c)=0$ が成り立つ．特に $I(z)=0$ となる $z\in\mathcal{X}$ が唯一つ存在する場合は，この事実から $\overline{\lim}_{\varepsilon\searrow 0}\mu_\varepsilon(F)\le\delta_z(F)$ が任意の閉集合 $F\subset\mathcal{X}$ に対して成立することがわかるので，$\lim_{\varepsilon\searrow 0}\mu_\varepsilon=\delta_z$ と弱収束している．[12]状況がさほど悪くなければ，次の命題にあるように「弱収束」を「概収束」に強めることができる．[13]

命題 1.2.15. $\mathcal{X}$ を距離空間とし，$\kappa\in(0,\infty)$ とする．確率空間 $(\Omega,\mathcal{F},\mathbb{P})$ 上で定義された $\mathcal{X}$ 値確率変数の列 $\{X_n\}_{n\in\mathbb{N}}$ が $n\to\infty$ のときに速度 n^κ で良い速度関数 I に対して大偏差原理を満たすとする．さらに $I(z)=0$ となる $z\in\mathcal{X}$ が唯一つ存在すると仮定する．このとき，$\mathbb{P}\left(\lim_{n\to\infty}X_n=z\right)=1$ である．

証明. 任意の $r\in(0,\infty)$ を取る．このとき開球 $B(z,r)$ に対して上と同様の議論をすると，$a_r:=\min_{x\notin B(z,r)}I(x)>0$ かつ $\overline{\lim}_{n\to\infty}n^{-\kappa}\log\mathbb{P}\left(X_n\notin B(z,r)\right)\le -a_r$ を得る．よって，n が十分大ならば $\mathbb{P}\left(X_n\notin B(z,r)\right)\le e^{-n^\kappa a_r/2}$ となり，特に $\sum_{n=1}^\infty\mathbb{P}\left(X_n\notin B(z,r)\right)<\infty$ となる．Borel-Cantelli の定理により，$C_r:=\{$ 無限個の n に対して $X_n\notin B(z,r)\}$ は零集合である．これらの合併 $\cup_{r\in\mathbb{Q}\cap(0,\infty)}C_r$ も零集合であり，その補集合上では $\lim_{n\to\infty}X_n=z$ が成立している． □

1.3　弱大偏差原理に関する基礎事項

本節では弱大偏差原理に関する基本的な事実を２つ紹介する．どちらも非常に有名なだけでなく実用上も頻出する議論であるため，大偏差原理を学ぶ者にとっては常識の範疇に属する事項である．

１つ目の事実を標語的にまとめると以下のようになる．

「弱大偏差原理」＋「指数的緊密性」$\Longrightarrow$ 「大偏差原理」

要するに，名前が示すとおり弱大偏差原理は大偏差原理より弱いのだが，指数的緊密性が成立している場合には両者は同値なのである．

命題 1.3.1. $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ を Hausdorff 空間 $\mathcal{X}$ 上の指数的に緊密な Borel 確率測度の族とする．もし $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ が $\varepsilon\searrow 0$ のときに速度関数 I に対して弱大偏差原理を満たすならば，実は $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のときに良い速度関数 I に対して大偏差原理を満たす．

---

[12]いわゆる Portmanteau の定理は一般の距離空間上で成立する．例えば [27, 定理 2.1] を参照せよ．

[13]非可算なパラメータ集合を持つ確率変数族は概収束との相性が良くない．よって，この命題だけパラメータ集合は $\mathbb{N}$ とする．

証明. まず指数的緊密性と弱大偏差原理の上からの評価から，大偏差原理の上からの評価が従うことを見る．F を $\mathcal{X}$ の閉集合とする．指数的緊密性により，各 $r \in (0,\infty)$ に対して $\overline{\lim}_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(K_r^c) < -r$ となるコンパクト集合 K_r が存在する．このとき $F\cap K_r$ はコンパクトなので，弱大偏差原理により

$$
\begin{aligned}
\overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(F) &\le \overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log[\mu_\varepsilon(F\cap K_r)+\mu_\varepsilon(K_r^c)] \\
&\le \left(\overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(F\cap K_r)\right)\vee\left(\overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(K_r^c)\right) \\
&\le \left(-\inf_{x\in F\cap K_r} I(x)\right)\vee(-r) \\
&\le \left(-\inf_{x\in F} I(x)\right)\vee(-r)
\end{aligned}
$$

を得る．２つ目の不等号で上極限に関する補題 1.2.2 も用いた．上式で $r\nearrow\infty$ として大偏差原理の上からの評価を得る．

次に I が良いことを示す．各 $r\in(0,\infty)$ に対して弱大偏差原理の下からの評価を用いると

$$
-\inf_{x\in K_r^c} I(x) \le \underline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(K_r^c) \le \overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(K_r^c) < -r
$$

を得る．この不等式から $I(x)\le r$ ならば $x\notin K_r^c$ となること，すなわち $I^{-1}([0,r])\subset K_r$ が従う．$I^{-1}([0,r])$ はコンパクト集合の閉部分集合なのでコンパクトである．これで I が良いことが示せた． □

２つ目の事実を少し単純化して説明する．大偏差原理の「局所版」または「各点版」だと解釈できる条件 (1.3.1) を導入すると，実はその条件から弱大偏差原理が従うことが証明できる.（命題 1.3.1 と合わせると，条件 (1.3.1) と指数的緊密性から大偏差原理が従うことになる.）

命題 1.3.2. $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ を距離空間 $\mathcal{X}$ 上の Borel 確率測度の族だとする．任意の $x\in\mathcal{X}$ に対して

$$
\lim_{\delta\searrow 0}\underline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(B(x,\delta)) = \lim_{\delta\searrow 0}\overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(B(x,\delta)) \ \in[-\infty,0] \tag{1.3.1}
$$

が成立することを仮定し，この値を $-I(x)$ と定める．このとき I は速度関数であり，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のときに I に対して弱大偏差原理を満たす.[14]

証明. 各 $x\in\mathcal{X}$ と $\delta>0$ に対して

$$
L(x,\delta) := -\underline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mu_\varepsilon(B(x,\delta))
$$

[14][33, 第 4.1.2 節] では Hausdorff 空間上でこの命題を証明しているが，本書ではそこまで一般化しない．

とおく．δ に関する広義単調性と (1.3.1) により

$$I(x)=\sup_{\delta>0}L(x,\delta)=\sup_{\delta>0}\Big[-\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))\Big] \tag{1.3.2}$$

が成り立つことに注意せよ．

$r\geq 0$ に対して $I(x)>r$ とすると，I の定義によりある $\delta>0$ に対して $L(x,\delta)>r$ となる．すると任意の $\widetilde{x}\in B(x,\delta)$ に対して，$\widetilde{\delta}>0$ を $B(\widetilde{x},\widetilde{\delta})\subset B(x,\delta)$ となるように取れば，$I(\widetilde{x})\geq L(\widetilde{x},\widetilde{\delta})\geq L(x,\delta)>r$ となる．これは $B(x,\delta)\subset\{x\in\mathcal{X}\mid I(x)>r\}$ を意味するので，$\{x\in\mathcal{X}\mid I(x)>r\}$ が $\mathcal{X}$ の開集合であることがわかる．これで I は下半連続であることがわかったので，I が速度関数であることが示せた．

次は O を $\mathcal{X}$ の開集合だとする．任意の $x\in O$ に対して，$\delta>0$ を十分小さく取り，$B(x,\delta)\subset O$ となるようにする．このとき

$$\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(O)\geq\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))=-L(x,\delta)\geq -I(x)$$

となる．この右辺の $x\in O$ に関する上限を取れば，開集合に対する下からの評価を得る．

最後に K を $\mathcal{X}$ のコンパクト集合だとする．$\kappa\in(0,1)$ に対して $I^\kappa(x):=(I(x)-\kappa)\wedge\kappa^{-1}$ とおき，I を下から近似する．(1.3.2) により，任意の $x\in K$ に対してある $\delta_x>0$ が存在して，

$$-\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta_x))\geq I^\kappa(x)$$

となる．K のコンパクト性により，有限個の $x_1,\dots,x_m\in K$ が存在し（それに対応する δ を $\delta_1,\dots,\delta_m$ と書くと），$K\subset\cup_{i=1}^m B(x_i,\delta_i)$ と被覆できる．したがって，

$$\begin{aligned}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(K)&\leq\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\Big(\sum_{i=1}^m\mu_\varepsilon(B(x_i,\delta_i))\Big)\\&\leq\max_{1\leq i\leq m}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x_i,\delta_i))\\&\leq-\min_{1\leq i\leq m}I^\kappa(x_i)\leq-\inf_{x\in K}I^\kappa(x)\end{aligned}$$

と評価できる．ここで上極限に関する補題 1.2.2 を用いた．最後に $\kappa\searrow 0$ とすると，右辺は $-\inf_{x\in K}I(x)$ に収束することがごく簡単に示せるので，コンパクト集合に対する上からの評価を得る．　□

上記の命題 1.3.2 のいわば「ほぼ逆」を主張するのが下記の命題 1.3.3 である（もちろん厳密には「逆」ではない）．命題 1.3.1, 命題 1.3.2, 命題 1.3.3 により，指数的緊密性の下では (i) 大偏差原理 (ii) 弱大偏差原理 (iii) 大偏差原理の局所版，すなわち条件 (1.3.1) または (1.3.3) の 3 者は全て同値である．

命題 1.3.3. 距離空間 $\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のときに速度関数 $I\colon\mathcal{X}\to[0,\infty]$ に対して大偏差原理を満たすとする．このとき，任意の $x\in\mathcal{X}$ に対して

$$\lim_{\delta\searrow 0}\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))=-I(x)=\lim_{\delta\searrow 0}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta)) \tag{1.3.3}$$

が成立する．[15]

証明. $x\in X$ に対して

$$g(x):=\lim_{\delta\searrow 0}\inf\{I(y)\mid y\in\overline{B(x,\delta)}\}=\lim_{\delta\searrow 0}\inf\{I(y)\mid y\in B(x,\delta)\}$$

とおく．右の等号は $B(x,\delta)\subset\overline{B(x,\delta)}\subset B(x,2\delta)$ による．このとき $g(x)=I(x)$ となることを見る．$g(x)\le I(x)$ は自明である．仮に $g(x)<I(x)\le\infty$ とする．$g(x)<r<I(x)$ のとき $x\in I^{-1}([0,r])^c$ である．$I^{-1}([0,r])^c$ は開なので，$\delta>0$ が十分小さいときは $B(x,\delta)\subset I^{-1}([0,r])^c$ となるが，これは $\inf\{I(y)\mid y\in B(x,\delta)\}\ge r$ を意味する．$\delta\searrow 0$ とすると $g(x)\ge r$ となり矛盾する．これで $g(x)=I(x)$ が得られた．

大偏差原理の下からの評価

$$-I(x)\le-\inf\{I(y)\mid y\in B(x,\delta)\}\le\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))$$

において，$\delta\searrow 0$ とすると $-I(x)\le\lim_{\delta\searrow 0}\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))$ を得る．大偏差原理の上からの評価

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))\le-\inf\{I(y)\mid y\in\overline{B(x,\delta)}\}$$

において，$\delta\searrow 0$ として $g(x)=I(x)$ を用いると，$\lim_{\delta\searrow 0}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))\le-I(x)$ を得る．以上をまとめて (1.3.3) を得る． □

注意 1.3.4. $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ が良い速度関数に対して大偏差原理を満たしている場合でも，指数的に緊密とは限らないことに注意せよ．[16]（ただし，距離空間 $\mathcal{X}$ が局所コンパクトの場合はこのことを比較的簡単に証明できる．）

1.4 Laplace 原理

測度論の重要な定理の多くが「部分集合形」と「関数形」の 2 つの形を持つ．前者は部分集合の重みを用いて定理の主張を記述し，後者は適切な試験関数の空間に属する関数の積分値を用いて記述する．大偏差原理の定義（定義 1.1.2）は明らかに前者に属する

[15] [33, 第 4.1.2 節] では正則空間上でこの命題を証明しているが，本書ではそこまで一般化しない．
[16] この件に関しては [33, P. 8, Remark] を見よ．

が，後者に相当するものはあるのだろうか．実は本節で解説する Laplace 原理がそれである．すなわち Laplace 原理とは大偏差原理のいわば関数形である．

本節では $(\mathcal{X}, d)$ を距離空間とし，$\mathcal{X}$ 上で定義された実数値有界連続関数の全体を $C_{\mathrm{b}}(\mathcal{X})$ と表す.

定義 1.4.1. $\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ と速度関数 $I\colon \mathcal{X} \to [0,\infty]$ に対して

$$\lim_{\varepsilon\searrow 0} \varepsilon \log \int_{\mathcal{X}} e^{h(x)/\varepsilon} \mu_\varepsilon(dx) = \sup_{x\in\mathcal{X}} \{h(x) - I(x)\}, \qquad h \in C_{\mathrm{b}}(\mathcal{X}) \tag{1.4.1}$$

が成立するとき，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに速度関数 I に対して（速度 $1/\varepsilon$ で）Laplace 原理を満たすという.[17][18]

明らかな言い換えであるが h の代わりに $-h$ を考えることにより，Laplace 原理は以下のように述べられる．$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ が $\varepsilon \searrow 0$ のときに速度関数 I に対して Laplace 原理を満たすとは，次が成立することである．

$$\lim_{\varepsilon\searrow 0} \varepsilon \log \int_{\mathcal{X}} e^{-h(x)/\varepsilon} \mu_\varepsilon(dx) = -\inf_{x\in\mathcal{X}} \{h(x) + I(x)\}, \qquad h \in C_{\mathrm{b}}(\mathcal{X}). \tag{1.4.2}$$

(1.4.1) の右辺は被積分関数と速度関数の差の上限になっており，(1.4.2) の右辺は被積分関数と速度関数の和の下限になっている．統計力学に関する問題では (1.4.1) が用いられ，弱収束法による大偏差原理の証明では (1.4.2) が用いられる．

まず Varadhan の補題を述べる．これは大偏差原理から無条件に Laplace 原理が従うことを主張する．

定理 1.4.2. 距離空間 $\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに速度関数 I に対して大偏差原理を満たすと仮定する．このとき，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は同じ速度関数 I に対して Laplace 原理を満たす.

証明. まず大偏差原理の上からの評価により，Laplace 原理 (1.4.1) の上からの評価が従うことを見る．少し一般化して $[-\infty,\infty)$ 値関数 h が上に有界かつ上半連続の場合に示す．$M > 0 \vee \sup_{x\in\mathcal{X}} h(x)$ を一つ取り，$N \in \mathbb{N}$ と $j = 1, \dots, N$ に対して

$$\begin{aligned} E_{N,j} &:= \left\{ x \in \mathcal{X} \;\middle|\; -M + \frac{2(j-1)M}{N} \le h(x) \right\}, \\ F_{N,j} &:= \left\{ x \in \mathcal{X} \;\middle|\; -M + \frac{2(j-1)M}{N} \le h(x) < -M + \frac{2jM}{N} \right\} \end{aligned}$$

[17] 正確には試験関数のクラスを明記して，「$C_{\mathrm{b}}(\mathcal{X})$ 上で Laplace 原理を満たす」と言うべきであろう．

[18] もちろん，速度を $1/\varepsilon$ でなく任意にした場合や，確率測度の列 $\{\mu_n\}_{n\in\mathbb{N}}$ の場合にも同様に Laplace 原理を定義できるが，ほぼ自明であろうから省略する．

とおき，$F_{N,0} := \{x \in \mathcal{X} \mid h(x) < -M\}$ とおく（これらの集合が M に依存することは明示しない）．h は上半連続なので $E_{N,j}$ は閉集合であり，$\cup_{0\le j\le N} F_{N,j} = \mathcal{X}$ となる．

補題 1.2.2 と大偏差原理の上からの評価により

$$\begin{aligned}
\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \int_{\mathcal{X}} e^{h/\varepsilon} d\mu_\varepsilon &= \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \Big(\sum_{j=0}^{N} \int_{F_{N,j}} e^{h/\varepsilon} d\mu_\varepsilon \Big) \\
&\le \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \Big(e^{-M/\varepsilon} + \sum_{j=1}^{N} e^{(-M+2jM/N)/\varepsilon} \mu_\varepsilon(F_{N,j}) \Big) \\
&\le \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \Big(e^{-M/\varepsilon} + \sum_{j=1}^{N} e^{(-M+2jM/N)/\varepsilon} \mu_\varepsilon(E_{N,j}) \Big) \\
&\le (-M) \vee \max_{1\le j\le N} \Big\{ -M + \frac{2jM}{N} - \inf_{x\in E_{N,j}} I(x) \Big\} \\
&\le (-M) \vee \max_{1\le j\le N} \Big\{ \sup_{x\in E_{N,j}} \{h(x) - I(x)\} + \frac{2M}{N} \Big\} \\
&\le (-M) \vee \Big[\sup_{x\in\mathcal{X}} \{h(x) - I(x)\} + \frac{2M}{N} \Big]
\end{aligned}$$

を得る．この式で $N \to \infty$ とした後に $M \to \infty$ とすれば，Laplace 原理 (1.4.1) の上からの評価を得る．

次に大偏差原理の下からの評価により，Laplace 原理 (1.4.1) の下からの評価が従うことを見る．少し一般化して $[-\infty, \infty)$ 値関数 h が下半連続の場合に示す．$z \in \mathcal{X}$ と $\delta > 0$ を任意とする．$h(z) = -\infty$ のときは後に示す評価が明らかに成立するので，$h(z) \in \mathbb{R}$ とする．h は下半連続なので $U := \{y \in \mathcal{X} \mid h(y) > h(z) - \delta\}$ は z の開近傍である．U に対する大偏差原理の下からの評価を用いると

$$\begin{aligned}
\varliminf_{\varepsilon\searrow 0} \varepsilon \log \int_{\mathcal{X}} e^{h/\varepsilon} d\mu_\varepsilon &\ge \varliminf_{\varepsilon\searrow 0} \varepsilon \log \int_{U} e^{h/\varepsilon} d\mu_\varepsilon \\
&\ge h(z) - \delta + \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(U) \\
&\ge h(z) - \delta - \inf_{x\in U} I(x) \\
&\ge h(z) - I(z) - \delta
\end{aligned}$$

を得る．この式は $h(z) = -\infty$ のときも明らかに成立する．$\delta \searrow 0$ とした後に z に関する上限を取ると

$$\varliminf_{\varepsilon\searrow 0} \varepsilon \log \int_{\mathcal{X}} e^{h/\varepsilon} d\mu_\varepsilon \ge \sup_{z\in\mathcal{X}} \{h(z) - I(z)\}$$

となり，Laplace 原理 (1.4.1) の下からの評価が得られた．これで定理 1.4.2 の証明が完成した．$\square$

注意 1.4.3. 定理 1.4.2 の証明において示したように，$[-\infty,\infty)$ 値関数 h が上に有界かつ上半連続であれば Laplace 原理 (1.4.1) の上からの評価が成立する．また $[-\infty,\infty)$ 値関数 h が下半連続であれば Laplace 原理 (1.4.1) の下からの評価が成立する．

Varadhan の補題の逆は速度関数が良ければ成り立つ．すなわち速度関数が良い場合には，大偏差原理と Laplace 原理は同値である．

定理 1.4.4. 距離空間 $\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のときに良い速度関数 I に対して Laplace 原理を満たすとする．このとき，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は同じ速度関数 I に対して大偏差原理を満たす．

証明. ここでは Laplace 原理として (1.4.2) を用いる．大偏差原理の下からの評価を先に示す．$O\subset\mathcal{X}$ を開集合とする．$\inf_{x\in O} I(x)=\infty$ のときは示すことはないので，$\inf_{x\in O} I(x)<\infty$ と仮定してよい．$I(x)<\infty$ となる $x\in O$ を任意に取る．さらに $M\in(I(x),\infty)$ を 1 つ取る．O は開なので，$B(x,\delta)\subset O$ を満たす $\delta>0$ が存在する．以上の x, M, δ に対して，$h\in C_{\mathrm{b}}(\mathcal{X})$ を次の式で定める．

$$h(y):=M\Big(\frac{d(y,x)}{\delta}\wedge 1\Big).$$

明らかに $h(x)=0$ であり，$y\notin B(x,\delta)$ であれば $h(y)=M$，さらに任意の $z\in\mathcal{X}$ に対して $0\le h(z)\le M$ を満たす．このとき

$$\int_{\mathcal{X}} e^{-h/\varepsilon}d\mu_\varepsilon=\int_{B(x,\delta)^c} e^{-h/\varepsilon}d\mu_\varepsilon+\int_{B(x,\delta)} e^{-h/\varepsilon}d\mu_\varepsilon\le e^{-M/\varepsilon}+\mu_\varepsilon(B(x,\delta)) \tag{1.4.3}$$

となる．

上の (1.4.3) から

$$\varliminf_{\varepsilon\searrow 0}\varepsilon\log\int_{\mathcal{X}} e^{-h/\varepsilon}d\mu_\varepsilon\le\max\Big\{-M,\,\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(B(x,\delta))\Big\} \tag{1.4.4}$$

が従うことを確認しよう．簡単のため $a_\varepsilon:=\mu_\varepsilon(B(x,\delta))$ および $\lambda:=\varliminf_{\varepsilon\searrow 0}\varepsilon\log a_\varepsilon$ と書く．ある 0 に収束する単調減少列 $\{\varepsilon_k\}_{k\in\mathbb{N}}\subset(0,1]$ に対して $\lambda=\lim_{k\to\infty}\varepsilon_k\log a_{\varepsilon_k}$ となることに注意せよ．この部分列に対して (1.4.3) を用いると

$$\begin{aligned}\varliminf_{\varepsilon\searrow 0}\varepsilon\log\int_{\mathcal{X}} e^{-h/\varepsilon}d\mu_\varepsilon&=\varliminf_{k\to\infty}\varepsilon_k\log\int_{\mathcal{X}} e^{-h/\varepsilon_k}d\mu_{\varepsilon_k}\\&\le\varlimsup_{k\to\infty}\varepsilon_k\log(e^{-M/\varepsilon_k}+a_{\varepsilon_k})\\&=\max\{\varlimsup_{k\to\infty}\varepsilon_k\log e^{-M/\varepsilon_k},\,\varlimsup_{k\to\infty}\varepsilon_k\log a_{\varepsilon_k}\}=\max\{-M,\lambda\}\end{aligned}$$

となり，(1.4.4) が確認できた．

一方，Laplace 原理を仮定したので

$$\lim_{\varepsilon\searrow 0}\varepsilon\log\int_{\mathcal{X}}e^{-h/\varepsilon}d\mu_\varepsilon=-\inf_{y\in\mathcal{X}}\{h(y)+I(y)\}\geq -h(x)-I(x)=-I(x)$$

となる．これら 2 式により $\max\{-M,\lambda\}\geq -I(x)$ を得るが，$-M<-I(x)$ であるため，結局

$$\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(O)\geq\lambda\geq -I(x)$$

を得る．x に関して上限を取ると

$$\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(O)\geq -\inf\{I(x)\mid x\in O, I(x)<\infty\}=-\inf\{I(x)\mid x\in O\}$$

となり，大偏差原理の下からの評価が得られた．

ここから大偏差原理の上からの評価の証明を始める．$F\subset\mathcal{X}$ を閉集合とする．この F に対して $\psi\restriction_F\equiv 0$ および $\psi\restriction_{F^c}\equiv\infty$ とおくと，ψ は $\mathcal{X}$ 上で下半連続である．$d(x,F):=\inf_{y\in F}d(x,y)$ を x から F への距離とすると，これは x に関して連続である．各 $m\in\mathbb{N}$ に対して

$$h_m(x):=m(d(x,F)\wedge 1)$$

と定めると，$h_m\in C_{\mathrm{b}}(\mathcal{X})$ であり，$m\to\infty$ のときに各 x で $h_m(x)\nearrow\psi(x)$ と広義単調収束する．よって明らかに

$$\varepsilon\log\mu_\varepsilon(F)=\varepsilon\log\int_{\mathcal{X}}e^{-\psi/\varepsilon}d\mu_\varepsilon\leq\varepsilon\log\int_{\mathcal{X}}e^{-h_m/\varepsilon}d\mu_\varepsilon$$

となるので，Laplace 原理を用いて次の評価を得る．

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(F)\leq\lim_{\varepsilon\searrow 0}\varepsilon\log\int_{\mathcal{X}}e^{-h_m/\varepsilon}d\mu_\varepsilon=-\inf_{x\in\mathcal{X}}\{h_m(x)+I(x)\}$$

したがって

$$\lim_{m\to\infty}\inf_{x\in\mathcal{X}}\{h_m(x)+I(x)\}=\inf_{x\in F}I(x)\tag{1.4.5}$$

を示せば証明が終わる．(1.4.5) の左辺が右辺を超えないことはほぼ自明である．実際

$$\inf_{x\in\mathcal{X}}\{h_m(x)+I(x)\}\leq\inf_{x\in\mathcal{X}}\{\psi(x)+I(x)\}=\inf_{x\in F}I(x)$$

に注意すればよい．よって逆向きの不等式

$$\lim_{m\to\infty}\inf_{x\in\mathcal{X}}\{h_m(x)+I(x)\}\geq\inf_{x\in F}I(x)\tag{1.4.6}$$

を示すことに問題は帰着された．$\inf_{x\in F}I(x)>0$ の場合のみ示せばよい．

(i) まず $\inf_{x\in F} I(x)=\infty$，すなわち F 上で $I\equiv\infty$ となる場合を考える．任意の十分大きな $R\in(1,\infty)$ を取る．仮定により $\{I\le R\}:=\{x\in\mathcal{X}\mid I(x)\le R\}$ はコンパクトである．

$$\begin{aligned}\inf_{x\in\mathcal{X}}\{h_m(x)+I(x)\}&=\inf_{x\in\{I\le R\}}\{h_m(x)+I(x)\}\wedge\inf_{x\in\{I\le R\}^c}\{h_m(x)+I(x)\}\\&\ge\inf_{x\in\{I\le R\}}\{m(d(x,F)\wedge 1)+I(x)\}\wedge R\end{aligned}$$

となる．コンパクト集合 $\{I\le R\}$ と閉集合 F は互いに素であるので，距離空間の一般論により 2 集合間の距離 $d(\{I\le R\},F):=\inf\{d(y,z)\mid y\in\{I\le R\},\,z\in F\}$ は真に正である．よって

$$\lim_{m\to\infty}\inf_{x\in\mathcal{X}}\{h_m(x)+I(x)\}\ge\lim_{m\to\infty}m\{d(\{I\le R\},F)\wedge 1\}\wedge R=R$$

となる．$R\nearrow\infty$ として，この場合について (1.4.6) を得る．

(ii) 次に $\inf_{x\in F} I(x)\in(0,\infty)$ となる場合を考える．簡単のために $c:=\inf_{x\in F} I(x)$ と表す．F 上で $h_m\equiv 0$ なので

$$\inf_{x\in\mathcal{X}}\{h_m(x)+I(x)\}=c\wedge\inf_{x\in F^c}\{h_m(x)+I(x)\}$$

が成り立つ．したがって，この場合に (1.4.6) を得るためには

$$\lim_{m\to\infty}\inf_{x\in F^c}\{h_m(x)+I(x)\}\ge c \tag{1.4.7}$$

を示せば十分である．

$\kappa\in(0,1)$ を任意に取り $O:=\{I>c-\kappa\}$ とおくと，O は F を含む開集合である．また $K:=\{I\le c+1\}$ とおくとこれはコンパクトで，$K\cap F\neq\emptyset$ を満たす．このとき，

$$F^c=(F^c\cap K^c)\cup(F^c\cap K)=(F^c\cap K^c)\cup\{(O\setminus F)\cap K\}\cup(O^c\cap K)$$

と 3 つの互いに素な部分集合に分割される．$(F^c\cap K^c)\cup\{(O\setminus F)\cap K\}$ 上では I の値は $c-\kappa$ を超えることに注意せよ．さらに

$$d(F,O^c\cap K):=\inf\{d(y,z)\mid y\in F,\,z\in O^c\cap K\}>0$$

が成り立つ．実際，F は閉で $O^c\cap K$ はコンパクトであり，$F\subset O$ によりそれらは互いに素であるため，距離空間の一般論が再び使える．

以上をまとめて

$$\begin{aligned}\inf_{x\in F^c}\{h_m(x)+I(x)\}&=\inf_{x\in F^c\cap K^c}\{h_m(x)+I(x)\}\\&\quad\wedge\inf_{x\in(O\setminus F)\cap K}\{h_m(x)+I(x)\}\wedge\inf_{x\in O^c\cap K}\{h_m(x)+I(x)\}\end{aligned}$$

$$
\begin{aligned}
&\geq (c-\kappa) \wedge \inf_{x \in O^c \cap K} \{m(d(x,F) \wedge 1) + I(x)\} \\
&\geq (c-\kappa) \wedge m(d(O^c \cap K, F) \wedge 1)
\end{aligned}
$$

を得る．ここで $m \to \infty$ とした後に $\kappa \searrow 0$ とすると (1.4.7) を得る．これで (ii) の場合も (1.4.6) が示せたので，定理 1.4.4 の証明が全て終わった． □

1.5　縮小原理と逆縮小原理

大偏差原理の定義が開集合と閉集合を用いて記述されていることから容易に想像がつくが，大偏差原理は連続写像と相性がよい．すなわち連続写像の定義域での大偏差原理はごく自然に値域に遺伝する．この事実は非常に有名で縮小原理と呼ばれている．ほぼ無条件に成り立つのだが，速度関数が良いことを仮定している点に注意せよ.（この仮定は値域に誘導される関数 J が本当に速度関数であることの証明中で使われている.）

命題 1.5.1. ψ を Hausdorff 空間 $\mathcal{X}$ から Hausdorff 空間 $\mathcal{Y}$ への連続写像とする．$\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon \in (0,1]}$ は $\varepsilon \searrow 0$ のときに良い速度関数 I に対して大偏差原理を満たすとする．このとき

$$
J(y) := \begin{cases} \inf\{I(x) \mid x \in \psi^{-1}(y)\}, & \psi^{-1}(y) \neq \emptyset \text{ のとき}, \\ \infty, & \psi^{-1}(y) = \emptyset \text{ のとき}, \end{cases}
$$

と定義すると，J は $\mathcal{Y}$ 上の良い速度関数であり，像測度の族 $\{\mu_\varepsilon \circ \psi^{-1}\}_{\varepsilon \in (0,1]}$ は $\varepsilon \searrow 0$ のときに J に対して大偏差原理を満たす.

証明. まず次の事実を示す.

$$
J^{-1}([0,\lambda]) = \psi(I^{-1}([0,\lambda])), \qquad \lambda \in [0,\infty). \tag{1.5.1}
$$

右辺はコンパクト集合の連続像なのでコンパクトである．よって，これから直ちに J が良い速度関数であることがわかる．$x \in \mathcal{X}$ のとき $x \in \psi^{-1}(\psi(x))$ なので，$J(\psi(x)) \leq I(x)$ である．特に $x \in I^{-1}([0,\lambda])$ のときは，$J(\psi(x)) \leq \lambda$ すなわち $\psi(x) \in J^{-1}([0,\lambda])$ がわかる．これで $J^{-1}([0,\lambda]) \supset \psi(I^{-1}([0,\lambda]))$ が示せた.

次は $J^{-1}([0,\lambda]) \subset \psi(I^{-1}([0,\lambda]))$ を示す．そのためには任意の $y \in J^{-1}([0,\lambda])$ に対して，$\psi^{-1}(y) \cap I^{-1}([0,\lambda]) \neq \emptyset$ を示せば十分である．$\mathcal{Y}$ の Hausdorff 性と ψ の連続性により $\psi^{-1}(y)$ は空でない閉集合であることに注意せよ．よって，I が良いことと補題 1.2.6 により $J(y)$ の定義に現れる $\psi^{-1}(y)$ 上の下限 inf は最小 min で置き換えてよい．特に $x \in \psi^{-1}(y)$ かつ $I(x) = J(y) \leq \lambda$ を満たす x が存在する．以上で (1.5.1) が示された.

次は大偏差原理を示す．まず J の定義の仕方から，任意の $A \subset \mathcal{Y}$ に対して

$$
\inf\{J(y) \mid y \in A\} = \inf\{I(x) \mid x \in \psi^{-1}(A)\}
$$

となることに注意せよ.（通常どおり $\inf\emptyset=\infty$ とする.）ψ が連続なので，$O\subset\mathcal{Y}$ が開集合のとき，$\psi^{-1}(O)$ も開集合である．I に対する大偏差原理を使うと

$$\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(\psi^{-1}(O))\geq -\inf\{I(x)\mid x\in\psi^{-1}(O)\}=-\inf\{J(y)\mid y\in O\}$$

を得る．同様の議論により，閉集合 $F\subset\mathcal{Y}$ に対して

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(\psi^{-1}(F))\leq -\inf\{J(y)\mid y\in F\}$$

も得られる．これで命題 1.5.1 の証明が終わった．□

$\mathcal{X}$ 上に 2 種類の位相が定義されていて，片方がもう片方より強いとする．また $\{\mu_\varepsilon\}$ は強い方の位相に関する Borel 確率測度の族だとする．仮に弱い方の位相に関して $\{\mu_\varepsilon\}$ が大偏差原理を満たすとしても，強い方の位相に関して大偏差原理を満たすことはもちろん一般には期待できない．しかし，実は強い方の位相に関する指数的緊密性を仮定すると強い方の位相に関する大偏差原理が成り立つ．この事実は逆縮小原理と呼ばれる．やや一般化して，連続な単射に対してこの原理を定式化する．

命題 1.5.2. ψ を Hausdorff 空間 $\mathcal{X}$ から Hausdorff 空間 $\mathcal{Y}$ への連続な単射とする．$\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のときに以下の 2 条件を満たすとする．

- $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{X}$ 上で指数的に緊密である．
- $\{\mu_\varepsilon\circ\psi^{-1}\}_{\varepsilon\in(0,1]}$ は $\mathcal{Y}$ 上で速度関数 J に対して弱大偏差原理を満たす．

このとき，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{X}$ 上で良い速度関数 $I:=J\circ\psi$ に対して大偏差原理を満たす．

証明．明らかに I は速度関数である．$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{X}$ 上で I に対して弱大偏差原理を満たすことを示せば十分である．なぜならば命題 1.3.1 により，これから所望の大偏差原理と I が良いことが共に従うからである．

$K\subset\mathcal{X}$ を任意のコンパクト集合だとする．ψ は連続なので，$\psi(K)$ もコンパクトである．像測度の族に対して弱大偏差原理を仮定したので，一般に $K\subset\psi^{-1}(\psi(K))$ であることと合わせて

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(K)\leq\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon\circ\psi^{-1}(\psi(K))\leq -\inf_{y\in\psi(K)}J(y)=-\inf_{x\in K}I(x)$$

と上からの評価を得る．

下からの評価を示す．$O\subset\mathcal{X}$ を任意の開集合とし，$I(x)<\infty$ となる $x\in O$ を任意に取る．指数的緊密性により，あるコンパクト集合 K で

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(K^c)<-I(x)\tag{1.5.2}$$

を満たすものが存在する．$\psi(K)$ は $\mathcal{Y}$ のコンパクト集合なので，像測度の族に対する弱大偏差原理により

$$\begin{aligned}
-\inf_{z\in K^c} I(z) &= -\inf_{y\in\psi(K^c)} J(y) \\
&\le -\inf_{y\in\psi(K)^c} J(y) \\
&\le \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon \circ \psi^{-1}(\psi(K)^c) \\
&\le \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(K^c) < -I(x)
\end{aligned}$$

を得る．特に $x \in K$ である．ψ の単射性により $\psi(K^c) \subset \psi(K)^c$ となることと，一般に $\psi^{-1}(\psi(K)^c) \subset K^c$ となることも用いている．

ψ を K に制限した $\psi\restriction_K$ はコンパクト集合 K から $\psi(K)$ への連続な全単射なので，実は同相写像であることに注意せよ．これは K の閉部分集合はコンパクトなので，その像は $\psi(K)$ のコンパクト部分集合，したがって $\psi(K)$ の閉部分集合になるためである．よって $\psi\restriction_K$ の逆写像は連続である．

$\psi(O\cap K)$ は $\psi(K)$ 内の開集合なので，相対位相の定義によりある開集合 $U \subset \mathcal{Y}$ が存在して $\psi(x) \in \psi(O\cap K) = U \cap \psi(K)$ となる．明らかに $U \cap \psi(O^c \cap K) = \emptyset$ であり，$\psi^{-1}(U) \cap (O^c \cap K) = \emptyset$ である．よって

$$\psi^{-1}(U) = (\psi^{-1}(U)\cap O) \cup (\psi^{-1}(U)\cap O^c \cap K^c) \cup (\psi^{-1}(U)\cap O^c\cap K) \subset O \cup K^c$$

という包含関係に注意すると，

$$\mu_\varepsilon \circ \psi^{-1}(U) \le \mu_\varepsilon(O) + \mu_\varepsilon(K^c)$$

という不等式を得る．弱大偏差原理と下極限に関する補題 1.2.3 を用いて

$$-I(x) = -J(\psi(x)) \le \varliminf_{\varepsilon\searrow 0} \varepsilon\log\mu_\varepsilon \circ \psi^{-1}(U) \le \left[\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(O)\right] \vee \left[\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(K^c)\right]$$

を得る．これを (1.5.2) と合わせると

$$-I(x) \le \varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon(O), \qquad x \in O \cap I^{-1}([0,\infty))$$

とわかる．この式で x に関する上限を取れば，所望の下からの評価を得る． □

1.6 指数的な近似

本節の目的は大偏差原理に対する指数的に良い近似を証明することである．これは連続写像に対する縮小原理を一般化したものだと解釈できる．

本節を通じて $(\mathcal{Y}, d)$ を可分距離空間とする.[19] $\mathcal{Y} \times \mathcal{Y}$ において，対角線の近傍の補集合を以下のように書く.

$$\Gamma_\delta := \{(\widetilde{y}, y) \in \mathcal{Y} \times \mathcal{Y} \mid d(\widetilde{y}, y) > \delta\}, \qquad \delta \in (0, \infty).$$

これは定め方から開集合である．通常どおり $\mathcal{Y}$ と $\mathcal{Y} \times \mathcal{Y}$ の可測構造としては Borel 加法族を考える．可分性により，Z_1 と Z_2 が共にある共通の確率空間上で定義された $\mathcal{Y}$ 値確率変数であれば，(Z_1, Z_2) は $\mathcal{Y} \times \mathcal{Y}$ 値確率変数であることがわかる.[20]

定義 1.6.1. 各 $\varepsilon \in (0,1]$ に対して，$(\Omega, \mathcal{B}_\varepsilon, \mathbb{Q}_\varepsilon)$ を確率空間とし，$\widetilde{Z}_\varepsilon$ と Z_ε をこの確率空間上で定義された $\mathcal{Y}$ 値確率変数だとする．任意の $\delta \in (0, \infty)$ に対して

$$\varlimsup_{\varepsilon \searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon((\widetilde{Z}_\varepsilon, Z_\varepsilon) \in \Gamma_\delta) = -\infty$$

となるとき，$\mathcal{Y}$ 値確率変数の族 $\{\widetilde{Z}_\varepsilon\}_{\varepsilon \in (0,1]}$ と $\{Z_\varepsilon\}_{\varepsilon \in (0,1]}$ は指数的に同値であるという.

またY上の Borel 確率測度の族 $\{\widetilde{\mu}_\varepsilon\}_{\varepsilon \in (0,1]}$ と $\{\mu_\varepsilon\}_{\varepsilon \in (0,1]}$ が指数的に同値であるとは，上記のような $\{(\Omega, \mathcal{B}_\varepsilon, \mathbb{Q}_\varepsilon)\}_{\varepsilon \in (0,1]}$ および $\{\widetilde{Z}_\varepsilon\}_{\varepsilon \in (0,1]}$，$\{Z_\varepsilon\}_{\varepsilon \in (0,1]}$ で，各 $\varepsilon \in (0,1]$ に対して $\mathbb{Q}_\varepsilon \circ \widetilde{Z}_\varepsilon^{-1} = \widetilde{\mu}_\varepsilon$ および $\mathbb{Q}_\varepsilon \circ Z_\varepsilon^{-1} = \mu_\varepsilon$ を満たすものが存在することをいう.

少なくとも速度関数が良い場合には，指数的に同値である 2 つの確率測度の族は大偏差原理の観点からは区別できない．このことを次の命題で確認する.

命題 1.6.2. 上記の定義 1.6.1 にある状況を仮定する．さらに $\varepsilon \searrow 0$ とするとき $\{\mu_\varepsilon\}_{\varepsilon \in (0,1]}$ は良い速度関数 I に対して大偏差原理を満たすとする．このとき $\{\widetilde{\mu}_\varepsilon\}_{\varepsilon \in (0,1]}$ も全く同じ大偏差原理を満たす.

証明. $\mathcal{Y}$ の閉集合 F と $\delta \in (0,1)$ を任意に取り，$F^\delta := \{y \in \mathcal{Y} \mid d(y, F) \le \delta\}$ とおく．明らかに

$$\begin{aligned} \widetilde{Z}_\varepsilon \in F \quad &\Longleftrightarrow \quad \text{“}\widetilde{Z}_\varepsilon \in F \text{ かつ } d(\widetilde{Z}_\varepsilon, Z_\varepsilon) \le \delta\text{”} \quad \text{または} \quad \text{“}\widetilde{Z}_\varepsilon \in F \text{ かつ } d(\widetilde{Z}_\varepsilon, Z_\varepsilon) > \delta\text{”} \\ &\Longrightarrow \qquad Z_\varepsilon \in F^\delta \quad \text{または} \quad d(\widetilde{Z}_\varepsilon, Z_\varepsilon) > \delta \end{aligned}$$

であるから，

$$\mathbb{Q}_\varepsilon(\widetilde{Z}_\varepsilon \in F) \le \mathbb{Q}_\varepsilon(Z_\varepsilon \in F^\delta) + \mathbb{Q}_\varepsilon(d(\widetilde{Z}_\varepsilon, Z_\varepsilon) > \delta)$$

が成り立つ．この両辺の $\varepsilon \log$ を取り，指数的に同値であることを使うと

$$\varlimsup_{\varepsilon \searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon(\widetilde{Z}_\varepsilon \in F) \le \left[\varlimsup_{\varepsilon \searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon(Z_\varepsilon \in F^\delta)\right] \vee \left[\varlimsup_{\varepsilon \searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon(d(\widetilde{Z}_\varepsilon, Z_\varepsilon) > \delta)\right]$$

[19] [33, 第 4.2.2 節] では一般の距離空間で議論しているが，本書では簡単のために可分性を仮定する.

[20] もちろん可測性を気にしている．距離空間に対しては可分性と Lindelöf 性が同値であることを使うと簡単に示せる.

$$\leq \left[-\inf_{y\in F^\delta} I(y)\right] \vee (-\infty) = -\inf_{y\in F^\delta} I(y)$$

を得る．ここで補題 1.2.2 も用いた．I が良いことと補題 1.2.8 により，右辺は $\delta \searrow 0$ のとき $-\inf_{y\in F} I(y)$ に収束する．これで上からの評価が示せた．

下からの評価を示す．$\mathcal{Y}$ の開集合 O および $x \in O$ を任意に取る．$\delta > 0$ を $B(x, 2\delta) \subset O$ となるように選ぶ．このとき

$$Z_\varepsilon \in B(x,\delta) \quad \Longrightarrow \quad \widetilde{Z}_\varepsilon \in O \ \text{ または } \ d(\widetilde{Z}_\varepsilon, Z_\varepsilon) > \delta$$

であるから，

$$\mathbb{Q}_\varepsilon(Z_\varepsilon \in B(x,\delta)) \leq \mathbb{Q}_\varepsilon(\widetilde{Z}_\varepsilon \in O) + \mathbb{Q}_\varepsilon(d(\widetilde{Z}_\varepsilon, Z_\varepsilon) > \delta)$$

となる．このとき指数的に同値であることを使うと，この式から

$$\begin{aligned}
-I(x) &\leq -\inf_{y\in B(x,\delta)} I(y) \\
&\leq \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon(Z_\varepsilon \in B(x,\delta)) \\
&\leq \left[\varliminf_{\varepsilon\searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon(\widetilde{Z}_\varepsilon \in O)\right] \vee \left[\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon(d(\widetilde{Z}_\varepsilon, Z_\varepsilon) > \delta)\right] \\
&= \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mathbb{Q}_\varepsilon(\widetilde{Z}_\varepsilon \in O)
\end{aligned}$$

が従う．なお下極限に関する補題 1.2.3 も用いた．最後に x に関する上限を取れば，求める下からの評価が得られる． □

ここで本節の主題である指数的に良い近似と呼ばれる概念を導入しよう．

定義 1.6.3. 各 $\varepsilon \in (0,1]$ と $m \in \mathbb{N}$ に対して $(\Omega, \mathcal{B}_\varepsilon, \mathbb{Q}_{\varepsilon,m})$ を確率空間とし，$\widetilde{Z}_\varepsilon$ と $Z_{\varepsilon,m}$ をこの確率空間上で定義された $\mathcal{Y}$ 値確率変数だとする．さらに任意の $\delta \in (0,\infty)$ に対して

$$\lim_{m\to\infty} \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{Q}_{\varepsilon,m}((\widetilde{Z}_\varepsilon, Z_{\varepsilon,m}) \in \Gamma_\delta) = -\infty$$

となるとき，$\{Z_{\varepsilon,m}\}_{\varepsilon\in(0,1],m\in\mathbb{N}}$ は $\{\widetilde{Z}_\varepsilon\}_{\varepsilon\in(0,1]}$ の指数的に良い近似であるという．

また $\mathcal{Y}$ 上の Borel 確率測度の族 $\{\mu_{\varepsilon,m}\}_{\varepsilon\in(0,1],m\in\mathbb{N}}$ と $\{\widetilde{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ に対して，前者が後者の指数的に良い近似であるとは，上記のような $\{(\Omega, \mathcal{B}_\varepsilon, \mathbb{Q}_{\varepsilon,m})\}_{\varepsilon\in(0,1],m\in\mathbb{N}}$ および $\{Z_{\varepsilon,m}\}_{\varepsilon\in(0,1],m\in\mathbb{N}}$, $\{\widetilde{Z}_\varepsilon\}_{\varepsilon\in(0,1]}$ で，任意の ε, m に対して，$\mathbb{Q}_{\varepsilon,m} \circ \widetilde{Z}_\varepsilon^{-1} = \widetilde{\mu}_\varepsilon$ および $\mathbb{Q}_{\varepsilon,m} \circ Z_{\varepsilon,m}^{-1} = \mu_{\varepsilon,m}$ を満たすものが存在することをいう．

次の命題が本節の主目的である．指数的に良い近似が成立している状況では，近似する確率測度の族に対する大偏差原理が，極限として得られる確率測度の族に遺伝する．なお本命題は前述の命題 1.6.2 を特殊な場合として含むことを注意せよ．($Z_{\varepsilon,m}$ や $\mu_{\varepsilon,m}$ が近似の番号 $m \in \mathbb{N}$ に依存しない場合だと思えばよい.)

命題 1.6.4. $\{\mu_{\varepsilon,m}\}_{\varepsilon\in(0,1],m\in\mathbb{N}}$ と $\{\widetilde{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{Y}$ 上の Borel 確率測度の族とし，前者は後者を指数的に良く近似すると仮定する．また任意の $m\in\mathbb{N}$ に対して $\{\mu_{\varepsilon,m}\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のとき速度関数 I_m に対して大偏差原理を満たすとする．ここで $I\colon\mathcal{Y}\to[0,\infty]$ を

$$I(x)=\sup_{\delta>0}\varliminf_{m\to\infty}\inf\{I_m(z)\mid z\in B(x,\delta)\} \tag{1.6.1}$$

と定める．このとき，次が成り立つ．

(1) $\{\widetilde{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のとき速度関数 I に対して弱大偏差原理を満たす．

(2) さらに I が良い速度関数であることと，任意の閉集合 $F\subset\mathcal{Y}$ に対して

$$\inf\{I(x)\mid x\in F\}\le\varliminf_{m\to\infty}\inf\{I_m(x)\mid x\in F\} \tag{1.6.2}$$

を満たすことを仮定すると，$\{\widetilde{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ のとき I に対して大偏差原理を満たす．

証明. 与えられた $\{\mu_{\varepsilon,m}\}$ と $\{\widetilde{\mu}_\varepsilon\}$ に対して，$\{\widetilde{Z}_\varepsilon\}$，$\{Z_{\varepsilon,m}\}$，$\{(\Omega,\mathcal{B}_\varepsilon,\mathbb{Q}_{\varepsilon,m})\}$ を定義 1.6.3 にあるように取る．本証明中では $\Gamma'_\delta=\{\omega\in\Omega\mid(\widetilde{Z}_\varepsilon,Z_{\varepsilon,m})\in\Gamma_\delta\}$ と略記する．(この事象が ε,m に依存することは明示しない．)

命題 1.3.2 により，(1) を証明するには任意の $x\in\mathcal{Y}$ に対して次の等式を示せば十分である．

$$I(x)=-\inf_{\delta>0}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(B(x,\delta))=-\inf_{\delta>0}\varliminf_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(B(x,\delta)) \tag{1.6.3}$$

まず δ,ε,m を固定するごとに，

$$\{Z_{\varepsilon,m}\in B(x,\delta)\}\subset\{\widetilde{Z}_\varepsilon\in B(x,2\delta)\}\cup\{(\widetilde{Z}_\varepsilon,Z_{\varepsilon,m})\in\Gamma_\delta\}$$

と包含されるため，

$$\mu_{\varepsilon,m}(B(x,\delta))\le\widetilde{\mu}_\varepsilon(B(x,2\delta))+\mathbb{Q}_{\varepsilon,m}(\Gamma'_\delta)$$

と評価できることに注意せよ．

次は $\{\mu_{\varepsilon,m}\}$ に対する大偏差原理の下からの評価を用いて

$$\begin{aligned}-\inf\{I_m(z)\mid z\in B(x,\delta)\}&\le\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mu_{\varepsilon,m}(B(x,\delta))\\&\le\varliminf_{\varepsilon\searrow 0}\varepsilon\log(\widetilde{\mu}_\varepsilon(B(x,2\delta))+\mathbb{Q}_{\varepsilon,m}(\Gamma'_\delta))\\&\le\varliminf_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(B(x,2\delta))\vee\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{Q}_{\varepsilon,m}(\Gamma'_\delta)\end{aligned}$$

とわかる．ここで補題 1.2.3 を用いた．ここで m に関する極限を取り，指数的に良い近似であることを使うと

$$\varlimsup_{m\to\infty}\Big\{-\inf\{I_m(z)\mid z\in B(x,\delta)\}\Big\}\le\varliminf_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(B(x,2\delta)).$$

を得る．

この一連の議論における $Z_{\varepsilon,m}$ と $\widetilde{Z}_\varepsilon$ の役割を交換して同様に評価すると，

$$\widetilde{\mu}_\varepsilon(B(x,\delta)) \leq \mu_{\varepsilon,m}(B(x,2\delta)) + \mathbb{Q}_{\varepsilon,m}(\Gamma'_\delta)$$

の極限を取ることになり，大偏差原理の上からの評価と補題 1.2.2 を用いて次の不等式も得る．

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(B(x,\delta)) \leq \varliminf_{m\to\infty}\Big\{-\inf\{I_m(z) \mid z\in\overline{B(x,2\delta)}\}\Big\}.$$

これら2つの不等式を合わせると，各 $\delta>0$ と $x\in\mathcal{Y}$ に対して

$$\begin{aligned}
\varlimsup_{m\to\infty}\Big\{-\inf\{I_m(z) \mid z\in B(x,\delta/2)\}\Big\} &\leq \varliminf_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(B(x,\delta))\\
&\leq \varlimsup_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(B(x,\delta))\\
&\leq \varliminf_{m\to\infty}\Big(-\inf\{I_m(z) \mid z\in\overline{B(x,2\delta)}\}\Big)\\
&\leq \varliminf_{m\to\infty}\Big(-\inf\{I_m(z) \mid z\in B(x,3\delta)\}\Big)\\
&\leq \varlimsup_{m\to\infty}\Big(-\inf\{I_m(z) \mid z\in B(x,3\delta)\}\Big)
\end{aligned}$$

となる．ここで $\delta>0$ に関する下限を取ると，最左辺と最右辺が一致するので，結局全てが等しくなる．これで (1.6.3) が示せた．

次は (2) を示す．閉集合 F を任意に取る．δ,ε,m を固定するごとに

$$\{\widetilde{Z}_\varepsilon\in F\}\subset\{Z_{\varepsilon,m}\in F^\delta\}\cup\{(\widetilde{Z}_\varepsilon,Z_{\varepsilon,m})\in\Gamma_\delta\}$$

となる．ここで $F^\delta := \{x\in\mathcal{Y} \mid d(x,F)\leq\delta\}$ も閉集合である．ここで大偏差原理の上からの評価と補題 1.2.2 を用いると次の評価を得る．

$$\begin{aligned}
\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(F) &\leq \varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_{\varepsilon,m}(F^\delta)\vee\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{Q}_{\varepsilon,m}(\Gamma'_\delta)\\
&\leq \Big(-\inf\{I_m(x) \mid x\in F^\delta\}\Big)\vee\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{Q}_{\varepsilon,m}(\Gamma'_\delta).
\end{aligned}$$

ここで $m\to\infty$ として，指数的に良い近似であることと仮定 (1.6.2) を用いると

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\widetilde{\mu}_\varepsilon(F)\leq -\varliminf_{m\to\infty}\inf\{I_m(x) \mid x\in F^\delta\}\leq -\inf\{I(x) \mid x\in F^\delta\}$$

となる．最後に I が良いことと補題 1.2.8 により

$$\inf\{I(x) \mid x\in F\} = \lim_{\delta\searrow 0}\inf\{I(x) \mid x\in F^\delta\}$$

が成り立つことに注意すれば，証明が終わる． □

縮小原理（命題1.5.1）において，連続写像の定義域で成り立つ大偏差原理は自然に値域に遺伝することを見た．もちろん写像が連続でないときはこの種のことは一般には期待できない．それでも適切な意味で連続写像に「近い」写像であれば，縮小原理が成り立つことを主張するのが次の命題である．値域に誘導される速度関数 I' は縮小原理に現れるものと同じ形をしていることに注意せよ．

命題 1.6.5. Hausdorff 空間 $\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は良い速度関数 I に対して，$\varepsilon \searrow 0$ のときに大偏差原理を満たすとする．各 $m \in \mathbb{N}$ に対して，$\psi_m\colon \mathcal{X} \to \mathcal{Y}$ は $\mathcal{X}$ から可分距離空間 $(\mathcal{Y}, d)$ への連続写像だとする．さらに $\psi\colon \mathcal{X} \to \mathcal{Y}$ は以下の条件を満たす可測写像だとする．

$$\varlimsup_{m\to\infty} \sup\{d(\psi_m(x), \psi(x)) \mid x \in \mathcal{X},\ I(x) \le \lambda\} = 0, \qquad \lambda \in [0,\infty). \tag{1.6.4}$$

$\mathcal{Y}$ 上の Borel 確率測度の族 $\{\widetilde{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ が $\{\mu_\varepsilon \circ \psi_m^{-1}\}_{\varepsilon\in(0,1],m\in\mathbb{N}}$ により指数的に良く近似されるならば，$\{\widetilde{\mu}_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のとき良い速度関数 I' に対して大偏差原理を満たす．ただし，$I'(y) = \inf\{I(x) \mid x \in \psi^{-1}(y)\}$ と定める．（通常どおりに $\inf\emptyset = \infty$ とおく．）

証明. 各 m に対して ψ_m は連続なので，縮小原理（命題1.5.1）により $\{\mu_\varepsilon \circ \psi_m^{-1}\}_{\varepsilon\in(0,1]}$ は大偏差原理を次の速度関数に対して満たす．

$$I'_m(y) = \inf\{I(x) \mid x \in \psi_m^{-1}(y)\}.$$

条件 (1.6.4) における一様収束により，ψ は集合 $I^{-1}([0,\lambda])$ 上で連続である．この事実を用いると，縮小原理の証明と同じ議論で

$$\psi(I^{-1}([0,\lambda])) = (I')^{-1}([0,\lambda]) \tag{1.6.5}$$

となることが示せるので，I' が $\mathcal{Y}$ 上の良い速度関数であることがわかる．

本証明はやや複雑なので，ここで証明の流れを説明しておく．もちろん上述の命題1.6.4を用いて大偏差原理を示すのだが，その命題中の速度関数 (1.6.1) に相当するのは，この場合は

$$\widetilde{I}(y) := \sup_{\delta>0} \varliminf_{m\to\infty} \inf\{I'_m(z) \mid z \in B(y,\delta)\}$$

であることに注意せよ．以下のように議論が進む．

- すでに最初の段落で示したように，I' が良いことを確認する．
- 命題1.6.4中の十分条件 (1.6.2) が（$\widetilde{I}$ ではなく）I' に対して成立することを見る．
- $I' = \widetilde{I}$ を示す．
- $\{\widetilde{\mu}_\varepsilon\}$ と $\{\mu_\varepsilon \circ \psi_m^{-1}\}$ に対して命題1.6.4を適用する．

ここから証明を再開する．閉集合 $F \subset \mathcal{Y}$ に対して，

$$\gamma_m := \inf\{I'_m(y) \mid y \in F\} = \inf\{I(x) \mid x \in \psi_m^{-1}(F)\}$$

および $\gamma := \underline{\lim}_{m\to\infty} \gamma_m$ とおく．このとき，任意の $\delta > 0$ に対して

$$\inf\{I'(y) \mid y \in F^\delta\} \le \gamma \tag{1.6.6}$$

となることを示す．ただし $F^\delta := \{y \in \mathcal{Y} \mid d(y, F) \le \delta\}$ である．$\gamma = \infty$ のときは自明なので，以下では $\gamma < \infty$ と仮定する．$\{\gamma_m\}$ の適当な部分列を選び，$\gamma_m \to \gamma$ とできる.（この部分列を再び同じ記号で表す.）さらに $\sup_{m\in\mathbb{N}} \gamma_m \le \gamma + 1$ と仮定してよい．$\psi_m^{-1}(F)$ が空でない閉集合であり I が良いため，補題 1.2.6 により $x_m \in \psi_m^{-1}(F)$ で $I(x_m) = \gamma_m$ となるものが存在することに注意せよ．

条件 (1.6.4) により $I^{-1}([0, \gamma+1])$ 上で ψ_m は ψ に一様収束するため，与えられた $\delta > 0$ に応じて m を十分大きく取ると $\psi(x_m) \in F^\delta$ となる．よって，

$$\inf\{I'(y) \mid y \in F^\delta\} \le I'(\psi(x_m)) \le I(x_m) = \gamma_m$$

であり，$m \to \infty$ として (1.6.6) が得られる．I' が良いので，補題 1.2.8 により

$$\inf\{I'(y) \mid y \in F\} = \lim_{\delta\searrow 0} \inf\{I'(y) \mid y \in F^\delta\} \tag{1.6.7}$$

である．これと (1.6.6) から命題 1.6.4 中の条件 (1.6.2) が I' に対して成立することがわかる．

さらに $y \in \mathcal{Y}$ のとき次の不等式が成り立つ．

$$\begin{aligned}
I'(y) &= \sup_{\delta>0} \inf\{I'(z) \mid z \in \overline{B}(y, 2\delta)\} \\
&\le \sup_{\delta>0} \inf\{I'(z) \mid z \in [\overline{B}(y, \delta)]^\delta\} \\
&\le \sup_{\delta>0} \underline{\lim}_{m\to\infty} \inf\{I'_m(z) \mid z \in B(y, \delta)\} = \widetilde{I}(y).
\end{aligned}$$

最初の等号では (1.6.7) を１点集合に対して用いた．次の不等号では $\overline{B}(y, 2\delta) \supset [\overline{B}(y, \delta)]^\delta$ を用いた．最後の不等号では (1.6.6) を $F = \overline{B}(y, \delta)$ に対して用いた．

次は逆向きの不等式 $\widetilde{I}(y) \le I'(y)$ を各 y に対して示す．簡単のため $\alpha := I'(y)$ とおく．$\alpha < \infty$ と仮定してよい．(1.6.5) により，$I(x) \le \alpha$ かつ $\psi(x) = y$ となる $x \in \mathcal{X}$ が存在する．定義の仕方から $I'_m(\psi_m(x)) \le \alpha$ である．再び (1.6.4) により，$m \to \infty$ のとき $\psi_m(x) \to \psi(x) = y$ となるため，$\widetilde{I}$ の定義により $\widetilde{I}(y) \le \underline{\lim}_{m\to\infty} I'_m(\psi_m(x)) \le \alpha$ となる．以上で $\widetilde{I}(y) = I'(y)$ が示せた．特に条件 (1.6.2) は $\widetilde{I}$ に対しても成立する．

以上の議論と指数的に良い近似であるという仮定により，$\{\widetilde{\mu}_\varepsilon\}$ と $\{\mu_\varepsilon \circ \psi_m^{-1}\}$ に対して命題 1.6.4 が使えて，$\{\widetilde{\mu}_\varepsilon\}$ は大偏差原理を $\widetilde{I}$ に対して満たすことがわかる．$\square$

1.7　大偏差評価から大偏差原理へ

本節では縮小原理（命題 1.5.1）を写像が連続でない場合に一般化する結果をもう１つ紹介する．この手法は大偏差評価と呼ばれる不等式を経由するもので，確率微分方程式に対する Freidlin-Wentzell の大偏差原理を証明するために Azencott により開発された.[21]（ここで言う大偏差評価とは下記の (1.7.1) のような不等式を意味する.）

以下では $(\mathcal{X}, d_{\mathcal{X}})$ と $(\mathcal{Y}, d_{\mathcal{Y}})$ を距離空間とし，確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上で定義された確率変数の族 $\{\Phi_\varepsilon\}_{\varepsilon\in(0,1]}$ と $\{\Lambda_\varepsilon\}_{\varepsilon\in(0,1]}$ はそれぞれ $\mathcal{X}$ と $\mathcal{Y}$ に値を取るとする．このような一般的な設定の下で，大偏差評価から大偏差原理が従うことを証明する．

命題 1.7.1. 上記の設定の下で，以下の３条件を仮定する．
(1) $\varepsilon \searrow 0$ とするとき，$\{\Phi_\varepsilon\}_{\varepsilon\in(0,1]}$ は良い速度関数 $I\colon \mathcal{X} \to [0,\infty]$ に対して大偏差原理を満たす.
(2) 写像 $\psi\colon I^{-1}([0,\infty)) \to \mathcal{Y}$ は次の性質を持つ．各 $r \in (0,\infty)$ に対して，ψ の $I^{-1}([0,r])$ への制限が連続である．
(3) 任意の $R, \delta, r \in (0,\infty)$ に対して，ある $\rho \in (0,\infty)$ と $\varepsilon_0 \in (0,1]$ が存在して

$$\mathbb{P}\left(d_{\mathcal{Y}}(\Lambda_\varepsilon, \psi(h)) > \delta,\ d_{\mathcal{X}}(\Phi_\varepsilon, h) \le \rho\right) \le e^{-R/\varepsilon}, \qquad h \in I^{-1}([0,r]), \varepsilon \in (0,\varepsilon_0] \tag{1.7.1}$$

を満たす.（ここで ρ と ε_0 は R, δ, r には依存するが，$h \in I^{-1}([0,r])$ と $\varepsilon \in (0,\varepsilon_0]$ には依存しない.）

このとき，$\{\Lambda_\varepsilon\}_{\varepsilon\in(0,1]}$ は次で定まる良い速度関数 $J\colon \mathcal{Y} \to [0,\infty]$ に対して，$\varepsilon \searrow 0$ のときに大偏差原理を満たす.

$$J(g) := \inf\{I(h) \mid h \in I^{-1}([0,\infty)),\ \psi(h) = g\}. \tag{1.7.2}$$

なお通常どおり $\inf \emptyset = \infty$ とおく．

注意 1.7.2. 命題 1.7.1 中の不等式 (1.7.1) にはパラメータが５つも現れるので，一見すると非常に複雑に見える．この点に関していくつか簡単な注意をする．

1. 実際には R と r は十分大きな場合のみ，δ は十分小さな場合のみ考えても同値である．

2. パラメータが多すぎて意味が取りづらい場合は，R と ρ にのみ着目するとよい．任意の十分大きな $R \gg 1$ に対して，(1.7.1) を成立させるような $\rho > 0$ が存在するかどうかを問題にしていると理解すべきである.（この項目の説明は厳密ではない.）

[21] おそらく [48] が最初の論文である．

3. この命題は連続写像に対する縮小原理（命題 1.5.1）の一般化だと思えることを説明する．ψ が $\mathcal{X}$ 全体から $\mathcal{Y}$ への連続写像に拡張できて，$\Lambda_\varepsilon = \psi(\Phi_\varepsilon)$ となっている場合を考えよう.（拡張した写像も再び ψ と表す.）この場合，縮小原理によれば $\{\Lambda_\varepsilon\}$ は大偏差原理を満たし，速度関数は (1.7.2) で与えられるのであった．命題 1.7.1 が縮小原理を含んでいることを見るには (1.7.1) を確認すればよいが，そのためには次に注意せよ.（ここで ρ が h に依存しないことが重要である.）

$$\begin{aligned}&\text{任意の } \delta, r \in (0,\infty) \text{ に対して，ある } \rho \in (0,\infty) \text{ が存在して}\\ &\qquad \widetilde{h} \in \mathcal{X},\ h \in I^{-1}([0,r]), d_{\mathcal{X}}(\widetilde{h}, h) \le \rho \implies d_{\mathcal{Y}}(\psi(\widetilde{h}), \psi(h)) \le \delta.\end{aligned}$$

この事実は明らかではないが,「コンパクト距離空間上の連続関数は一様連続である」ことの証明と同様の論法で証明できるので，ここでは認めることにする．所与の r と δ に対してこの ρ を選べば，(1.7.1) の左辺は 0 になるため，任意の R に対して (1.7.1) は成立する．

4. ψ が $\mathcal{X}$ 全体から $\mathcal{Y}$ への可測写像に拡張できて，$\Lambda_\varepsilon = \psi(\Phi_\varepsilon)$ となっている場合を考える．この場合は (1.7.1) の左辺は点 h における「不連続性の程度」を，ある確率論的な意味で測定していると解釈できる．その程度が「指数的な尺度で計っても任意に小さくあるべき」という要求が条件 (1.7.1) の直感的な内容である．

以上で注意 1.7.2 を終える．

命題 1.7.1 の証明. まず $J(g) < \infty$ のとき，$g = \psi(h)$ かつ $J(g) = I(h)$ を満たす $h \in \mathcal{X}$ が存在することを見る.（つまりこの場合は (1.7.2) の右辺の inf は min でおき換えられる.）このとき $J(g) = \inf\{I(h) \mid h \in I^{-1}([0, J(g)+1]) \cap \psi^{-1}(g)\}$ である．ψ の連続性に関する仮定と I が良いことにより，$I^{-1}([0, J(g)+1]) \cap \psi^{-1}(g)$ は空でないコンパクト集合であるので，補題 1.2.6 により $J(g) = \min\{I(h) \mid h \in I^{-1}([0, J(g)+1]) \cap \psi^{-1}(g)\}$ とわかる．この最小値を実現する h が求めるものである．

次は J が良い速度関数であることを示す．そのためには，任意の $r \in (0,\infty)$ に対して $J^{-1}([0,r]) = \psi(I^{-1}([0,r]))$ を示せば十分である．ここで ψ が $I^{-1}([0,r])$ 上で連続であることと I が良いことを使っている．$J^{-1}([0,r]) \supset \psi(I^{-1}([0,r]))$ は J の定義から明らかであり，$J^{-1}([0,r]) \subset \psi(I^{-1}([0,r]))$ は上の段落で示した事実から明らかである．

下からの評価を示す．$O \subset \mathcal{Y}$ を開集合とする．$\inf_{g' \in O} J(g') < \infty$ と仮定してよい（そうでない場合は自明である）．このとき任意の $n \in \mathbb{N}$ に対して，$J(g) < \inf_{g' \in O} J(g') + 1/n$ となる $g = g_n \in O$ を取り，さらに $g = \psi(h)$ かつ $J(g) = I(h)$ となる $h = h_n \in I^{-1}([0,\infty))$ を取る．$\delta > 0$ を $\overline{B}_{\mathcal{Y}}(g,\delta) \subset O$ となるように 1 つ選ぶ．すると任意の $\rho > 0$ に対して

$$\begin{aligned}\mathbb{P}(\Lambda_\varepsilon \in O) &\ge \mathbb{P}\left(\Lambda_\varepsilon \in \overline{B}_{\mathcal{Y}}(g,\delta)\right)\\ &\ge \mathbb{P}(d_{\mathcal{X}}(\Phi_\varepsilon, h) \le \rho) - \mathbb{P}(d_{\mathcal{Y}}(\Lambda_\varepsilon, \psi(h)) > \delta,\ d_{\mathcal{X}}(\Phi_\varepsilon, h) \le \rho)\end{aligned}$$

となることが簡単に示せる．

第２項については条件 (1.7.1) が使えるため，任意の $R>0$ に対して，$\rho>0$ と $\varepsilon_0\in(0,1]$ を十分小さく取ると

$$\mathbb{P}(\Lambda_\varepsilon\in O)\geq\mathbb{P}(d_{\mathcal{X}}(\Phi_\varepsilon,h)\leq\rho)-e^{-R/\varepsilon},\qquad \varepsilon\in(0,\varepsilon_0]$$

となる．第１項については $\{\Phi_\varepsilon\}$ に対する大偏差原理により，任意の $\rho>0$ に対して

$$\begin{aligned}\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}(d_{\mathcal{X}}(\Phi_\varepsilon,h)\leq\rho)&\geq\varliminf_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}(d_{\mathcal{X}}(\Phi_\varepsilon,h)<\rho)\\&\geq-\inf\{I(k)\mid k\in\mathcal{X},\,d_{\mathcal{X}}(k,h)<\rho\}\\&\geq-I(h)\\&\geq-\inf_{g'\in O}J(g')-\frac{1}{n}\end{aligned}\tag{1.7.3}$$

と評価できる．右辺は ρ に依存しない．

後は $R>\inf_{g'\in O}J(g')+1/n$ のとき

$$\varliminf_{\varepsilon\searrow 0}\varepsilon\log(\mathbb{P}(d_{\mathcal{X}}(\Phi_\varepsilon,h)\leq\rho)-e^{-R/\varepsilon})\geq-\inf_{g'\in O}J(g')-\frac{1}{n}\tag{1.7.4}$$

となることを確認すれば，n は任意なので大偏差原理の下からの評価が得られる．(以下，記号を簡単にするため $a_\varepsilon:=\mathbb{P}(d_{\mathcal{X}}(\Phi_\varepsilon,h)\leq\rho)$ および $c:=\inf_{g'\in O}J(g')+1/n$ と書く．) 任意の $\kappa\in(0,R-c)$ を取る．(1.7.3) により，$\varepsilon_0>0$ を小さく取り直せば，$\varepsilon\in(0,\varepsilon_0]$ のときに $\varepsilon\log a_\varepsilon\geq-c-\kappa$ となり，特に

$$a_\varepsilon-e^{-R/\varepsilon}\geq e^{-(c+\kappa)/\varepsilon}-e^{-R/\varepsilon}=e^{-(c+\kappa)/\varepsilon}(1-e^{-(R-c-\kappa)/\varepsilon})$$

となる．よって，$\varliminf_{\varepsilon\searrow 0}\varepsilon\log(a_\varepsilon-e^{-R/\varepsilon})\geq-c-\kappa$ がわかり，最後に $\kappa\searrow 0$ として (1.7.4) を得る．

上からの評価を示す．$F\subset\mathcal{Y}$ を閉集合とする．$\inf_{g'\in F}J(g')>0$ と仮定してよい（そうでない場合は自明である）．$0<r<\inf_{g'\in F}J(g')$ とする．$J(g)\leq r$ であれば $g\in F^c$ なので，$J^{-1}([0,r])=\psi(I^{-1}([0,r]))\subset F^c$ である．F が閉集合なので，任意の $g\in F^c$ に対して $d_{\mathcal{Y}}(g,F):=\inf_{l\in F}d_{\mathcal{Y}}(g,l)>0$ であるが，$d_{\mathcal{Y}}(g,F)$ の g に関する連続性と $J^{-1}([0,r])$ のコンパクト性により $\inf\{d_{\mathcal{Y}}(g,F)\mid g\in J^{-1}([0,r])\}>0$ である．よって，ある $\delta>0$ が存在して，任意の $h\in I^{-1}([0,r])$ に対して $B_{\mathcal{Y}}(\psi(h),\delta)\subset F^c$ となる（この δ は h に依存しない）．上記の r,δ と任意の $R>0$ に対して，ある $\rho>0$ と $\varepsilon_0\in(0,1]$ が存在して不等式 (1.7.1) が成立することを以下で用いる．

明らかに $\{B_{\mathcal{X}}(h,\rho)\mid h\in I^{-1}([0,r])\}$ はコンパクト集合 $I^{-1}([0,r])$ の開被覆なので，有限個の $h_1,\dots,h_m\in I^{-1}([0,r])$ が存在して $I^{-1}([0,r])\subset\cup_{i=1}^m B_{\mathcal{X}}(h_i,\rho)$ となる．このとき，ほぼ明らかに

$$\mathbb{P}(\Lambda_\varepsilon\in F)\leq\mathbb{P}(\Lambda_\varepsilon\in F,\,\Phi_\varepsilon\in\cup_{i=1}^m B_{\mathcal{X}}(h_i,\rho))+\mathbb{P}(\Phi_\varepsilon\in[\cup_{i=1}^m B_{\mathcal{X}}(h_i,\rho)]^c)$$

$$=: A_1^\varepsilon + A_2^\varepsilon$$

である．$[\cup_{i=1}^m B_{\mathcal{X}}(h_i,\rho)]^c$ は閉集合であり，$I^{-1}((r,\infty])$ に含まれるので，$\{\Phi_\varepsilon\}$ に対する大偏差原理により $\overline{\lim}_{\varepsilon\searrow 0}\varepsilon\log A_2^\varepsilon \le -r$ が成り立つ．一方で

$$A_1^\varepsilon \le \sum_{i=1}^m \mathbb{P}\left(\Lambda_\varepsilon\in F,\ \Phi_\varepsilon\in B_{\mathcal{X}}(h_i,\rho)\right) \le \sum_{i=1}^m \mathbb{P}\left(d_{\mathcal{Y}}(\Lambda_\varepsilon,\psi(h_i))>\delta,\ d_{\mathcal{X}}(\Phi_\varepsilon,h_i)\le\rho\right)$$

であるので，(1.7.1) と補題 1.2.2 を用いて

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log A_1^\varepsilon \le \max_{1\le i\le m}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}\left(d_{\mathcal{Y}}(\Lambda_\varepsilon,\psi(h_i))>\delta,\ d_{\mathcal{X}}(\Phi_\varepsilon,h_i)\le\rho\right)\le -R$$

を得る．以上をまとめて，$\overline{\lim}_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}(\Lambda_\varepsilon\in F)\le\max\{-R,\,-r\}$ を得る．$R\nearrow\infty$ とした後に $r\nearrow\inf_{g'\in F}J(g')$ として，上からの評価を得る． □

1.8　積測度の大偏差原理

２つの確率測度からそれらの積測度を作ることは非常に自然な操作であり，実際に確率論の各分野において頻出する．これはもちろん２つの独立な確率変数を組にしたものを作ることと同値である．そこで自然な疑問が湧くが，２つの確率測度の族がそれぞれ大偏差原理を満たしているとき，それらの積測度の族は大偏差原理を満たすだろうか．この問題に関しては次の事実が知られている．

命題 1.8.1. $(\mathcal{X},d_{\mathcal{X}})$ と $(\mathcal{Y},d_{\mathcal{Y}})$ を可分距離空間とし，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ と $\{\nu_\varepsilon\}_{\varepsilon\in(0,1]}$ をそれぞれ $\mathcal{X}$ 上と $\mathcal{Y}$ 上の Borel 確率測度の族とする．$\varepsilon\searrow 0$ とするとき，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ と $\{\nu_\varepsilon\}_{\varepsilon\in(0,1]}$ はそれぞれ $\mathcal{X}$ 上と $\mathcal{Y}$ 上で指数的に緊密で，速度関数 $I\colon\mathcal{X}\to[0,\infty]$ と $J\colon\mathcal{Y}\to[0,\infty]$ に対して大偏差原理を満たすとする．このとき積測度の族 $\{\mu_\varepsilon\otimes\nu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon\searrow 0$ とするとき $\mathcal{X}\times\mathcal{Y}$ 上で指数的に緊密であり，さらに次で定まる良い速度関数 $K\colon\mathcal{X}\times\mathcal{Y}\to[0,\infty]$ に対して大偏差原理を満たす.

$$K(x,y)=I(x)+J(y),\qquad (x,y)\in\mathcal{X}\times\mathcal{Y}.$$

証明. まず K が速度関数であることから確認する．$\mathcal{X}\times\mathcal{Y}$ 内の点列 $\{(x_n,y_n)\}_{n\in\mathbb{N}}$ が (x,y) に収束すると仮定する．このとき下極限の基本性質により,

$$\varliminf_{n\to\infty}K(x_n,y_n)=\varliminf_{n\to\infty}(I(x_n)+J(y_n))\ge\varliminf_{n\to\infty}I(x_n)+\varliminf_{n\to\infty}J(y_n)\ge I(x)+J(y)=K(x,y)$$

となり，K の下半連続性がわかる．よって K は速度関数である．任意の $r\in[0,\infty)$ に対して

$$\{(x,y)\mid K(x,y)\le r\}\subset\{(x,y)\mid I(x)\le r,\ J(y)\le r\}=I^{-1}([0,r])\times J^{-1}([0,r])$$

である．命題 1.3.1 により実は I と J は良いため，右辺はコンパクトである．このとき左辺もコンパクト集合内の閉集合なのでコンパクトである．これで K が良いことが示せた．

指数的緊密性により，各 $r \in (0,\infty)$ に対してコンパクト部分集合 $A_r \subset \mathcal{X}$ と $B_r \subset \mathcal{Y}$ で

$$[\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(A_r^c)] \vee [\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \nu_\varepsilon(B_r^c)] \le -r$$

を満たすものが存在する．このとき $A_r \times B_r$ はコンパクトで，$(A_r \times B_r)^c \subset (A_r^c \times \mathcal{Y}) \cup (\mathcal{X} \times B_r^c)$ となるので，

$$\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon \otimes \nu_\varepsilon((A_r \times B_r)^c) \le \varlimsup_{\varepsilon\searrow 0} \varepsilon \log[\mu_\varepsilon(A_r^c) + \nu_\varepsilon(B_r^c)] \le -r$$

となる．ここで補題 1.2.2 を用いた．これで $\{\mu_\varepsilon \otimes \nu_\varepsilon\}$ も指数的に緊密であることがわかった．命題 1.3.1 により，後は $\{\mu_\varepsilon \otimes \nu_\varepsilon\}$ が弱大偏差原理を満たすことを示せば十分である．

下からの評価を示す．$O \subset \mathcal{X} \times \mathcal{Y}$ を開集合とする．そのためには，任意の $(x,y) \in O$ に対して $\varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon \otimes \nu_\varepsilon(O) \ge -K(x,y)$ を示せばよい．$K(x,y) = \infty$ のときには示すことは何もないので，$K(x,y) < \infty$ の場合を示す．$\{B_{\mathcal{X}}(x,r) \times B_{\mathcal{Y}}(y,r)\}_{r>0}$ が (x,y) の積位相に関する基本近傍系をなすので，十分小さな $r > 0$ に対して $B_{\mathcal{X}}(x,r) \times B_{\mathcal{Y}}(y,r) \subset O$ となる．積測度と下極限の基本性質により，

$$\begin{aligned}
\varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon \otimes \nu_\varepsilon(O) &\ge \varliminf_{\varepsilon\searrow 0} \{\varepsilon \log \mu_\varepsilon(B_{\mathcal{X}}(x,r)) + \varepsilon \log \nu_\varepsilon(B_{\mathcal{Y}}(y,r))\} \\
&\ge \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(B_{\mathcal{X}}(x,r)) + \varliminf_{\varepsilon\searrow 0} \varepsilon \log \nu_\varepsilon(B_{\mathcal{Y}}(y,r)) \\
&\ge -\inf_{x' \in B_{\mathcal{X}}(x,r)} I(x') - \inf_{y' \in B_{\mathcal{Y}}(y,r)} J(y')
\end{aligned}$$

となる．最後の不等式は $\{\mu_\varepsilon\}$ と $\{\nu_\varepsilon\}$ に対する大偏差原理の下からの評価による．ここで $r \searrow 0$ とすると，(1.2.5) または補題 1.2.8 により，右辺は $-I(x) - J(y)$ に収束する．これで下からの評価が得られた．

最後に上からの評価を示す．$L \subset \mathcal{X} \times \mathcal{Y}$ を任意のコンパクト集合とし，この L に対して $a := \min_{(x,y)\in L} K(x,y)$ とおく．$\delta \in (0,1)$ を任意とする．再び (1.2.5) により，任意の $(x,y) \in L$ に対して，ある $r = r_{(x,y)} > 0$ が存在して，

$$\inf_{x' \in B_{\mathcal{X}}(x,2r)} I(x') \ge (I(x) - \delta) \wedge \delta^{-1}, \qquad \inf_{y' \in B_{\mathcal{Y}}(y,2r)} J(y') \ge (J(y) - \delta) \wedge \delta^{-1},$$

を満たす．$\{B_{\mathcal{X}}(x, r_{(x,y)}) \times B_{\mathcal{Y}}(y, r_{(x,y)}) \mid (x,y) \in L\}$ はコンパクト集合 L の開被覆なので，有限個の $(x_i, y_i) \in L$ $(1 \le i \le m)$ が存在して，$\{B_{\mathcal{X}}(x_i, r_i) \times B_{\mathcal{Y}}(y_i, r_i) \mid 1 \le i \le m\}$ が L を被覆する．ただし $r_i = r_{(x_i,y_i)}$ とおいた．すると

$$\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon \otimes \nu_\varepsilon(L) \le \varlimsup_{\varepsilon\searrow 0} \varepsilon \log\Big[\sum_{1\le i\le m} \mu_\varepsilon \otimes \nu_\varepsilon\big(B_{\mathcal{X}}(x_i, r_i) \times B_{\mathcal{Y}}(y_i, r_i)\big)\Big]$$

$$
\begin{aligned}
&\le \max_{1\le i\le m} \varlimsup_{\varepsilon\searrow 0}\{\varepsilon\log\mu_\varepsilon\big(B_{\mathcal{X}}(x_i,r_i)\big)+\varepsilon\log\nu_\varepsilon\big(B_{\mathcal{Y}}(y_i,r_i)\big)\}\\
&\le \max_{1\le i\le m}\Big\{\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon\big(\overline{B}_{\mathcal{X}}(x_i,r_i)\big)+\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\nu_\varepsilon\big(\overline{B}_{\mathcal{Y}}(y_i,r_i)\big)\Big\}\\
&\le -\min_{1\le i\le m}\Big\{(I(x_i)-\delta)\wedge\delta^{-1}+(J(y_i)-\delta)\wedge\delta^{-1}\Big\}\\
&\le -\inf_{(x,y)\in L}\Big\{(I(x)-\delta)\wedge\delta^{-1}+(J(y)-\delta)\wedge\delta^{-1}\Big\}
\end{aligned}
$$

を得る．ここで積測度の基本性質，補題 1.2.2 および $\{\mu_\varepsilon\}$ と $\{\nu_\varepsilon\}$ に対する大偏差原理の上からの評価を用いた．

まず $a=\infty$ の場合を示す．このとき各 $(x,y)\in L$ に対して $I(x)=\infty$ または $J(y)=\infty$ であるため，$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon\otimes\nu_\varepsilon(L)\le-(\delta^{-1}-\delta)$ を得る．ここで $\delta\searrow 0$ として，求める上からの評価 $\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon\otimes\nu_\varepsilon(L)=-\infty$ を得る．

次に $a<\infty$ の場合を示す．十分小さな δ のみ考えれば十分なので，$\delta^{-1}>a+1$ と仮定する．すると明らかに

$$
\begin{aligned}
&\inf_{(x,y)\in L}\Big\{(I(x)-\delta)\wedge\delta^{-1}+(J(y)-\delta)\wedge\delta^{-1}\Big\}\\
&\ge \inf_{(x,y)\in L}\Big\{(I(x)-\delta)\wedge(a+1)+(J(y)-\delta)\wedge(a+1)\Big\} \qquad (1.8.1)
\end{aligned}
$$

である．$K(\widetilde{x},\widetilde{y})=I(\widetilde{x})+J(\widetilde{y})=a$ を満たす $(\widetilde{x},\widetilde{y})\in L$ に対しては，明らかに $I(\widetilde{x})\vee J(\widetilde{y})\le a$ なので $(I(\widetilde{x})-\delta)\wedge(a+1)+(J(\widetilde{y})-\delta)\wedge(a+1)=K(\widetilde{x},\widetilde{y})-2\delta=a-2\delta$ が成り立つ．$I(x)-\delta\ge a+1$ または $J(y)-\delta\ge a+1$ のときは $(I(x)-\delta)\wedge\delta^{-1}+(J(y)-\delta)\wedge\delta^{-1}\ge a$ なので，(1.8.1) の右辺の下限に寄与しない．$I(x)-\delta<a+1$ かつ $J(y)-\delta<a+1$ のときは $(I(x)-\delta)\wedge\delta^{-1}+(J(y)-\delta)=K(x,y)-2\delta\ge a-2\delta$ である．以上により (1.8.1) の右辺の下限は $a-2\delta$ である．よって $\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon\otimes\nu_\varepsilon(L)\le-(a-2\delta)$ となるので，最後に $\delta\searrow 0$ として，この場合も求める上からの評価 $\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mu_\varepsilon\otimes\nu_\varepsilon(L)\le -a$ となる．これで上からの評価も完成した． □

1.9　射影極限法

本節では $(\Lambda,\prec)$ を有向集合とする．すなわち $(\Lambda,\prec)$ は順序集合であり,[22] さらに任意の $i,j\in\Lambda$ に対して $i\prec k$ かつ $j\prec k$ を満たす $k\in\Lambda$ が存在する（このような k は i,j の上界と呼ばれる.）有向集合の具体例をいくつか与える．

- 全順序集合は明らかに有向集合である．特に $\mathbb{N}$ や $\mathbb{R}$ に通常の順序を与えたものは有向集合である．

[22]すなわち，$\prec$ は Λ 上の 2 項関係で，反射律，反対称律，推移律を満たすものである．ただし，全順序だとは仮定していないので，いかなる大小関係もつかない 2 つの元が存在しうる．

- 空でない集合 Z の部分集合の全体 $2^Z := \{A \mid A \subset Z\}$ は包含関係を順序とみなすと有向集合である．Z の有限部分集合の全体 $2^Z_{<\infty} := \{A \mid A \subset Z, |A| < \infty\}$ も同様に有向集合である．ここで $|A|$ は A の濃度（A の元の個数）を表す．当然ながら，$2^Z_{<\infty} \setminus \{\emptyset\} = \{A \mid A \subset Z, 0 < |A| < \infty\}$ も有向集合である．

- x は位相空間 W の点だとする．x の開近傍全体 $\mathcal{N}_x := \{U \subset W \mid U \text{ は開かつ } x \in U\}$ は通常の大小関係を逆にして「$U \prec V \iff U \supset V$」と定義すると有向集合になる．また x の基本近傍系も同様の考え方で有向集合になる．

Hausdorff 空間の射影極限を思い出そう．[23] $\{\mathcal{Y}_j\}_{j\in\Lambda}$ を Λ で添字付けされた Hausdorff 空間の族とする．また $\{p_{ij} \colon \mathcal{Y}_j \to \mathcal{Y}_i\}_{i,j\in\Lambda, i\prec j}$ を連続写像の族とし，さらに以下の条件を満たすとする．

(1) $i \prec j \prec k$ のとき, $p_{ik} = p_{ij} \circ p_{jk}$.

(2) 任意の j に対して p_{jj} は $\mathcal{Y}_j$ の恒等写像.

このような $(\{\mathcal{Y}_j\}_j, \{p_{ij}\}_{i\prec j})$ を射影系と呼ぶ．

上記の射影系に対して，位相空間としての直積空間を $\mathcal{Y} := \prod_{j\in\Lambda} \mathcal{Y}_j$ と書き，その一般的な元を $\mathbf{y} = (y_j)_{j\in\Lambda}$ のように表す．[24] この射影系の射影極限 $\varprojlim \mathcal{Y}_j$ は以下で定義される $\mathcal{Y}$ の部分集合で，$\mathcal{Y}$ からの相対位相（つまり部分集合としての位相）を備えている．

$$\varprojlim \mathcal{Y}_j := \{\mathbf{y} = (y_j)_{j\in\Lambda} \in \mathcal{Y} \mid i \prec k \implies y_i = p_{ik}(y_k)\}.$$

$\mathcal{Y}$ は Hausdorff なので，その部分空間である $\varprojlim \mathcal{Y}_j$ も Hausdorff である．直積空間 $\mathcal{Y}$ の i 番目の座標成分を与える標準的な射影を $\varprojlim \mathcal{Y}_j$ に制限して得られる連続写像を $p_i \colon \varprojlim \mathcal{Y}_j \to \mathcal{Y}_i$ と書く．明らかに $i \prec j$ ならば $p_i = p_{ij} \circ p_j$ である．(各 $\mathcal{Y}_j$ が非空でも，$\varprojlim \mathcal{Y}_j$ が空になることがある．本書では $\varprojlim \mathcal{Y}_j \neq \emptyset$ である場合のみを扱う.) $i \prec k$ のとき, $p_{ik}(A_k) \subset A_i$ を満たす部分集合の族 $\{A_j \subset \mathcal{Y}_j \mid j \in \Lambda\}$ に対しては $\varprojlim \mathcal{Y}_j$ の部分集合 $\varprojlim A_j$ が自然に定まることを注意しておく．

例 1.9.1. 空でない集合 Z 上の実数値関数の全体 $\mathbb{R}^Z$ を射影極限の形で表してみよう．[25] 以下では Z の部分集合は空でないとする．$A \subset Z$ に対して $\mathcal{Y}_A := \mathbb{R}^A$ と定める．$A \subset B \subset Z$ のとき，関数の定義域を B から A に制限することにより自然な射影 $p_{AB} \colon \mathbb{R}^B \to \mathbb{R}^A$ が定まる．

有向集合として $\Lambda := 2^Z_{<\infty} \setminus \{\emptyset\} = \{A \mid A \subset Z, 0 < |A| < \infty\}$ を取る．$A \in \Lambda$ に対して $\mathcal{Y}_A$ には $|A|$ 次元 Euclid 空間としての位相を入れる．$A, B \in \Lambda$ かつ $A \subset B$ のとき

---
[23] 逆極限とも呼ばれる．
[24] 本書では通常どおりに選択公理を仮定するので，各 j に対して $\mathcal{Y}_j \neq \emptyset$ ならば $\mathcal{Y} \neq \emptyset$ である．
[25] この例では集合 X から集合 Y への写像の全体を Y^X と書くことにする．

は，p_{AB} は結局「$B \setminus A$ に属する座標成分を捨てる」という連続射影のことである．この $(\{\mathcal{Y}_A\}, \{p_{AB}\}_{A\subset B})$ は明らかに上述した射影系の一例である．

このとき射影極限 $\varprojlim \mathcal{Y}_A$ が $\mathbb{R}^Z$ に各点収束の位相を入れたものと位相同型であることを確認する．$F\colon \varprojlim \mathcal{Y}_A \to \mathbb{R}^Z$ を $F(\mathbf{y})_s = y_{\{s\}}$ $(s \in Z)$ と定め，$G\colon \mathbb{R}^Z \to \varprojlim \mathcal{Y}_A$ を $G(\psi)_A = p_{AZ}\psi$ $(A \in \Lambda)$ と定めると，作り方により F と G は互いに逆写像になっている．また「各点収束の位相」と「各有限集合上での一様収束の位相」は同値なので，F と G は同相写像である．これで射影極限が $\mathbb{R}^Z$ であることがわかった．

次の補題で射影極限の基本性質をいくつか示すが，これらは本節の主目的であるDawson-Gärtnerの定理（定理1.9.3）の証明中で用いる．

補題 1.9.2. 射影極限 $\varprojlim \mathcal{Y}_j$ を有向集合 Λ を添字に持つ射影系 $(\{\mathcal{Y}_j\}_{j\in\Lambda}, \{p_{ij}\}_{i,j\in\Lambda,i\prec j})$ から決まるものとする．

(1) O を $\varprojlim \mathcal{Y}_j$ における点 $\mathbf{z} = (z_j)_{j\in\Lambda} \in \varprojlim \mathcal{Y}_j$ の開近傍とする．このとき，ある $k \in \Lambda$ と $\mathcal{Y}_k$ における z_k の十分小さな開近傍 U_k が存在して $\mathbf{z} \in p_k^{-1}(U_k) \subset O$ となる．

(2) F を $\varprojlim \mathcal{Y}_j$ における閉集合とし，各 $j \in \Lambda$ に対して $F_j := p_j(F)$ とおく．このとき $i \prec k$ ならば $p_{ik}(\overline{F_k}) \subset \overline{F_i}$ が成り立ち，さらに $F = \varprojlim \overline{F_j}$ となる．

(3) 各 $j \in \Lambda$ に対して $\mathcal{Y}_j$ は空でないコンパクト集合だとする．このとき $\varprojlim \mathcal{Y}_j$ も空でないコンパクト集合である．

証明. (1) 直積位相の定義により，適当な $m \in \mathbb{N}$, $i_1, \ldots, i_m \in \Lambda$ および各 $s = 1, \ldots, m$ に対して z_{i_s} の十分小さな開近傍 $U_{i_s} \subset \mathcal{Y}_{i_s}$ が存在して

$$\{\mathbf{y} = (y_j)_{j\in\Lambda} \in \varprojlim \mathcal{Y}_j \mid y_{i_s} \in U_{i_s}\ (1 \le s \le m)\} \subset O \tag{1.9.1}$$

を満たす．k を $i_1, \ldots, i_m$ の１つの上界とする．$U_k := \cap_{s=1}^m p_{i_s k}^{-1}(U_{i_s})$ とおくとこれは $\mathcal{Y}_k$ において開であり，

$$p_k^{-1}(U_k) = \{\mathbf{y} = (y_j)_{j\in\Lambda} \in \varprojlim \mathcal{Y}_j \mid y_k \in U_k\}$$

は明らかに $\mathbf{z}$ を含みかつ (1.9.1) の左辺の開集合に含まれる．

(2) $i \prec k$ のとき $p_{ik}(F_k) = F_i$ となることに注意すると，p_{ik} の連続性と合わせて，$p_{ik}(\overline{F_k}) \subset \overline{F_i}$ が成り立つことがわかる．よって $(\{\overline{F_j}\}_j, \{p_{ij}\}_{i\prec j})$ は射影系であり，作り方により $\varprojlim \overline{F_j} \supset F$ であるが，実は逆向きの包含も成立することを見よう．$\mathbf{z} = (z_j)_{j\in\Lambda} \in F^c$ を任意に取る．F^c は開なので (1) により，ある $k \in \Lambda$ と $\mathcal{Y}_k$ の開集合 U_k が存在して $\mathbf{z} \in p_k^{-1}(U_k) \subset F^c$ となる．仮に $U_k \cap F_k \neq \emptyset$ だとすると，ある $\mathbf{y} \in F$ とある $u \in U_k$ に対して $u = p_k(\mathbf{y})$ となるが，これは $\mathbf{y} \in p_k^{-1}(u) \subset p_k^{-1}(U_k)$ を意味するので，$p_k^{-1}(U_k) \cap F = \emptyset$ に矛盾する．よって $p_k(\mathbf{z}) \in U_k \subset (F_k)^c$，特に $p_k(\mathbf{z}) \notin \overline{F_k}$ が示された．したがって，$\mathbf{z} \in (\varprojlim \overline{F_j})^c$ であり，これで $\varprojlim \overline{F_j} \subset F$ が示された．

(3) 直積空間 $\mathcal{Y} := \prod_{j\in\Lambda} \mathcal{Y}_j \neq \emptyset$ は Tychonoff の定理によりコンパクトなので，$\varprojlim \mathcal{Y}_j$ がその閉部分集合であることを見れば十分である．$i \prec j$ とする．まず $(p_i,\, p_{ij} \circ p_j)(\mathbf{y}) := (p_i(\mathbf{y}),\, p_{ij} \circ p_j(\mathbf{y}))$ とおくと $(p_i,\, p_{ij} \circ p_j)\colon \mathcal{Y} \to \mathcal{Y}_i \times \mathcal{Y}_i$ は連続写像である．また $\mathcal{Y}_i$ は Hausdorff 空間であるが，これは対角線集合 $\mathrm{Diag}(\mathcal{Y}_i) := \{(v,v) \mid v \in \mathcal{Y}_i\}$ が $\mathcal{Y}_i \times \mathcal{Y}_i$ において閉であることと同値である．したがって

$$A_{ij} := \{\mathbf{y} = (y_l)_{l\in\Lambda} \in \mathcal{Y} \mid y_i = p_{ij}(y_j)\} = (p_i,\, p_{ij} \circ p_j)^{-1}(\mathrm{Diag}(\mathcal{Y}_i))$$

は $\mathcal{Y}$ の閉部分集合である．定義により

$$\varprojlim \mathcal{Y}_j = \bigcap \{A_{ij} \mid i, j \in \Lambda,\ i \prec j\}$$

であるが右辺は閉である．これで $\mathcal{Y}$ のコンパクト性が確認できた．

有限交叉性を用いて射影極限が空でないことを示そう．$k \in \Lambda$ を任意に取り，$\Lambda_{\prec k} := \{i \in \Lambda \mid i \prec k\}$ とおく．また $\mathcal{Y}_{\prec k} := \prod_{j\in\Lambda_{\prec k}} \mathcal{Y}_j$ とおき，$q_k\colon \mathcal{Y} \to \mathcal{Y}_{\prec k}$ を自然な連続射影とし，

$$B_k := \bigcap \{A_{ij} \mid i, j \in \Lambda_{\prec k},\ i \prec j\} = \{\mathbf{y} = (y_l)_{l\in\Lambda} \in \mathcal{Y} \mid i \prec j \prec k \implies y_i = p_{ij}(y_j)\}$$

と定ると，上と同じ理由で B_k は $\mathcal{Y}$ のコンパクト集合である．$\cap_{k\in\Lambda} B_k = \varprojlim \mathcal{Y}_j$ なので，後は $\{B_k \mid k \in \Lambda\}$ が有限交叉性を持つことを見れば十分である．$N \in \mathbb{N}$ と $k_1, \ldots, k_N \in \Lambda$ を任意とする．このとき，これら N 個の元の上界 $m \in \Lambda$ が存在する．$\xi \in \mathcal{Y}_m\ (\neq \emptyset)$ を１つ選び，$i \prec m$ のとき $z_i := p_{im}(\xi)$ とおく．明らかに $i \prec j \prec m$ であれば $z_i = p_{ij}(z_j)$ を満たす．$\mathcal{Y}_{\prec m}$ の１点集合の引き戻し $q_m^{-1}(\{z_i\}_{i\in\Lambda_{\prec m}}) \subset \mathcal{Y}$ は空ではなく，作り方により各 B_{k_r} $(1 \leq r \leq N)$ に含まれている．これで $\cap_{r=1}^{N} B_{k_r} \neq \emptyset$ が示せたので証明が終わった．　□

Dawson-Gärtner の定理と呼ばれる次の定理が本節の主目標である．射影系の「各成分」上で大偏差原理が成立していれば，射影極限上に大偏差原理が遺伝することを主張する．

定理 1.9.3. 有向集合 Λ で添字付けられた射影系 $(\{\mathcal{Y}_j\}_{j\in\Lambda}, \{p_{ij}\}_{i,j\in\Lambda, i\prec j})$ から（空でない）射影極限 $\varprojlim \mathcal{Y}_j$ が定まっているとする．さらに各 $j \in \Lambda$ に対して $\mathcal{Y}_j$ が正則空間であることも仮定する．$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varprojlim \mathcal{Y}_j$ 上の Borel 確率測度の族で，各 $j \in \Lambda$ に対して $\{\mu_\varepsilon \circ p_j^{-1}\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに $\mathcal{Y}_j$ 上で良い速度関数 I_j に対して大偏差原理を満たすと仮定する．このとき，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに $\varprojlim \mathcal{Y}_j$ 上で次で定義される良い速度関数 I に対して大偏差原理を満たす．

$$I(\mathbf{x}) := \sup_{j\in\Lambda} I_j(p_j(\mathbf{x})), \qquad \mathbf{x} \in \varprojlim \mathcal{Y}_j.$$

証明. $I \geq 0$ は自明である．$\alpha \in [0,\infty)$ のとき $I^{-1}([0,\alpha])$ がコンパクトであることを証明する．$i \prec j$ とする．$\mu_\varepsilon \circ p_i^{-1} = (\mu_\varepsilon \circ p_j^{-1}) \circ p_{ij}^{-1}$ かつ p_{ij} は連続なので，$\{\mu_\varepsilon \circ p_i^{-1}\}$ に対する大偏差原理は $\{\mu_\varepsilon \circ p_j^{-1}\}$ に対する大偏差原理に縮小原理（命題 1.5.1）を適用して得られる．$\mathcal{Y}_i$ の正則性により大偏差原理の速度関数は一意的に決まる（命題 1.2.11）．よって $I_i(y_i) = \inf\{I_j(y_j) \mid y_j \in p_{ij}^{-1}(y_i)\}$ を得る．また縮小原理の証明中で示したように $I_i^{-1}([0,\alpha]) = p_{ij}(I_j^{-1}([0,\alpha]))$ も成り立つ．以上により

$$I^{-1}([0,\alpha]) = (\varprojlim \mathcal{Y}_j) \bigcap \prod_{j\in\Lambda} I_j^{-1}([0,\alpha]) = \varprojlim I_j^{-1}([0,\alpha]) \tag{1.9.2}$$

を得るが，補題 1.9.2(3) により最右辺はコンパクト集合である．これで I が良い速度関数であることがわかった．

次は大偏差原理の下からの評価を示す．$O \subset \varprojlim \mathcal{Y}_j$ を開集合とする．補題 1.9.2(1) により，任意の $\mathbf{z} = (z_j)_{j\in\Lambda} \in O$ に対して，ある $k \in \Lambda$ と $\mathcal{Y}_k$ における z_k の十分小さな開近傍 U_k が存在して $\mathbf{z} \in p_k^{-1}(U_k) \subset O$ となる．$\{\mu_\varepsilon \circ p_k^{-1}\}$ に対する大偏差原理の下からの評価を用いると

$$\begin{aligned}\varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(O) &\geq \varliminf_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon \circ p_k^{-1}(U_k) \\ &\geq -\inf_{\xi\in U_k} I_k(\xi) \geq -I_k(p_k(\mathbf{z})) \geq -\sup_{j\in\Lambda} I_j(p_j(\mathbf{z})) = -I(\mathbf{z})\end{aligned}$$

となる．$\mathbf{z}$ に関する上限を取ると，求める評価を得る．

最後に大偏差原理の上からの評価を示す．$F \subset \varprojlim \mathcal{Y}_j$ を閉集合とする．補題 1.9.2(2) により $F = \varprojlim \overline{F_j}$ なので，(1.9.2) と合わると

$$F \cap I^{-1}([0,\alpha]) = \varprojlim \{\overline{F_j} \cap I_j^{-1}([0,\alpha])\}, \qquad \alpha \in [0,\infty)$$

となることがわかる．以下では $\alpha < \inf_{\mathbf{x}\in F} I(\mathbf{x})$ とする．明らかに $F \cap I^{-1}([0,\alpha]) = \emptyset$ である．このとき補題 1.9.2(3) により，ある $k \in \Lambda$ に対して $\overline{F_k} \cap I_k^{-1}([0,\alpha]) = \emptyset$ である．$F \subset p_k^{-1}(\overline{F_k})$ に注意すると，$\{\mu_\varepsilon \circ p_k^{-1}\}$ に対する大偏差原理の上からの評価により

$$\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon(F) \leq \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mu_\varepsilon \circ p_k^{-1}(\overline{F_k}) \leq -\inf_{\xi\in\overline{F_k}} I_k(\xi) \ \leq -\alpha$$

を得る．$\alpha \nearrow \inf_{\mathbf{x}\in F} I(\mathbf{x})$ として求める評価を得る．これで定理 1.9.3 の証明が終わった． □

Dawson-Gärtner の定理の簡単な応用として，次の例 1.9.4 において可算直積測度に対する大偏差原理を紹介しよう．この例は以下の設定の下で議論する．

各 $i \in \mathbb{N}$ に対して $\mathcal{X}_i$ を可分距離空間とし，$\mathcal{X} := \prod_{i\in\mathbb{N}} \mathcal{X}_i$ をそれらの直積空間とする．$\mathcal{X}$ も再び可分距離空間になる．$\mathcal{X}$ の一般的な元を $\mathbf{x} = (x_i)_{i\in\mathbb{N}}$ と書く．各 $n \in \mathbb{N}$ に対し $\mathcal{Y}_n := \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ とおき，$n \leq m$ のとき $p_{nm}\colon \mathcal{Y}_m \to \mathcal{Y}_n$ を最後の $m-n$ 個の成分を

捨てる射影とする．すると $(\{\mathcal{Y}_n\}_n, \{p_{nm}\}_{n\le m})$ は $\mathbb{N}$ を添字集合とする射影系をなす．これから決まる射影極限 $\varprojlim \mathcal{Y}_n$ は $\mathcal{X}$ と同相であることが簡単にわかるので，これらを同一視する．また $j \in \mathbb{N}$ のとき，射影 $p_{j\infty}\colon \mathcal{X} = \varprojlim \mathcal{Y}_n \to \mathcal{Y}_j$ を $p_{j\infty}(\mathbf{x}) = (x_1, \dots, x_j)$ として定める．

例 1.9.4. 各 $i \in \mathbb{N}$ に対して，$\{\nu_{i,\varepsilon}\}_{\varepsilon\in(0,1]}$ は $\mathcal{X}_i$ 上の Borel 確率測度の族で，$\varepsilon \searrow 0$ のときに指数的に緊密であり，さらに速度関数 $J_i\colon \mathcal{X}_i \to [0,\infty]$ に対して大偏差原理を満たすと仮定する．可算直積測度 $\mu_\varepsilon := \otimes_{i\in\mathbb{N}} \nu_{i,\varepsilon}$ は $\mathcal{X}$ 上の Borel 確率測度である．このとき，$\{\mu_\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{X}$ 上において次で定まる良い速度関数 I に対して大偏差原理を満たす．

$$I(\mathbf{x}) := \sum_{i=1}^{\infty} J_i(x_i), \qquad \mathbf{x} = (x_i)_{i\in\mathbb{N}} \in \mathcal{X}.$$

上で述べた事実を確認しよう．各 $n \in \mathbb{N}$ に対して $\mu_{n,\varepsilon} := \nu_{1,\varepsilon} \otimes \cdots \otimes \nu_{n,\varepsilon}$ とおくと，命題 1.8.1 により $\{\mu_{n,\varepsilon}\}_{\varepsilon\in(0,1]}$ は次で定まる良い速度関数 $I_n\colon \mathcal{Y}_n \to [0,\infty]$ に対して $\mathcal{Y}_n$ 上で大偏差原理を満たす．

$$I_n(x_1, \dots, x_n) := \sum_{i=1}^{n} J_i(x_i), \qquad (x_1, \dots, x_n) \in \mathcal{Y}_n.$$

明らかに $\sup_{n\in\mathbb{N}} I_n(p_{n\infty}(\mathbf{x})) = I(\mathbf{x})$ である．以上により，Dawson-Gärtner の定理（定理 1.9.3）を用いて，可算直積測度に対する大偏差原理が証明できた．

第2章　Cramérの大偏差原理

本章では Cramér の大偏差原理について論ずる．Cramér の大偏差原理は最も基本的な大偏差原理であるが，その証明は大偏差原理を示すための基本的な手法を多く含んでいる．

次元 $d \in \mathbb{N}$ を任意とする．$\{X_i\}_{i\in\mathbb{N}}$ を独立同分布な $\mathbb{R}^d$ 値確率変数の列とし，X_1 は平均 $m \in \mathbb{R}^d$ を持つと仮定する．各 $n \in \mathbb{N}$ に対して $S_n = X_1 + \cdots + X_n$ とおくと，大数の弱法則により任意の $\varepsilon > 0$ に対して

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) = 0$$

が成立する．この確率の指数減衰の速さを記述するのが Cramér の大偏差原理である．

Cramér の大偏差原理を正確に述べるために必要な記法と仮定を用意する．積率母関数 $M : \mathbb{R}^d \to (0, \infty]$ と対数積率母関数 $\Lambda : \mathbb{R}^d \to (-\infty, \infty]$ をそれぞれ次で定義する．

$$M(\theta) := \mathbb{E}\left[e^{\langle\theta, X_1\rangle}\right], \quad \Lambda(\theta) := \log M(\theta), \quad \theta \in \mathbb{R}^d.$$

また Λ の Fenchel-Legendre 変換 $I : \mathbb{R}^d \to [0, \infty]$ を次で定義する．

$$I(x) := \sup_{\theta\in\mathbb{R}^d}\{\langle\theta, x\rangle - \Lambda(\theta)\}, \quad x \in \mathbb{R}^d. \tag{2.0.1}$$

I が非負であることは $\theta = 0$ を代入すればわかる．本章で重要な役割を果たす条件 **(FEM)** を導入する．

(FEM) 任意の $\theta \in \mathbb{R}^d$ に対して $M(\theta) < \infty$ である．[1]

これで Cramér の大偏差原理を正確に述べる準備ができた．本章では大偏差原理の基本事項に触れつつ次の定理を証明する．命題 2.2.1 で示すように I は良い速度関数になっている．

定理 2.0.1. 条件 **(FEM)** を仮定する．$\mathbb{R}^d$ 値確率変数の列 $\{S_n/n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 I に対して速度 n で大偏差原理を満たす．

[1] FEM は Finite Exponential Moment の略である．

本章では以下の条件 **(SLG)** も重要な役割を果たす.[2]

(SLG) 対数積率母関数 Λ に対して次が成立する.

$$\lim_{|\theta|\to\infty}\frac{\Lambda(\theta)}{|\theta|}=\infty.$$

定理 2.0.1 の証明は第 2.5 節で行うが，第 2.3 節と第 2.4 節ではひとまず条件 **(SLG)** も仮定して大偏差原理を示す.

2.1 Cramér の大偏差原理の速度関数の例

本節では積率母関数 M と速度関数 I を，正規分布・Bernoulli 分布・指数分布の場合に簡単に計算する．以下の例では S_n/n の分布の具体形がわかる例になっており，その具体形により大偏差原理の成立が自然に期待されるので，このことも各例で簡単に見る．また以下の例は全て良い速度関数になっている.

次の例では S_n/n の分布が再生性により直ちにわかる．単純な例ではあるものの，第 9 章で扱う Schilder の大偏差原理に関連していることを注意しておく.

例 2.1.1 (1 次元正規分布). $d=1$ とする．X_1 を平均 0，分散 $\sigma^2>0$ の 1 次元正規分布とする，つまり X_1 の分布の密度関数は $g(x):=(2\pi\sigma^2)^{-1/2}\exp\left(-\frac{x^2}{2\sigma^2}\right), x\in\mathbb{R}$ であるとする．まず M は次で与えられる.

$$\begin{aligned}M(\theta)&=\frac{1}{(2\pi\sigma^2)^{1/2}}\int_{\mathbb{R}}\exp\left(\theta x-\frac{x^2}{2\sigma^2}\right)dx\\&=\frac{e^{(\sigma^2\theta^2)/2}}{(2\pi\sigma^2)^{1/2}}\int_{\mathbb{R}}\exp\left(-\frac{(x-\sigma^2\theta)^2}{2\sigma^2}\right)dx=e^{(\sigma^2\theta^2)/2}.\end{aligned}$$

ただし最後の等式では変数変換 $x-\sigma^2\theta=\sqrt{2}\sigma y$ を行い，Gauss 積分 $\int_{\mathbb{R}}e^{-x^2}dx=\sqrt{\pi}$ を用いた．再び平方完成を行えば

$$I(x)=\frac{x^2}{2\sigma^2}$$

となる.

また正規分布の再生性により S_n の分布は平均 0，分散 $n\sigma^2$ の 1 次元正規分布になるので，S_n/n の分布は

$$\mathbb{P}\left(\frac{S_n}{n}\in A\right)=\left(\frac{n}{2\pi\sigma^2}\right)^{1/2}\int_A\exp\left(-\frac{nx^2}{2\sigma^2}\right)dx,\quad A\in\mathcal{B}(\mathbb{R})$$

となる．この式を用いて直接大偏差原理の定義式の成立を確認することができるが，詳細には立ち入らない.

---

[2] SLG は Super Linear Growth の略である.

例 2.1.2 (多次元正規分布)**.** X_1 を平均 0，共分散行列 Σ の d 次元正規分布とし，Σ は非退化とする．つまり X_1 の分布の密度関数は

$$f(x) = \frac{1}{(2\pi)^{d/2}\,(\det\Sigma)^{1/2}} \exp\left(-\frac{1}{2}\langle x, \Sigma^{-1}x\rangle\right), \quad x \in \mathbb{R}^d$$

で与えられる．$\langle\theta, X_1\rangle$ の分布は平均 0，分散 $\langle\theta, \Sigma\theta\rangle$ の 1 次元正規分布なので，例 2.1.1 により

$$M(\theta) = e^{\langle\theta,\Sigma\theta\rangle/2}$$

となる．また

$$\begin{aligned}\langle\theta, x\rangle - \Lambda(\theta) &= \langle\Sigma^{1/2}\theta, \Sigma^{-1/2}x\rangle - \frac{1}{2}\langle\Sigma^{1/2}\theta, \Sigma^{1/2}\theta\rangle \\ &= -\frac{1}{2}\langle\Sigma^{1/2}\theta - \Sigma^{-1/2}x, \Sigma^{1/2}\theta - \Sigma^{-1/2}x\rangle + \frac{1}{2}\langle\Sigma^{-1/2}x, \Sigma^{-1/2}x\rangle\end{aligned}$$

であるから，

$$I(x) = \frac{1}{2}\langle\Sigma^{-1/2}x, \Sigma^{-1/2}x\rangle$$

となる．

本節で現れる例では Λ の具体形がわかるので，例 2.1.1 と例 2.1.2 は **(FEM)** かつ **(SLG)** を満たすことがわかる．しかしながら，次の例は **(FEM)** は満たすが **(SLG)** を満たさないことが Λ の具体形からわかる．

例 2.1.3 (Bernoulli 分布)**.** $0 < p < 1$ とし，X_1 は平均 p の Bernoulli 分布に従うとする，すなわち

$$\mathbb{P}(X_1 = 1) = 1 - \mathbb{P}(X_1 = 0) = p$$

である．このとき M と I は次で与えられる．

$$M(\theta) = pe^\theta + 1 - p, \quad I(x) = \begin{cases} x\log\dfrac{x}{p} + (1-x)\log\dfrac{1-x}{1-p}, & x \in [0,1], \\ \infty, & x \notin [0,1]. \end{cases}$$

$x \in (0,1)$ の場合に I の表示式を確認する．$x \in (0,1)$ なので $\theta \in \mathbb{R} \mapsto \theta x - \Lambda(\theta)$ はある $\theta_x \in \mathbb{R}$ で最大になることがわかる．θ_x における微分は 0 になるので，具体的に微分を実行することにより

$$(\theta x - \Lambda(\theta))'\,|_{\theta=\theta_x} = 0 \quad\Longleftrightarrow\quad \theta_x = \log\frac{x}{p} + \log\frac{1-p}{1-x}$$

となる．この θ_x を $I(x)$ の定義式に代入して，式を整理することにより所望の式を得る．

$n \to \infty$ のとき Stirling の公式 $n! \sim \sqrt{2\pi n}(n/e)^n, n \to \infty$ を用いて，$0 \le k \le n$ に対して確率 $\mathbb{P}(S_n = k)$ を計算する．以下の計算を厳密化することにより大偏差原理を示すことは可能であるが，詳細には立ち入らない．2 項定理により

$$
\begin{aligned}
\mathbb{P}(S_n = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&\sim \left(\frac{n}{2\pi k(n-k)}\right)^{1/2} \frac{n^n}{k^k (n-k)^{n-k}} p^k (1-p)^{n-k} \\
&= \left(\frac{n}{2\pi k(n-k)}\right)^{1/2} \frac{1}{\left(\frac{(k/n)}{p}\right)^k \left(\frac{1-(k/n)}{1-p}\right)^{n-k}}
\end{aligned}
$$

となる.「$\sim$」は比の極限が 1 に収束するという意味である．よって $k = \lfloor xn \rfloor, x \in (0,1)$ とすれば，$n \to \infty$ のとき

$$
\begin{aligned}
\frac{1}{n} \log \mathbb{P}(S_n = k) &\sim -\left\{ \frac{k}{n} \log \left(\frac{(k/n)}{p}\right) + (1 - \frac{k}{n}) \log \left(\frac{1-(k/n)}{1-p}\right) \right\} \\
&\to -\left\{ x \log \frac{x}{p} + (1-x) \log \frac{1-x}{1-p} \right\} = -I(x)
\end{aligned}
$$

となる．

例 2.1.4. Bernoulli 分布は $\{-1, +1\}$ 上の分布として考えることも多い．実際，後に見る Curie-Weiss 模型では $\{-1, +1\}$ 上のパラメーター $p = 1/2$ の Bernoulli 分布を考える．この場合の速度関数は

$$
I(z) := \frac{1}{2}(1-z)\log(1-z) + \frac{1}{2}(1+z)\log(1+z), \quad z \in [-1, 1]
$$

となる．

注意 2.1.5. 例 2.1.3 において，$x \in (0,1)$ のときは (2.0.1) における上限が達成されている．条件 **(SLG)** は任意の $x \in \mathbb{R}^d$ に対して $I(x)$ を定める上限が，ある $\theta_x \in \mathbb{R}^d$ で達成されるための十分条件である．このことは定理 2.3.1 の証明中で確認する．またその点における微分は 0 である．

次の例は条件 **(FEM)** を満たさない例である．しかしこの例では S_n/n の分布が計算できるため，速度関数 I も直接計算できる．

例 2.1.6 (指数分布). X_1 をパラメータ 1 の指数分布とする，すなわち X_1 の分布の密度関数は $f(x) = \mathbf{1}_{\{x\geq 0\}}e^{-x}, x \in \mathbb{R}$ を持つ．このとき M と I は次で与えられることが初等的な計算により確認できる．

$$M(\theta) = \begin{cases} \dfrac{1}{1-\theta}, & \theta \in (-\infty, 1), \\ \infty, & \theta \in [1, \infty), \end{cases} \qquad I(x) = \begin{cases} x - 1 - \log x, & x \in (0, \infty), \\ \infty, & x \notin (0, \infty). \end{cases}$$

よって条件 **(FEM)** を満たさない．

またS S_n/n の分布は

$$\mathbb{P}\left(\frac{S_n}{n} \in A\right) = \frac{n^n}{(n-1)!}\int_A e^{-nx}x^{n-1}dx, \quad A \in \mathcal{B}(\mathbb{R})$$

となることが帰納法で確認できる．再び Stirling の公式により

$$\frac{n^n}{(n-1)!}e^{-nx}x^{n-1} \sim \left(\frac{n}{2\pi}\right)^{1/2} e^{n-nx}x^{n-1}$$

であるから，

$$\frac{1}{n}\log \mathbb{P}\left(\frac{S_n}{n} \neq x\right) \sim -(x - 1 - \log x) = -I(x)$$

となり速度関数 I が現れることが確認できた．これらの計算の厳密化も可能であるが，詳細には立ち入らない．

2.2　積率母関数と速度関数の初歩的な性質

本節では積率母関数 M と速度関数 I の基本性質をいくつか見る．**(FEM)** を仮定せずとも，一般的に $M(0) = 1$ かつ M は $(0, \infty]$ 値である．よって $\Lambda(0) = 0$ かつ Λ は $(-\infty, \infty]$ 値である．さらに Λ は下半連続かつ凸である．下半連続であることは Fatou の補題による．任意の $\theta_1, \theta_2 \in \mathbb{R}^d, t \in (0, 1)$ に対して，Hölder の不等式により

$$\mathbb{E}\left[e^{\langle t\theta_1 + (1-t)\theta_2, X_1\rangle}\right] \leq \left(\mathbb{E}\left[e^{\langle \theta_1, X_1\rangle}\right]\right)^t \left(\mathbb{E}\left[e^{\langle \theta_2, X_1\rangle}\right]\right)^{1-t}$$

が成立する．この式で対数を取ることにより，Λ が凸であることがわかる．また $\Lambda(0) = 0$ なので，I は $[0, \infty]$ 値である．次に M は $0 \in \mathbb{R}^d$ の近傍上で有限であると仮定する．これは **(FEM)** より弱い条件である．このとき，十分小さい $s > 0$ に対して $\mathbb{E}\left[e^{s|X_1|}\right] < \infty$ となる．このことから簡単にわかるが，任意の $p \in [1, \infty)$ に対して $\mathbb{E}[|X_1|^p] < \infty$ が成立しており，さらに M は 0 で連続である．(命題 2.2.1(1) の証明を参照せよ.)

条件 **(FEM)** の下ではさらに M や I の性質がわかる．これを次の命題で述べる．

命題 2.2.1. 条件 **(FEM)** を仮定する．このとき次が成立する．

(1) M は $\mathbb{R}^d$ 上 C^1 級であり，任意の $\theta \in \mathbb{R}^d$ に対して $\nabla M(\theta) = \mathbb{E}\left[X_1 e^{\langle \theta, X_1 \rangle}\right]$ である．[3]

(2) I は下半連続かつ凸である．

(3) I は良い速度関数である．

証明. まず (1) を示す．M は $\theta \in \mathbb{R}^d$ において偏微分可能であることを示す．$h \in \mathbb{R}, h \neq 0$ を任意に取る．$\mathbb{R}^d$ における標準基底を $\{e_1, \ldots, e_d\}$ と書くと，$i = 1, \ldots, d$ に対して

$$\frac{1}{h}\left(\mathbb{E}\left[e^{\langle \theta + he_i, X_1 \rangle}\right] - \mathbb{E}\left[e^{\langle \theta, X_1 \rangle}\right]\right) = \mathbb{E}\left[\frac{e^{hX_1^{(i)}} - 1}{h} e^{\langle \theta, X_1 \rangle}\right]$$

と書ける．ただし $X_1^{(i)}$ は X_1 の第 i 成分である．不等式 $|e^a - 1| \leq |a| e^{|a|}$ により，$|h| \leq 1$ のとき

$$\left|\frac{e^{hX_1^{(i)}} - 1}{h} e^{\langle \theta, X_1 \rangle}\right| \leq \left|X_1^{(i)}\right| e^{\left|hX_1^{(i)}\right|} \left|e^{\langle \theta, X_1 \rangle}\right| \leq |X_1| e^{(|\theta|+1)|X_1|} \leq e^{(|\theta|+2)|X_1|}$$

である．ただし最後の不等式では自明な不等式 $|a| \leq e^{|a|}$ を用いた．条件 **(FEM)** により最右辺の積分は期待値は有限なので，Lebesgue の優収束定理により

$$\lim_{h \to 0} \frac{1}{h}\left(\mathbb{E}\left[e^{\langle \theta + he_i, X_1 \rangle}\right] - \mathbb{E}\left[e^{\langle \theta, X_1 \rangle}\right]\right) = \mathbb{E}\left[X_1^{(i)} e^{\langle \theta, X_1 \rangle}\right]$$

となるので，M は θ において偏微分可能であり，$\nabla M(\theta) = \mathbb{E}\left[X_1 e^{\langle \theta, X_1 \rangle}\right]$ を得る．また ∇M の表示式と Lebesgue の収束定理により，M が C^1 級であることがわかる．

次に (2) を示す．この証明では $f(\theta, x) = \langle \theta, x \rangle - \Lambda(\theta)$ と書く．各 $\theta \in \mathbb{R}^d$ に対して $f(\theta, \cdot)$ は連続（特に下半連続）なので，補題 1.2.5 により，I は下半連続である．I が凸であることを示すために，$x, y \in \mathbb{R}^d, t \in [0, 1]$ とする．各 $\theta \in \mathbb{R}^d$ に対して $f(\theta, \cdot)$ はアファインなので

$$f(\theta, tx + (1-t)y) = tf(\theta, x) + (1-t)f(\theta, y) \leq tI(x) + (1-t)I(y)$$

である．よって $\theta \in \mathbb{R}^d$ についての上限を取れば $I(tx + (1-t)y) \leq tI(x) + (1-t)I(y)$ となり I が凸であることが示された．

最後に (3) を示す．任意の $a \in [0, \infty)$ に対して $\{x \in \mathbb{R}^d \mid I(x) \leq a\}$ がコンパクトであることを示す．I は下半連続なので，$\{x \in \mathbb{R}^d \mid I(x) \leq a\}$ は閉集合である．なので $\{x \in \mathbb{R}^d \mid I(x) \leq a\}$ がコンパクトであることを示すためには

$$\lim_{|x| \to \infty} I(x) = \infty \tag{2.2.1}$$

[3] ∇ は $\mathbb{R}^d$ 上の勾配作用素である．

を示せば十分である．I の定義 (2.0.1) において $\theta = x/|x|, x \neq 0$ とすれば

$$I(x) \geq |x| - \log M\left(\frac{x}{|x|}\right) \geq |x| - \sup_{e \in \mathbb{R}^d : |e|=1} \log M(e)$$

が得られる．(1) より特に **(FEM)** の下では M は連続なので，最後の上限は有限である．よって (2.2.1) が示されたので，I は良い速度関数である．□

次に条件 **(SLG)** について補足的な命題を述べる．X_1 の法則 μ が与えられたとき，**(SLG)** が成立するかどうかは非自明である．そのため次の命題で **(SLG)** の成立・不成立に関して，μ の台に関する十分条件を与えておく．

命題 2.2.2. $\mathbb{R}^d$ 値確率変数 X_1 の法則を μ とおく．このとき次が成立する．

(1) μ の台が $\mathbb{R}^d$ のとき，**(SLG)** が成立する．

(2) μ の台がコンパクトのとき，**(SLG)** は成立しない．

証明. まず μ の台が $\mathbb{R}^d$ だと仮定する．$\mathbb{R}^d$ の単位球面を $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d \mid |x| = 1\}$ とおくと，内積の連続性と $\mathbb{S}^{d-1}$ のコンパクト性により，$\mathbb{S}^{d-1}$ の有限個の空でない開集合 $O_1, \ldots, O_N$ が存在して，$\mathbb{S}^{d-1} = \cup_{i=1}^N O_i$ かつ

$$x, y \in O_i \Longrightarrow \langle x, y \rangle \geq \frac{1}{2}$$

を満たす．$r \geq 1$ を任意に取り，各 $i = 1, \ldots, N$ に対して

$$O_i^r := \{\rho x \mid x \in O_i, 2r < \rho < 2r+1\}$$

と定める．$\theta \in \mathbb{R}^d \setminus \{0\}$ を任意に取ると，定義によりある $i_\theta = 1, \ldots, N$ が存在して $\theta/|\theta| \in O_{i_\theta}$ となる．よって任意の $x \in O_{i_\theta}^r$ に対して $x/|x| \in O_{i_\theta}$ となるので，

$$\langle \theta, x \rangle = |\theta||x|\langle \frac{\theta}{|\theta|}, \frac{x}{|x|} \rangle \geq r|\theta|$$

が成立する．これにより

$$M(\theta) \geq \int_{O_{i_\theta}^r} e^{\langle \theta, x \rangle} \mu(dx) \geq e^{r|\theta|} \mu(O_{i_\theta}^r) \geq e^{r|\theta|} \min_{i=1,\ldots,N} \mu(O_i^r)$$

を得るので，

$$\frac{\Lambda(\theta)}{|\theta|} \geq r + \frac{1}{|\theta|} \log \min_{i=1,\ldots,N} \mu(O_i^r)$$

を得るが，右辺の最小値は θ によらず，台についての仮定により正なので

$$\varliminf_{|\theta|\to\infty} \frac{\Lambda(\theta)}{|\theta|} \geq r$$

となる．$r \geq 1$ は任意なので **(SLG)** が成立する．

一方 μ の台がコンパクトだと仮定する．このときある定数 $C_0 > 0$ が存在して，ほとんど確実に $|X_1| \leq C_0$ である．よって $|\Lambda(\theta)| \leq C_0|\theta|$ となるので，**(SLG)** は成立しない． □

注意 2.2.3. 命題 2.2.2 の証明から簡単にわかるように，$d = 1$ のときは条件 **(SLG)** と μ の台が上からも下からも非有界であることは同値である．

2.3　1 次元の場合

独立確率変数の列に対する大偏差事象の確率は歴史的には，X_1 の分布が Bernoulli 分布の場合に 1929 年に Khinchin により調べられ，その後 1938 年に Cramér により一般の分布の場合に調べられた．本節では $d = 1$ の場合を考え，定義 1.1.2 の意味の大偏差原理ではなく，まず Khinchin と Cramér にならって次の定理を示す.[4]

定理 2.3.1. 条件 **(FEM)** と **(SLG)** を仮定する．任意の $x > m$ に対して次が成立する．

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \geq x\right) = -I(x). \tag{2.3.1}$$

本節では $x > m$ に対する末尾確率 $\mathbb{P}(S_n/n \geq x)$ を扱うが，$x < m$ に対する末尾確率 $\mathbb{P}(S_n/n \leq x)$ についても同じ主張が成立する，すなわち，任意の $x < m$ に対して次が成立する．

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \leq x\right) = -I(x).$$

証明は定理 2.3.1 の証明と同じなので省略する．

定理 2.3.1 を証明する前に，(2.3.1) の左辺の極限は常に存在することを見る．これを見るために，優加法的な数列は収束するという事実を用いる．ここで，実数列 $\{a_n\}_{n\in\mathbb{N}}$ が優加法的であるとは，任意の $m, n \in \mathbb{N}$ に対して $a_{m+n} \geq a_m + a_n$ が成立することである．この事実は Fekete の補題として知られており，確率論において様々な場面で現れる．

[4] 本節の記述は [10, 第 1.5 節] と [42, 第 2.4 節] を参考にした．

補題 2.3.2 (Feketeの補題). 実数列 $\{a_n\}_{n\in\mathbb{N}}$ は優加法的であると仮定する．このとき数列 $\{a_n/n\}_{n\in\mathbb{N}}$ は $\overline{\mathbb{R}}$ において収束し，次が成立する．

$$\lim_{n\to\infty}\frac{a_n}{n}=\sup_{m\in\mathbb{N}}\frac{a_m}{m}.$$

証明. 明らかに $\overline{\lim}_{n\to\infty}(a_n/n)\leq\sup_{m\in\mathbb{N}}(a_m/m)$ なので，逆向きの不等式 $\underline{\lim}_{n\to\infty}(a_n/n)\geq\sup_{m\in\mathbb{N}}(a_m/m)$ を示せばよい．そのためには，任意の $m\in\mathbb{N}$ に対して $\underline{\lim}_{n\to\infty}(a_n/n)\geq(a_m/m)$ を示せばよい．以下 $m\in\mathbb{N}$ を固定する．

任意に $n\geq m$ を取り，$n=km+l$ $(k,l\in\mathbb{N},0\leq l<m)$ と書く．優加法性の性質を繰り返し用いることにより，$a_n\geq ka_m+a_l$ を得る．ただし $a_0:=0$ とおいた．よって $n=km+l$ で両辺を割ることにより

$$\frac{a_n}{n}\geq\frac{km}{km+l}\cdot\frac{a_m}{m}+\frac{a_l}{n}$$

を得る．m は固定されているので，$n\to\infty$ のとき

$$\left|\frac{a_l}{n}\right|\leq\frac{\max_{0\leq i<m}|a_i|}{n}\longrightarrow 0$$

である．よって $n\to\infty$ とすれば $k\to\infty$ となるので，$\underline{\lim}_{n\to\infty}(a_n/n)\geq(a_m/m)$ が得られ，補題を得る．□

大偏差事象の確率に戻る．$x\in\mathbb{R}$ を任意に取り，(2.3.1) の左辺の極限が存在することを示す．各 $n\in\mathbb{N}$ に対して $a_n:=\log\mathbb{P}(S_n\geq nx)$ とおくと，数列 $\{a_n\}_{n\in\mathbb{N}}$ は優加法的である．実際，任意の $m,n\in\mathbb{N}$ に対して

$$\begin{aligned}a_m+a_n&=\log\left(\mathbb{P}(S_m\geq mx)\mathbb{P}(S_n\geq nx)\right)\\&=\log\left(\mathbb{P}(S_m\geq mx)\mathbb{P}(S_{m+n}-S_m\geq nx)\right)\\&=\log\left(\mathbb{P}(S_m\geq mx,S_{m+n}-S_m\geq nx)\right)\\&\leq\mathbb{P}(S_{m+n}\geq(m+n)x)=a_{m+n}\end{aligned}$$

となる．ただし，２つ目の等号では S_n と $S_{m+n}-S_m$ が同分布であることを，３つ目の等号では S_n と $S_{m+n}-S_m$ が独立であることをそれぞれ用いた．よって Feketeの補題により次の極限が存在する．

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\geq x\right).$$

Feketeの補題のみでは，この極限の具体形や性質などは分からない．定理 2.3.1 はこの極限が対数積率母関数の Fenchel-Legendre 変換であることを特定する．

それでは定理 2.3.1 を証明する．

定理 2.3.1 の証明. $x > m$ を固定する．仮定 **(SLG)** と M の連続性により，I の定義式 (2.0.1) の右辺はある $\theta_x \in \mathbb{R}$ により達成されることがわかる，つまりある $\theta_x \in \mathbb{R}$ が存在して

$$I(x) = \theta_x x - \Lambda(\theta_x) \tag{2.3.2}$$

を満たす．また $\theta \mapsto \theta x - \Lambda(\theta)$ は $\theta = \theta_x$ で極大値を達成するので

$$(\theta x - \Lambda(\theta))' |_{\theta=\theta_x} = 0 \quad \Longleftrightarrow \quad x = \frac{M'(\theta_x)}{M(\theta_x)} \tag{2.3.3}$$

が成立する．

$(\Omega, \mathcal{F}, \mathbb{P})$ を確率変数 $\{X_i\}_{i=1}^{\infty}$ が定義されている確率空間とする．$(\Omega, \mathcal{F})$ 上の確率測度 $\mathbb{P}_{x,n}$ を，$\mathbb{P}$ に絶対連続であってその Radon-Nikodym 微分が

$$\frac{d\mathbb{P}_{x,n}}{d\mathbb{P}} = \frac{e^{\theta_x S_n}}{\mathbb{E}\left[e^{\theta_x S_n}\right]} \tag{2.3.4}$$

で与えられるものとする．確率測度 $\mathbb{P}_{x,n}$ に関する期待値は $\mathbb{E}_{x,n}[\cdot]$ と書く．このとき確率変数の列 $\{X_i\}_{i=1}^{n}$ は $\mathbb{P}_{x,n}$ の下で独立同分布かつ X_1 の平均は x であり，X_1 の $\mathbb{P}_{x,n}$ の下での分布は n に依らない．実際，任意の Borel 集合 $A_k \in \mathcal{B}(\mathbb{R}), k = 1, \ldots, n$ に対して次が成立する．

$$\begin{aligned}
\mathbb{P}_{x,n}\left(\bigcap_{k=1}^{n}\{X_k \in A_k\}\right) &= \mathbb{E}\left[\left(\prod_{k=1}^{n} \mathbf{1}_{\{X_k \in A_k\}}\right) \frac{e^{\theta_x S_n}}{\mathbb{E}\left[e^{\theta_x S_n}\right]}\right] \\
&= \prod_{k=1}^{n} \mathbb{E}\left[\mathbf{1}_{\{X_k \in A_k\}} \frac{e^{\theta_x X_k}}{\mathbb{E}\left[e^{\theta_x X_k}\right]}\right] \\
&= \prod_{k=1}^{n} \mathbb{E}\left[\mathbf{1}_{\{X_k \in A_k\}} \frac{e^{\theta_x S_n}}{\mathbb{E}\left[e^{\theta_x S_n}\right]}\right] \\
&= \prod_{k=1}^{n} \mathbb{P}_{x,n}\left(\{X_k \in A_k\}\right).
\end{aligned}$$

ただし１つ目と４つ目の等号は $\mathbb{P}_{x,n}$ の定義から従い，２つ目と３つ目の等号は $\{X_i\}_{i=1}^{n}$ が $\mathbb{P}$ の下で独立同分布であることから従う．これにより $\{X_i\}_{i=1}^{n}$ は独立である．また３つ目の等号で見たように

$$\mathbb{P}_{x,n}\left(X_k \in A\right) = \mathbb{E}\left[\mathbf{1}_{\{X_k \in A\}} \frac{e^{\theta_x X_k}}{\mathbb{E}\left[e^{\theta_x X_k}\right]}\right] = \mathbb{E}\left[\mathbf{1}_{\{X_1 \in A\}} \frac{e^{\theta_x X_1}}{\mathbb{E}\left[e^{\theta_x X_1}\right]}\right] \tag{2.3.5}$$

であるから，$\{X_i\}_{i=1}^{n}$ は同分布である．また $\mathbb{P}_{x,n}$ の下での X_1 の平均は

$$\mathbb{E}_{x,n}\left[X_1\right] = \frac{\mathbb{E}\left[X_1 e^{\theta_x S_n}\right]}{\mathbb{E}\left[e^{\theta_x S_n}\right]} = \frac{\mathbb{E}\left[X_1 e^{\theta_x X_1}\right]}{\mathbb{E}\left[e^{\theta_x X_1}\right]} = \frac{M'(\theta_x)}{M(\theta_x)} = x$$

となる．ただし2つ目の等号は $\{X_i\}_{i=1}^n$ が $\mathbb{P}$ の下で独立同分布であることから従い，4つ目の等号では (2.3.3) を用いた．

ここで $\overline{S}_n := S_n - nx$ とおき，再び Radon-Nikodym 微分の定義を用いると，

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{n} \geq x\right) &= \mathbb{E}_{x,n}\left[\mathbf{1}_{\{S_n \geq nx\}} \frac{\mathbb{E}\left[e^{\theta_x S_n}\right]}{e^{\theta_x S_n}}\right] \\ &= e^{-n\{\theta_x x - \log M(\theta_x)\}} \mathbb{E}_{x,n}\left[\mathbf{1}_{\{\overline{S}_n \geq 0\}} e^{-\theta_x \overline{S}_n}\right] \\ &= e^{-nI(x)} \mathbb{E}_{x,n}\left[\mathbf{1}_{\{\overline{S}_n \geq 0\}} e^{-\theta_x \overline{S}_n}\right]\end{aligned}$$

を得る．ただし最後の等号では (2.3.2) を用いた．よって次を示せば定理 2.3.1 の証明が完了する．

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{x,n}\left[\mathbf{1}_{\{\overline{S}_n \geq 0\}} e^{-\theta_x \overline{S}_n}\right] = 0. \tag{2.3.6}$$

(2.3.6) を示す．Jensen の不等式により，任意の $\theta \in \mathbb{R}$ に対して $e^{\theta m} \leq M(\theta)$ が成立するので，

$$\theta x - \Lambda(\theta) \leq \theta x - \theta m$$

である．$I(x)$ は非負であり $x > m$ なので $\theta_x > 0$ であることがわかる．また X_1 の $\mathbb{P}_{x,n}$ の下での分散を σ^2 と書く．(2.3.5) からわかるように σ^2 は n には依存せず，条件 **(FEM)** により $\sigma^2 < \infty$ である．$\sigma = 0$ のときは (2.3.6) は明らかに成立するので，$\sigma^2 > 0$ のときに (2.3.6) を示す．$\theta_x > 0$ なので Chebyshev の不等式により

$$e^{-\theta_x\sqrt{n}} \mathbb{P}_{x,n}\left(\frac{\overline{S}_n}{\sqrt{n}} \in [0,1]\right) \leq \mathbb{E}_{x,n}\left[\mathbf{1}_{\{\overline{S}_n \geq 0\}} e^{-\theta_x \overline{S}_n}\right] \leq 1$$

を得る．中心極限定理により

$$\lim_{n\to\infty} \mathbb{P}_{x,n}\left(\frac{\overline{S}_n}{\sqrt{n}} \in [0,1]\right) = \int_0^1 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx > 0$$

となる．(2.3.6) が示されたので，定理 2.3.1 が示された． □

定理 2.3.1 の証明で重要な点は，新たな確率測度 $\mathbb{P}_{x,n}$ を導入したことである．大数の法則により元の確率測度 $\mathbb{P}$ の下では平均的に m が観測されるが，新たな確率測度 $\mathbb{P}_{x,n}$ の下では平均的に x が観測される．x は平均的には観測されにくい稀な値であるため，平均を m から x に変更するには指数的に大きい代償がかかる．そのような変更の方法はいくつもあるが，一番良い変更の方法が確率測度を $\mathbb{P}$ から $\mathbb{P}_{x,n}$ にすることであり，その際かかる最良の代償が $e^{nI(x)}$ となっており，その逆数が極限に現れる．大偏差事象の確率を調べる際には多くの場合に，稀な値を達成させる状況を特定する必要がある．今回は

$\mathbb{P}_{x,n}$ を導入することがそれに対応する．このような測度の変更は Cramér の手法と呼ばれており，大偏差原理を調べる際に典型的に行われる手法である．次の節で見るように，Cramér の手法は大偏差原理の下からの評価で実行される．

本節の最後の内容として，定理 2.3.1 の収束が大偏差原理を意味することを示す．

命題 2.3.3. $d=1$ とし，条件 **(FEM)** と **(SLG)** を仮定する．このとき $\mathbb{R}$ 値確率変数の列 $\{S_n/n\}_{n\in\mathbb{N}}$ は $n\to\infty$ のときに良い速度関数 I に対して速度 n で大偏差原理を満たす．すなわち任意の閉集合 $C\subset\mathbb{R}$ に対して

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\in C\right)\le-\inf_{x\in C}I(x)$$

が成立し，任意の開集合 $O\subset\mathbb{R}$ に対して

$$\varliminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\in O\right)\ge-\inf_{x\in O}I(x)$$

が成立する．

証明. 証明の本筋に入る前に，I の性質についていくつか述べる．$m:=\mathbb{E}\left[X_1\right]$ とおく．条件 **(SLG)** により $\{x\in\mathbb{R}\mid I(x)<\infty\}=\mathbb{R}$ であり，I は凸なので命題 A.1.3 により $\mathbb{R}$ 上連続である．また m は I の唯一の零点である．m が唯一の零点であることを示すために，$x>m$ を任意に取る．ここで

$$\lim_{\theta\searrow 0}\frac{\Lambda(\theta)}{\theta}=\Lambda'(0)=m<x$$

であるから，$\theta>0$ が小さければ $\theta x-\Lambda(\theta)>0$ となり，$I(x)>0$ がわかる．同様に $x<m$ のときも $I(x)>0$ となるので，m は唯一の零点である．m は I の唯一の零点であるので，系 A.1.4 により I は $(-\infty,m]$ 上で狭義単調減少かつ $[m,\infty)$ 上で狭義単調増加である．

$C\subset\mathbb{R}$ を閉集合とする．$C=\emptyset$ または $m\in C$ のときは，大偏差原理の上からの評価は明らかなので，$C\neq\emptyset$ かつ $m\notin C$ のときに示す．x_C^- と x_C^+ をそれぞれ次で定義する．

$$x_C^-:=\sup\{x\in C\mid x<m\},\quad x_C^+:=\inf\{x\in C\mid x>m\}.$$

上限と下限に現れる集合のどちらか一方は空集合になり得るが，その場合は空でないもののみを考える．$m\notin C$ かつ C は閉集合であるので，$x_C^-<m<x_C^+$ かつ $x_C^\pm\in C$ が成立する．さらに，$x_C^\pm$ の定義により

$$\left\{\frac{S_n}{n}\in C\right\}\subset\left\{\frac{S_n}{n}\le x_C^-\right\}\cup\left\{\frac{S_n}{n}\ge x_C^+\right\}$$

が成立する．よって定理 2.3.1 により

$$\begin{aligned}\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\in C\right) &\le \max\left\{\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\le x_C^-\right), \varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\ge x_C^+\right)\right\}\\ &= \max\{-I(x_C^-), -I(x_C^+)\}\\ &= -\min\{I(x_C^-), I(x_C^+)\}\end{aligned}$$

が成立する．ただし最初の不等式では補題 1.2.2 を用いた．一方 I の単調性と $x_C^\pm \in C$ により

$$\min\{I(x_C^-), I(x_C^+)\} = \inf_{x\in C} I(x)$$

が成立するので，大偏差原理の上からの評価を得る．

$O\subset\mathbb{R}$ を開集合とし，$x\in O$ を任意に取る．大偏差原理の下からの評価を示すためには,

$$\varliminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\in O\right)\ge -I(x) \tag{2.3.7}$$

を示せば十分である．$x>m$ のときに (2.3.7) を示す．$x\le m$ のときは証明を少し修正すればよい．O は開集合なので，$\varepsilon>0$ を小さく取ると $m<x-\varepsilon$ かつ $[x-\varepsilon, x+\varepsilon)\subset O$ となる．よって

$$\mathbb{P}\left(\frac{S_n}{n}\in O\right)\ge\mathbb{P}\left(\frac{S_n}{n}\ge x-\varepsilon\right)-\mathbb{P}\left(\frac{S_n}{n}\ge x+\varepsilon\right)$$

を得る．ここで δ_0 を

$$\delta_0 := \frac{1}{2}\left(I(x+\varepsilon)-I(x-\varepsilon)\right)$$

と定義すると，I の狭義単調性により $\delta_0>0$ である．ここで $0<\delta<\delta_0$ を任意に取ると，定理 2.3.1 により，十分大きい任意の $n\in\mathbb{N}$ に対して

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{n}\ge x-\varepsilon\right)&\ge\exp\left\{-n(I(x-\varepsilon)+\delta)\right\},\\ \mathbb{P}\left(\frac{S_n}{n}\ge x+\varepsilon\right)&\le\exp\left\{-n(I(x+\varepsilon)-\delta)\right\}\end{aligned}$$

を得る．よって

$$\mathbb{P}\left(\frac{S_n}{n}\in O\right)\ge\exp\left\{-n(I(x-\varepsilon)+\delta)\right\}\left(1-e^{-nc_\delta}\right)$$

である．ただし $c_\delta>0$ を

$$c_\delta := I(x+\varepsilon)-I(x-\varepsilon)-2\delta$$

とおいた．簡単にわかるように

$$\lim_{n\to\infty} \frac{\log\left(1-e^{-nc_\delta}\right)}{n} = 0$$

となるので，

$$\varliminf_{n\to\infty} \mathbb{P}\left(\frac{S_n}{n} \in O\right) \geq -I(x-\varepsilon) - \delta$$

を得る．よって $\delta \searrow 0$ とした後に $\varepsilon \searrow 0$ とすれば，(2.3.7) を得るので本命題の証明が終わる． □

2.4　多次元の場合

本節では次元 $d \in \mathbb{N}$ が一般のとき，条件 **(SLG)** も仮定に追加して定理 2.0.1 を証明する．大偏差原理の定義に基づき上からの評価と下からの評価をそれぞれを示すわけだが，証明に進む前に概略を述べておく．上からの評価ではまずコンパクト凸集合に対して評価を示す．指数型の Jensen の不等式を用いた後に minimax 定理を用いる．コンパクト凸を仮定する理由は，この定理を用いるためである．その後，指数的緊密性を示すことで上からの評価は一般の閉集合へと拡張される．下からの評価は本質的に定理 2.3.1 の証明と同じである．平均値から離れた稀な値に対して，その値の小さい近傍の確率を評価する．定理 2.3.1 の証明と同様に，その稀な値が平均値として現れるように確率測度の変更を行うが，その際に支払う代償として速度関数が現れる．新たな測度の下でのその近傍の確率も評価する必要があるが，それは導入した確率測度の下での大数の法則により極限がわかる．[5]

上からの評価に進む前に，次の von Neumann の minimax 定理を紹介する．$\mathbb{R}^d$ 内のコンパクト凸集合という設定で述べているが，より一般の場合にも拡張されることが知られている．[6]

定理 2.4.1. $C, \Theta \subset \mathbb{R}^d$ をコンパクト凸集合とし，関数 $f : C \times \Theta \to \mathbb{R}$ は次の 4 条件を満たすとする．

(i) 任意の $\theta \in \Theta$ に対して $f(\cdot, \theta)$ は凸である．

(ii) 任意の $\theta \in \Theta$ に対して $f(\cdot, \theta)$ は下半連続である．

(iii) 任意の $x \in C$ に対して $f(x, \cdot)$ は凹である．

[5] 本節の記述は [42, 第 2.4 節] を参考にした．
[6] 詳細については例えば [42, Appendix A.5] を参照せよ．

(iv) 任意の $x \in C$ に対して $f(x, \cdot)$ は上半連続である．

このとき

$$\sup_{\theta \in \Theta} \inf_{x \in C} f(x, \theta) = \inf_{x \in C} \sup_{\theta \in \Theta} f(x, \theta)$$

が成立する．

証明. まず左辺より右辺の方が大きいことは，f, C, Θ に関する条件無しに成立することに注意する．実際，任意の $y \in C$ と $\theta \in \Theta$ に対して

$$\inf_{x \in C} f(x, \theta) \leq f(y, \theta) \leq \sup_{\psi \in \Theta} f(y, \psi)$$

が成立するので，$\theta \in \Theta$ についての上限と $y \in C$ についての下限を取り，y と ψ をそれぞれ x と θ に書き直すことにより

$$\sup_{\theta \in \Theta} \inf_{x \in C} f(x, \theta) \leq \inf_{x \in C} \sup_{\theta \in \Theta} f(x, \theta)$$

が得られる．

f に対する次の条件 (i)' を導入する．

(i)' 任意の $\theta \in \Theta$ に対して $f(\cdot, \theta)$ は真に凸である．[7]

定理を示す前段として，条件 (i)' と (ii)–(iv) の下ではある $(x_*, \theta_*) \in C \times \Theta$ が存在して

$$f(x_*, \theta) \leq f(x_*, \theta_*) \leq f(x, \theta_*), \quad (x, \theta) \in C \times \Theta \tag{2.4.1}$$

が成立することを示す．このような (x_*, θ_*) が存在すれば，(2.4.1) の左側の不等式において $\theta \in \Theta$ についての上限を取れば

$$f(x_*, \theta_*) \geq \sup_{\theta \in \Theta} f(x_*, \theta) \geq \inf_{x \in C} \sup_{\theta \in \Theta} f(x, \theta)$$

を得る．一方，(2.4.1) の右側の不等式において $x \in C$ について下限を取れば

$$f(x_*, \theta_*) \leq \inf_{x \in C} f(x, \theta_*) \leq \sup_{\theta \in \Theta} \inf_{x \in C} f(x, \theta)$$

[7] 任意の $x_1, x_2 \in C, x_1 \neq x_2, t \in (0, 1)$ に対して

$$tf(x_1, \theta) + (1-t)f(x_2, \theta) > f(tx_1 + (1-t)x_2, \theta)$$

が成立することをいう．

を得る．よって逆向きの不等式

$$\sup_{\theta\in\Theta}\inf_{x\in C} f(x,\theta) \geq \inf_{x\in C}\sup_{\theta\in\Theta} f(x,\theta) \tag{2.4.2}$$

が得られる．

ある $(x_*, \theta_*) \in C \times \Theta$ が存在して (2.4.1) が成立することを示す．C はコンパクトなので，(ii) により各 $\theta \in \Theta$ に対してある $x(\theta) \in C$ が一意的に存在して

$$m(\theta) := \min_{x\in C} f(x,\theta) = f(x(\theta),\theta)$$

となることがわかる．存在性は補題 1.2.6 により，一意性は (i)' による.[8]また各 $x \in C$ に対して，(iv) により $f(x,\cdot)$ は Θ 上で上半連続なので，m は上半連続である．よって Θ はコンパクトなので m はある $\theta_* \in \Theta$ で最大値

$$m(\theta_*) = \max_{\theta\in\Theta} m(\theta)$$

を取る．

ここで (2.4.1) の右側の不等式を示す．$\theta \in \Theta$ に対して，$x(t,\theta) := x((1-t)\theta_* + t\theta)$ とおく．$\{x(t,\theta)\}_{t>0}$ はコンパクト集合 C に含まれるので，0 に収束する単調減少列 $\{t_n\}_{n\in\mathbb{N}}$ を適当に取れば $\{x(t_n,\theta)\}_{n\in\mathbb{N}}$ はある $y(\theta) \in C$ に収束する.（$\{t_n\}_{n\in\mathbb{N}}$ も θ の取り方に依存する.）実は $y(\theta)$ は θ に依らず，$y(\theta) = x(\theta_*)$ であることを次に見る．$x(t_n,\theta)$ の定義により任意の $x \in C$ と $n \in \mathbb{N}$ に対して

$$f(x(t_n,\theta), (1-t_n)\theta_* + t_n\theta) \leq f(x, (1-t_n)\theta_* + t_n\theta)$$

である．一方 (iii) により

$$f(x(t_n,\theta), (1-t_n)\theta_* + t_n\theta) \geq (1-t_n)f(x(t_n,\theta),\theta_*) + t_n f(x(t_n,\theta),\theta)$$

である．また $f(x(t_n,\theta),\theta) \geq m(\theta)$ と (ii) および (iv) により，$n \to \infty$ とすれば任意の $x \in C$ に対して $f(y(\theta),\theta_*) \leq f(x,\theta_*)$ が成立する．よって $x(\theta_*)$ の一意性から $y(\theta) = x(\theta_*)$ がわかり，x_* を $x(\theta_*)$ とおけば (2.4.1) の右側の不等式が得られる．

次に (2.4.1) の左側の不等式を示すために，任意の $\theta \in \Theta, t \in (0,1)$ に対して

$$f(x(t,\theta),\theta) \leq m(\theta_*) \tag{2.4.3}$$

を示す．再び (iii) により

$$(1-t)f(x(t,\theta),\theta_*) + tf(x(t,\theta),\theta) \leq f(x(t,\theta),(1-t)\theta_* + t\theta) \tag{2.4.4}$$

[8] 補題 1.2.6 では閉集合上で下から有界と仮定しているが，コンパクトであれば同じ結論を得られるので，ここではこの事実を用いている．

を得る．ここで右辺は $m((1-t)\theta_*+t\theta)$ と書けるので，θ_* の定義から

$$f(x(t,\theta),(1-t)\theta_*+t\theta)\le m(\theta_*)$$

である．一方 $x(\theta_*)$ と $m(\theta_*)$ の定義から

$$f(x(t,\theta),\theta_*)\ge f(x(\theta_*),\theta_*)=m(\theta_*)$$

である．これらの式を (2.4.4) に代入して整理すれば (2.4.3) を得る．最後に (2.4.3) において $t=t_n$ とおき $\to\infty$ とすれば，(2.4.1) の左側の不等式が得られる．これで (i) を (i)' でにおき替えた場合に定理が示された．

最後に (i)–(iv) を満たす f に対して (2.4.2) を示す．$\varepsilon>0$ に対して

$$f_\varepsilon(x,\theta):=f(x,\theta)+\varepsilon|x|^2$$

とおく．f_ε は (i)' 及び (ii)–(iv) を満たすので，

$$\sup_{\theta\in\Theta}\inf_{x\in C}f_\varepsilon(x,\theta)\ge\inf_{x\in C}\sup_{\theta\in\Theta}f_\varepsilon(x,\theta)$$

が成立する．明らかに

$$\inf_{x\in C}\sup_{\theta\in\Theta}f_\varepsilon(x,\theta)\ge\inf_{x\in C}\sup_{\theta\in\Theta}f(x,\theta)$$

である．一方 $C_0:=\sup_{x\in C}|x|^2<\infty$ とおくと，

$$\sup_{\theta\in\Theta}\inf_{x\in C}f_\varepsilon(x,\theta)\le\sup_{\theta\in\Theta}\inf_{x\in C}f(x,\theta)+\varepsilon C_0$$

となっているので，$\varepsilon\to0$ とすれば (2.4.2) を得る．以上により定理 2.4.1 の証明が完結した．$\square$

それでは一般の次元 $d\in\mathbb{N}$ に対して，条件 **(FEM)** と **(SLG)** の下で定理 2.0.1 を証明する．

条件 (FEM) と (SLG) の下での定理 2.0.1 の証明.

条件 **(FEM)** と **(SLG)** を仮定する．まず大偏差原理の上からの評価を示す．$C\subset\mathbb{R}^d$ をコンパクト凸集合とし，$\theta\in\mathbb{R}^d$ とする．Chebyshev の不等式から

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{n}\in C\right)&\le\mathbb{P}\left(\langle\theta,S_n\rangle\ge\inf_{x\in C}n\langle\theta,x\rangle\right)\\&\le\exp\left(-\inf_{x\in C}n\langle\theta,x\rangle\right)\mathbb{E}\left[e^{\langle\theta,S_n\rangle}\right]\end{aligned}$$

$$= \exp\left(-\inf_{x\in C} n\langle\theta, x\rangle\right) M(\theta)^n$$

を得る．よって $\theta \in \mathbb{R}^d$ について下限を取れば

$$\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\in C\right) \le -\sup_{\theta\in\mathbb{R}^d}\inf_{x\in C}\{\langle\theta, x\rangle - \Lambda(\theta)\} \tag{2.4.5}$$

がわかる．C は有界なので仮定 **(SLG)** により，ある $R>0$ が存在して任意の $x\in C$ と $\theta\notin\overline{B}(0,R)$ に対して

$$\langle\theta, x\rangle - \Lambda(\theta) \le -1$$

が成立する．また任意の $x \in C$ に対して

$$\sup_{\theta\in\mathbb{R}^d}\{\langle\theta, x\rangle - \Lambda(\theta)\} \ge 0$$

となるので，(2.4.5) の θ についての上限を取る範囲は $\overline{B}(0,R)$ に制限してよい．また θ と x についての関数 $\langle\theta, x\rangle - \Lambda(\theta)$ は定理 2.4.1 の仮定を満たすので，$\theta\in\mathbb{R}^d$ についての上限と $x\in C$ についての下限の順番を入れ替えられる．よって速度関数 I の定義 (2.0.1) により，コンパクト凸集合 $C\subset\mathbb{R}^d$ に対して上からの評価

$$\frac{1}{n}\log\mathbb{P}\left(\frac{S_n}{n}\in C\right) \le -\inf_{x\in C} I(x)$$

が得られる．[9]

次にこの評価を $C\subset\mathbb{R}^d$ がコンパクトである場合に拡張する．$\alpha < \inf_{x\in C} I(x)$ を任意に取る．I は下半連続なので $I^{-1}((\alpha,\infty])$ は開集合である．よって各 $x\in C$ に対して x を中心とする閉球 $\overline{B}(x,\delta_x)$ であって，$\overline{B}(x,\delta_x)\subset I^{-1}((\alpha,\infty])$ となるものが取れる．C がコンパクト集合なので開被覆 $\{B(x,\delta_x)\}_{x\in C}$ から有限開被覆 $\{B(x_1,\delta_{x_1}),\dots,B(x_N,\delta_{x_N})\}$ が取れる．閉球はコンパクト凸であるから，$\mathbb{P}$ の劣加法性とコンパクト凸集合に対する上からの評価により

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{n}\in C\right) &\le \sum_{i=1}^{N}\mathbb{P}\left(\frac{S_n}{n}\in\overline{B}(x_i,\delta_{x_i})\right)\\ &\le \sum_{i=1}^{N}\exp\{-n\inf_{x\in\overline{B}(x_i,\delta_{x_i})} I(x)\} \le Ne^{-n\alpha}\end{aligned}$$

を得る．よって $n\to\infty$ とした後に $\alpha\nearrow\inf_{x\in C} I(x)$ とすれば，C がコンパクト集合の場合に上からの評価が得られる．

[9] この場合には任意の $n\in\mathbb{N}$ に対して，大偏差原理の上からの評価が成立することに注意する．

最後に一般の閉集合 $C \subset \mathbb{R}^d$ の場合に上からの評価を示す．仮定 **(FEM)** により確率変数の列 $\{S_n/n\}_{n\in\mathbb{N}}$ は指数的に緊密である．実際，$x \in \mathbb{R}^d$ の第 i 成分を $x^{(i)}$ で書くと，$\alpha > 0$ に対して

$$\mathbb{P}\left(\frac{|S_n|}{n} \geq \alpha\right) \leq \mathbb{P}\left(\sum_{i=1}^d |S_n^{(i)}| \geq n\alpha\right) \leq \sum_{i=1}^d \mathbb{P}\left(|S_n^{(i)}| \geq \frac{n\alpha}{d}\right)$$

となる．各 i に対して Chebyshev の不等式から

$$\mathbb{P}\left(|S_n^{(i)}| \geq \frac{n\alpha}{d}\right) \leq e^{-(n\alpha/d)}\mathbb{E}\left[e^{S_n^{(i)}} + e^{-S_n^{(i)}}\right] = e^{-(n\alpha/d)}\left(M(e_i)^n + M(-e_i)^n\right)$$

がわかる．よって任意に $\beta > 0$ が与えられたとき，$\alpha = d\{\beta + \max_{i=1,\dots,d}(\Lambda(e_i) \vee \Lambda(-e_i))\}$ と取れば

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\left(\frac{|S_n|}{n} \geq \alpha\right) \leq -\beta$$

が得られるので，$\{S_n/n\}_{n\in\mathbb{N}}$ は指数的に緊密である．コンパクト集合に対する上からの評価と指数的緊密性を示したので，命題 1.3.1 により一般の閉集合に対する上からの評価が得られる．指数的緊密性の証明では仮定 **(SLG)** を用いていないことに注意する．

次に大偏差原理の下からの評価を示す．下からの評価は定理 2.3.1 の証明と基本的に同様である．条件 **(SLG)** を仮定しているので $I(x)$ を達成する $\theta_x \in \mathbb{R}^d$ を考えることや確率測度 $\mathbb{P}_{x,n}$ の導入は同様である．ただし今は多次元の設定なので (2.3.3) と (2.3.4) がそれぞれ

$$\nabla(\langle\theta, x\rangle - \Lambda(\theta))|_{\theta=\theta_x} = 0 \quad \Longleftrightarrow \quad x = \frac{\nabla M(\theta_x)}{M(\theta_x)}$$

$$\frac{d\mathbb{P}_{x,n}}{d\mathbb{P}} = \frac{e^{\langle\theta_x, S_n\rangle}}{\mathbb{E}\left[e^{\langle\theta_x, S_n\rangle}\right]}$$

になる．確率測度 $\mathbb{P}_{x,n}$ の下で $\{X_i\}_{i=1}^n$ が独立同分布かつ平均が x になることも同様に成立する．上記の事実の証明は省略して，これらを用いて下からの評価を示す．

$O \subset \mathbb{R}^d$ を開集合とする．下からの評価を示すためには，任意の $x \in O$ に対して

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \in O\right) \geq -I(x) \tag{2.4.6}$$

を示せば十分である．実際，(2.4.6) で $x \in O$ についての上限を取れば下からの評価が得られる．

(2.4.6) を示すために，$x \in O$ と $B(x,\varepsilon) \subset O$ となる $\varepsilon > 0$ を任意に取る．このとき

$$\mathbb{P}\left(\frac{S_n}{n} \in O\right) \geq \mathbb{P}\left(\frac{S_n}{n} \in B(x,\varepsilon)\right)$$

$$= \mathbb{E}_{x,n}\left[\mathbf{1}_{\{|S_n - nx| < n\varepsilon\}} \frac{\mathbb{E}\left[e^{\langle \theta_x, S_n\rangle}\right]}{e^{\langle \theta_x, S_n\rangle}}\right]$$
$$= e^{n \log M(\theta_x)} \mathbb{E}_{x,n}\left[\mathbf{1}_{\{|S_n - nx| < n\varepsilon\}} e^{-\langle \theta_x, S_n\rangle}\right]$$

となる．また $\{|S_n - nx| < n\varepsilon\}$ 上では

$$\langle \theta_x, S_n\rangle = \langle \theta_x, S_n - nx\rangle + n\langle \theta_x, x\rangle \le n(\langle \theta_x, x\rangle + |\theta_x|\varepsilon)$$

となっているので，

$$\mathbb{P}\left(\frac{S_n}{n} \in O\right) \ge \exp\left\{-n(\langle \theta_x, x\rangle - \log M(\theta_x) + |\theta_x|\varepsilon)\right\} \mathbb{P}_{x,n}\left(\left|\frac{S_n}{n} - x\right| < \varepsilon\right)$$

を得る．大数の法則により

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{x,n}\left(\left|\frac{S_n}{n} - x\right| < \varepsilon\right) = 0$$

であるから，$\langle \theta_x, x\rangle - \log M(\theta_x) = I(x)$ により

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \in O\right) \ge -I(x) - \varepsilon|\theta_x|$$

を得る．最後に $\varepsilon \searrow 0$ とすれば (2.4.6) が得られるので，下からの評価の証明が終了する． □

上記の証明に関していくつか注意を述べる．指数型の Chebyshev の不等式は大偏差原理の研究で最も基本的な不等式の一つである．本証明では minimax 定理を用いたが，1 次元の場合には速度関数 I の単調性を用いれば minimax 定理に相当する内容を示すことが可能であり，さらにこの議論を多次元の場合にも一般化することができる．本章で minimax 定理をあえて用いたのは，一般的な手法を紹介するためである．コンパクト集合についての上からの評価と指数的緊密性を示す議論は，大偏差原理の証明では常套手段である．下からの評価で重要なのは，証明で行った Cramér の手法と呼ばれる確率測度の変更であり，これも大偏差原理の証明では常套手段である．その際，新たな測度の下での大数の法則が必要になる．各種の大偏差原理を示すときに，これらの議論は詳細は違うものの多くの確率模型において行われる．

2.5　定理 2.0.1 の証明

本節では条件 **(SLG)** を仮定せずに Cramér の定理を示す．証明の鍵は指数的に良い近似（命題 1.6.4）である．[10]

[10] 本節の記述は [33, 第 2.2 節] を参考にした．

定理 2.0.1 の証明. $\{V_i\}_{i\in\mathbb{N}}$ を独立同分布な $\mathbb{R}^d$ 値確率変数の列であって，V_1 は d 次元標準正規分布に従うとする．さらに $\{V_i\}_{i\in\mathbb{N}}$ と $\{X_i\}_{i\in\mathbb{N}}$ は独立とする．$m\in\mathbb{N}$ を任意に取る．各 $i\in\mathbb{N}$ に対して $X_i^{(m)} := X_i + V_i/m$ とおくと，$X_1^{(m)}$ に対する対数積率母関数 Λ_m とその Fenchel-Legendre 変換 Λ_m^* は

$$\begin{aligned}\Lambda_m(\theta) &:= \log\mathbb{E}\left[e^{\langle\theta, X_1^{(m)}\rangle}\right] = \Lambda(\theta) + \frac{|\theta|^2}{2m^2}, \quad \theta\in\mathbb{R}^d,\\ \Lambda_m^*(x) &:= \sup_{\theta\in\mathbb{R}^d}\{\langle\theta,x\rangle - \Lambda_m(\theta)\}, \quad x\in\mathbb{R}^d,\end{aligned}$$

で与えられる．X_1 に対する条件 **(FEM)** により，$X_1^{(m)}$ に対して条件 **(FEM)** が成立する．また Jensen の不等式により任意の $\theta\in\mathbb{R}^d$ に対して $\Lambda(\theta)\geq\langle\theta,\mathbb{E}[X_1]\rangle$ が成立するので，$X_1^{(m)}$ に対して条件 **(SLG)** が成立する．よって各 $n\in\mathbb{N}$ に対して

$$S_n^{(m)} := \sum_{i=1}^n\left(X_i + \frac{V_i}{m}\right)$$

とおくと，$\mathbb{R}^d$ 値確率変数の列 $\{S_n^{(m)}/n\}_{n\in\mathbb{N}}$ は速度を n，速度関数を Λ_m^* として大偏差原理が成立する．

次に $\{S_n^{(m)}/n\}_{n\in\mathbb{N},m\in\mathbb{N}}$ は $\{S_n/n\}_{n\in\mathbb{N}}$ の指数的に良い近似であることを示す．定義により

$$\frac{S_n}{n} - \frac{S_n^{(m)}}{n} = -\frac{1}{mn}\sum_{i=1}^n V_i$$

である．$\{V_i\}_{i\in\mathbb{N}}$ は独立同分布な確率変数の列であって，その積率母関数は **(FEM)** と **(SLG)** を満たす．また対応する速度関数は $I_V(x) := |x|^2/2$ $(x\in\mathbb{R}^d)$ である（例 2.1.2 を参照）．よって条件 **(SLG)** の下での大偏差原理が適用できて，任意の $\delta>0$ に対して

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\left|\frac{S_n}{n} - \frac{S_n^{(m)}}{n}\right|\geq\delta\right)\leq -\frac{m^2\delta^2}{2}$$

が成立するので，$\{S_n^{(m)}/n\}_{n\in\mathbb{N},m\in\mathbb{N}}$ は $\{S_n/n\}_{n\in\mathbb{N}}$ の指数的に良い近似である．よって命題 1.6.4(1) により，$\{S_n/n\}_{n\in\mathbb{N}}$ は次で定義される速度関数 $J:\mathbb{R}^d\to[0,\infty]$ に対して弱大偏差原理を満たす.

$$J(x) := \sup_{\delta>0}\varliminf_{m\to\infty}\inf\{\Lambda_m^*(z)\mid z\in B(x,\delta)\}.$$

また多次元の場合の定理 2.0.1 の証明中で示したように，条件 **(FEM)** により $\{S_n/n\}_{n\in\mathbb{N}}$ は指数的に緊密なので，$I=J$ を示せば命題 1.3.1 により証明が終わる．

最後に $I=J$ を示す．任意の $m\in\mathbb{N}$ に対して明らかに $\Lambda_m\geq\Lambda$ なので，$\Lambda_m^*\leq I$ となり $J\leq I$ が成立する．各 $x\in\mathbb{R}^d$ に対して逆向きの不等式 $J(x)\geq I(x)$ を示す．$J(x)=\infty$ の

ときは明らかに成立するので，$J(x) < \infty$ のときに示す．$J(x)$ の定義により任意の $\delta > 0$ に対して

$$\begin{aligned} J(x) &\geq \varliminf_{m\to\infty} \inf\{\Lambda_m^*(z) \mid z \in B(x,\delta)\} \\ &\geq \varliminf_{m\to\infty} \inf\{\Lambda_m^*(z) \mid z \in \overline{B}(x,\delta), \Lambda_m^*(z) \leq J(x)+1\} \end{aligned} \tag{2.5.1}$$

が成立する．Λ_m^* は良い速度関数なので，ある $x_m^{(\delta)} \in \overline{B}(x,\delta)$ が存在して

$$\inf\{\Lambda_m^*(z) \mid z \in \overline{B}(x,\delta), \Lambda_m^*(z) \leq J(x)+1\} = \Lambda_m^*(x_m^{(\delta)})$$

が成立する．(2.5.1) の右辺の下極限を実現する部分列を取り，さらに部分列に沿って $x_m^{(\delta)} \to x^{(\delta)} \in \overline{B}(x,\delta)$ と収束したとする．(記号の簡単化のために部分列を同じ記号で表す.）このとき任意の $\theta \in \mathbb{R}^d$ に対して

$$\begin{aligned} \Lambda_m^*(x_m^{(\delta)}) &\geq \langle\theta, x_m^{(\delta)}\rangle - \Lambda_m(\theta) \\ &= \langle\theta, x_m^{(\delta)}\rangle - \Lambda(\theta) - \frac{|\theta|^2}{2m^2} \\ &\to \langle\theta, x^{(\delta)}\rangle - \Lambda(\theta) \end{aligned}$$

が成立する．よって $J(x) \geq \langle\theta, x^{(\delta)}\rangle - \Lambda(\theta)$ が得られる．$\theta \in \mathbb{R}^d$ についての上限を取れば $J(x) \geq I(x^{(\delta)})$ が得られる．$\delta \searrow 0$ とすると $x^{(\delta)} \to x$ なので，I の下半連続性により $J(x) \geq I(x)$ が得られるので，定理 2.0.1 の証明が終わる． □

2.6 Curie-Weiss 模型への応用

本節では Cramér の大偏差原理を，Curie-Weiss 模型と呼ばれる磁石の確率模型に応用する.[11]

Curie-Weiss 模型を導入する．$n \in \mathbb{N}$ に対し，以下で記述される n 個の原子からなる確率模型を考える．各原子は上向きスピン $+1$ か下向きスピン -1 を持つ．つまり原子のスピン配置は $\Omega_n = \{-1,+1\}^n$ の元として表現される．Ω_n は配置空間と呼ばれる．その元は $\omega = (\omega_i)_{i=1}^n$ と書くことにする．見本 ω の意味として，$\omega_i = +1$ のとき i 番目の原子は上向きスピンを持ち，$\omega_i = -1$ のとき下向きスピンを持つと解釈する．またスピン配置 $\omega \in \Omega_n$ のエネルギーは次で与えられる．

$$\mathcal{H}_n(\omega) := -\frac{1}{2n} \sum_{1\leq i,j\leq n} \omega_i\omega_j.$$

[11] 本節の記述は [42, 第 3.4 節] を参考にした．

$\mathcal{H}_n(\omega)$ は統計物理学においてハミルトニアンと呼ばれる．また P_n を Ω_n 上の一様分布，すなわち $P_n(\omega) = 2^{-n}$ $(\omega \in \Omega_n)$ とし，Ω_n 上の Gibbs 測度と呼ばれる確率測度を次で定義する．

$$\gamma_n^\beta(\omega) := \frac{1}{Z_n^\beta} e^{-\beta\mathcal{H}_n(\omega)} P_n(\omega).$$

ただし $\beta \geq 0$ は与えられたパラメーターであり，逆温度と呼ばれる．Z_n^β は γ_n^β を確率測度にするための正規化定数であり，分配関数と呼ばれる．

$$Z_n^\beta := \sum_{\omega\in\Omega_n} e^{-\beta\mathcal{H}_n(\omega)} P_n(\omega).$$

Curie-Weiss 模型の意味を簡単に説明する．P_n は Ω_n 上の一様分布であるが，言い換えるとパラメータ 1/2 の直積 Bernoulli 分布である．そのため温度が無限大（$\beta = 0$）の場合には原子間に相互作用はなく，各原子は独立に上向きスピンか下向きスピンを取る．一方温度が有限（$\beta > 0$）のときには，各原子間に相互作用が働きハミルトニアンの値に応じたスピン配置を確率的に取る．ハミルトニアンの定義に負の符号が含まれていることに注意すると，ハミルトニアンが小さくなるような配置が確率的に現れやすい．ここでハミルトニアンの定義から，スピンの値が揃っているほどハミルトニアンは小さくなる．しかしながら P_n の影響により各原子は独立にスピンの値を取るので,「スピンの値を揃える効果」と「独立にスピンを取る効果」の２つの効果の競合が生まれる．これら２つの競合効果により，以下で見るように Curie-Weiss 模型では相転移現象が観測される.[12]

スピン配置 $\omega \in \Omega_n$ に対してスピンの総和を S_n と書くことにする，つまり $S_n = \sum_{i=1}^n \omega_i$ である．このとき「スピン平均 S_n/n は γ_n^β の下でどのように振る舞うか」という問題を考える．Cramér の大偏差原理の応用として次の定理を得ることができる．

定理 2.6.1. (1) $\beta \leq 1$ とする．このとき γ_n^β の下で S_n/n は 0 に確率収束する．すなわち，任意の $\varepsilon > 0$ に対して次が成立する．

$$\lim_{n\to\infty} \gamma_n^\beta \left(\left| \frac{S_n}{n} \right| \leq \varepsilon \right) = 1.$$

(2) $\beta > 1$ とする．このときある定数 $m(\beta) \in (0,1)$ が存在して，γ_n^β の下で S_n/n は $(\delta_{m(\beta)} + \delta_{-m(\beta)})/2$ に法則収束する．すなわち，$0 < \varepsilon < m(\beta)$ となる任意の $\varepsilon > 0$ に対して次が成立する．

$$\lim_{n\to\infty} \gamma_n^\beta \left(\left| \frac{S_n}{n} - m(\beta) \right| \leq \varepsilon \right) = \lim_{n\to\infty} \gamma_n^\beta \left(\left| \frac{S_n}{n} + m(\beta) \right| \leq \varepsilon \right) = \frac{1}{2}.$$

[12] 相転移現象とは，考えている系の構造がある値を境に劇的に変化する現象のことである．

証明. まずハミルトニアン $\mathcal{H}_n(\omega)$ は S_n/n の関数であることに注意する．実際

$$\begin{aligned}\mathcal{H}_n(\omega) &= -\frac{1}{2n}\sum_{1\le i,j\le n}\omega_i\omega_j \\ &= -\frac{n}{2}\left(\frac{1}{n}\sum_{1\le i\le n}\omega_i\right)\left(\frac{1}{n}\sum_{1\le j\le n}\omega_j\right) = -\frac{n}{2}\left(\frac{S_n}{n}\right)^2\end{aligned}$$

となっている．ここで P_n の下での S_n/n の分布を μ_n, γ_n^β の下での S_n/n の分布を ν_n^β と書くことにする．そのとき，$A\in\mathcal{B}(\mathbb{R})$ に対して

$$\begin{aligned}\nu_n^\beta(A) = \gamma_n^\beta\left(\frac{S_n}{n}\in A\right) &= \frac{1}{Z_n^\beta}\int_{\Omega_n}\mathbf{1}_A\left(\frac{S_n}{n}\right)e^{-\beta\mathcal{H}_n}P_n(d\omega) \\ &= \frac{1}{Z_n^\beta}\int_{\Omega_n}\mathbf{1}_A\left(\frac{S_n}{n}\right)\exp\left[\frac{n\beta}{2}\left(\frac{S_n}{n}\right)^2\right]P_n(d\omega) \\ &= \frac{1}{Z_n^\beta}\int_{\mathbb{R}}\mathbf{1}_A(z)\exp\left[\frac{n\beta z^2}{2}\right]\mu_n(dz)\end{aligned}$$

となる．また $\nu_n^\beta(\mathbb{R})=1$ であるから

$$Z_n^\beta = \int_{\mathbb{R}}\exp\left[\frac{n\beta z^2}{2}\right]\mu_n(dz)$$

を得る．よって

$$\nu_n^\beta(A) = \frac{\int_{\mathbb{R}}\mathbf{1}_A(z)\exp\left[\frac{n\beta z^2}{2}\right]\mu_n(dz)}{\int_{\mathbb{R}}\exp\left[\frac{n\beta z^2}{2}\right]\mu_n(dz)} \tag{2.6.1}$$

を得る．

ν_n^β は $\mathbb{R}$ 上の確率測度であるが，μ_n の台が $[-1,1]$ に含まれているので $[-1,1]$ 上の確率測度と見なせることに注意する．$[-1,1]$ 上の関数 I と J_β をそれぞれ次で定義する．

$$\begin{aligned}I(z) &:= \frac{1}{2}(1-z)\log(1-z)+\frac{1}{2}(1+z)\log(1+z),\quad z\in[-1,1], \\ J_\beta(z) &:= I(z)-\frac{\beta z^2}{2}-\inf_{w\in[-1,1]}\left\{I(w)-\frac{\beta w^2}{2}\right\},\quad z\in[-1,1].\end{aligned}$$

このとき，$[-1,1]$ 値確率測度の列 $\{\nu_n^\beta\}_{n\in\mathbb{N}}$ は $n\to\infty$ のときに良い速度関数 J_β に対して速度 n で大偏差原理を満たすことを示す．まず (2.6.1) の分母が収束することを見る．Cramér の大偏差原理と命題 1.2.12 により，$[-1,1]$ 値確率測度の列 $\{\mu_n\}_{n\in\mathbb{N}}$ は $n\to\infty$ のときに良い速度関数 I に対して速度 n で大偏差原理を満たす（例 2.1.4 を参照せよ）．$[-1,1]\ni w\mapsto\beta w^2/2$ は明らかに有界連続なので，定理 1.4.2 により

$$\lim_{n\to\infty}\frac{1}{n}\log\int_{\mathbb{R}}\exp\left[\frac{n\beta z^2}{2}\right]\mu_n(dz) = \sup_{w\in[-1,1]}\left\{\frac{\beta w^2}{2}-I(w)\right\}$$

が成立する．次に (2.6.1) の分子を考える．ν_n^β に対する上からの評価を考えるために，任意に閉集合 $F \subset [-1,1]$ を取る．f を F 上で 0，F^c 上で $-\infty$ となる関数とすれば f は上半連続であり，明らかに

$$\int_{[-1,1]} \mathbf{1}_F(z) \exp\left[\frac{n\beta z^2}{2}\right] \mu_n(dz) = \int_{[-1,1]} \exp\left[nf(z) + \frac{n\beta z^2}{2}\right] \mu_n(dz)$$

である．f は下に有界ではないが，上に有界なので定理 1.4.2 が適用できる．よって定理 1.4.2 の上からの評価により

$$\begin{aligned}\varlimsup_{n\to\infty} \frac{1}{n} \log \int_{[-1,1]} \mathbf{1}_F(z) \exp\left[\frac{n\beta z^2}{2}\right] \mu_n(dz) &\le \sup_{z\in[-1,1]} \left\{ f(z) + \frac{\beta z^2}{2} - I(z) \right\} \\ &= -\inf_{z\in F} \left\{ I(z) - \frac{\beta z^2}{2} \right\}\end{aligned}$$

を得る．

一方，ν_n^β に対する下からの評価を考えるために，相対位相に関する開集合 $G \subset [-1,1]$ を任意に取る．以下本証明では $B(\cdot,\cdot)$ は $[-1,1]$ における開球とする．任意の $B(w,\delta) \subset G$ となる $w \in [-1,1]$ と $\delta > 0$ に対して，μ_n に対する大偏差原理の下からの評価により

$$\begin{aligned}\varliminf_{n\to\infty} \frac{1}{n} \log \int_{[-1,1]} \mathbf{1}_G(z) \exp\left[\frac{n\beta z^2}{2}\right] \mu_n(dz) &\ge \inf_{z\in B(w,\delta)} \frac{\beta z^2}{2} + \varliminf_{n\to\infty} \frac{1}{n} \log \mu_n(B(w,\delta)) \\ &\ge \inf_{z\in B(w,\delta)} \frac{\beta z^2}{2} - \inf_{z\in B(w,\delta)} I(z)\end{aligned}$$

を得る．$\delta \searrow 0$ とした後に $w \in G$ についての上限を取れば

$$\varliminf_{n\to\infty} \frac{1}{n} \log \int_{[-1,1]} \mathbf{1}_G(z) \exp\left[\frac{n\beta z^2}{2}\right] \mu_n(dz) \ge -\inf_{w\in G} \left\{ I(w) - \frac{\beta w^2}{2} \right\}$$

を得る．以上により $\{\nu_n^\beta\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 J_β に対して速度 n で大偏差原理を満たすことがわかった．

J_β を微分すれば

$$J'_\beta(z) = \frac{1}{2} \log \frac{1+z}{1-z} - \beta z, \quad J''_\beta(z) = \frac{1}{1-z^2} - \beta$$

となる．よって

$$\lim_{z\to-1+0} J'_\beta(z) = -\infty, \quad \lim_{z\to1-0} J'_\beta(z) = \infty, \quad J''_\beta(0) = 1-\beta$$

となるので，J_β のグラフの概形は $\beta = 1$ を境として変わることがわかる．実際，$\beta \le 1$ のとき J_β は唯一の零点 0 をもち，$\beta > 1$ のときある $m(\beta) \in (0,1)$ が存在して，2 点 $\pm m(\beta)$ においてのみ J_β は 0 になる．よって $\beta \le 1$ のとき J_β の連続性から任意の $\varepsilon > 0$ に対して

$$\alpha_\varepsilon := \inf_{|z|\ge\varepsilon} J_\beta(z) > 0$$

が成立する．そのため ν_n^β に対する大偏差原理により，任意の $\varepsilon > 0$ に対して確率

$$\gamma_n^\beta\left(\left|\frac{S_n}{n}\right| > \varepsilon\right)$$

は $n \to \infty$ のとき指数的に 0 に収束する．よって

$$\lim_{n\to\infty}\gamma_n^\beta\left(\left|\frac{S_n}{n}\right| \le \varepsilon\right) = 1$$

を得る．また $\beta > 1$ のときも同じ議論により任意の $\varepsilon > 0$ に対して

$$\lim_{n\to\infty}\gamma_n^\beta\left(\frac{S_n}{n} \in B(m(\beta),\varepsilon)\cup B(-m(\beta),\varepsilon)\right) = 1$$

が成立する．一方，任意の $A \in \mathcal{B}([-1,1])$ に対して $\nu_n^\beta(A) = \nu_n^\beta(-A)$ が成立するので，$0 < \varepsilon < m(\beta)$ となる任意の $\varepsilon > 0$ に対して

$$\lim_{n\to\infty}\gamma_n^\beta\left(\left|\frac{S_n}{n} - m(\beta)\right| \le \varepsilon\right) = \lim_{n\to\infty}\gamma_n^\beta\left(\left|\frac{S_n}{n} + m(\beta)\right| \le \varepsilon\right) = \frac{1}{2}.$$

が成立する．これで定理の証明が終了する．　$\square$

第3章　Gärtner-Ellisの大偏差原理とその応用

本章ではGärtner-Ellisの大偏差原理について述べる．簡単に述べると,「$\mathbb{R}^d$ 上のBorel確率測度の列 $\{\mu_n\}_{n\in\mathbb{N}}$ が対数積率母関数について適切な条件を満たせば，$\{\mu_n\}_{n\in\mathbb{N}}$ は大偏差原理を満たす」という主張である．その条件は対数積率母関数（の極限）の存在と正則性についての仮定であり，Cramérの大偏差原理のように単純な設定であれば簡単に確認することができる．そのため注意3.0.2で見る様に条件 **(FEM)** の下でのCramérの大偏差原理は，Gärtner-Ellisの大偏差原理により直ちに従う．その条件（特に正則性についての条件）を確認することは難しい問題であるものの，Gärtner-Ellisの大偏差原理は多くの問題に適用できる一般性の高い結果として知られている．

次元 $d\in\mathbb{N}$ を任意とする．$\{\mu_n\}_{n\in\mathbb{N}}$ を $\mathbb{R}^d$ 上のBorel確率測度の列とし，各 $n\in\mathbb{N}$ に対して対数積率母関数 $\Lambda_n:\mathbb{R}^d\to(-\infty,\infty]$ を次で定義する．

$$\Lambda_n(\lambda):=\log\int_{\mathbb{R}^d}e^{\langle\lambda,x\rangle}\mu_n(dx),\quad \lambda\in\mathbb{R}^d.$$

本節では次の条件 **(GE)** を常に仮定する．

(GE) 任意の $\lambda\in\mathbb{R}^d$ に対して次の極限 $\Lambda(\lambda)$ が $[-\infty,\infty]$ において存在する．

$$\Lambda(\lambda):=\lim_{n\to\infty}\frac{1}{n}\Lambda_n(n\lambda).$$

さらに，0は $\operatorname{dom}\Lambda:=\{\lambda\in\mathbb{R}^d\mid\Lambda(\lambda)\in\mathbb{R}\}$ の内点である．

ここで $\Lambda^*:\mathbb{R}^d\to[0,\infty]$ を Λ のFenchel-Legendre変換とする．

$$\Lambda^*(x):=\sup_{\lambda\in\mathbb{R}^d:\Lambda(\lambda)<\infty}\{\langle\lambda,x\rangle-\Lambda(\lambda)\},\quad x\in\mathbb{R}^d.$$

Λ^* が非負であることは $\Lambda(0)=0$ であることから従う．また補題1.2.5により Λ^* は下半連続である．一般の $\Gamma:\mathbb{R}^d\to[-\infty,\infty]$ についても，Fenchel-Legendre変換を Γ^* と書くことにして，$\operatorname{dom}\Gamma:=\{x\in\mathbb{R}^d\mid\Gamma(x)\in\mathbb{R}\}$ と定義しておく．Γ^* は凸関数の上限として表せているので凸である．

次に Gärtner-Ellis の定理に現れる $\mathbb{R}^d$ 上の関数の性質を定義する．$\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ が $y \in \mathbb{R}^d$ で狭義劣微分可能とは次を満たすことをいう．ある $\lambda \in \mathbb{R}^d$ が存在して，任意の $x \neq y$ に対して

$$\Gamma(x) > \langle \lambda, x - y \rangle + \Gamma(y).$$

が成立する．この不等式を満たす $\lambda \in \mathbb{R}^d$ 全体がなす集合は Γ の y における狭義劣微分と呼ばれる．狭義劣微分は 1 点集合とは限らないことを注意する．実際，$\Gamma(x) = |x|, x \in \mathbb{R}$ の 0 における狭義劣微分は $(-1, 1)$ である．また $\mathcal{F}_{\Gamma^*}$ を Γ^* の狭義劣微分可能な点 $y \in \mathbb{R}^d$ であって，y における狭義劣微分と $(\mathrm{dom}\,\Gamma)^\circ$ の共通部分が空でないもの全体がなす集合とする．

$\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ は次の 3 条件を満たすときに本質的に滑らかであるといわれる．

- $(\mathrm{dom}\,\Gamma)^\circ \subset \mathbb{R}^d$ は空でない．
- Γ は $(\mathrm{dom}\,\Gamma)^\circ$ 上で微分可能である．
- $\lambda_n \to \lambda \in \partial(\mathrm{dom}\,\Gamma)$ なる任意の数列 $\{\lambda_n\}_{n\in\mathbb{N}} \subset (\mathrm{dom}\,\Gamma)^\circ$ に対して次が成立する．

$$\lim_{n\to\infty} |\nabla\Gamma(\lambda_n)| = \infty.$$

次の定理が本章の主定理であり，Gärtner-Ellis の大偏差原理と呼ばれる．

定理 3.0.1. 条件 **(GE)** を仮定する．このとき次が成立する．

(1) 任意の閉集合 $C \subset \mathbb{R}^d$ に対して次が成立する．

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mu_n(C) \leq -\inf_{x\in C} \Lambda^*(x).$$

(2) 任意の開集合 $O \subset \mathbb{R}^d$ に対して次が成立する．

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mu_n(O) \geq -\inf_{x\in O\cap\mathcal{F}_{\Lambda^*}} \Lambda^*(x).$$

(3) さらに Λ は下半連続かつ本質的に滑らかだと仮定する．このとき任意の開集合 $O \subset \mathbb{R}^d$ に対して次が成立する．

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mu_n(O) \geq -\inf_{x\in O} \Lambda^*(x).$$

すなわち，$\mathbb{R}^d$ 上の確率測度の列 $\{\mu_n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 Λ^* に対して速度 n で大偏差原理を満たす．

注意 3.0.2. Cramér の大偏差原理（定理 2.0.1）は定理 3.0.1 の系として得られる．実際，μ_n を S_n/n の分布と定義すれば，本節で定義した Λ は第 2 節で定義したものに他ならない．また仮定 **(FEM)** の下では条件 **(GE)** は明らかに成立し，さらに命題 2.2.1(1) により Λ は C^1 級なので定理 3.0.1(3) の仮定も明らかに成立する．よって定理 3.0.1 は定理 2.0.1 の一般化になっている．

注意 3.0.3. 本節では記法の単純化のため大偏差原理の速度が n である場合を取り扱ったが，∞ に発散する数列 $\{a_n\}_{n\in\mathbb{N}}$ を速度とした場合も同様の定理が成立する．

注意 3.0.4. Gärtner-Ellis の定理は一般的な位相ベクトル空間上で成立することが知られている．証明は与えないが主張のみ述べておく．X を局所凸 Hausdorff 位相ベクトル空間とし，X^* をその位相的双対とする．つまり X^* は X 上の連続線型汎函数全体である．$\{\mu_\varepsilon\}_{\varepsilon>0}$ を X 上の指数的に緊密な確率測度の族とし，$\Lambda_\varepsilon : X^* \to (-\infty, \infty]$ を μ_ε から決まる対数積率母関数とする．

$$\Lambda_\varepsilon(\lambda) := \log \int_X e^{\lambda(x)} \mu_\varepsilon(dx), \quad \lambda \in X^*.$$

ここで任意の $\lambda \in X^*$ に対して極限 $\lim_{\varepsilon\searrow 0} \varepsilon\Lambda_\varepsilon(\lambda/\varepsilon) =: \Lambda(\lambda)$ が存在して，X^* 上の関数 Λ は有限値を取り下半連続かつ Gateaux 微分可能であると仮定する.[1] このとき，X 上の確率測度の族 $\{\mu_\varepsilon\}_{\varepsilon>0}$ は $\varepsilon \searrow 0$ のときに良い速度関数 Λ^* に対して速度 ε^{-1} で大偏差原理を満たす.[2] ここで Λ^* は Λ の Fenchel-Legendre 変換である．

$$\Lambda^*(x) := \sup_{\lambda\in X^*} \{\lambda(x) - \Lambda(\lambda)\}, \quad x \in X.$$

本節の残りでは条件 **(GE)** の下で Λ^* が良い速度関数であることと，$\{\mu_n\}_{n\in\mathbb{N}}$ が指数的に緊密であることを示す．その次に第 3.1 節において Fenchel-Legendre 変換や劣微分に関するいくつかの補題を示した後，第 3.2 節において定理 3.0.1 を示す．第 3.3 節と第 3.4 節では Gärtner-Ellis の定理を用いて独立同分布な確率変数の列に対する中偏差原理と有限状態 Markov 連鎖に対する大偏差原理をそれぞれ調べる．

まず Λ^* が良い速度関数であることを示す．次の命題や本章を通して，$\mathbb{R}^d$ 上の凸関数に関する基本事項が用いられる．いくつかの事項は付録第 A 章にまとめられているので，適宜参照せよ．

命題 3.0.5. 条件 **(GE)** を仮定する．このとき次が成立する．

(1) Λ は凸であり任意の $\lambda \in \mathbb{R}^d$ に対して $\Lambda(\lambda) > -\infty$ である．

---

[1] $\lambda : X \to \mathbb{R}$ が Gateaux 微分可能であるとは，任意の $x, y \in X$ に対して $\lambda(x + ty)$ が $t = 0$ で微分可能であることをいう．

[2] 詳しくは [33, Corollary 4.5.16] を参照せよ．

(2) Λ^* は良い速度関数である．

証明. (1) 明らかに $\Lambda_n(0)=0$ なので，$\Lambda(0)=0$ である．また任意の $\lambda_1,\lambda_2\in\mathbb{R}^d, t\in(0,1)$ に対して，Hölder の不等式により

$$\int_{\mathbb{R}^d} e^{\langle t\lambda_1+(1-t)\lambda_2,x\rangle}\mu_n(dx)\le\left(\int_{\mathbb{R}^d}e^{\langle\lambda_1,x\rangle}\mu_n(dx)\right)^t\left(\int_{\mathbb{R}^d}e^{\langle\lambda_2,x\rangle}\mu_n(dx)\right)^{1-t}$$

が成立する．この式で対数を取ることにより，Λ_n が凸であることがわかる．よって Λ は凸関数の各点極限になっているので，Λ も凸である．ここである $\lambda\in\mathbb{R}^d$ に対して $\Lambda(\lambda)=-\infty$ となったとすると，任意の $0<\alpha\le 1$ に対して

$$\Lambda(\alpha\lambda)\le\alpha\Lambda(\lambda)+(1-\alpha)\Lambda(0)=\alpha\Lambda(\lambda)=-\infty$$

が成立するが，これは **(GE)** に矛盾する．よって $\Lambda(\lambda)>-\infty$ である．

(2) 各 $\alpha\ge 0$ を固定して $\{x\in\mathbb{R}^d\mid\Lambda^*(x)\le\alpha\}$ がコンパクト集合であることを示す．**(GE)** により $\delta>0$ を小さくとれば $\overline{B}(0,\delta)\subset(\mathrm{dom}\,\Lambda)^\circ$ となるようにできる．Λ は凸なので系 A.2.4 により $(\mathrm{dom}\,\Lambda)^\circ$ 上で連続である．よって $c:=\sup_{\lambda\in\overline{B}(0,\delta)}\Lambda(\lambda)<\infty$ である．ここで $x\in\mathbb{R}^d$ が $\Lambda^*(x)\le\alpha$ とすると

$$\alpha\ge\Lambda^*(x)\ge\sup_{\lambda\in\overline{B}(0,\delta)}\{\langle\lambda,x\rangle-\Lambda(\lambda)\}\ge\sup_{\lambda\in\overline{B}(0,\delta)}\langle\lambda,x\rangle-c=\delta|x|-c$$

となっている．よって $\{x\in\mathbb{R}^d\mid\Lambda^*(x)\le\alpha\}$ は有界である．また Λ^* は先に述べたように下半連続なので，$\{x\in\mathbb{R}^d\mid\Lambda^*(x)\le\alpha\}$ は閉集合である．よって Λ^* は良い速度関数である． □

次に $\{\mu_n\}_{n\in\mathbb{N}}$ が指数的に緊密であることを示す．

補題 3.0.6. 条件 **(GE)** を仮定する．このとき $\{\mu_n\}_{n\in\mathbb{N}}$ は指数的に緊密である．

証明. **(GE)** により $a>0$ を小さく取ると $\Lambda(\pm ae_i)<\infty,(i=1,\dots,d)$ となるようにできる．任意に $K>0$ を固定する．$\mathbb{R}^{i-1}\times(K,\infty)\times\mathbb{R}^{d-i}$ 上では $\langle e_i,x\rangle-K\ge 0$ なので

$$\mu_n\left(\mathbb{R}^{i-1}\times(K,\infty)\times\mathbb{R}^{d-i}\right)\le\int_{\mathbb{R}^d}e^{n(\langle ae_i,x\rangle-aK)}\mu_n(dx)=\exp\left(-naK+\Lambda_n(nae_i)\right)$$

である．よって

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mu_n\left(\mathbb{R}^{i-1}\times(K,\infty)\times\mathbb{R}^{d-i}\right)\le -aK+\Lambda(ae_i)$$

が成立する．同様にして

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mu_n\left(\mathbb{R}^{i-1}\times(-\infty,-K)\times\mathbb{R}^{d-i}\right)\le -aK+\Lambda(-ae_i)$$

が示される．よって補題 1.2.2 により

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mu_n\left(\mathbb{R}^d\setminus[-K,K]^d\right)\le -aK+\max_{i=1,\dots,d}\{\Lambda(ae_i),\Lambda(-ae_i)\}$$

がわかるので，$K\to\infty$ とすれば $\{\mu_n\}_{n\in\mathbb{N}}$ は指数的に緊密であることが示される． □

3.1 Fenchel-Legendre 変換の性質

本節では Fenchel-Legendre 変換に関する性質をいくつか述べる．それらは Gärtner-Ellis の定理 (3) の証明において基本的な役割を果たす．本節の内容は凸解析に関するものであり，確率測度とは無関係であることを注意しておく.[3]

命題 3.1.1. $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ を凸関数とする．$\eta \in (\mathrm{dom}\,\Gamma)^\circ$ を固定して $y = \nabla\Gamma(\eta)$ とおく．

(1) $\Gamma^*(y) = \langle \eta, y\rangle - \Gamma(\eta)$ が成立する．

(2) Γ^* は y において狭義劣微分可能であり，η は y における狭義劣微分に属す．特に，$y \in \mathcal{F}_{\Gamma^*}$ である．

証明. (1) $\lambda \in \mathbb{R}^d$ を固定する．$\alpha \in [0,1]$ に対して $g(\alpha) \in [-\infty, \infty)$ を次で定める．

$$g(\alpha) := (1-\alpha)\langle \eta, y\rangle + \alpha\langle \lambda, y\rangle - \Gamma((1-\alpha)\eta + \alpha\lambda).$$

$\Gamma(\eta) < \infty$ なので $g(0)$ は有限である．また Γ は凸なので g は凹である．よって

$$g(1) - g(0) \leq \lim_{\alpha\searrow 0} \frac{g(\alpha) - g(0)}{\alpha} = \langle \lambda - \eta, y - \nabla\Gamma(\eta)\rangle = 0$$

である．よって任意の $\lambda \in \mathbb{R}^d$ に対して

$$\langle \lambda, y\rangle - \Gamma(\lambda) \leq \langle \eta, y\rangle - \Gamma(\eta) \leq \Gamma^*(y)$$

が成立する．最左辺の $\lambda \in \mathbb{R}^d$ を渡る上限を取ることにより (1) が得られる．

(2) Γ^* は y において狭義劣微分可能であることを示そう．そのために，ある $x \neq y$ に対して

$$\langle \eta, y\rangle - \Gamma^*(y) \leq \langle \eta, x\rangle - \Gamma^*(x)$$

となったと仮定する．なお $\Gamma^*(y) \in \mathbb{R}$ であるから $\Gamma^*(x) \in \mathbb{R}$ である．(1) より左辺は $\Gamma(\eta)$ である．一方任意の $\theta \in \mathbb{R}^d$ に対して

$$\Gamma^*(x) \geq \langle \eta + \theta, x\rangle - \Gamma(\eta + \theta)$$

であることを組み合わせると，

$$\langle \theta, x\rangle \leq \Gamma(\eta + \theta) - \Gamma(\eta)$$

[3] 本節の記述は [33] を参考にした．

が得られる.（ここでの式変形においても $\infty - \infty$ という状況は発生していない.）

θ の代わりに $\varepsilon\theta$ を考え，$\varepsilon \to 0$ とすれば

$$\langle \theta, x \rangle \leq \lim_{\varepsilon \to 0} \frac{\Gamma(\eta + \varepsilon\theta) - \Gamma(\eta)}{\varepsilon} = \langle \theta, \nabla\Gamma(\eta) \rangle$$

が得られる．特に $\theta = x - \nabla\Gamma(\eta)$ とすれば $x = \nabla\Gamma(\eta)$ が得られる．これは $x \neq y$ と矛盾する．

以上により $x \neq y = \nabla\Gamma(\eta)$ のとき

$$\langle \eta, y \rangle - \Gamma^*(y) > \langle \eta, x \rangle - \Gamma^*(x)$$

が成立する．よって Γ^* は y において狭義劣微分可能であり，η は y における狭義劣微分に属す. □

本節の残りでは補題 3.1.5 を示すことが主目的である．そのために技術的な補題が 3 つ必要になるので，それらを順に示す.

命題 3.1.2. $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ は下半連続かつ凸であるとする．このとき次が成立する.

$$\Gamma(x) = \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, x \rangle - \Gamma^*(\lambda)\}, \quad x \in \mathbb{R}^d. \tag{3.1.1}$$

すなわち，$\Gamma^{**} = \Gamma$ が成立する.

証明. Γ が恒等的に ∞ ならば Γ^* は恒等的に $-\infty$ になるので，(3.1.1) は成立している．以下 Γ が恒等的に ∞ ではない場合に (3.1.1) を示す．$\mathbb{R}^{d+1}$ の部分集合 $\mathcal{E}$ と $\mathcal{E}^*$ をそれぞれ次で定義する.

$$\mathcal{E} := \{(x, \alpha) \in \mathbb{R}^d \times \mathbb{R} \mid \Gamma(x) \leq \alpha\}, \quad \mathcal{E}^* := \{(\lambda, \beta) \in \mathbb{R}^d \times \mathbb{R} \mid \Gamma^*(\lambda) \leq \beta\}.$$

Γ が恒等的に ∞ ではないので，$\mathcal{E} \neq \emptyset$ である.

本段落では $\mathcal{E}^* \neq \emptyset$ であることを示す．Γ は恒等的に ∞ ではないので，ある $x_0 \in \mathbb{R}^d$ に対して $\Gamma(x_0) \in \mathbb{R}$ となる．$\alpha_0 < \Gamma(x_0)$ となる α_0 を固定する．仮定により Γ は下半連続かつ凸であるので，$\mathcal{E}$ は $\mathbb{R}^{d+1}$ において閉かつ凸であることがわかる．よって Hahn-Banach の定理により，$\mathbb{R}^{d+1}$ の超平面で $\mathcal{E}$ と $(x_0, \alpha_0) \notin \mathcal{E}$ を狭義に分離するものが存在する．つまり，ある $(\mu_0, -\rho_0) \in \mathbb{R}^d \times \mathbb{R}$ と $\gamma_0 \in \mathbb{R}$ が存在して次が成立する.

$$\sup_{(y,\xi) \in \mathcal{E}} \{\langle \mu_0, y \rangle - \rho_0 \xi\} \leq \gamma_0 < \langle \mu_0, x_0 \rangle - \rho_0 \alpha_0. \tag{3.1.2}$$

(3.1.2) において $(y, \xi) = (x_0, \Gamma(x_0)) \in \mathcal{E}$ とすることにより，$-\rho_0\Gamma(x_0) < -\rho_0\alpha_0$ を得る．よって $\rho_0 > 0$ である．任意に $y \in \mathbb{R}^d$ を取る．$\Gamma(y) = \infty$ のときは明らかに

$$\langle \mu_0/\rho_0, y \rangle - \Gamma(y) \leq \gamma_0/\rho_0 \tag{3.1.3}$$

が成立する．一方 $\Gamma(y) < \infty$ のときも，(3.1.2) の左側の不等式により (3.1.3) は成立している．よって任意の $y \in \mathbb{R}^d$ に対して (3.1.3) が成立する．(3.1.3) において $y \in \mathbb{R}^d$ についての上限を取れば，$\Gamma^*(\mu_0/\rho_0) \le \gamma_0/\rho_0$ となるので，$(\mu_0/\rho_0, \gamma_0/\rho_0) \in \mathcal{E}^*$ である．よって $\mathcal{E}^* \neq \emptyset$ が示された．

任意に $x \in \mathbb{R}^d$ を固定する．ここで $(\lambda, \beta) \in \mathcal{E}^*$ とすると，$\Gamma^*(\lambda) \ge \langle \lambda, x\rangle - \Gamma(x)$ と $\Gamma^*(\lambda) \le \beta$ により $\Gamma(x) \ge \langle \lambda, x\rangle - \beta$ を得る．この式で $(\lambda, \beta) \in \mathcal{E}^*$ についての上限を取れば

$$\Gamma(x) \ge \sup_{(\lambda,\beta)\in\mathcal{E}^*}\{\langle \lambda, x\rangle - \beta\} = \sup_{\lambda\in\mathbb{R}^d}\{\langle \lambda, x\rangle - \Gamma^*(\lambda)\}$$

が得られる．残るは逆向きの不等式の証明である．それを示すためには，任意の $\alpha < \Gamma(x)$ に対してある $(\lambda, \beta) \in \mathcal{E}^*$ が存在して

$$\langle \lambda, x\rangle - \beta > \alpha \tag{3.1.4}$$

となることを示せば十分である．

$x \in \mathbb{R}^d$ と $\alpha < \Gamma(x)$ となる $\alpha \in \mathbb{R}$ を任意に固定する．このとき $(x, \alpha) \notin \mathcal{E}$ であるから再び Hahn-Banach の定理により，ある $(\mu, -\rho) \in \mathbb{R}^d \times \mathbb{R}$ と $\gamma \in \mathbb{R}$ が存在して次が成立する．

$$\sup_{(y,\xi)\in\mathcal{E}}\{\langle \mu, y\rangle - \rho\xi\} \le \gamma < \langle \mu, x\rangle - \rho\alpha. \tag{3.1.5}$$

Γ は恒等的に ∞ ではないので，ある $x_0 \in \mathbb{R}^d$ に対して $\Gamma(x_0) \in \mathbb{R}$ である．よって任意の $\xi \ge \Gamma(x_0)$ に対して $\langle \mu, x_0\rangle - \rho\xi \le \gamma$ が成立する．$\xi \to \infty$ とすることにより $\rho \ge 0$ がわかる．

まず $\rho > 0$ のときを考える．先程と同じ議論で $(\mu/\rho, \gamma/\rho) \in \mathcal{E}^*$ である．また (3.1.5) の右側の不等式により，$(\lambda, \beta) = (\mu/\rho, \gamma/\rho)$ は (3.1.4) を満たしている．次に $\rho = 0$ のときを考える．このとき (3.1.5) により

$$\sup_{y:\Gamma(y)<\infty}\{\langle \mu, y\rangle - \gamma\} \le 0 < \langle \mu, x\rangle - \gamma \tag{3.1.6}$$

が成立している．任意に $(\lambda_0, \beta_0) \in \mathcal{E}^*$ を取る．このとき $\delta > 0$ に対して $\lambda_\delta := \lambda_0 + \mu/\delta$ および $\beta_\delta := \beta_0 + \gamma/\delta$ と定義すると，$\Gamma(y) < \infty$ となる任意の $y \in \mathbb{R}^d$ に対して

$$\begin{aligned}\langle \lambda_\delta, y\rangle - \beta_\delta &= \frac{1}{\delta}(\langle \mu, y\rangle - \gamma) + \langle \lambda_0, y\rangle - \beta_0 \\ &\le \langle \lambda_0, y\rangle - \Gamma^*(\lambda_0) \le \Gamma(y)\end{aligned}$$

が成立する．ただし最初の不等号では (3.1.6) の左側の不等式と $(\lambda_0, \beta_0) \in \mathcal{E}^*$ であることを用いた．また $\Gamma(y) = \infty$ となるときも同式は明らかに成立している．これから

$\Gamma^*(\lambda_\delta) \le \beta_\delta$ が得られるので，$(\lambda_\delta, \beta_\delta) \in \mathcal{E}^*$ がわかった．最後に (3.1.6) の右側の不等式により，$\delta \searrow 0$ のとき

$$\langle \lambda_\delta, x \rangle - \beta_\delta = \frac{1}{\delta}(\langle \mu, x \rangle - \gamma) + \langle \lambda_0, x \rangle - \beta_0 \nearrow \infty$$

となっているので，$\delta > 0$ が十分小さければ $(\lambda, \beta) = (\lambda_\delta, \beta_\delta)$ は (3.1.4) を満たしている．これで本命題の証明が終わった． □

次の補題では凸集合に対する相対内部が現れるので，補題に進む前に相対内部の定義を導入する．空でない凸集合 $C \subset \mathbb{R}^d$ に対して C の相対内部 $\mathrm{ri}\,(C)$ を次で定義する．

$$\mathrm{ri}\,(C) := \{y \in C \mid \text{任意の } x \in C \text{ に対してある } \varepsilon > 0 \text{ が存在して } y - \varepsilon(x - y) \in C\}.$$

相対内部の定義から $C^\circ \subset \mathrm{ri}\,(C)$ だが，一般に逆向きの包含関係は成立しない．凸関数 $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ に対して，$\mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ は $\mathrm{dom}\,\Gamma$ の凸集合としての内部を表し，次元が退化する場合もあり得る．例として $\Gamma : \mathbb{R}^2 \to (-\infty, \infty]$ を

$$\Gamma(x) := \begin{cases} 0, & x \text{ の第 2 成分が } 0 \text{ のとき}, \\ \infty, & \text{それ以外}, \end{cases}$$

として定義すると，$\mathrm{dom}\,\Gamma = \mathrm{ri}\,(\mathrm{dom}\,\Gamma) = \mathbb{R} \times \{0\}$ であるが $(\mathrm{dom}\,\Gamma)^\circ = \emptyset$ である．

補題 3.1.3. $\Gamma : \mathbb{R}^d \to [0, \infty]$ は下半連続な凸関数で，$\inf_{\lambda \in \mathbb{R}^d} \Gamma(\lambda) = 0$ かつ $0 \in \mathrm{ri}\,(\mathrm{dom}\,\Gamma^*)$ を満たすとする．このとき，ある $\eta \in \mathbb{R}^d$ に対して $\Gamma(\eta) = 0$ となる．

証明. まず $\Gamma^*(0) = -\inf_{\lambda \in \mathbb{R}^d} \Gamma(\lambda) = 0$ であることを注意しておく．また $0 < \delta_1 < \delta_2$ および $y \in \mathbb{R}^d$ とすると，Γ^* の凸性により $\Gamma^*(\delta_1 y) \le (\delta_1/\delta_2)\Gamma^*(\delta_2 y)$ がわかる．よって $\{\Gamma^*(\delta y)/\delta\}_{\delta > 0}$ は広義単調増加である．ここで関数 $g : \mathbb{R}^d \to [-\infty, \infty]$ を次で定める．

$$g(y) := \inf_{\delta > 0} \frac{\Gamma^*(\delta y)}{\delta} = \lim_{\delta \searrow 0} \frac{\Gamma^*(\delta y)}{\delta}.$$

本補題を示すために，ある $\eta \in \mathbb{R}^d$ が存在して

$$g(y) \ge \langle \eta, y \rangle, \quad y \in \mathbb{R}^d \tag{3.1.7}$$

が成立することを後で示す．まずはこれを認めて本補題の証明を終了させる．(3.1.7) と g の定義により，任意の $y \in \mathbb{R}^d$ に対して $\Gamma^*(y) \ge \langle \eta, y \rangle$ が成立する．ここで $\Gamma : \mathbb{R}^d \to [0, \infty]$ は下半連続かつ凸関数なので，命題 3.1.2 により

$$\Gamma(\eta) = \sup_{y \in \mathbb{R}^d} \{\langle \eta, y \rangle - \Gamma^*(y)\}$$

であるから，$\Gamma(\eta) \leq 0$ である．一方，$\Gamma^*(0) = 0$ なので $\Gamma(\eta) \geq 0$ である．よって本補題の主張である $\Gamma(\eta) = 0$ が示された．

本補題の証明を完結させるために，ある $\eta \in \mathbb{R}^d$ に対して (3.1.7) が成立することを示す．そのために g の性質をいくつか見ておく．g は凸関数の各点極限として表されるので凸関数である．また g の定義により $g(0) = 0$ であり，任意の $\alpha \geq 0$ に対して $g(\alpha y) = \alpha g(y)$ である．次に任意の $y \in \mathbb{R}^d$ に対して $g(y) > -\infty$ であることを示そう．まず $\{ty \mid t > 0\}$ と $\mathrm{dom}\,\Gamma^*$ が共通部分をもたないときは，任意の $\delta > 0$ に対して $\Gamma^*(\delta y) = \infty$ となるので，明らかに $g(y) = \infty$ である．次に $ty \in \mathrm{dom}\,\Gamma^*$ となる $t > 0$ が存在する場合を考える．このとき仮定 $0 \in \mathrm{ri}\,(\mathrm{dom}\,\Gamma^*)$ により，ある $\varepsilon > 0$ に対して $\Gamma^*(-\varepsilon y) < \infty$ となる．従って Γ^* の凸性により

$$0 = \Gamma^*(0) \leq \left(\frac{1}{\delta} + \frac{1}{\varepsilon}\right)^{-1} \left(\frac{\Gamma^*(\delta y)}{\delta} + \frac{\Gamma^*(-\varepsilon y)}{\varepsilon}\right)$$

が任意の $\delta > 0$ に対して成立する．よって $\delta > 0$ について下限を取ることにより，$g(y) \geq -\Gamma^*(-\varepsilon y)/\varepsilon > -\infty$ となる．この場合も $g(y) > -\infty$ が示された．

g に対する性質として最後に $\underline{\lim}_{y\to 0}\, g(y) \geq 0$ を示そう．そのために $0 \in \mathrm{ri}\,(\mathrm{dom}\, g)$ を示す．$x \in \mathrm{dom}\, g$ を任意に取ると，$g(x) < \infty$ なのである $\delta > 0$ に対して $\Gamma^*(\delta x) < \infty$ となる．よって仮定 $0 \in \mathrm{ri}\,(\mathrm{dom}\,\Gamma^*)$ により，ある $\varepsilon > 0$ に対して $\Gamma^*(-\varepsilon\delta x) < \infty$ がわかる．これから $g(-\varepsilon x) < \infty$，つまり $-\varepsilon x \in \mathrm{dom}\, g$ となるので，$0 \in \mathrm{ri}\,(\mathrm{dom}\, g)$ が示された．次に $\mathrm{ri}\,(\mathrm{dom}\, g) = \mathrm{dom}\, g$ を示そう．定義により $\mathrm{ri}\,(\mathrm{dom}\, g) \subset \mathrm{dom}\, g$ なので逆向きの包含関係を示せばよい．$x \in \mathrm{dom}\, g$ を任意に取る．$g(2x) = 2g(x)$ なので $2x \in \mathrm{dom}\, g$ である．先に $0 \in \mathrm{ri}\,(\mathrm{dom}\, g)$ を示したので，命題 A.2.1(2) により $x \in \mathrm{ri}\,(\mathrm{dom}\, g)$ となり，$\mathrm{ri}\,(\mathrm{dom}\, g) = \mathrm{dom}\, g$ が示された．また g は凸関数なので命題 A.2.3 により，g は $\mathrm{ri}\,(\mathrm{dom}\, g) = \mathrm{dom}\, g$ 上連続である．$g(0) = 0$ であるから，$y \in \mathbb{R}^d$ が $\mathrm{dom}\, g$ から 0 に近づくときは $g(y) \to 0$ となる．一方 $(\mathrm{dom}\, g)^c$ から 0 に近づくときは明らかに $g(y) \to \infty$ である．以上により $\underline{\lim}_{y\to 0}\, g(y) \geq 0$ が示された．

ここで $\mathbb{R}^{d+1}$ の部分集合を $\mathcal{E} := \{(y, \xi) \in \mathbb{R}^{d+1} \mid \xi \geq g(y)\}$ と定める．g は凸なので $\mathcal{E}$ も凸であり，$g(0) = 0$ なので $(0, 0) \in \mathcal{E}$ であるから $\mathcal{E}$ は空でない．また $\underline{\lim}_{y\to 0}\, g(y) \geq 0 > -1$ なので $(0, -1) \notin \overline{\mathcal{E}}$ である．よって Hahn-Banach の定理[4]により，$\mathbb{R}^{d+1}$ 内の超平面で $(0, -1)$ と $\overline{\mathcal{E}}$ を狭義に分離するものが存在する．具体的に述べると，ある $\lambda \in \mathbb{R}^d$ と $\rho \in \mathbb{R}$ が存在して任意の $(y, \xi) \in \mathcal{E}$ に対して

$$\rho > \langle \lambda, y \rangle - \rho\xi \tag{3.1.8}$$

が成立する．(3.1.8) において $(y, \xi) = (0, 0) \in \mathcal{E}$ とすることにより $\rho > 0$ がわかる．ここで $\eta = \lambda/\rho$ とおけば，(3.1.7) を満たすことを示す．$g(y) = \infty$ のとき (3.1.7) は明らかに

[4]例えば [33, Theorem B.6] を参照せよ．

成立するので，以下 $g(y) < \infty$ とする．$\alpha > 0$ を任意に取る．$(\alpha y, g(\alpha y)) \in \mathcal{E}$ に対して (3.1.8) を用いると

$$1/\alpha > \langle \eta, y \rangle - g(y)$$

が得られる．$g(y) > -\infty$ であったから，$g(y)$ を左辺に移項して $\alpha \to \infty$ とすることにより，(3.1.7) が得られる．これで本補題の証明が終わる． □

補題 3.1.4. $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ は本質的に滑らかな凸関数であり，$\Gamma(0) = 0$ かつある $x \in \mathbb{R}^d$ に対して $\Gamma^*(x) = 0$ となると仮定する．このとき $0 \in (\mathrm{dom}\,\Gamma)^\circ$ が成立する．

証明. $\Gamma(0) = 0$ と Γ の凸性から，任意の $t \in [0,1]$ と $\lambda \in \mathbb{R}^d$ に対して

$$\Gamma(t\lambda) \le t\Gamma(\lambda) \tag{3.1.9}$$

が成立する．また $\Gamma^*(x) = 0$ なので

$$\Gamma(t\lambda) \ge \langle t\lambda, x \rangle \ge -t|\lambda||x|$$

が成立する．Γ は本質的に滑らかなので $(\mathrm{dom}\,\Gamma)^\circ \neq \emptyset$ であるから，ある $z \in \mathbb{R}^d$ と $r > 0$ に対して Γ は $\overline{B}(z,r) \subset (\mathrm{dom}\,\Gamma)^\circ$ 上微分可能である．特に Γ は $\overline{B}(z,r)$ 上連続なので

$$M := \sup_{\lambda \in \overline{B}(z,r)} \{|\Gamma(\lambda)| + |\lambda||x|\} < \infty$$

である．よって任意の $t \in (0,1]$ に対して $\theta \in \overline{B}(tz, tr) \Leftrightarrow \lambda = (\theta/t) \in \overline{B}(z,r)$ であることと (3.1.9) から

$$\sup_{\theta \in \overline{B}(tz,tr)} |\Gamma(\theta)| \le tM$$

である．

次に任意の $t \in (0,1]$ と $\theta \in \overline{B}(tz,tr)$ に対して

$$|\Gamma(\theta) - \Gamma(tz)| \le \frac{2M}{r}|\theta - tz| \tag{3.1.10}$$

を示そう．$\theta = tz$ のときは明らかなので $\theta \neq tz$ とする．ここで

$$y = tz + \frac{tr}{|\theta - tz|}(\theta - tz) \quad \Longleftrightarrow \quad \theta = \frac{|\theta - tz|}{tr}y + \left(1 - \frac{|\theta - tz|}{tr}\right)tz$$

とおくと，$y \in \overline{B}(tz,tr)$ である．よって Γ の凸性により

$$\Gamma(\theta) - \Gamma(tz) \le \frac{|\theta - tz|}{tr}(\Gamma(y) - \Gamma(tz)) \le \frac{2M}{r}|\theta - tz|$$

が得られる．同様に

$$y' = tz - \frac{tr}{|\theta - tz|}(\theta - tz) \iff tz = \frac{|\theta - tz|}{|\theta - tz| + tr}y' + \frac{tr}{|\theta - tz| + tr}\theta$$

とおくと，$y' \in \overline{B}(tz, tr)$ である．よって Γ の凸性により

$$\Gamma(tz) - \Gamma(\theta) \leq \frac{|\theta - tz|}{|\theta - tz| + tr}(\Gamma(y') - \Gamma(\theta)) \leq \frac{2M}{r}|\theta - tz|$$

が得られるので，(3.1.10) が示される．

再び (3.1.9) により $tz \in (\mathrm{dom}\,\Gamma)^\circ$ であることと Γ が本質的に滑らかであることから，Γ は tz で微分可能であることがわかる．よって (3.1.10) により $|\nabla\Gamma(tz)| \leq 2M/r$ がわかる．$t \searrow 0$ として，再び Γ が本質的に滑らかであることを用いると $0 \in (\mathrm{dom}\,\Gamma)^\circ$ が得られる． □

補題 3.1.5. 下半連続な凸関数 $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ が本質的に滑らかであれば,

$$\mathrm{ri}\,(\mathrm{dom}\,\Gamma^*) \subset \mathcal{F}_{\Gamma^*}.$$

証明. $\mathrm{dom}\,\Gamma^* \neq \emptyset$ としてよい．このとき命題 A.2.1(1) により $\mathrm{ri}\,(\mathrm{dom}\,\Gamma^*) \neq \emptyset$ である．$x \in \mathrm{ri}\,(\mathrm{dom}\,\Gamma^*)$ を固定し，関数 $f : \mathbb{R}^d \to [0, \infty]$ を次で定める．

$$f(\lambda) = \Gamma(\lambda) - \langle \lambda, x \rangle + \Gamma^*(x), \quad \lambda \in \mathbb{R}^d.$$

証明は $\nabla f(\eta) = 0$ となる $\eta \in \mathbb{R}^d$ を見つけることにより，Γ^* が x で劣微分可能であることを示すという方針で進む.

f は下半連続かつ凸で $\inf_{\lambda\in\mathbb{R}^d} f(\lambda) = 0$ である．また f の Fenchel-Legendre 変換は

$$\begin{aligned} f^*(z) &= \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda, z \rangle - f(\lambda)\} \\ &= \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda, z + x \rangle - \Gamma(\lambda)\} - \Gamma^*(x) \\ &= \Gamma^*(z + x) - \Gamma^*(x) \end{aligned}$$

となっている．この式を用いると $x \in \mathrm{ri}\,(\mathrm{dom}\,\Gamma^*)$ なので $0 \in \mathrm{ri}\,(\mathrm{dom}\,f^*)$ であることがわかる．よって補題 3.1.3 によりある $\eta \in \mathbb{R}^d$ に対して $f(\eta) = 0$ となるが，特に $\eta \in \mathrm{dom}\,\Gamma$ である．

$\Gamma(\eta) \in \mathbb{R}$ に注意して，$\widetilde{\Gamma} : \mathbb{R}^d \to (-\infty, \infty]$ を次で定義する．

$$\widetilde{\Gamma}(\xi) = \Gamma(\xi + \eta) - \Gamma(\eta), \quad \xi \in \mathbb{R}^d.$$

Γ に対する仮定から $\widetilde{\Gamma}$ は本質的に滑らかで下半連続かつ凸で，さらに $\widetilde{\Gamma}(0) = 0$ である．また

$$\widetilde{\Gamma}^*(x) = \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda, x \rangle - \Gamma(\lambda + \eta) + \Gamma(\eta)\}$$

$$
\begin{aligned}
&= \Gamma(\eta) - \langle \eta, x \rangle + \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda + \eta, x \rangle - \Gamma(\lambda + \eta)\} \\
&= \Gamma(\eta) - \langle \eta, x \rangle + \Gamma^*(x) = f(\eta) = 0
\end{aligned}
$$

となっている．補題 3.1.4 により $0 \in (\mathrm{dom}\,\widetilde{\Gamma})^\circ$ なので $\eta \in (\mathrm{dom}\,\Gamma)^\circ$ がわかる．また Γ は本質的に滑らかなので Γ は η で微分可能である．さらに $\inf_{\lambda \in \mathbb{R}^d} f(\lambda) = 0 = f(\eta)$ なので $\nabla f(\eta) = 0$，よって $x = \nabla \Gamma(\eta)$ がわかる．最後に命題 3.1.1 により，Γ^* は x において狭義劣微分可能で $\eta \in (\mathrm{dom}\,\Gamma)^\circ$ は x における狭義劣微分に属すので，証明は終了する． □

3.2 Gärtner-Ellis の定理の証明

本節では定理 3.0.1 を証明する．証明の基本的な方針は Cramér の定理と同じであるが，一般的な状況では速度関数に対する性質が分からないために，証明は技術的に複雑になる．その困難を克服するために，前節までに証明した凸解析についての結果を援用する.[5]

定理 3.0.1(1) の証明. 補題 3.0.6 により $\{\mu_n\}_{n\in\mathbb{N}}$ は指数的に緊密なので，補題 1.3.1 によりコンパクト集合 $C \subset \mathbb{R}^d$ に対して上からの評価を示せば十分である．$\delta > 0$ を固定して，

$$
I^\delta(x) := \min\{\Lambda^*(x) - \delta, 1/\delta\}, \quad x \in \mathbb{R}^d
$$

とする．このとき $\delta \searrow 0$ とすると $\inf_{x\in C} I^\delta(x) \nearrow \inf_{x\in C} \Lambda^*(x)$ となっている（$\inf_{x\in C} \Lambda^*(x)$ が ∞ のときは自明，有限のときはその下限を達成する $x_0 \in C$ を考えればわかる）．

各 $x \in C$ に対して I^δ の定義によりある $\lambda_x \in \mathbb{R}^d$ に対して

$$
\langle \lambda_x, x \rangle - \Lambda(\lambda_x) \geq I^\delta(x)
$$

が成立する．また $r_x > 0$ を $r_x|\lambda_x| \leq \delta$ となるようにとっておく．このとき

$$
\inf_{y \in B(x, r_x)} \langle \lambda_x, y \rangle = \langle \lambda_x, x \rangle - r_x|\lambda_x| \geq \langle \lambda_x, x \rangle - \delta
$$

に注意すれば，指数型の Chebyshev の不等式により

$$
\begin{aligned}
\mu_n(B(x, r_x)) &\leq \int_{\mathbb{R}^d} \exp\left(\langle n\lambda_x, \widetilde{x} \rangle - n \inf_{y \in B(x, r_x)} \langle \lambda_x, y \rangle\right) \mu_n(d\widetilde{x}) \\
&\leq \exp\left(\Lambda_n(n\lambda_x) - n\langle \lambda_x, x \rangle + n\delta\right)
\end{aligned}
$$

が成立することがわかる．C の有限開被覆 $B(x_1, r_{x_1}), \ldots, B(x_N, r_{x_N})$ を取れば

$$
\varlimsup_{n\to\infty} \frac{1}{n} \log \mu_n(C) \leq \max_{i=1,\ldots,N} (\Lambda(\lambda_{x_i}) - \langle \lambda_{x_i}, x_i \rangle + \delta)
$$

[5] 本節の記述は [51, 第 3 章] を参考にした.

$$\leq \delta - \min_{i=1,\dots,N} I^{\delta}(x_i) \leq \delta - \inf_{x\in C} I^{\delta}(x)$$

が得られる．ここで C のコンパクト性と補題 1.2.2 を用いた．最後に $\delta \searrow 0$ とすることにより所望の評価

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(C) \leq -\inf_{x\in C} \Lambda^*(x)$$

が得られる．□

定理 3.0.1(2) の証明． (2) を示すためには任意の $y \in \mathcal{F}_{\Lambda^*}$ と $\delta > 0$ に対して

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(B(y,\delta)) \geq -\Lambda^*(y) - \delta|\eta| \tag{3.2.1}$$

を示せば十分である．ここで $\eta \in \mathbb{R}^d$ は Λ^* の y における狭義劣微分と $(\mathrm{dom}\,\Lambda)^\circ$ に属す任意の元である．（１つ固定しておく．）実際，(3.2.1) が示されたとすると，任意の開集合 $O \subset \mathbb{R}^d$ と任意の $y \in O \cap \mathcal{F}_{\Lambda^*}$ に対して $B(y,\delta) \subset O$ となるように δ を取ると

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(O) \geq \underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(B(y,\delta)) \geq -\Lambda^*(y) - \delta|\eta|$$

が成立することがわかる．よって $\delta \searrow 0$ として $y \in O \cap \mathcal{F}_{\Lambda^*}$ に関する上限を取れば所望の評価が得られる．

(3.2.1) を示すために，Cramér の手法による確率測度の変更を行う．まず $\eta \in (\mathrm{dom}\,\Lambda)^\circ$ であるから十分大きい $n \in \mathbb{N}$ に対して $\Lambda_n(n\eta) < \infty$ であることを注意しておく．確率測度 $\widetilde{\mu}_n \in \mathcal{M}_1(\mathbb{R}^d)$ を μ_n に絶対連続であってその Radon-Nikodym 微分が

$$\frac{d\widetilde{\mu}_n}{d\mu_n}(x) = e^{n\langle \eta, x\rangle - \Lambda_n(n\eta)}$$

であるものとする．このとき

$$\begin{aligned}
\frac{1}{n} \log \mu_n(B(y,\delta)) &= \frac{1}{n} \log \int_{B(y,\delta)} e^{\Lambda_n(n\eta) - n\langle \eta, x\rangle} \widetilde{\mu}_n(dx) \\
&= \frac{1}{n}\Lambda_n(n\eta) - \langle \eta, y\rangle + \frac{1}{n} \log \int_{B(y,\delta)} e^{-n\langle \eta, x-y\rangle} \widetilde{\mu}_n(dx) \\
&\geq \frac{1}{n}\Lambda_n(n\eta) - \langle \eta, y\rangle - \delta|\eta| + \frac{1}{n} \log \widetilde{\mu}_n(B(y,\delta))
\end{aligned}$$

が成立する．ここで

$$\lim_{n\to\infty} \widetilde{\mu}_n(B(y,\delta)) = 1 \tag{3.2.2}$$

を認めると，Λ と Λ^* の定義から

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mu_n(B(y,\delta)) \geq \Lambda(\eta) - \langle \eta, y\rangle - \delta|\eta| \geq -\Lambda^*(y) - \delta|\eta|$$

が分かるので，(3.2.1) が示された．

次に (3.2.2) を示そう．そのためには

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \widetilde{\mu}_n(B(y,\delta)^c) < 0 \tag{3.2.3}$$

を示せば十分である．(3.2.3) は $\{\widetilde{\mu}_n\}_{n\in\mathbb{N}}$ に対する大偏差原理の上からの評価により従う．それを見るために $\{\widetilde{\mu}_n\}_{n\in\mathbb{N}}$ に対する条件 **(GE)** が成立していることを確認する．そのために

$$\widetilde{\Lambda}_n(\lambda) := \log \int_{\mathbb{R}^d} e^{\langle \lambda, x\rangle} \widetilde{\mu}_n(dx), \quad \lambda \in \mathbb{R}^d$$

と定義しておく．このとき，

$$\widetilde{\Lambda}_n(n\lambda) = \log \int_{\mathbb{R}^d} e^{n\langle \lambda+\eta, x\rangle - \Lambda_n(n\eta)} \mu_n(dx) = \Lambda_n(n(\lambda+\eta)) - \Lambda_n(n\eta)$$

となっている．よって $\Lambda(\eta) \in \mathbb{R}$ に注意して

$$\widetilde{\Lambda}(\lambda) := \lim_{n\to\infty} \frac{1}{n} \widetilde{\Lambda}_n(n\lambda) = \Lambda(\lambda+\eta) - \Lambda(\eta)$$

となっているので，条件 **(GE)** が成立していることが分かる．よって $\{\widetilde{\mu}_n\}_{n\in\mathbb{N}}$ に対して，

$$\begin{aligned}
\widetilde{\Lambda}^*(x) :&= \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda, x\rangle - \widetilde{\Lambda}(\lambda)\} \\
&= \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda+\eta, x\rangle - \Lambda(\lambda+\eta)\} - \langle \eta, x\rangle + \Lambda(\eta) \\
&= \Lambda^*(x) - \langle \eta, x\rangle + \Lambda(\eta)
\end{aligned}$$

を良い速度関数として定理 3.0.1 の (1) を適用できる．そのため

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \widetilde{\mu}_n(B(y,\delta)^c) \leq -\inf_{x\in B(y,\delta)^c} \widetilde{\Lambda}^*(x)$$

が得られる．よって (3.2.3) を示すためには右辺の下限が正であることを示せばよい．

Λ^* は良い速度関数であるので，補題 1.2.6 によりある $x_0 \in B(y,\delta)^c$ に対して

$$\inf_{x\in B(y,\delta)^c} \widetilde{\Lambda}^*(x) = \widetilde{\Lambda}^*(x_0)$$

となる．$x_0 \neq y$ であり，Λ^* は y において狭義劣微分可能で η は y における狭義劣微分に属すので

$$\begin{aligned}\widetilde{\Lambda}^*(x_0) &= \Lambda^*(x_0) - \langle \eta, x_0 \rangle + \Lambda(\eta) \\ &> \langle \eta, x_0 - y \rangle + \Lambda^*(y) - \langle \eta, x_0 \rangle + \Lambda(\eta) \\ &= \Lambda^*(y) - \langle \eta, y \rangle + \Lambda(\eta) \geq 0\end{aligned}$$

が得られる．最後の不等式は Λ^* の定義により得られる．よって $\widetilde{\Lambda}^*(x_0) > 0$ が得られたので，(2) の証明が終了する．□

定理 3.0.1(3) の証明. Λ は凸であり，下半連続かつ本質的に滑らかであるので，補題 3.1.5 により $\mathrm{ri}\,(\mathrm{dom}\,\Lambda^*) \subset \mathcal{F}_{\Lambda^*}$ が成立する．よって

$$\inf_{x \in O \cap \mathrm{ri}\,(\mathrm{dom}\,\Lambda^*)} \Lambda^*(x) \geq \inf_{x \in O \cap \mathcal{F}_{\Lambda^*}} \Lambda^*(x) \geq \inf_{x \in O} \Lambda^*(x)$$

であるから，定理 3.0.1(3) を示すためには任意の開集合 $O \subset \mathbb{R}^d$ に対して

$$\inf_{x \in O \cap \mathrm{ri}\,(\mathrm{dom}\,\Lambda^*)} \Lambda^*(x) = \inf_{x \in O} \Lambda^*(x)$$

を示せばよい．そのためには任意の $\widetilde{x} \in O$ に対して

$$\Lambda^*(\widetilde{x}) \geq \inf_{x \in O \cap \mathrm{ri}\,(\mathrm{dom}\,\Lambda^*)} \Lambda^*(x) \tag{3.2.4}$$

を示せば十分である．

$\Lambda^*(\widetilde{x}) = \infty$ の場合には (3.2.4) は明らかなので，$\Lambda^*(\widetilde{x}) < \infty$ とする．このとき $\mathrm{dom}\,\Lambda^* \neq \emptyset$ なので命題 A.2.1(1) により $\mathrm{ri}\,(\mathrm{dom}\,\Lambda^*) \neq \emptyset$ である．また命題 A.2.1(2) により任意の $y \in \mathrm{ri}\,(\mathrm{dom}\,\Lambda^*)$ と十分小さい $\alpha > 0$ に対して $(1-\alpha)\widetilde{x} + \alpha y \in O \cap \mathrm{ri}\,(\mathrm{dom}\,\Lambda^*)$ となる．よって Λ^* は凸であるので

$$\begin{aligned}\Lambda^*(\widetilde{x}) &= \lim_{\alpha \searrow 0} \{(1-\alpha)\Lambda^*(\widetilde{x}) + \alpha\Lambda^*(y)\} \\ &\geq \varliminf_{\alpha \searrow 0} \Lambda^*\left((1-\alpha)\widetilde{x} + \alpha y\right) \geq \inf_{x \in O \cap \mathrm{ri}\,(\mathrm{dom}\,\Lambda^*)} \Lambda^*(x)\end{aligned}$$

となり (3.2.4) が示された．□

3.3　独立同分布な確率変数の列に対する中偏差原理

本節では Gärtner-Ellis の大偏差原理の応用として，独立同分布な確率変数の列に対する中偏差原理を調べる．[6]中偏差原理は定義としては大偏差原理の一種であるが，考えて

[6]中偏差原理は英語では Moderate deviation principle と呼ばれる．

いる確率模型において大数の法則（またはそれに付随した大偏差原理）と中心極限定理の間の尺度を調べる大偏差原理を特に中偏差原理と呼ぶ．また重複対数の法則と密接に関連することが知られており，多くの確率模型において調べられている．[7]

$\{X_i\}_{i\in\mathbb{N}}$ を独立同分布な $\mathbb{R}^d$ 値確率変数の列とし，S_n を $S_n := X_1 + \cdots + X_n$ とする．$X_1 = (X_1^1, \ldots, X_1^d)$ に対して次の 3 条件を仮定する．

(MDP1) 原点のある近傍に属する λ に対して $\mathbb{E}\left[e^{\langle \lambda, X_1\rangle}\right] < \infty$.
(MDP2) $\mathbb{E}[X_1] = 0$.
(MDP3) X_1 の共分散行列 $\Gamma = (\Gamma_{ij})_{i,j=1}^d$ は正定値である．ここで通常どおり $\Gamma_{ij} = \mathbb{E}\left[X_1^i X_1^j\right]$ とおく．

ここで関数 $\Lambda : \mathbb{R}^d \to \mathbb{R}$ を

$$\Lambda(\lambda) = \frac{1}{2}\mathbb{E}\left[\langle \lambda, X_1\rangle^2\right] = \frac{1}{2}\langle \lambda, \Gamma\lambda\rangle$$

とし，Λ^* を Λ の Fenchel-Legendre 変換とする．

$$\Lambda^*(x) := \sup_{\lambda\in\mathbb{R}^d}\{\langle \lambda, x\rangle - \Lambda(\lambda)\}, \quad x \in \mathbb{R}^d.$$

Γ は正定値対称なので平方完成をすることにより，各 $x \in \mathbb{R}^d$ に対して $\lambda \mapsto \langle \lambda, x\rangle - \langle \lambda, \Gamma\lambda\rangle/2$ は $\lambda = \Gamma^{-1}x$ において最大になることがわかる．よって

$$\Lambda^*(x) = \frac{1}{2}\langle x, \Gamma^{-1}x\rangle, \quad x \in \mathbb{R}^d$$

である．また $n \to \infty$ のとき $a_n \to 0$ かつ $na_n \to \infty$ となる数列 $\{a_n\}_{n\in\mathbb{N}}$ を任意に選ぶ．

定理 3.3.1. 条件 (MDP1)–(MDP3) を仮定する．$\mathbb{R}^d$ 値確率変数の列 $\{\sqrt{a_n/n}S_n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 Λ^* に対して速度 a_n で大偏差原理を満たす．

注意 3.3.2. $a_n = n^{-a}, a > 0$ の場合を考えると，数列 $\{a_n\}_{n\in\mathbb{N}}$ の増大度 a に対する条件は $0 < a < 1$ であり，$\sqrt{a_n/n} = n^{-(a+1)/2}$ となっている．$a = 1$ は大数の法則（と Cramér の大偏差原理），$a = 0$ は中心極限定理がそれぞれ対応することから，中偏差原理はそれらの間の尺度を調べていることに相当する．

証明. 証明は定理 3.0.1 に基づく．Λ_n を $\sqrt{a_n/n}S_n$ の対数積率母関数

$$\Lambda_n(\lambda) = \log \mathbb{E}\left[e^{\langle \lambda, \sqrt{a_n/n}S_n\rangle}\right]$$

[7] 本節の記述は [33, 第 3.7 節] を参考にした．

とする．定理 3.3.1 を示すためには，任意の $\lambda \in \mathbb{R}^d$ に対して

$$\lim_{n\to\infty} a_n \Lambda_n(a_n^{-1}\lambda) = \Lambda(\lambda) \tag{3.3.1}$$

を示せばよい．実際，Λ は明らかに下半連続かつ本質的に滑らかかつ $\operatorname{dom}\Lambda = \mathbb{R}^d$ なので，定理 3.0.1（の n を a_n^{-1} に取り替えた版）により定理 3.3.1 が得られる．

(3.3.1) を示すために $\lambda \in \mathbb{R}^d$ を固定する．$\{X_i\}_{i\in\mathbb{N}}$ の独立同分布性により

$$\Lambda_n(a_n^{-1}\lambda) = n \log \mathbb{E}\left[e^{(na_n)^{-1/2}\langle \lambda, X_1\rangle}\right]$$

が成立する．ここで

$$\begin{aligned}\varepsilon_n &= \mathbb{E}\left[e^{(na_n)^{-1/2}\langle \lambda, X_1\rangle}\right] - 1 - (na_n)^{-1/2}\mathbb{E}\left[\langle \lambda, X_1\rangle\right] - \frac{1}{2}(na_n)^{-1}\mathbb{E}\left[\langle \lambda, X_1\rangle^2\right] \\ &= \mathbb{E}\left[e^{(na_n)^{-1/2}\langle \lambda, X_1\rangle}\right] - 1 - (na_n)^{-1}\Lambda(\lambda)\end{aligned}$$

とおけば,

$$\Lambda_n(a_n^{-1}\lambda) = n\log\left(1 + (na_n)^{-1}\Lambda(\lambda) + \varepsilon_n\right)$$

である．Taylor の定理による不等式 $|e^x - 1 - x - x^2/2| \le |x|^3 e^{|x|}$ と，仮定 **(MDP)** により $\mathbb{E}\left[e^{|\langle \lambda, X_1\rangle|}\right]$ は λ が原点に近いとき有限であることにより，$n \to \infty$ のとき $na_n\varepsilon_n \to 0$ となる．よって $n \to \infty$ のとき

$$a_n\Lambda_n(a_n^{-1}\lambda) = na_n \log\left(1 + (na_n)^{-1}\Lambda(\lambda) + \varepsilon_n\right) \longrightarrow \Lambda(\lambda)$$

となる．ここで $u \to 0$ のときに $u^{-1}\log(1+u) \to 1$ となることを用いた．よって (3.3.1) が得られたので定理 3.3.1 が示された．　□

以下本節で述べる内容は第 4 章で扱う Sanov の定理に関連する事柄なので，Sanov の定理をまだ学習していない読者はいったん読み飛ばすことをお勧めする．

Cramér の定理に対して中偏差原理（定理 3.3.1）が対応するように，Sanov の定理に対してある中偏差原理が成立することが知られている．以下で紹介するが証明は与えない．S を可分完備距離空間として，$\mathcal{M}_1(S)$ を S 上の Borel 確率測度全体とする．また $\mathcal{M}^{\mathrm{sgn}}(S)$ を S 上の有限符号付き測度全体とし $\mathcal{M}^{\mathrm{sgn}}(S)$ に弱位相を導入すると，$\mathcal{M}^{\mathrm{sgn}}(S)$ は局所凸 Hausdorff 位相ベクトル空間になる．またその位相的双対は

$$\{\nu \mapsto \int_S \varphi d\nu \mid \varphi \in C_{\mathrm{b}}(S)\}$$

となるので,[8] この集合と $C_{\mathrm{b}}(S)$ を同一視する．$\mathcal{M}_0^{\mathrm{sgn}}(S)$ を $\nu \in \mathcal{M}^{\mathrm{sgn}}(S)$ で $\nu(S) = 0$ となるもの全体とする．

[8] この事実については [33, p261] を参照せよ．

$\{X_i\}_{i\in\mathbb{N}}$ を独立同分布な S 値確率変数の列とし，$\mu \in \mathcal{M}_1(S)$ を X_1 の分布とする．各 $n \in \mathbb{N}$ に対して $L_n = n^{-1}(\delta_{X_1} + \cdots + \delta_{X_n}) \in \mathcal{M}_1(S)$ を経験測度とする．経験測度に対する大数の法則により，$\mathcal{M}^{\mathrm{sgn}}(S)$ 値確率変数の列 $\{L_n\}_{n\in\mathbb{N}}$ が μ に確率 1 で弱収束することを思い出しておく．中偏差原理は大数の法則からのずれを調べるものであったので，経験測度の設定では $\mathcal{M}^{\mathrm{sgn}}(S)$ 値確率変数の列 $\{\sqrt{na_n}(L_n-\mu)\}_{n\in\mathbb{N}}$ の大偏差原理を調べることに相当する．ここで数列 $\{a_n\}_{n\in\mathbb{N}}$ は $n\to\infty$ のとき $a_n \to 0$ かつ $na_n \to \infty$ となるものとする．

次に $\{\sqrt{na_n}(L_n-\mu)\}_{n\in\mathbb{N}}$ に対する中偏差原理の速度関数を定義する．$I : \mathcal{M}^{\mathrm{sgn}}(S) \to [0,\infty]$ を次で定義する．

$$I(\nu) = \begin{cases} \dfrac{1}{2}\displaystyle\int_S \left(\frac{d\nu}{d\mu}\right)^2 d\mu, & \nu \in \mathcal{M}_0^{\mathrm{sgn}}(S) \text{ が } \mu \text{ に対して絶対連続のとき}, \\ \infty, & \text{それ以外のとき}. \end{cases}$$

このとき次の大偏差原理が知られている．[9]

定理 3.3.3. $\mathcal{M}^{\mathrm{sgn}}(S)$ 値確率変数の列 $\{\sqrt{na_n}(L_n-\mu)\}_{n\in\mathbb{N}}$ は $n\to\infty$ のときに良い速度関数 I に対して速度 n で大偏差原理を満たす．

注意 3.3.4. 定理 3.3.3 に現れる速度関数 I は，注意 3.0.4 で述べた大偏差原理に現れる Fenchel-Legendre 変換になることを確認する．$\Lambda_n : C_{\mathrm{b}}(S) \to \mathbb{R}$ を次で定義する．

$$\Lambda_n(\varphi) = \log \mathbb{E}\left[e^{\sqrt{na_n}\int_S \varphi d(L_n-\mu)}\right], \quad \varphi \in C_{\mathrm{b}}(S).$$

$\varphi \in C_{\mathrm{b}}(S)$ に対して $a_n\Lambda_n(a_n^{-1}\varphi)$ の $n\to\infty$ における極限を計算する．$\{X_i\}_{i\in\mathbb{N}}$ の独立同分布性により

$$\Lambda_n(a_n^{-1}\varphi) = n\log\mathbb{E}\left[e^{(na_n)^{-1/2}(\varphi(X_1)-\int_S \varphi d\mu)}\right]$$

である．定理 3.3.1 における証明と同様にして

$$\begin{aligned}\lim_{n\to\infty} a_n\Lambda_n(a_n^{-1}\varphi) &= \frac{1}{2}\mathbb{E}\left[\left(\varphi(X_1) - \int_S \varphi d\mu\right)^2\right] \\ &= \frac{1}{2}\left(\int_S \varphi^2 d\mu - \left(\int_S \varphi d\mu\right)^2\right) =: \Lambda(\varphi)\end{aligned}$$

が得られる（$\Lambda(\varphi)$ は μ の下での φ の分散の半分である）．

[9]例えば [53] を参照せよ．

簡単のために $\nu \in \mathcal{M}_0^{\mathrm{sgn}}(S)$ が μ に絶対連続であって，その Radon-Nikodym 微分が $C_{\mathrm{b}}(S)$ に属す場合に $\Lambda^*(\nu)$ を計算する．定義により

$$\Lambda^*(\nu) = \sup_{\varphi \in C_{\mathrm{b}}(S)} \left\{ \int_S \varphi d\nu - \Lambda(\varphi) \right\} \tag{3.3.2}$$

であるが，右辺の上限の中身は

$$\frac{1}{2}\int_S \left(\frac{d\nu}{d\mu}\right)^2 d\mu + \frac{1}{2}\left(\int_S \varphi d\mu\right)^2 - \frac{1}{2}\int_S \left(\varphi - \frac{d\nu}{d\mu}\right)^2 d\mu$$

と書ける．$\varphi = d\nu/d\mu \in C_{\mathrm{b}}(S)$ とすると，第 3 項は 0 になる．また $\nu \in \mathcal{M}_0^{\mathrm{sgn}}(S)$ であったから第 2 項は 0 になる．よって (3.3.2) の右辺の上限は $\varphi = d\nu/d\mu \in C_{\mathrm{b}}(S)$ で達成されるので，

$$\Lambda^*(\nu) = \frac{1}{2}\int_S \left(\frac{d\nu}{d\mu}\right)^2 d\mu$$

とわかる．

3.4 有限状態 Markov 連鎖に対する大偏差原理

本節では Gärtner-Ellis の定理の応用として，有限状態 Markov 連鎖に対する大偏差原理を調べる．これは大偏差原理の研究において最も重要な研究の一つに挙げられる，Donsker-Varadhan の大偏差原理の離散版と解釈できる．有限集合上の Markov 連鎖については例えば [1, 第 7 節] を参照せよ.[10]

本節では S を空でない有限集合とする．$\{X_k\}_{k\in\mathbb{N}}$ を推移確率行列 $P = \{p(x,y)\}_{x,y\in S}$ から決まる S 上の Markov 連鎖とする.（本節では P を $|S|$ 次正方行列と見なす.）すなわち，任意の $n \in \mathbb{N}$ と任意の $x_0, x_1, \ldots, x_n \in S$ に対して

$$\mathbb{P}\left(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n\right) = \mathbb{P}\left(X_0 = x_0\right) \prod_{i=1}^n p(x_{i-1}, x_i)$$

により決定される確率過程である．初期分布 $\mathbb{P}\left(X_0 = \cdot\right)$ は本節の議論に影響しないので，任意に 1 つ固定して話を進める．また P は既約行列であると仮定する．つまり，任意の $x, y \in S$ に対してある $P^m(x,y) > 0$ となる $m = m(x,y) \in \mathbb{N}$ が存在することである.（本節では正方行列 A に対して行列 A^n の (x,y) 成分を $A^n(x,y)$ と書く.）任意に $d \in \mathbb{N}$ と $f : S \to \mathbb{R}^d$ を固定し，$\mathbb{R}^d$ 値確率変数の列 $\{Z_n\}_{n\in\mathbb{N}}$ を

$$Z_n := \frac{1}{n}\sum_{k=1}^n f(X_k)$$

[10] 本節の記述は [33] を参考にした.

により定義する．

本節で現れる大偏差原理では既約な非負行列の Perron-Frobenius 固有値が重要な役割を果たす.[11]次の命題で Perron-Frobenius 固有値の定義と性質を述べておく．

命題 3.4.1. A を既約な非負行列とする．このとき以下の性質を満たす A の正の固有値 ρ が存在する．

(1) 固有値 ρ は単純である．

(2) A の任意の固有値 λ に対して $|\lambda| \leq \rho$ である．

(3) 固有値 ρ に対する右・左固有ベクトルで，全ての成分が正になるものが存在する．さらに，固有値 ρ に対する右・左固有ベクトルは，定数倍を除いてそれぞれ一意的である．

命題 3.4.1 の証明は [1, 定理 7-15] を参照せよ．命題 3.4.1 に現れる固有値 ρ は A の Perron-Frobenius 固有値と呼ばれる．以下既約な非負行列 A の Perron-Frobenius 固有値を $\rho(A)$ と書く．また $\lambda \in \mathbb{R}^d$ に対して

$$p_\lambda(x,y) := p(x,y)e^{\langle \lambda, f(y)\rangle}, \quad P_\lambda := \{p_\lambda(x,y)\}_{x,y\in S}$$

と定義する．P が既約なので P_λ も既約であることに注意する．$\lambda \in \mathbb{R}^d$ に対して $\Lambda(\lambda) := \log \rho(P_\lambda)$ と定義して，Λ^* を Λ の Fenchel-Legendre 変換とする．

$$\Lambda^*(x) := \sup_{\lambda\in\mathbb{R}^d}\{\langle \lambda, x\rangle - \Lambda(\lambda)\}, \quad x \in \mathbb{R}^d.$$

定理 3.4.2. $\mathbb{R}^d$ 値確率変数の列 $\{Z_n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 Λ^* に対して速度 n で大偏差原理を満たす．

証明. Gärtner-Ellis の定理から，

$$\Lambda(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left[e^{n\langle \lambda, Z_n\rangle}\right], \quad \lambda \in \mathbb{R}^d \tag{3.4.1}$$

が成立することと，Λ が $\mathbb{R}^d$ 上微分可能であることを示せばよい．(Λ^* が良い速度関数であることも，Gärtner-Ellis の定理の帰結として従う．）Perron-Frobenius 固有値 $\rho(P_\lambda)$ は特性多項式

$$\det(zI - P_\lambda) = 0$$

[11]全ての成分が非負である正方行列を非負行列という．

の単純根であり，特性多項式の係数は λ について滑らかである．よって陰関数定理により，$\lambda \in \mathbb{R}^d \mapsto \log \rho(P_\lambda)$ は $\mathbb{R}^d$ 上微分可能であることがわかる．

(3.4.1) を示す．$\lambda \in \mathbb{R}^d$ に対して

$$\begin{aligned}
\frac{1}{n}\log \mathbb{E}\left[e^{n\langle \lambda, Z_n\rangle}\right] &= \frac{1}{n}\log\left[\sum_{x_0,x_1,\dots,x_n\in S} \mathbb{P}\left(X_0=x_0, X_1=x_1,\dots,X_n=x_n\right)\prod_{k=1}^{n} e^{\langle \lambda, f(x_k)\rangle}\right]\\
&= \frac{1}{n}\log\left[\sum_{x_0,x_1,\dots,x_n\in S} \mathbb{P}\left(X_0=x_0\right)p(x_0,x_1)\cdots p(x_{n-1},x_n)\prod_{k=1}^{n} e^{\langle \lambda, f(x_k)\rangle}\right]\\
&= \frac{1}{n}\log\left[\sum_{x_0,x_1,\dots,x_n\in S} \mathbb{P}\left(X_0=x_0\right)\left\{\prod_{k=1}^{n} p(x_{k-1},x_k)e^{\langle \lambda, f(x_k)\rangle}\right\}\right]\\
&= \frac{1}{n}\log\left[\sum_{x_0,x_n\in S} \mathbb{P}\left(X_0=x_0\right)P_\lambda^n(x_0,x_n)\right] \qquad (3.4.2)
\end{aligned}$$

である．$\varphi : S \to [1,\infty)$ を P_λ の $\rho(P_\lambda)$ に対する左固有ベクトルとする（命題 3.4.1(3) により $\varphi : S \to [1,\infty)$ としてよい）．このとき (3.4.2) により

$$\begin{aligned}
\frac{1}{n}\log \mathbb{E}\left[e^{n\langle \lambda, Z_n\rangle}\right] &\le \frac{1}{n}\log\left[\sum_{x_0,x_n\in S}\varphi(x_0)P_\lambda^n(x_0,x_n)\right]\\
&= \frac{1}{n}\log\left[\sum_{x_n\in S}\rho(P_\lambda)^n\varphi(x_n)\right]\\
&= \log\rho(P_\lambda) + \frac{1}{n}\log\left[\sum_{x\in S}\varphi(x)\right]
\end{aligned}$$

がわかる．同様に，$\psi : S \to (0,1]$ を P_λ の $\rho(P_\lambda)$ に対する右固有ベクトルとすれば，(3.4.2) により

$$\begin{aligned}
\frac{1}{n}\log \mathbb{E}\left[e^{n\langle \lambda, Z_n\rangle}\right] &\ge \frac{1}{n}\log\left[\sum_{x_0,x_n\in S}\mathbb{P}\left(X_0=x_0\right)P_\lambda^n(x_0,x_n)\psi(x_n)\right]\\
&= \frac{1}{n}\log\left[\sum_{x_0\in S}\mathbb{P}\left(X_0=x_0\right)\rho(P_\lambda)^n\psi(x_0)\right]\\
&= \log\rho(P_\lambda) + \frac{1}{n}\log\left[\sum_{x\in S}\mathbb{P}\left(X_0=x\right)\psi(x)\right]
\end{aligned}$$

がわかる．これらより (3.4.1) が成立する．　□

本節の残りでは $S=\{1,\dots,|S|\}$ と書くことにし，$d=|S|$ の場合を考える．S は有限

集合なので，S 上の確率測度全体 $\mathcal{M}_1(S)$ は $\mathbb{R}^d$ 内の単体

$$\left\{ x = (x_i)_{i=1}^d \in [0,\infty)^d \,\middle|\, \sum_{i=1}^d x_i = 1 \right\}$$

と同一視できるので，両者を同じ記号 $\mathcal{M}_1(S)$ で書くことにする．この同一視を通じて経験測度

$$L_n = \frac{1}{n}\sum_{k=1}^n \delta_{X_k}$$

は $\mathbb{R}^d$ の単体 $L_n = (L_n(1), \ldots, L_n(d))$ と同一視される．ここで

$$L_n(i) = \frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{X_k = i\}}, \quad i = 1, \ldots, d,$$

である．これは定理 3.4.2 における Z_n が

$$f(i) = (\mathbf{1}_{\{i=1\}}, \ldots, \mathbf{1}_{\{i=d\}}), \quad i = 1, \ldots, d$$

で与えられる場合に対応するので，$\mathcal{M}_1(S)$ 値確率変数の列 $\{L_n\}_{n\in\mathbb{N}}$ は大偏差原理を満たし速度関数は

$$I(\alpha) = \sup_{\lambda\in\mathbb{R}^d}\{\langle \lambda, \alpha\rangle - \log\rho(P_\lambda)\}, \quad \alpha \in \mathcal{M}_1(S) \tag{3.4.3}$$

で与えられる．(ここでも先述の同一視を行っている．）ただし

$$P_\lambda = \{p_\lambda(i,j)\}_{i,j\in S}, \quad p_\lambda(i,j) = p(i,j)e^{\lambda_j}$$

で与えられる．

速度関数 (3.4.3) はある変分公式で書けることが知られている．$u \in \mathbb{R}^d$ に対して u の成分が全て正であるとき $u > 0$ と書くことにする．(u は行ベクトルと思っている．）関数 $J : \mathcal{M}_1(S) \to [0,\infty]$ を次で定義する．

$$J(\alpha) := \sup_{u>0} \sum_{i=1}^d \alpha_i \log\left(\frac{u_i}{(uP)_i}\right), \quad \alpha \in \mathcal{M}_1(S).$$

次の命題で $I = J$ であることを示す．

定理 3.4.3. 任意の $\alpha \in \mathcal{M}_1(S)$ に対して $I(\alpha) = J(\alpha)$ が成立する．

証明. $\alpha \in \mathcal{M}_1(S)$ を固定して $I(\alpha) \geq J(\alpha)$ を示す．$u \in \mathbb{R}^d$ を $u > 0$ なるものとする．P は既約なので任意の $i = 1, \ldots, d$ に対して $(uP)_i > 0$ である．よって

$$\lambda_i = \log\left(\frac{u_i}{(uP)_i}\right)$$

とおくと λ_i は実数である．また任意の $j = 1, \ldots, d$ に対して

$$\sum_{i=1}^d u_i p_\lambda(i,j) = \sum_{i=1}^d u_i p(i,j) e^{\lambda_j} = u_j$$

であるから $uP_\lambda = u$ である．また (3.4.2) の極限を求めたときと同じ議論により

$$\log \rho(P_\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{i=1}^d u_i P_\lambda^n(i,j)$$

がわかるが，$uP_\lambda = u$ なので右辺の極限は 0 である．よって $\rho(P_\lambda) = 1$ となるので，I の定義により

$$I(\alpha) \geq \sum_{i=1}^d \alpha_i \log\left(\frac{u_i}{(uP)_i}\right)$$

である．$u > 0$ は任意なので $I(\alpha) \geq J(\alpha)$ を得る．

次に $I(\alpha) \leq J(\alpha)$ を示す．$\lambda \in \mathbb{R}^d$ を固定して，$v \in \mathbb{R}^d$ を P_λ の Perron-Frobenius 固有値 $\rho(P_\lambda)$ に対する左固有ベクトルとする．つまり，$vP_\lambda = \rho(P_\lambda)v$ を満たし $v > 0$ となるもののひとつである．ここで

$$\langle \lambda, \alpha \rangle - \sum_{i=1}^d \alpha_i \log\left(\frac{v_i}{(vP)_i}\right) = \sum_{i=1}^d \alpha_i \log\left(\frac{e^{\lambda_i}(vP)_i}{v_i}\right) = \sum_{i=1}^d \alpha_i \log \rho(P_\lambda) = \log \rho(P_\lambda)$$

である．よって

$$\langle \lambda, \alpha \rangle - \log \rho(P_\lambda) \leq \sup_{u>0} \sum_{i=1}^d \alpha_i \log\left(\frac{u_i}{(uP)_i}\right) = J(\alpha)$$

が成立する．$\lambda \in \mathbb{R}^d$ は任意なので $I(\alpha) \leq J(\alpha)$ が示された．　□

経験測度 L_n は状態空間 S に値を取る確率変数の列 $\{X_k\}_{k\in\mathbb{N}}$ を用いて定義されたが，路の空間に値を取る確率変数の列に対して類似物を定義することができる．その類似物はしばしば経験過程と呼ばれる．本節の残りでは Markov 連鎖の経験過程に対する大偏差原理を示す．経験平均・経験測度・経験過程に対する大偏差原理は，それぞれレベル

1・レベル2・レベル3の大偏差原理とよばれ，数学的には番号順に結果が強くなっている.[12]

まず経験過程を定義する．$i \in \mathbb{N}$ に対して $\theta_i : S^{\mathbb{N}} \to S^{\mathbb{N}}$ を $S^{\mathbb{N}}$ 上のシフト作用素とする．つまり $\mathbf{x} = (x_1, x_2, \ldots) \in S^{\mathbb{N}}$ に対して，

$$\theta_i \mathbf{x} := (x_{i+1}, x_{i+2}, \ldots) \in S^{\mathbb{N}}$$

である．各 $n \in \mathbb{N}$ に対して経験過程と呼ばれる，$\mathcal{M}_1(S^{\mathbb{N}})$ 値確率変数 $L_{n,\infty}$ を次で定義する．

$$\mathbf{X} := (X_1, X_2, \ldots) \in S^{\mathbb{N}}, \quad L_{n,\infty} := \frac{1}{n}\sum_{i=1}^{n} \delta_{\theta_i \mathbf{X}} \in \mathcal{M}_1(S^{\mathbb{N}}).$$

本節の残りの主定理は経験過程 $\{L_{n,\infty}\}_{n\in\mathbb{N}}$ に対する大偏差原理である．正確な主張は定理3.4.5で述べられる．証明の鍵はこれまでに述べたMarkov連鎖に対する大偏差原理とDawson-Gärtnerの定理である．

Dawson-Gärnerの定理を適用するためには，射影極限を取る前の空間に値を取る確率変数に対する大偏差原理が必要になる．次はこれについて述べる．$k \geq 2$ を固定して，S^k 値確率変数の列 $\{(X_i, X_{i+1}, \ldots, X_{i+k-1})\}_{i\in\mathbb{N}}$ に対する経験測度を

$$L_{n,k} := \frac{1}{n}\sum_{i=1}^{n} \delta_{(X_i, X_{i+1}, \ldots, X_{i+k-1})} \in \mathcal{M}_1(S^k)$$

により定義する．$\{(X_i, X_{i+1}, \ldots, X_{i+k-1})\}_{i\in\mathbb{N}}$ は S^k 上のMarkov連鎖であり，その推移確率は

$$p_k(x,y) = \prod_{i=1}^{k-1} \mathbf{1}_{\{x_{i+1}=y_i\}} p(x_k, y_k),$$

$x = (x_1, \ldots, x_k), y = (y_1, \ldots, y_k) \in S^k$ で与えられる．対応する推移確率行列を $P^{(k)} = \{p_k(x,y)\}_{x,y\in S^k}$ と表すことにする．p が既約であっても，一般に $P^{(k)}$ が既約にならないことに注意する．そのため以下では任意の $x, y \in S$ に対して $p(x,y) > 0$ となることを仮定しておく．この仮定により $P^{(k)}$ は既約行列になる．

$\{(X_i, X_{i+1}, \ldots, X_{i+k-1})\}_{i\in\mathbb{N}}$ は有限状態空間上の既約なMarkov連鎖なので，$\mathcal{M}_1(S^k)$ 値確率変数の列 $\{L_{n,k}\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに速度 n で大偏差原理を満たす（以前と同様に定理3.4.2の特殊な場合である）．その速度関数は命題3.4.3により

$$I_k(\alpha) := \sup_{u\in(\mathbb{R}^d)^k, u>0} \int_{S^k} \log\left(\frac{u}{(uP^{(k)})}\right) d\alpha, \quad \alpha \in \mathcal{M}_1(S^k)$$

[12] Cramérの大偏差原理はレベル1，Sanovの大偏差原理はレベル2に相当する．

で与えられる．

次に Dawson-Gärtner の定理を適用する際に現れる射影系について述べる．$k \le l$ に対して射影 $q_{k,l} : S^l \to S^k$ を

$$q_{k,l}(x_1, \ldots, x_l) := (x_1, \ldots, x_k)$$

により定義する．$((S^k)_k, (q_{k,l})_{k \le l})$ は射影系でありその射影極限は $S^{\mathbb{N}}$ である．また $q_k : S^{\mathbb{N}} \to S^k$ を自然な射影とする．$S^{\mathbb{N}}$ には S に離散位相を導入した際の直積位相が導入されていることに注意しておく．次に $\pi_{k,l} : \mathcal{M}_1(S^l) \to \mathcal{M}_1(S^k)$ を $q_{k,l}$ から決まる射影としておく．つまり $\alpha \in \mathcal{M}_1(S^l)$ に対して，$\pi_{k,l}(\alpha)$ は S^k 上の確率測度 $\alpha \circ q_{k,l}^{-1}$ である．射影系 $((\mathcal{M}_1(S^k))_k, (\pi_{k,l})_{k \le l})$ の射影極限に対して次が成立する．

命題 3.4.4. 射影系 $((\mathcal{M}_1(S^k))_k, (\pi_{k,l})_{k \le l})$ の射影極限は $\mathcal{M}_1(S^{\mathbb{N}})$ に同相である．

命題 3.4.4 は本節の最後に示すことにして，経験過程に対する大偏差原理を正確に述べる．命題 3.4.4 による同一視の下，$k \in \mathbb{N}$ に対して π_k を $\mathcal{M}_1(S^{\mathbb{N}})$ から $\mathcal{M}_1(S^k)$ への射影とする．つまり $\alpha \in \mathcal{M}_1(S^{\mathbb{N}})$ に対して，$\pi_k(\alpha)$ は $\mathcal{M}_1(S^k)$ 上の確率測度 $\alpha \circ q_k^{-1}$ である．このとき経験過程 $L_{n,\infty}$ の定義により

$$\pi_k(L_{n,\infty}) = L_{n,k}$$

が成立している．よって $I_\infty : \mathcal{M}_1(S^{\mathbb{N}}) \to [0, \infty]$ を

$$I_\infty(\alpha) := \sup_{k \in \mathbb{N}} I_k(\pi_k(\alpha)), \quad \alpha \in \mathcal{M}_1(S^{\mathbb{N}})$$

と定義すれば，Dawson-Gärtner の定理から次を得る．

定理 3.4.5. 任意の $x, y \in S$ に対して $p(x, y) > 0$ であることを仮定する．このとき，$\mathcal{M}_1(S^{\mathbb{N}})$ 値確率変数の列 $\{L_{n,\infty}\}_{n \in \mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 I_∞ に対して速度 n で大偏差原理を満たす.

最後に命題 3.4.4 を証明して本節を終える．

命題 3.4.4 の証明. 射影系 $((\mathcal{M}_1(S^k))_k, (\pi_{k,l})_{k \le l})$ の射影極限を $\mathcal{X}$ と書くことにする．つまり $\mathcal{X}$ は $\prod_{k=1}^\infty \mathcal{M}_1(S^k)$ の部分集合

$$\left\{ (\nu_k)_{k \in \mathbb{N}} \in \prod_{k=1}^\infty \mathcal{M}_1(S^k) \;\middle|\; \pi_{k,l}(\nu_l) = \nu_k, k \le l \right\}$$

である．$\mathcal{X}$ の元 $(\nu_k)_{k \in \mathbb{N}}$ が与えられたとすると，Kolmogorov の拡張定理により $S^{\mathbb{N}}$ 上の確率測度 ν_∞ であって任意の $k \in \mathbb{N}$ に対して $\pi_k(\nu_\infty) = \nu_k$ を満たすものが唯一存在する.[13]

[13]Kolmogorov の拡張定理については，例えば [9, 定理 4.22] を参照せよ．

逆に $S^{\mathbb{N}}$ 上の確率測度 ν_∞ が与えられたとき，$\nu_k := \pi_k(\nu_\infty)$ と定義すれば任意の $k \le l$ に対して $\pi_{l,k}(\nu_l) = \nu_k$ が成立する．よって $\mathcal{X}$ と $\mathcal{M}_1(S^{\mathbb{N}})$ には全単射が存在する．

この全単射が同相写像であることを示そう．$\mathcal{X}$ と $\mathcal{M}_1(S^{\mathbb{N}})$ はコンパクト距離空間であるから，点列連続性を示せば十分である．$S^{\mathbb{N}}$ 上の確率測度の列 $\{\nu^n\}_{n\in\mathbb{N}}$ は $\nu \in \mathcal{M}_1(S^{\mathbb{N}})$ に弱収束すると仮定する．各 $k \in \mathbb{N}$ に対し，S^k 上の確率測度をそれぞれ

$$\nu_k^n := \pi_k(\nu^n), \quad \nu_k := \pi_k(\nu)$$

と定義する．ここで f_k を S^k 上の（有界連続）関数とすると，

$$\langle \nu_k^n, f_k \rangle = \langle \nu^n, f_k \circ q_k \rangle$$

である．よって $f_k \circ q_k$ は $S^{\mathbb{N}}$ 上の有界連続関数なので，

$$\lim_{n\to\infty} \langle \nu_k^n, f_k \rangle = \langle \nu, f_k \circ q_k \rangle = \langle \nu_k, f_k \rangle$$

がわかった．よって任意の $k \in \mathbb{N}$ に対して $\{\nu_k^n\}_{n\in\mathbb{N}}$ は ν_k に弱収束する．これにより $\mathcal{X}$ の点列 $\{(\nu_k^n)_{k\in\mathbb{N}}\}_{n\in\mathbb{N}}$ が $(\nu_k)_{k\in\mathbb{N}}$ に収束することが示された．

逆に $\mathcal{X}$ の点列 $\{(\nu_k^n)_{k\in\mathbb{N}}\}_{n\in\mathbb{N}}$ が与えられたとし，$(\nu_k)_{k\in\mathbb{N}}$ に収束すると仮定する．Kolmogorov の拡張定理により対応する $\mathcal{M}_1(S^{\mathbb{N}})$ 上の確率測度をそれぞれ ν_∞^n, ν_∞ とする．$\varepsilon > 0$ と $S^{\mathbb{N}}$ 上の有界連続関数 f を固定しておく．$S^{\mathbb{N}}$ 上の距離を

$$d_\infty(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} 2^{-k} \mathbf{1}_{\{x_k = y_k\}}$$

により定義しておく．このとき任意の $\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2 \in S^{\mathbb{N}}$ に対して

$$d_\infty((q_N(\mathbf{x}), \mathbf{y}_1), (q_N(\mathbf{x}), \mathbf{y}_2)) \le 2^{-N}$$

である（ここで $x \in S^N$ と $\mathbf{x} \in S^{\mathbb{N}}$ に対して，$(x, \mathbf{x})$ は $S^{\mathbb{N}}$ の元であって $q_N(x, \mathbf{x}) = x$ かつ $\theta_N(x, \mathbf{x}) = \mathbf{x}$ となるものである）．$S^{\mathbb{N}}$ はコンパクトなので f は $S^{\mathbb{N}}$ 上一様連続である．よって $N \in \mathbb{N}$ が十分大きければ任意の $\mathbf{x}, \mathbf{y} \in S^{\mathbb{N}}$ に対して

$$|f(\mathbf{x}) - f(q_N(\mathbf{x}), \mathbf{y})| \le \varepsilon$$

となる．$\mathbf{y}_* \in S^{\mathbb{N}}$ を任意に取って固定して，S^N 上の関数 f_N を

$$f_N(x) = f(x, \mathbf{y}_*), \quad x \in S^N$$

により定義すれば，f_N は S^N 上有界連続である．また $n \ge N$ であれば三角不等式により

$$|\langle \nu_\infty^n, f \rangle - \langle \nu_\infty, f \rangle| \le 2\varepsilon + \left| \int_{S^{\mathbb{N}}} f(q_N(\mathbf{x}), \mathbf{y}_*) \nu_\infty^n(d\mathbf{x}) - \int_{S^{\mathbb{N}}} f(q_N(\mathbf{x}), \mathbf{y}_*) \nu_\infty(d\mathbf{x}) \right|$$

$$= 2\varepsilon + \left| \int_{S^N} f_N(x)\nu_N^n(dx) - \int_{S^N} f_N(x)\nu_N(dx) \right|$$

が成立する．よって $n \to \infty$ とした後，$\varepsilon \searrow 0$ とすれば

$$\lim_{n\to\infty} \langle \nu_\infty^n, f\rangle = \langle \nu_\infty, f\rangle$$

が得られるので，$\{\nu_\infty^n\}_{n\in\mathbb{N}}$ は ν_∞ に弱収束することがわかった． $\square$

第4章　Sanovの大偏差原理

本章ではSanovの大偏差原理について論じる.[1] Sを可分完備距離空間とし，$C_{\mathrm{b}}(S)$をS上の有界連続関数全体，$\mathcal{M}_1(S)$をS上の確率測度全体とする．Sは可分なので，有界連続関数の列$\{f_n\}_{n\in\mathbb{N}}\subset C_{\mathrm{b}}(S)$であって，$\mathcal{M}_1(S)$を分離するものが存在する.[2] このとき$\mathcal{M}_1(S)$上の距離を

$$d_{\mathcal{M}_1(S)}(\mu,\nu):=\sum_{n=1}^{\infty}\frac{1}{2^n}\left\{1\wedge\left|\int_S f_n d\mu-\int_S f_n d\nu\right|\right\},\quad \mu,\nu\in\mathcal{M}_1(S) \tag{4.0.1}$$

により定義すると，$d_{\mathcal{M}_1(S)}$は弱収束の位相を誘導して，$\mathcal{M}_1(S)$は可分完備距離空間になる．

$\lambda\in\mathcal{M}_1(S)$を任意に取り，$\{X_i\}_{i\in\mathbb{N}}$を$S$値独立同分布な確率変数の列で，$X_1$の分布が$\lambda$であるものとする．ここで経験測度と呼ばれる$S$上のランダムな確率測度を

$$L_n:=\frac{1}{n}\sum_{k=1}^{n}\delta_{X_k}\in\mathcal{M}_1(S)$$

により定義する．L_nは$\mathcal{M}_1(S)$値確率変数なのでその法則は$\mathcal{M}_1(\mathcal{M}_1(S))$の元である．

各$f\in C_{\mathrm{b}}(S)$に対して大数の強法則により$n\to\infty$のとき

$$\int_S f dL_n=\frac{1}{n}\sum_{k=1}^{n}f(X_k)\to\mathbb{E}\left[f(X_1)\right]=\int_S f d\lambda,\quad \text{a.s.}$$

が成立する．距離$d_{\mathcal{M}_1(S)}$の定義により，この収束は経験測度に対する弱収束を意味する．つまり$\mathcal{M}_1(S)$において$n\to\infty$のとき

$$L_n\to\lambda,\quad \text{a.s.}$$

が成立する．特に$\mathcal{M}_1(S)$値確率変数の列$\{L_n\}_{n\in\mathbb{N}}$は$n\to\infty$のときδ_λに法則収束している．(δ_λは$\lambda\in\mathcal{M}_1(S)$に集中するDirac測度である.）これは大偏差原理を論じる典型的な状況になっており，対応する大偏差原理はSanovの大偏差原理と呼ばれる．

---

[1] 本章の記述は [42, 第5章] を参考にした．

[2] 有界連続関数の列$\{f_n\}_{n\in\mathbb{N}}\subset C_{\mathrm{b}}(S)$が$\mathcal{M}_1(S)$を分離するとは，$\mu,\nu\in\mathcal{M}_1(S)$が任意の$n\in\mathbb{N}$に対して$\int_S f_n d\mu=\int_S f_n d\nu$を満たすならば，$\mu=\nu$となることである．

Sanov の大偏差原理を述べるために，相対エントロピー $H : \mathcal{M}_1(S)^2 \to [0,\infty]$ を定義する．$\mu,\nu \in \mathcal{M}_1(S)$ に対して，μ の ν に対する相対エントロピー $H(\mu|\nu)$ を次で定義する．

$$H(\mu|\nu) := \begin{cases} \displaystyle\int_S \frac{d\mu}{d\nu} \log\left(\frac{d\mu}{d\nu}\right) d\nu, & \mu \text{ が } \nu \text{ に対して絶対連続なとき}, \\ \infty, & \text{それ以外のとき}. \end{cases}$$

ただし $0\log 0 := 0$ と定める．相対エントロピーが非負であることは命題 4.1.1 で見る．

これで Sanov の大偏差原理を述べる準備ができた．速度関数が良いことは命題 4.1.4 で見る．定理 4.0.1 の証明は第 4.2 節で行う．

定理 4.0.1. $\mathcal{M}_1(S)$ 値確率変数の列 $\{L_n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 $H(\cdot|\lambda)$ に対して速度 n で大偏差原理を満たす．

本論に入る前に相対エントロピーの定義についていくつかの注意を述べる．

注意 4.0.2. $h : [0,\infty) \to \mathbb{R}$ を $h(u) := u\log u$ $(u \geq 0)$ と定義する．(先に述べたように $h(0) = 0$ である．) このとき h は $[0,\infty)$ 上で連続であり，$(0,\infty)$ 上で $h'' > 0$ なので真に凸である．また h は $[0,\infty)$ 上の最小値 $-e^{-1}$ を $u = e^{-1}$ において達成する．

次に $\mu,\nu \in \mathcal{M}_1(S)$ として，μ は ν に絶対連続であると仮定する．$\varphi := d\mu/d\nu$ と書く．φ はほとんど確実に ν に対して定義されるが，μ は ν に絶対連続なので，ほとんど確実に μ に対しても定義されている．また

$$\int_S \log\varphi d\mu, \quad \int_{\{\varphi>0\}} \log\varphi d\mu \tag{4.0.2}$$

はどちらも積分確定で相対エントロピー $H(\mu|\nu)$ に等しい．(ただし $\log 0 := -\infty$ と定義して $\log$ は $[0,\infty)$ 上の関数と見なす．) これを見るために，S を 3 つの集合 $\{\varphi = 0\}, \{0 < \varphi < 1\}, \{\varphi \geq 1\}$ に分解して，各集合上での積分をそれぞれ考える．まず，$\mu(\{\varphi = 0\}) = 0$ なので，積分の定義と $0 \times \infty = 0$ という測度論の規約により

$$\int_{\{\varphi=0\}} \log\varphi d\mu = (-\infty) \times \mu(\{\varphi = 0\}) = 0 \tag{4.0.3}$$

となる．また $\log\varphi$ は $\{0 < \varphi < 1\}, \{\varphi \geq 1\}$ 上でそれぞれ定符号なので，Radon-Nikodym の定理により

$$\int_{\{0<\varphi<1\}} \log\varphi d\mu = \int_{\{0<\varphi<1\}} \varphi\log\varphi d\nu \in [-e^{-1}, 0]$$
$$\int_{\{\varphi\geq 1\}} \log\varphi d\mu = \int_{\{\varphi\geq 1\}} \varphi\log\varphi d\nu \in [0,\infty]$$

となるので，これらの和は積分確定である．また (4.0.3) により，(4.0.2) の 2 つの積分は一致して

$$\int_{\{\varphi>0\}} \varphi \log \varphi d\nu = \int_S \varphi \log \varphi d\nu = H(\mu|\nu)$$

に等しい．

4.1 相対エントロピーの性質

Sanov の定理を示すときに必要になる，相対エントロピーの基本性質を本節で示す．本節でも引き続き $\lambda \in \mathcal{M}_1(S)$ は任意に選び，固定しておく．

命題 4.1.1. 任意の $\mu \in \mathcal{M}_1(S)$ に対して $H(\mu|\lambda) \geq 0$ であり，$H(\mu|\lambda) = 0$ となるのは $\mu = \lambda$ のときに限る．

証明. 注意 4.0.2 で述べたように $h(u) := u \log u$ とおくと，h は凸なので Jensen の不等式により $H(\mu|\lambda) \geq 0$ がわかる．また $H(\mu|\lambda) = 0$ となるのは,（h が真に凸なので）Jensen の不等式で等号が達成されるとき，つまり λ についてほとんど至る所 $d\mu/d\lambda = 1$ となるときである．よって $H(\mu|\lambda) = 0$ となるのは $\mu = \lambda$ のときに限る． □

いくつか記号を用意する．$M_{\mathrm{b}}(S)$ を S 上の有界 Borel 可測関数全体とし，関数 $p : M_{\mathrm{b}}(S) \to \mathbb{R}$ を次で定義する．

$$p(f) := \log \int_S e^f d\lambda, \quad f \in M_{\mathrm{b}}(S).$$

また $\mathcal{M}^{\mathrm{sgn}}(S)$ を S 上の符号付き有限測度全体とし，$\mathcal{M}^{\mathrm{sgn}}(S) \backslash \mathcal{M}_1(S)$ 上では $H(\cdot|\lambda) := \infty$ として拡張する.[3]さらに関数 $G : \mathcal{M}^{\mathrm{sgn}}(S) \to [-\infty, \infty]$ と $q : M_{\mathrm{b}}(S) \to [-\infty, \infty]$ に対して，Fenchel-Legendre 変換 $G^* : M_{\mathrm{b}}(S) \to [-\infty, \infty]$ と $q^* : \mathcal{M}^{\mathrm{sgn}}(S) \to [-\infty, \infty]$ をそれぞれ次で定義する．

$$G^*(f) := \sup_{\nu \in \mathcal{M}^{\mathrm{sgn}}(S)} \left\{ \int_S f d\nu - G(\nu) \right\}, \quad f \in M_{\mathrm{b}}(S),$$
$$q^*(\mu) := \sup_{f \in M_{\mathrm{b}}(S)} \left\{ \int_S f d\mu - q(f) \right\}, \quad \mu \in \mathcal{M}^{\mathrm{sgn}}(S).$$

相対エントロピーに対する変分公式として，Donsker-Varadhan の変分公式と呼ばれる次の公式が知られている．

[3]符号付き測度については，例えば [22, 第 4.1 節] を参照せよ．

定理 4.1.2. $H(\cdot|\lambda) = p^*(\cdot)$ が成立する．すなわち任意の $\mu \in \mathcal{M}_1(S)$ に対して次が成立する．

$$H(\mu|\lambda) = \sup_{f \in M_\mathrm{b}(S)} \left\{ \int_S f d\mu - \log \int_S e^f d\lambda \right\}.$$

証明. 本証明中では $I(\cdot) := H(\cdot|\lambda)$ と書く．まず $p = I^*$ を示す．そのために $f \in M_\mathrm{b}(S)$ を固定する．ν を S 上の確率測度であって

$$\frac{d\nu}{d\lambda} = \frac{e^f}{\int_S e^f d\lambda}$$

となるものとすると，I^* の定義により

$$\begin{aligned} I^*(f) \geq \int_S f d\nu - I(\nu) &= \frac{\int_S f e^f d\lambda}{\int_S e^f d\lambda} - \int_S \left[\frac{e^f}{\int_S e^f d\lambda} \left(f - \log \int_S e^f d\lambda \right) \right] d\lambda \\ &= \log \int_S e^f d\lambda = p(f) \end{aligned}$$

が成立する．よって $p(f) \leq I^*(f)$ である．逆向きの不等式を見るために，$\nu \in \mathcal{M}_1(S)$ を $I(\nu) < \infty$ となるものとする．このとき相対エントロピーの定義により，ν は λ に対して絶対連続であり，Radon-Nikodym 微分 $d\nu/d\lambda$ が存在する．$\varphi := d\nu/d\lambda$ と書くと，$\log \varphi$ は ν について可積分である．このとき

$$\int_S f d\nu - I(\nu) = \int_S \log \frac{e^f}{\varphi} d\nu \leq \log \int_S \frac{e^f}{\varphi} d\nu$$

が成立する．ただし最後の不等式では Jensen の不等式を用いた．注意 4.0.2 に注意すると φ の定義により

$$\log \int_S \frac{e^f}{\varphi} d\nu = \log \int_S e^f \mathbf{1}_{\{\varphi > 0\}} d\lambda \leq \log \int_S e^f d\lambda = p(f)$$

を得る．よって $\nu \in \mathcal{M}_1(S)$ について上限を取ると $I^*(f) \leq p(f)$ が得られたので，$p(f) = I^*(f)$ が示せた．

以下 $\nu \in \mathcal{M}^{\mathrm{sgn}}(S)$ とし $I(\nu) \leq p^*(\nu)$ を示す．まず $\nu \in \mathcal{M}^{\mathrm{sgn}}(S) \setminus \mathcal{M}_1(S)$ のときに $p^*(\nu) = \infty$ となることを確かめる．まず符号付き測度 ν が負の部分をもつと仮定する．つまり，Borel 集合 $A \subset S$ を ν の Hahn 分解における負集合としたとき，$\nu(A) < 0$ になると仮定する．[4] このとき p^* の定義において $f = -c\mathbf{1}_A$ $(c > 0)$ とすれば

$$p^*(\nu) \geq \int_S -c\mathbf{1}_A d\nu - \log \int_S e^{-c\mathbf{1}_A} d\lambda \geq -c\nu(A)$$

[4] つまり任意の Borel 集合 $B \subset S$ に対して，$\nu(A \cap B) \leq 0$ かつ $\nu(A^c \cap B) \geq 0$ が成立する．Hahn 分解については，例えば [22, 第 4.2 節] を参照せよ．

が成立する．よって $c \nearrow \infty$ とすることにより $p^*(\nu) = \infty$ がわかる．次に ν が正の測度である場合は，p^* の定義において $f = c$ $(c \in \mathbb{R})$ とすれば

$$p^*(\nu) \geq \int_S c d\nu - \log \int_S e^c d\lambda = c(\nu(S) - 1)$$

が得られる．よって $\nu(S) \neq 1$ のときも $p^*(\nu) = \infty$ である．本段落の最後に $\nu \in \mathcal{M}_1(S)$ が λ に対して絶対連続でない場合にも $p^*(\nu) = \infty$ となることを確かめる．このとき S の可測部分集合 A で $\lambda(A) = 0, \nu(A) > 0$ となるものが存在する．p^* の定義において $f = c\mathbf{1}_A$ $(c > 0)$ とすれば

$$p^*(\nu) \geq \int_S c\mathbf{1}_A d\nu - \log \int_S e^{c\mathbf{1}_A} d\lambda = c\nu(A)$$

が得られる．よって $p^*(\nu) = \infty$ が得られる．

本段落で $\nu \in \mathcal{M}_1(S)$ が λ に対して絶対連続であるときに $I(\nu) \leq p^*(\nu)$ を示す．φ を Radon-Nikodym 微分 $d\nu/d\lambda$ とし，$0 < a < 1 < b$ に対して $\varphi_a^b := a \vee (\varphi \wedge b)$ とおく．p^* の定義で $f = \log \varphi_a^b$ とおき，$\varphi_a^b \leq a \vee \varphi$ を用いれば

$$p^*(\nu) \geq \int_S \log \varphi_a^b d\nu - \log \int_S \varphi_a^b d\lambda \geq \int_S \log \varphi_a^b d\nu - \log \int_S a \vee \varphi d\lambda$$

が得られる．$a > 0$ を固定すれば φ_a^b は下から一様に有界なので，$b \to \infty$ とするとき単調収束定理が使えて

$$\begin{aligned} p^*(\nu) &\geq \int_S \log(a \vee \varphi) d\nu - \log \int_S a \vee \varphi d\lambda \\ &\geq \int_S \log \varphi d\nu - \log \int_S a \vee \varphi d\lambda \\ &= I(\nu) - \log \int_S a \vee \varphi d\lambda \end{aligned}$$

が得られる．$0 \leq a \vee \varphi \leq \varphi + 1$ なので Lebesgue の優収束定理により $a \searrow 0$ とすれば $\log \int_S a \vee \varphi d\lambda \to \log \int_S \varphi d\lambda = 0$ となり，$I(\nu) \leq p^*(\nu)$ が示された．

これまで $I \leq p^* = I^{**}$ を示した．$I = p^*$ を得るには $I^{**} \leq I$ を示せばよい．任意の $\mu \in \mathcal{M}_1(S)$ と $f \in M_{\mathrm{b}}(S)$ に対して

$$I^*(f) \geq \int_S f d\mu - I(\mu)$$

が成立するので，

$$I^{**}(\mu) = \sup_{f \in M_{\mathrm{b}}(S)} \left\{ \int_S f d\mu - I^*(f) \right\} \leq I(\mu)$$

となり $I^{**} \leq I$ も得られた．以上により $I = p^*$ が示された．　□

単調族定理を用いた議論により，定理 4.1.2 において上限を取る範囲を $C_{\mathrm{b}}(S)$ に替えることができる．その形の変分公式も大偏差原理の理論において頻出するので，定理として述べておく．

定理 4.1.3. 任意の $\mu \in \mathcal{M}_1(S)$ に対して次が成立する．

$$H(\mu|\lambda) = \sup_{f \in C_{\mathrm{b}}(S)} \left\{ \int_S f d\mu - \log \int_S e^f d\lambda \right\}.$$

証明. 右辺を α とおく，つまり

$$\alpha := \sup_{f \in C_{\mathrm{b}}(S)} \left\{ \int_S f d\mu - \log \int_S e^f d\lambda \right\}$$

である．$M_{\mathrm{b}}(S)$ の部分集合 N_α を

$$N_\alpha := \left\{ f \in M_{\mathrm{b}}(S) \,\middle|\, \int_S f d\mu - \log \int_S e^f d\lambda \leq \alpha \right\}$$

と定めると，$N_\alpha = M_{\mathrm{b}}(S)$ を示せば定理の証明が終わる．まず Lebesgue の収束定理により次の条件が成立することに注意する．

(♣) $f \in M_{\mathrm{b}}(S)$ がある一様有界な関数列 $\{f_n\}_{n\in\mathbb{N}} \subset N_\alpha$ の各点収束極限になっていれば，$f \in N_\alpha$ である．

条件 **(♣)** により，単関数の線型結合が N_α の元であることを示せば，$N_\alpha = M_{\mathrm{b}}(S)$ が示される．以下単関数の線型結合が N_α の元であることを示す．

条件「任意の $f \in C_{\mathrm{b}}(S)$ に対して $g + f \in N_\alpha$」を満たす $g \in N_\alpha$ を任意に選び固定する．$C_{\mathrm{b}}(S) \subset N_\alpha$ なので，定数関数 $g \equiv 0$ は条件を満たす．S の Borel 集合族 $\mathscr{B}(S)$ の部分集合を次で定義する．

$$\mathscr{C}_g := \{A \in \mathscr{B}(S) \mid \text{任意の } a \in \mathbb{R}, f \in C_{\mathrm{b}}(S) \text{ に対して } g + a\mathbf{1}_A + f \in N_\alpha\}.$$

$\{A_n\}_{n\in\mathbb{N}} \subset \mathscr{C}_g$ を単調増大な集合列とすると，任意の $a \in \mathbb{R}, f \in C_{\mathrm{b}}(S), n \in \mathbb{N}$ に対して $g + a\mathbf{1}_{A_n} + f \in N_\alpha$ なので，条件 **(♣)** により $g + a\mathbf{1}_{\cup_{n\in\mathbb{N}} A_n} + f \in N_\alpha$ である．単調減少な集合列の場合も同様である．よって $\mathscr{C}_g$ は単調族である．次に有限加法族 $\mathscr{A}$ を

$$\mathscr{A} := \{A \in \mathscr{B}(S) \mid \text{ある一様有界な } \{f_n\}_{n\in\mathbb{N}} \subset C_{\mathrm{b}}(S) \text{ に対して各点で } \mathbf{1}_A = \lim_{n\to\infty} f_n\}$$

とおくと，再び条件 **(♣)** により $\mathscr{A} \subset \mathscr{C}_g$ がわかる．さらに，$C \subset S$ を閉集合とすると，

$$\mathbf{1}_C(x) = \lim_{n\to\infty} \frac{1}{1 + n d_S(x, C)}, \quad x \in S$$

となるので，$C \in \mathscr{A}$ である．ただし，d_S は S の位相を誘導する距離である．よって単調族定理により $\mathscr{C}_g = \mathscr{B}(S)$ を得る．

g を定数関数 0 と取ると $\mathscr{C}_0 = \mathscr{B}(S)$ となるので，任意の $a \in \mathbb{R}, A \in \mathscr{B}(S), f \in C_{\mathrm{b}}(S)$ に対して $a\mathbf{1}_A + f \in N_\alpha$ であることがわかる．次に $g = a\mathbf{1}_A$ として再び $\mathscr{C}_g = \mathscr{B}(S)$ を用いれば，任意の $a, b \in \mathbb{R}, A, B \in \mathscr{B}(S), f \in C_{\mathrm{b}}(S)$ に対して $a\mathbf{1}_A + b\mathbf{1}_B + f \in N_\alpha$ を得る．これを繰り返すことにより，単関数の線型結合は N_α の元であることが示されるので，定理の証明が終わる． □

任意の $f \in C_{\mathrm{b}}(S)$ に対して，$\mu \mapsto \int_S f d\mu - \log \int_S e^f d\lambda$ は連続かつ凸である．よって補題 1.2.5 と定理 4.1.3 により，$H(\cdot|\lambda)$ は下半連続かつ凸である．さらに，相対エントロピーは良い速度関数であることを次の命題で示す．

命題 4.1.4. $H(\cdot|\lambda) : \mathcal{M}_1(S) \to [0, \infty]$ は良い速度関数である．すなわち，任意の $A \in [0, \infty)$ に対して $\{\mu \in \mathcal{M}_1(S) \mid H(\mu|\lambda) \leq A\}$ はコンパクト集合である．

証明. $A \geq 0$ を固定する．$H(\cdot|\lambda)$ は下半連続なので，$\{\mu \in \mathcal{M}_1(S) \mid H(\mu|\lambda) \leq A\}$ は閉集合である．よってこの集合がコンパクト集合であることを示すには，緊密であることを示せばよい．そのために，確率測度 $\mu \in \mathcal{M}_1(S)$ は $H(\mu|\lambda) \leq A$ を満たすとする．定理 4.1.2 により任意の有界可測関数 $f : S \to \mathbb{R}$ に対して

$$\int_S f d\mu - \log \int_S e^f d\lambda \leq H(\mu|\lambda) \leq A$$

が成立するので，

$$\int_S f d\mu \leq A + \log \int_S e^f d\lambda \tag{4.1.1}$$

を得る．$\varepsilon > 0$ を任意に固定して，コンパクト集合 $K_\varepsilon \subset S$ を $\lambda(K_\varepsilon^c) \leq \varepsilon$ が成立するように選んでおく．(4.1.1) において $f = \log(1 + 1/\varepsilon)\mathbf{1}_{K_\varepsilon^c}$ とすると

$$\begin{aligned}\mu_n(K_\varepsilon^c) &\leq \frac{1}{\log(1+1/\varepsilon)} \left(A + \log \left[\lambda(K_\varepsilon) + \left(1 + \frac{1}{\varepsilon}\right) \lambda(K_\varepsilon^c) \right] \right) \\ &\leq \frac{1}{\log(1+1/\varepsilon)} (A + \log 2)\end{aligned}$$

を得る．$\varepsilon \searrow 0$ とすれば，$\{\mu \in \mathcal{M}_1(S) \mid H(\mu|\lambda) \leq A\}$ が緊密であることがわかる． □

4.2 Sanov の定理の証明

本節において定理 4.0.1 を証明する．証明中では $I(\cdot) := H(\cdot|\lambda)$ と書くことにし，S の部分集合はローマン体，$\mathcal{M}_1(S)$ の部分集合は筆記体により表すことにする．また本節では積分 $\int_S f dL_n$ を $\langle L_n, f \rangle$ と書く．

定理 4.0.1 の上からの評価の証明． まずコンパクト集合に対する上からの評価を示す．$\mathcal{C} \subset \mathcal{M}_1(S)$ をコンパクト集合とし，$\alpha < \inf_{\mu\in\mathcal{C}} I(\mu)$ を固定する．定理 4.1.3 により任意の $\mu \in \mathcal{C}$ に対して

$$\int_S f_\mu d\mu - \log \int_S e^{f_\mu} d\lambda > \alpha$$

を満たす $f_\mu \in C_{\mathrm{b}}(S)$ が存在する．また

$$\mathcal{O}(\mu,\delta) = \left\{ \nu \in \mathcal{M}_1(S) \;\middle|\; \left| \int_S f_\mu d\nu - \int_S f_\mu d\mu \right| < \delta \right\}$$

とおくと，弱位相の定義により $\mathcal{O}(\mu,\delta)$ は μ の開近傍である．$\mathcal{C}$ はコンパクトなので $\mathcal{C}$ の開被覆 $\{\mathcal{O}(\mu,\delta)\}_{\mu\in\mathcal{C}}$ から有限開被覆 $\{\mathcal{O}(\mu_1,\delta),\ldots,\mathcal{O}(\mu_N,\delta)\}$ が取れる．

一方，任意の $\mu \in \mathcal{M}_1(S)$ に対して

$$\begin{aligned}\mathbb{P}\left(L_n \in \mathcal{O}(\mu,\delta)\right) &= \mathbb{E}\left[\mathbf{1}_{\{L_n\in\mathcal{O}(\mu,\delta)\}} \exp\left\{n\langle L_n, f_\mu\rangle - n\langle L_n, f_\mu\rangle\right\}\right] \\ &\le \exp\left\{-n\left(\int_S f_\mu d\mu - \delta\right)\right\} \mathbb{E}\left[\exp\left\{n\langle L_n, f_\mu\rangle\right\}\right]\end{aligned}$$

が成立する．また任意の $f \in C_{\mathrm{b}}(S)$ に対して

$$\mathbb{E}\left[\exp\left\{n\langle L_n, f\rangle\right\}\right] = \mathbb{E}\left[\exp\left\{\sum_{k=1}^n f(X_k)\right\}\right] = \left(\int_S e^f d\lambda\right)^n = \exp\left\{n \log \int_S e^f d\lambda\right\}$$

が成立する．これらを用いれば

$$\begin{aligned}\mathbb{P}\left(L_n \in \mathcal{C}\right) &\le \sum_{i=1}^N \mathbb{P}\left(L_n \in \mathcal{O}(\mu_i,\delta)\right) \\ &\le \sum_{i=1}^N \exp\left\{-n\left(\int_S f_{\mu_i} d\mu_i - \log\left(\int_S e^{f_{\mu_i}} d\lambda\right) - \delta\right)\right\} \\ &\le N e^{-n(\alpha-\delta)}\end{aligned}$$

が成立する．よって順番に $n\to\infty, \delta \searrow 0, \alpha \nearrow \inf_{\mu\in\mathcal{C}} I(\mu)$ とすれば，

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(L_n \in \mathcal{C}\right) \le -\inf_{\mu\in\mathcal{C}} I(\mu)$$

となり，$\mathcal{C}$ がコンパクト集合の場合に大偏差原理の上からの評価が得られる．

上からの評価を閉集合に拡張するために $\{L_n\}_{n\in\mathbb{N}}$ の指数的緊密性を示す．各 $r \in \mathbb{N}$ に対してコンパクト集合 $A_r \subset S$ を $\lambda(A_r^c) \le e^{-2r^2}$ となるように取る．ここで $\mathcal{M}_1(S)$ の部

分集合を $\mathcal{A}_r = \{\nu \in \mathcal{M}_1(S) \mid \nu(A_r) \geq 1 - 1/r\}$ とおくと，$\mathcal{A}_r$ は $\mathcal{M}_1(S)$ において閉集合である．一方，$\{L_n \in \mathcal{A}_r^c\} = \{L_n(A_r^c) > 1/r\} = \{\sum_{k=1}^n \mathbf{1}_{\{X_k \in A_r^c\}} > n/r\}$ であるから，

$$\begin{aligned}
\mathbb{P}\left(L_n \in \mathcal{A}_r^c\right) &= \mathbb{P}\left(L_n(A_r^c) > \frac{1}{r}\right) \\
&= \mathbb{P}\left(\sum_{k=1}^n \mathbf{1}_{\{X_k \in A_r^c\}} > \frac{n}{r}\right) \\
&\leq e^{-2nr}\mathbb{E}\left[\exp\left\{2r^2 \sum_{k=1}^n \mathbf{1}_{\{X_k \in A_r^c\}}\right\}\right] \\
&= e^{-2nr}\left(e^{2r^2}\mathbb{P}\left(X_1 \in A_r^c\right) + \mathbb{P}\left(X_1 \in A_r\right)\right)^n \\
&\leq e^{-2nr}2^n \leq e^{-nr}
\end{aligned}$$

を得る．ただし，3 行目では Chebyshev の不等式を，4 行目では $\{X_k\}_{k\in\mathbb{N}}$ の独立同分布性を，5 行目では $\mathbb{P}\left(X_1 \in A_r^c\right) = \lambda(A_r^c) \leq e^{-2r^2}$ をそれぞれ用いた．ここで各 $R \in \mathbb{N}$ に対して $\mathcal{K}_R := \cap_{r\geq R}\mathcal{A}_r$ とおくと，$\mathcal{K}_R$ は $\mathcal{M}_1(S)$ において緊密かつ閉なので，コンパクトである．

$$\mathbb{P}\left(L_n \in \mathcal{K}_R^c\right) \leq \sum_{r\geq R}\mathbb{P}\left(L_n \in \mathcal{A}_r^c\right) \leq \sum_{r\geq R} e^{-nr} \leq 2e^{-nR}$$

となり $\{L_n\}_{n\in\mathbb{N}}$ の指数的緊密性が示されたので，一般の閉集合に対する大偏差原理の上からの評価が得られた． □

定理 4.0.1 の下からの評価の証明． $\mathcal{O} \subset \mathcal{M}_1(S)$ を任意の開集合とする．下からの評価を示すためには，$I(\mu) < \infty$ となる任意の $\mu \in \mathcal{O}$ に対して

$$\varliminf_{n\to\infty} \frac{1}{n}\log\mathbb{P}\left(L_n \in \mathcal{O}\right) \geq -I(\mu) \tag{4.2.1}$$

を示せば十分である．そのために $I(\mu) < \infty$ なる $\mu \in \mathcal{O}$ を固定しておく．

$I(\mu) < \infty$ なので μ は λ に対して絶対連続である．その Radon-Nikodym 微分を $\varphi = d\mu/d\lambda$ と書くことにする．$(\Omega, \mathcal{F}, \mathbb{P})$ を確率変数 $\{X_k\}_{k\in\mathbb{N}}$ が定義されている確率空間とする．$(\Omega, \mathcal{F})$ 上の確率測度 $\mathbb{P}_{\mu,n}$ を，$\mathbb{P}$ に対して絶対連続であってその Radon-Nikodym 微分が

$$\frac{d\mathbb{P}_{\mu,n}}{d\mathbb{P}} = \prod_{k=1}^n \varphi(X_k)$$

で与えられるものとする．（φ は λ についてほとんど確実にしか定義されないが，X_k の $\mathbb{P}$ の下での分布が λ なので右辺は問題なく定義される．）この $d\mathbb{P}_{\mu,n}/d\mathbb{P}$ を Φ_n と書く．$\mathbb{P}_{\mu,n}$ に関する期待値は $\mathbb{E}_{\mu,n}$ で書くことにする．

$\mathbb{P}_{\mu,n}$ の下で S 値確率変数の列 $\{X_k\}_{k=1}^n$ は独立同分布であり，X_1 の分布は n に依らず μ である．よって確率変数の列 $\{L_n\}_{n\in\mathbb{N}}$ に対する大数の弱法則により

$$\lim_{n\to\infty} \mathbb{P}_{\mu,n}\left(L_n \in \mathcal{O}\right) = 1 \tag{4.2.2}$$

が成立する．特に n が十分大きいとき $\mathbb{P}_{\mu,n}\left(L_n \in \mathcal{O}\right) > 0$ となる．また φ の定義と $\mathbb{P}_{\mu,n}$ の下で X_k の分布が μ であることから，$\mathbb{P}_{\mu,n}(\Phi_n > 0) = 1$ である．よって

$$\begin{aligned}\frac{1}{n}\log\mathbb{P}\left(L_n \in \mathcal{O}\right) &\geq \frac{1}{n}\log\mathbb{E}\left[\mathbf{1}_{\{L_n\in\mathcal{O}\}}\mathbf{1}_{\{\Phi_n>0\}}\right] = \frac{1}{n}\log\mathbb{E}_{\mu,n}\left[\mathbf{1}_{\{L_n\in\mathcal{O}\}}\Phi_n^{-1}\right] \\ &= \frac{1}{n}\log\left[\frac{\mathbb{E}_{\mu,n}\left[\mathbf{1}_{\{L_n\in\mathcal{O}\}}\Phi_n^{-1}\right]}{\mathbb{P}_{\mu,n}\left(L_n\in\mathcal{O}\right)}\right] + \frac{1}{n}\log\mathbb{P}_{\mu,n}\left(L_n\in\mathcal{O}\right)\end{aligned}$$

と評価できる．よって Jensen の不等式により

$$\frac{1}{n}\log\mathbb{P}\left(L_n \in \mathcal{O}\right) \geq -\frac{\mathbb{E}_{\mu,n}\left[\mathbf{1}_{\{L_n\in\mathcal{O}\}}\log\Phi_n\right]}{n\mathbb{P}_{\mu,n}\left(L_n\in\mathcal{O}\right)} + \frac{1}{n}\log\mathbb{P}_{\mu,n}\left(L_n\in\mathcal{O}\right) \tag{4.2.3}$$

を得る．($\log\Phi_n$ の扱いについては注意 4.0.2 と同様である．)

(4.2.3) の右辺の項をそれぞれ考える．まず

$$\begin{aligned}\mathbb{E}_{\mu,n}\left[\mathbf{1}_{\{L_n\in\mathcal{O}\}}\log\Phi_n\right] &= \mathbb{E}_{\mu,n}\left[\log\Phi_n\right] - \mathbb{E}_{\mu,n}\left[\mathbf{1}_{\{L_n\in\mathcal{O}^c\}}\log\Phi_n\right] \\ &= \mathbb{E}\left[\Phi_n\log\Phi_n\right] - \mathbb{E}\left[\mathbf{1}_{\{L_n\in\mathcal{O}^c\}}\Phi_n\log\Phi_n\right]\end{aligned}$$

である．また相対エントロピーの定義により $\mathbb{E}\left[\Phi_n\log\Phi_n\right] = nH(\mu|\lambda) = nI(\mu)$ である．一方，任意の $u \geq 0$ に対して $u\log u \geq -1/e$ であることを用いることにより，(4.2.3) の右辺は

$$\frac{1}{n}\log\mathbb{P}\left(L_n\in\mathcal{O}\right) \geq \frac{-1}{\mathbb{P}_{\mu,n}\left(L_n\in\mathcal{O}\right)}\left\{I(\mu) + \frac{1}{ne}\right\} + \frac{1}{n}\log\mathbb{P}_{\mu,n}\left(L_n\in\mathcal{O}\right)$$

と評価される．(4.2.2) により (4.2.1) が示されるので，大偏差原理の下からの評価の証明が終了する． □

第5章　弱収束法によるCramérの定理とSanovの定理の証明

本章では弱収束法を用いて Cramér の定理と Sanov の定理を示す.[1] S を可分完備距離空間とし, $C_{\mathrm{b}}(S)$ を S 上の実数値有界連続関数全体, $\mathcal{M}_1(S)$ を S 上の Borel 確率測度全体とする. 第 4 章で見たように, S は可分なので有界連続関数の列 $\{f_n\}_{n\in\mathbb{N}} \subset C_{\mathrm{b}}(S)$ で $\mathcal{M}_1(S)$ を分離するものが存在する. $\mathcal{M}_1(S)$ 上の距離を

$$d_{\mathcal{M}_1(S)}(\mu,\nu) := \sum_{n=1}^{\infty} \frac{1}{2^n} \left\{ 1 \wedge \left| \int_S f_n \, d\mu - \int_S f_n \, d\nu \right| \right\}, \quad \mu, \nu \in \mathcal{M}_1(S)$$

により定義すると, $d_{\mathcal{M}_1(S)}$ は弱収束の位相を誘導して, $\mathcal{M}_1(S)$ は可分完備距離空間になる.

$\{X_i\}_{i\in\mathbb{N}}$ を独立同分布な S 値確率変数の列とし, X_1 の分布を $\lambda \in \mathcal{M}_1(S)$ とする. また L^n を経験測度とする, すなわち

$$L^n := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$$

とする.[2] 次に相対エントロピーの定義を思い出す. $\mu, \nu \in \mathcal{M}_1(S)$ に対して, μ の ν に関する相対エントロピー $H(\mu|\nu)$ は次で定義される.

$$H(\mu|\nu) := \begin{cases} \displaystyle\int_S \frac{d\mu}{d\nu} \log\left(\frac{d\mu}{d\nu}\right) d\nu, & \mu \text{ が } \nu \text{ に対して絶対連続なとき}, \\ \infty, & \text{それ以外のとき}. \end{cases}$$

命題 4.1.4 で示したように $H(\cdot|\lambda)$ は良い速度関数である.

次の定理（Sanov の大偏差原理）は第 4 章で示したが, 本章では弱収束法による別証明を与える.

定理 5.0.1. $\mathcal{M}_1(S)$ 値確率変数の列 $\{L^n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数 $H(\cdot|\lambda)$ に対して速度 n で大偏差原理を満たす.

[1] 本章の記述は [29, 第 3.1 節] と [34, 第 C 章] を参考にした.

[2] 第 4 章では経験測度を L_n と書いたが, 本章では記法の都合上 L^n と書く.

次に Cramér の大偏差原理を述べる．記号の濫用であるが，本章において Cramér の定理に関する命題を考えるときには $S = \mathbb{R}^d$ $(d \in \mathbb{N})$ として同じ記号を用いることにする．つまり，$\{X_i\}_{i\in\mathbb{N}}$ は独立同分布な $\mathbb{R}^d$ 値確率変数の列であって，X_1 の分布を $\lambda \in \mathcal{M}_1(\mathbb{R}^d)$ と書く．本章においても第 2 章の冒頭で導入した条件 **(FEM)** を仮定する．各 $n \in \mathbb{N}$ に対して Z^n を経験平均とする．

$$Z^n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

定理 5.0.1 の証明と同様に，本章では次の Cramér の定理も弱収束法により証明する．

定理 5.0.2. 条件 **(FEM)** を仮定する．このとき，$\mathbb{R}^d$ 値確率変数の列 $\{Z^n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のときに良い速度関数

$$I(x) := \inf\left\{ H(\mu|\lambda) \;\middle|\; \mu \in \mathcal{M}_1(\mathbb{R}^d), \int_{\mathbb{R}^d} |y|\mu(dy) < \infty, \int_{\mathbb{R}^d} y\mu(dy) = x \right\}, \quad x \in \mathbb{R}^d$$

に対して速度 n で大偏差原理を満たす.

注意 5.0.3. 速度関数の一意性（命題1.2.11）により，$I(x)$ は積率母関数 $\Lambda(\theta) := \log \mathbb{E}\left[e^{\langle \theta, X_1\rangle}\right]$ $(\theta \in \mathbb{R}^d)$ の Fenchel-Legendre 変換

$$\Lambda^*(x) := \sup_{\theta\in\mathbb{R}^d}\{\langle\theta, x\rangle - \Lambda(\theta)\}, \quad x \in \mathbb{R}^d$$

に一致することが分かる．しかしながら速度関数の一意性を用いずとも，仮定 **(SLG)** の下では以下で見るように直接示すことができる．そのために任意の $x \in \mathbb{R}^d$ に対して $I(x) = \Lambda^*(x)$ を示す．定理 2.0.1 の証明中に示したように，仮定 **(SLG)** の下ではある $\theta_x \in \mathbb{R}^d$ が存在して

$$\Lambda^*(x) = \langle\theta_x, x\rangle - \Lambda(\theta_x)$$

が成立することに注意する．

まず $I(x) \le \Lambda^*(x)$ を示す．$\mathbb{R}^d$ 上の確率測度 μ_* を，λ に対して絶対連続でその Radon-Nikodym 微分が

$$\frac{d\mu_*}{d\lambda}(y) = \frac{e^{\langle\theta_x, y\rangle}}{\mathbb{E}\left[e^{\langle\theta_x, X_1\rangle}\right]}, \quad y \in \mathbb{R}^d$$

であるものとする．仮定 **(FEM)** により右辺の期待値は有限なので μ_* は確かに確率測度である．このとき

$$\int_{\mathbb{R}^d} y\mu_*(dy) = x, \quad H(\mu_*|\lambda) = \langle\theta_x, x\rangle - \Lambda(\theta_x)$$

が成立することがわかる．実際，１つ目の式は定理 2.3.1 の証明中に示しており（可積分性は **(FEM)** から従う），２つ目の式は１つ目の式と相対エントロピーの定義により簡単に示される．よって $I(x) \le \Lambda^*(x)$ である．

次に逆向きの不等式 $I(x) \ge \Lambda^*(x)$ を示す．$I(x) = \infty$ のときは自明に $I(x) \ge \Lambda^*(x)$ が成立するので，$I(x) < \infty$ のときに示す．$\mathbb{R}^d$ 上の Borel 確率測度 μ であって

$$\int_{\mathbb{R}^d} |y| \mu(dy) < \infty, \quad \int_{\mathbb{R}^d} y \mu(dy) = x \tag{5.0.1}$$

を満たすものを任意に固定する．$n \in \mathbb{N}$ と $\theta \in \mathbb{R}^d$ に対して $f_n(y) := [\langle \theta, y \rangle \vee (-n)] \wedge n$ とおくと，f_n は有界連続なので定理 4.1.3 により

$$H(\mu|\lambda) \ge \int_{\mathbb{R}^d} f_n \, d\mu - \log \int_{\mathbb{R}^d} e^{f_n} \, d\lambda$$

が成立する．仮定 **(FEM)** と (5.0.1) および Lebesgue の収束定理により，$n \to \infty$ とすれば

$$H(\mu|\lambda) \ge \langle \theta, x \rangle - \Lambda(\theta)$$

を得る．$\theta \in \mathbb{R}^d$ について上限を取ることにより $H(\mu|\lambda) \ge \Lambda^*(x)$ が成立する．μ の任意性により $I(x) \ge \Lambda^*(x)$ を得る．以上で本注意を終わる．

本章の主題に入る前に，弱収束する確率変数の列に対する Fatou の補題を示す．これは通常の Fatou の補題と Skorokhod の表現定理から直ちに従うが，弱収束法による大偏差原理の証明では有用であるため，独立した補題として述べる．そのために Skorokhod の表現定理を思い出す.[3] S を可分完備距離空間とし，S 値確率変数の列 $\{X_n\}_{n \in \mathbb{N}}$ は $n \to \infty$ のとき S 値確率変数 X_∞ に弱収束すると仮定する.（これらの確率変数は同一の確率空間上に定義されている必要はない.）このとき，ある確率空間 $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ とこの上で定義された S 値確率変数の列 $\{\widehat{X}_n\}_{n \in \mathbb{N} \cup \{\infty\}}$ が存在して，次を満たす．

(i) 任意の $n \in \mathbb{N} \cup \{\infty\}$ に対して X_n と $\widehat{X}_n$ は同分布である．

(ii) S 値確率変数の列 $\{\widehat{X}_n\}_{n \in \mathbb{N}}$ は $n \to \infty$ のとき $\widehat{X}_\infty$ に概収束する．

$\widehat{\mathbb{P}}$ に対する期待値は $\widehat{\mathbb{E}}\,[\cdot]$ で表すことにする．弱収束する確率変数の列に対する Fatou の補題とは次の主張である．

補題 5.0.4. S を可分完備距離空間として，$f : S \to [0, \infty]$ を下半連続とする．また S 値確率変数の列 $\{X_n\}_{n \in \mathbb{N}}$ は $n \to \infty$ のとき S 値確率変数 X_∞ に弱収束すると仮定する．このとき次が成立する．

$$\mathbb{E}\,[f(X_\infty)] \le \varliminf_{n \to \infty} \mathbb{E}\,[f(X_n)].$$

[3] Skorokhod の表現定理の証明は [23, P. 163, 定理 1.7] を参照せよ．

証明. 確率空間 $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ と S 値確率変数 $\{\widehat{X}_n\}_{n\in\mathbb{N}\cup\{\infty\}}$ を Skorokhod の表現定理で与えられるものとする．(i) により任意の $n \in \mathbb{N}\cup\{\infty\}$ に対して $\mathbb{E}\,[f(X_n)] = \widehat{\mathbb{E}}\left[f(\widehat{X}_n)\right]$ が成立するので，

$$\widehat{\mathbb{E}}\left[f(\widehat{X}_\infty)\right] \leq \varliminf_{n\to\infty} \widehat{\mathbb{E}}\left[f(\widehat{X}_n)\right]$$

を示せば十分である．(ii) と f の下半連続性により確率 1 で

$$f(\widehat{X}_\infty) = f(\lim_{n\to\infty} \widehat{X}_n) \leq \varliminf_{n\to\infty} f(\widehat{X}_n)$$

が成立する．よって通常の Fatou の補題により

$$\widehat{\mathbb{E}}\left[\varliminf_{n\to\infty} f(\widehat{X}_n)\right] \leq \varliminf_{n\to\infty} \widehat{\mathbb{E}}\left[f(\widehat{X}_n)\right]$$

を得るので，証明が終わる． □

注意 5.0.5. 補題 5.0.4 における S は可分完備距離空間であればよいので，弱収束する $\mathcal{M}_1(S)$ 値確率変数の列に対して，補題 5.0.4 が適用できる．

5.1　弱収束法による証明の概説

本節では弱収束法による定理 5.0.1（Sanov の定理）の証明を概説する．弱収束法による証明では，大偏差原理と同値な命題である Laplace 原理を示す．Laplace 原理を示す上では，指数型の積率に対する変分公式と，相対エントロピーに対する連鎖律の公式が基本的な役割を果たす．本節ではこれらについて述べた後，定理 5.0.1 の証明を概説する．実際の証明は第 5.4 節において行い，定理 5.0.2（Cramér の定理）の証明は第 5.5 節において行う．

まず指数型の積率に対する変分公式を述べる．

補題 5.1.1. S を可分完備距離空間とし，$\theta \in \mathcal{M}_1(S)$ とする．任意の有界 Borel 可測関数 $f : S \to \mathbb{R}$ に対して次が成立する．

$$-\log\left(\int_S e^{-f} d\theta\right) = \inf_{\mu\in\mathcal{M}_1(S)} \left[\int_S f\, d\mu + H(\mu|\theta)\right].$$

証明. f は有界なので右辺の下限において $\mu = \theta$ と取ることにより，右辺の下限は有限であることがわかる．よって右辺の下限では $H(\mu|\theta) < \infty$ となる $\mu \in \mathcal{M}_1(S)$ のみを考えればよい．$H(\mu|\theta) < \infty$ となる $\mu \in \mathcal{M}_1(S)$ を固定する．このとき μ は θ に対して絶対連続である．また $\mu^* \in \mathcal{M}_1(S)$ を θ について絶対連続で，その Radon-Nikodym 微分が

$$\frac{d\mu^*}{d\theta} = \frac{e^{-f}}{\int_S e^{-f} d\theta} \in (0, \infty)$$

で与えられるものとする．f は有界なので θ は μ^* に対して絶対連続である．よって μ は μ^* に対して絶対連続であり

$$\log\frac{d\mu}{d\theta} = \log\frac{d\mu}{d\mu^*} + \log\frac{d\mu^*}{d\theta}$$

が成立する．この等式により

$$\begin{aligned}\int_S f\,d\mu + H(\mu|\theta) &= \int_S f\,d\mu + \int_S \log\left(\frac{d\mu}{d\theta}\right)d\mu \\ &= \int_S f\,d\mu + \int_S \log\left(\frac{d\mu}{d\mu^*}\right)d\mu + \int_S \log\left(\frac{d\mu^*}{d\theta}\right)d\mu \\ &= -\log\int_S e^{-f}d\theta + H(\mu|\mu^*)\end{aligned}$$

が成立する．命題 4.1.1 により $H(\mu|\mu^*) \geq 0$ であり，等号が成立するのは $\mu = \mu^*$ のときなので補題は示された． □

次の公式は相対エントロピーに対する連鎖律として知られている．また正則条件付き確率については命題 B.0.1 を参照せよ．

補題 5.1.2. S_1 と S_2 を可分完備距離空間とする．μ^2, θ^2 を $S_1 \times S_2$ 上の Borel 確率測度とし，μ_1^2 と θ_1^2 をそれぞれ μ^2 と θ^2 の S_1 への周辺分布（第１成分への標準射影が誘導する法則）とする．また μ^2 と θ^2 の第１成分への射影を与えたときの正則条件付き確率をそれぞれ

$$\{\mu^2_{2|1}(x_1, A_2) \mid x_1 \in S_1, A_2 \in \mathcal{B}(S_2)\}, \quad \{\theta^2_{2|1}(x_1, A_2) \mid x_1 \in S_1, A_2 \in \mathcal{B}(S_2)\}$$

とおく.[4] このとき次が成立する.[5]

$$H(\mu^2|\theta^2) = H(\mu_1^2|\theta_1^2) + \int_{S_1} H(\mu^2_{2|1}|\theta^2_{2|1})d\mu_1^2. \tag{5.1.1}$$

---

[4] 正確には正則条件付き確率は

$$\{\mu^2_{2|1}(x_1, \{x_1\}\times A_2) \mid x_1 \in S_1, \{x_1\}\times A_2 \in \mathcal{B}(\{x_1\}\times S_2)\}$$

と書くべきだが，$\{x_1\}\times S_2$ と S_2 は同相（特に可測空間として同型）なので，$\{x_1\}\times S_2$ 上の測度を S_2 上の測度と見なしている．

[5] (5.1.1) の右辺の積分は

$$\int_{S_1} H(\mu^2_{2|1}(x_1,\cdot)|\theta^2_{2|1}(x_1,\cdot))\mu_1^2(dx_1)$$

の意味である．

注意 5.1.3. (5.1.1) の右辺が正則条件付き確率の取り方によらないこと，すなわち

$$\{\widetilde{\mu}^2_{2|1}(x_1, A_2) \mid x_1 \in S_1, A_2 \in \mathcal{B}(S_2)\}, \quad \{\widetilde{\theta}^2_{2|1}(x_1, A_2) \mid x_1 \in S_1, A_2 \in \mathcal{B}(S_2)\}$$

もそれぞれ μ^2 と θ^2 の第1成分への射影を与えたときの正則条件付き確率とすると，

$$H(\mu^2_1|\theta^2_1) + \int_{S_1} H(\mu^2_{2|1}|\theta^2_{2|1})d\mu^2_1 = H(\mu^2_1|\theta^2_1) + \int_{S_1} H(\widetilde{\mu}^2_{2|1}|\widetilde{\theta}^2_{2|1})d\mu^2_1 \tag{5.1.2}$$

が成立することを確認する．

$H(\mu^2_1|\theta^2_1) = \infty$ のときは明らかに (5.1.2) の両辺ともに ∞ である．次に $H(\mu^2_1|\theta^2_1) < \infty$ とする．このとき μ^2_1 は θ^2_1 に対して絶対連続であり，正則条件付き確率の一意性により

$$\theta^2_1 \left(x_1 \in S_1 \;\middle|\; \theta^2_{2|1}(x_1, \cdot) \neq \widetilde{\theta}^2_{2|1}(x_1, \cdot) \right) = 0$$

が成立するので，

$$\mu^2_1 \left(x_1 \in S_1 \;\middle|\; \theta^2_{2|1}(x_1, \cdot) \neq \widetilde{\theta}^2_{2|1}(x_1, \cdot) \right) = 0$$

を得る．再び正則条件付き確率の一意性により

$$\mu^2_1 \left(x_1 \in S_1 \mid \mu^2_{2|1}(x_1, \cdot) \neq \widetilde{\mu}^2_{2|1}(x_1, \cdot) \right) = 0$$

が成立するので，(5.1.2) を得る．よって (5.1.1) の右辺は正則条件付き確率の取り方によらない．

補題 5.1.2 の証明. まず (5.1.1) の左辺が有限のときに示す．相対エントロピーの定義により μ^2 は θ^2 に対して絶対連続である．このとき μ^2_1 は θ^2_1 に対して絶対連続である．Radon-Nikodym 微分をそれぞれ

$$\varphi := \frac{d\mu^2}{d\theta^2}, \quad \psi := \frac{d\mu^2_1}{d\theta^2_1}$$

と書く．ただしここでは各点で有限な修正を一つ選び固定する．Radon-Nikodym 微分と正則条件付き確率の定義により，任意の $A_1 \in \mathcal{B}(S_1)$ と $A_2 \in \mathcal{B}(S_2)$ に対して

$$\begin{aligned}
\int_{A_1} \mu^2_{2|1}(x_1, A_2)\psi(x_1)\theta^2_1(dx_1) &= \int_{A_1} \mu^2_{2|1}(x_1, A_2)\mu^2_1(dx_1) \\
&= \mu^2(A_1 \times A_2) \\
&= \int_{A_1 \times A_2} \varphi(x_1, x_2)\theta^2(dx_1 dx_2) \\
&= \int_{A_1} \left(\int_{A_2} \varphi(x_1, x_2)\theta^2_{2|1}(x_1, dx_2) \right) \theta^2_1(dx_1)
\end{aligned}$$

が成立する．よって任意の $A_2 \in \mathcal{B}(S_2)$ に対して，ある θ_1^2 零集合 $\Gamma_1(A_2) \in \mathcal{B}(S_1)$ が存在して任意の $x_1 \in \Gamma_1(A_2)^c$ に対して

$$\mu_{2|1}^2(x_1, A_2)\psi(x_1) = \int_{A_2} \varphi(x_1, x_2)\theta_{2|1}^2(x_1, dx_2) \tag{5.1.3}$$

が成立する．

実はある θ_1^2 零集合 $\Gamma_1 \in \mathcal{B}(S_1)$ が存在して，任意の $x_1 \in \Gamma_1^c$ と任意の $A_2 \in \mathcal{B}(S_2)$ に対して (5.1.3) が成立することを次に示す．S_2 は可分な距離空間なので，可算開基 $\mathcal{U} = \{U_i\}_{i\in\mathbb{N}}$ かつ π 系になるものが存在する．実際，$\mathcal{V} = \{V_i\}_{i\in\mathbb{N}}$ を S_2 の可算開基とすると

$$\left\{ \bigcap_{k=1}^{m} V_{i_k} \;\middle|\; m \in \mathbb{N}, i_1, \ldots, i_m \in \mathbb{N} \right\}$$

は S_2 の可算開基かつ π 系である．まず $A_2 = S_2$ として，θ_1^2 零集合 $\Gamma_1(S_2) \in \mathcal{B}(S_1)$ で任意の $x_1 \in \Gamma_1(S_2)^c$ に対して (5.1.3) が成立するものを取る．次に $A_2 = U_i$ $(i \in \mathbb{N})$ として，θ_1^2 零集合 $\Gamma_1(U_i) \in \mathcal{B}(S_1)$ で任意の $x_1 \in \Gamma_1(U_i)^c$ に対して (5.1.3) が成立するものを取り，$\Gamma_1 \in \mathcal{B}(S_1)$ を

$$\Gamma_1 := \Gamma_1(S_2) \cup \left(\bigcup_{i\in\mathbb{N}} \Gamma_1(U_i) \right)$$

と定義する．このとき Γ_1 は θ_1^2 零集合であり，π-λ 定理（Dynkin 族定理）により任意の $x_1 \in \Gamma_1^c$ と任意の $A_2 \in \mathcal{B}(S_2)$ に対して (5.1.3) が成立する．

特に任意の $x_1 \in \Gamma_1^c \cap \{\psi > 0\}$ に対して $\mu_{2|1}^2(x_1, \cdot)$ は $\theta_{2|1}^2(x_1, \cdot)$ に対して絶対連続であり，$\theta_{2|1}^2(x_1, \cdot)$ についてほとんど全ての $x_2 \in S_2$ に対して

$$\zeta(x_1, x_2) := \frac{d\mu_{2|1}^2(x_1, \cdot)}{d\theta_{2|1}^2(x_1, \cdot)}(x_2) = \frac{\varphi(x_1, x_2)}{\psi(x_1)}$$

が成立する．$\mu_{2|1}^2(x_1, \cdot)$ は $\theta_{2|1}^2(x_1, \cdot)$ に対して絶対連続なので，上式は $\mu_{2|1}^2(x_1, \cdot)$ についてほとんど全ての $x_2 \in S_2$ に対しても成立する．つまり $x_1 \in \Gamma_1^c \cap \{\psi > 0\}$ ならば，$\mu_{2|1}^2(x_1, \cdot)$ についてほとんど至る所 $\varphi = \psi\zeta$ が成立する．一方 μ_1^2 は θ_1^2 に対して絶対連続なので，Γ_1 は μ_1^2 零集合である．さらに明らかに $\{\psi = 0\}$ も μ_1^2 零集合なので，$\mu_1^2(\Gamma_1^c \cap \{\psi > 0\}) = 1$ である．以上により

$$\begin{aligned}
H(\mu^2|\theta^2) &= \int_{S_1\times S_2} \log\varphi \; d\mu^2 \\
&= \int_{S_1} \left(\int_{S_2} \log\varphi(x_1, x_2)\mu_{2|1}^2(x_1, dx_2) \right) \mu_1^2(dx_1) \\
&= \int_{\Gamma_1^c\cap\{\psi>0\}} \left(\int_{S_2} \log\left(\psi(x_1)\zeta(x_1, x_2)\right) \mu_{2|1}^2(x_1, dx_2) \right) \mu_1^2(dx_1)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\Gamma_1^c \cap \{\psi>0\}} \log\psi \; d\mu_1^2 + \int_{\Gamma_1^c \cap \{\psi>0\}} \left(\int_{S_2} \log\zeta(x_1,x_2)\mu_{2|1}^2(x_1,dx_2) \right) \mu_1^2(dx_1) \\
&= H(\mu_1^2|\theta_1^2) + \int_{S_1} H(\mu_{2|1}^2|\theta_{2|1}^2) d\mu_1^2
\end{aligned}
$$

を得る．

次に (5.1.1) の右辺が有限のときに示す．このとき μ_1^2 は θ_1^2 に対して絶対連続なので，その Radon-Nikodym 微分を ψ と書く．またある μ_1^2 零集合 $\Gamma_1 \in \mathcal{B}(S_1)$ が存在して，任意の $x_1 \in \Gamma_1^c$ に対して $H(\mu_{2|1}^2(x_1,\cdot)|\theta_{2|1}^2(x_1,\cdot))$ は有限である．特に任意の $x_1 \in \Gamma_1^c$ に対して $\mu_{2|1}^2(x_1,\cdot)$ は $\theta_{2|1}^2(x_1,\cdot)$ に対して絶対連続である．$x_1 \in \Gamma_1$ に対して $\mu_{2|1}^2(x_1,\cdot)$ を定義し直すことにより，任意の $x_1 \in S_1$ に対して $\mu_{2|1}^2(x_1,\cdot)$ は $\theta_{2|1}^2(x_1,\cdot)$ に対して絶対連続になる．ここである Borel 可測関数 $\zeta : S_1 \times S_2 \to [0,\infty)$ が存在して

$$
\theta_{2|1}^2 \left(x_1, \left\{ x_2 \in S_2 \;\middle|\; \zeta(x_1,x_2) = \frac{d\mu_{2|1}^2(x_1,\cdot)}{d\theta_{2|1}^2(x_1,\cdot)}(x_2) \right\} \right) = 1, \quad x_1 \in S_1 \tag{5.1.4}
$$

が成立することを次に示す．

証明の前半で取った様に，S_2 の可算開基 $\mathcal{U} = \{U_i\}_{i\in\mathbb{N}}$ かつ π 系になるものを一つ選ぶ．各 $n \in \mathbb{N}$ に対して，$U_1,\dots,U_n$ により生成される S_2 上の σ 加法族を $\mathcal{G}_n$ と書く．また $i \in \mathbb{N}$ に対して $U_i^1 = U_i$ と $U_i^2 = U_i^c$ とおき，S_2 の有限分割 $\mathcal{P}_n$ を

$$
\mathcal{P}_n := \left\{ \bigcap_{i=1}^n U_i^{s_i} \;\middle|\; s_i \in \{1,2\}, i = 1,\dots,n \right\}
$$

と定義する．このとき明らかに $\mathcal{P}_{n+1}$ は $\mathcal{P}_n$ の細分であり，$\sigma(\mathcal{P}_n) = \mathcal{G}_n$ が成立する．ここで $n \in \mathbb{N}$ に対して

$$
\zeta_n(x_1,x_2) := \sum_{A\in\mathcal{P}_n} \frac{\mu_{2|1}^2(x_1,A)}{\theta_{2|1}^2(x_1,A)} \mathbf{1}_A(x_2), \quad (x_1,x_2) \in S_1 \times S_2
$$

と定義する．ただし $0/0 := 0$ とする．このとき

$$
\zeta(x_1,x_2) := \begin{cases} \lim_{n\to\infty} \zeta_n(x_1,x_2), & \text{右辺が収束するとき}, \\ 0, & \text{それ以外のとき}, \end{cases}
$$

と定義すると，これが所望の関数になることを以下示す．

ζ は明らかに Borel 可測かつ非負値なので，(5.1.4) が成立することを示せばよい．各 $x_1 \in S_1$ に対して，$\{\zeta_n(x_1,\cdot)\}_{n\in\mathbb{N}}$ を確率空間 $(S_2, \mathcal{B}(S_2), \theta_{2|1}^2(x_1,\cdot))$ 上で定義された $\mathbb{R}$ 値確率過程と見なすと，簡単な計算により $\{\zeta_n(x_1,\cdot)\}_{n\in\mathbb{N}}$ が情報系 $\{\mathcal{G}_n\}_{n\in\mathbb{N}}$ について非負値マルチンゲールであることがわかる．[6]さらに $\{\zeta_n(x_1,\cdot)\}_{n\in\mathbb{N}}$ は一様可積分である．一様可

[6]「情報系」は “filtration” の訳である．

積分性を示す前にまず「任意の $\varepsilon>0$ に対してある $\delta=\delta_{x_1}>0$ が存在して，$A\in\mathcal{B}(S_2)$ が $\theta^2_{2|1}(x_1,A)\leq\delta$ を満たすならば $\mu^2_{2|1}(x_1,A)\leq\varepsilon$」となることを示す．これが成立しないとすると，ある $\varepsilon>0$ が存在して，任意の $n\in\mathbb{N}$ に対して $\theta^2_{2|1}(x_1,A_n)\leq n^{-2}$ かつ $\mu^2_{2|1}(x_1,A_n)\geq\varepsilon$ を満たす $A_n\in\mathcal{B}(S_2)$ が存在する．A を $\{A_n\}_{n\in\mathbb{N}}$ の上極限集合とする．Fatou の補題を関数列 $\{\mathbf{1}_{A_n^c}\}_{n\in\mathbb{N}}$ に適用することにより $\mu^2_{2|1}(x_1,A)\geq\varepsilon$ を得る．一方 $\sum_{n=1}^\infty\theta^2_{2|1}(x_1,A_n)<\infty$ なので，Borel-Cantelli の補題により $\theta^2_{2|1}(x_1,A)=0$ を得る．これは $\mu^2_{2|1}(x_1,\cdot)$ が $\theta^2_{2|1}(x_1,\cdot)$ に対して絶対連続ということに矛盾する．よって鉤括弧内の主張が成立する．

$\{\zeta_n(x_1,\cdot)\}_{n\in\mathbb{N}}$ の一様可積分性を示すために，任意に $\varepsilon>0$ を取り鉤括弧内で与えられる $\delta>0$ を取る．このとき $M=\delta^{-1}$ とおくと Chebyshev の不等式により

$$\theta^2_{2|1}\left(x_1,\{x_2\in S_2\mid\zeta_n(x_1,x_2)\geq M\}\right)\leq\frac{1}{M}\int_{S_2}\zeta_n(x_1,x_2)\theta^2_{2|1}(x_1,dx_2)=\delta$$

が成立するので

$$\int_{\{x_2\in S_2|\zeta_n(x_1,x_2)\geq M\}}\zeta_n(x_1,x_2)\theta^2_{2|1}(x_1,dx_2)=\mu^2_{2|1}\left(x_1,\{x_2\in S_2\mid\zeta_n(x_1,x_2)\geq M\}\right)\leq\varepsilon$$

を得る．よって $\{\zeta_n(x_1,\cdot)\}_{n\in\mathbb{N}}$ は一様可積分性である．

$\{\zeta_n(x_1,\cdot)\}_{n\in\mathbb{N}}$ は一様可積分な非負値マルチンゲールなので，マルチンゲール収束定理により $n\to\infty$ のとき $\{\zeta_n(x_1,\cdot)\}_{n\in\mathbb{N}}$ は $\zeta(x_1,\cdot)$ に L^1 かつ概収束するので,[7]任意の $m\in\mathbb{N}$ と $B\in\mathcal{P}_m$ に対して

$$\begin{aligned}\int_B\zeta(x_1,x_2)\theta^2_{2|1}(x_1,dx_2)&=\lim_{n\to\infty}\int_B\zeta_n(x_1,x_2)\theta^2_{2|1}(x_1,dx_2)\\&=\mu^2_{2|1}(x_1,B)\\&=\int_B\frac{d\mu^2_{2|1}(x_1,\cdot)}{d\theta^2_{2|1}(x_1,\cdot)}(x_2)\theta^2_{2|1}(x_1,dx_2)\end{aligned}$$

を得る．$\cup_{m\in\mathbb{N}}\mathcal{P}_m$ は π 系なので，π-λ 定理（Dynkin 族定理）により，この式は任意の $B\in\mathcal{B}(S_2)$ に対して成立する．これで (5.1.4) が示された．

Radon-Nikodym 微分の定義と正則条件付き確率の定義により，任意の $A_1\in\mathcal{B}(S_1)$ と $A_2\in\mathcal{B}(S_2)$ に対して

$$\begin{aligned}\mu^2(A_1\times A_2)&=\int_{A_1}\mu^2_{2|1}(x_1,A_2)\mu^2_1(dx_1)\\&=\int_{A_1}\left(\int_{A_2}\zeta(x_1,x_2)\theta^2_{2|1}(x_1,dx_2)\right)\psi(x_1)\theta^2_1(dx_1)\end{aligned}$$

[7]マルチンゲール収束定理については，[1, 定理 5-52]，[15, 定理 8.3.8]，[20, 定理 4.6.4] などを参照せよ．

$$= \int_{A_1 \times A_2} \zeta(x_1, x_2)\psi(x_1)\theta^2(dx_1 dx_2)$$

が成立する．ただし，２つ目の等号で (5.1.4) を用いた．よって μ^2 は θ^2 に対して絶対連続であり，θ^2 についてほとんど全ての $(x_1, x_2) \in S_1 \times S_2$ に対して

$$\frac{d\mu^2}{d\theta^2}(x_1, x_2) = \psi(x_1)\zeta(x_1, x_2)$$

が成立する．μ^2 は θ^2 に対して絶対連続なので，上式は μ^2 についてほとんど全ての $(x_1, x_2) \in S_1 \times S_2$ に対しても成立する．以上により

$$\begin{aligned}
&H(\mu_1^2|\theta_1^2) + \int_{S_1} H(\mu_{2|1}^2|\theta_{2|1}^2) d\mu_1^2 \\
&= \int_{S_1} \log\psi d\mu_1^2 + \int_{S_1} \left(\int_{S_2} \log\zeta(x_1, x_2)\mu_{2|1}^2(x_1, dx_2) \right) \mu_1^2(dx_1) \\
&= \int_{S_1 \times S_2} \log\psi d\mu^2 + \int_{S_1 \times S_2} \log\zeta(x_1, x_2)\mu^2(dx_1 dx_2) \\
&= \int_{S_1 \times S_2} \log\left(\psi(x_1)\zeta(x_1, x_2)\right)\mu^2(dx_1 dx_2) \\
&= H(\mu^2|\theta^2)
\end{aligned}$$

を得る．これで補題の証明が終了する．　□

次に補題 5.1.1 の下限の中身は，ある確率変数の期待値として書けることについて説明する．$\mu^2 \in \mathcal{M}_1(S^2)$ を任意に固定する．[8]以下では確率測度 μ^2 の下での標準射影を

$$\bar{X}_1^2(x_1, x_2) := x_1, \quad \bar{X}_2^2(x_1, x_2) := x_2, \quad (x_1, x_2) \in S^2$$

と書き，$\bar{X}_1^2$ と $\bar{X}_2^2$ を確率空間 $(S^2, \mathcal{B}(S^2), \mu^2)$ 上の S 値確率変数とそれぞれ見る．このとき $f : S^2 \to \mathbb{R}$ を有界可測関数とすると

$$\int_{S^2} f \, d\mu^2 = \mathbb{E}\left[f(\bar{X}_1^2, \bar{X}_2^2)\right]$$

である．また $\mu_{2|1}^2$ を補題 5.1.2 に現れる正則条件付き確率とすると，補題 5.1.2 において θ^2 として直積測度 $\lambda^{\otimes 2}$ と取ると

$$H(\mu^2|\lambda^{\otimes 2}) = H(\mu_1^2|\lambda) + \int_S H(\mu_{2|1}^2|\lambda)\mu_1^2(dx_1)$$

[8] S^2 は直積空間 $S \times S$ を表すが，μ^2 の「2」は単に上付き添字である．S^n と μ^n に関しても同様である．

を得る.[9] ここで $\mathcal{M}_1(S)$ 値確率変数 $\bar{\mu}_1^2$ と $\bar{\mu}_2^2$ をそれぞれ $\bar{\mu}_1^2(\cdot) := \mu_1^2(\cdot)$（$\bar{\mu}_1^2$ はランダムでないことに注意），$\bar{\mu}_2^2(\cdot) := \mu_{2|1}^2(\bar{X}_1^2, \cdot)$ と定義すれば，定義より明らかに

$$H(\mu_1^2|\lambda) = \mathbb{E}\left[H(\bar{\mu}_1^2|\lambda)\right], \quad \int_S H(\mu_{2|1}^2|\lambda)\mu_1^2(dx_1) = \mathbb{E}\left[H(\bar{\mu}_2^2|\lambda)\right]$$

が成立する．以上をまとめると，$\mu^2 \in \mathcal{M}_1(S^2)$ が与えられるごとに S 値確率変数 $\bar{X}_1^2, \bar{X}_2^2$ と $\mathcal{M}_1(S)$ 値確率変数 $\bar{\mu}_1^2, \bar{\mu}_2^2$ が上記の要領で定められて

$$\int_{S^2} f\,d\mu^2 + H(\mu^2|\lambda^{\otimes 2}) = \mathbb{E}\left[f(\bar{X}_1^2, \bar{X}_2^2) + \sum_{i=1}^{2} H(\bar{\mu}_i^2|\lambda)\right]$$

が成立する．この式と補題 5.1.1 により次の変分公式を得る．

$$-\log \mathbb{E}\left[e^{-f(X_1,X_2)}\right] = \inf_{\mu^2 \in \mathcal{M}_1(S^2)} \mathbb{E}\left[f(\bar{X}_1^2, \bar{X}_2^2) + \sum_{i=1}^{2} H(\bar{\mu}_i^2|\lambda)\right].$$

上の議論を $n \in \mathbb{N}$ 変数の場合に一般化する．$\mu^n \in \mathcal{M}_1(S^n)$ を任意に固定する．$i = 1, \ldots, n$ に対して確率測度 μ^n の下での標準射影を

$$\bar{X}_i^n(x_1, \ldots, x_n) := x_i, \quad (x_1, \ldots, x_n) \in S^n$$

と書き，$(\bar{X}_1^n, \ldots, \bar{X}_n^n)$ を確率空間 $(S^n, \mathcal{B}(S^n), \mu^n)$ 上の S^n 値確率変数と見る．明らかに $(\bar{X}_1^n, \ldots, \bar{X}_n^n)$ の分布は μ^n なので，有界可測関数 $f : S^n \to \mathbb{R}$ に対して

$$\int_{S^n} f\,d\mu^n = \mathbb{E}\left[f(\bar{X}_1^n, \ldots, \bar{X}_n^n)\right]$$

が成立する．また $i = 1, \ldots, n-1$ に対して，μ_i^n を $(\bar{X}_1^n, \ldots, \bar{X}_i^n)$ が S^i 上に誘導する分布とし，確率変数 $(\bar{X}_1^n, \ldots, \bar{X}_{i+1}^n)$ の $(\bar{X}_1^n, \ldots, \bar{X}_i^n)$ が与えられたときの正則条件付き確率（つまり S^{i+1} の最後の成分を取り除く射影を与えたときの μ_{i+1}^n の正則条件付き確率）を

$$\{\mu_{i+1|i}^n((x_1, \ldots, x_i), A) \mid (x_1, \ldots, x_i) \in S^i, A \in \mathcal{B}(S)\}$$

と書くことにする．このとき補題 5.1.2 を繰り返し用いることにより

$$\begin{aligned} H(\mu^n|\lambda^{\otimes n}) &= H(\mu_{n-1}^n|\lambda^{\otimes n-1}) + \int_{S^{n-1}} H(\mu_{n|n-1}^n|\lambda)\mu_{n-1}^n(dx_1 \cdots dx_{n-1}) \\ &= H(\mu_1^n|\lambda) + \sum_{i=1}^{n-1} \int_{S^i} H(\mu_{i+1|i}^n|\lambda)\mu_i^n(dx_1 \cdots dx_i) \end{aligned}$$

---

[9] $n \in \mathbb{N}$ に対して $\mu^{\otimes n}$ は μ の n 重直積測度を表す.

を得る．ここで $i=1,\dots,n$ に対して $\mathcal{M}_1(S)$ 値確率変数 $\bar{\mu}_i^n$ を

$$\bar{\mu}_1^n(\cdot) := \mu_1^n(\cdot),\quad \bar{\mu}_{i+1}^n(\cdot) := \mu_{i+1|i}^n((\bar{X}_1^n,\dots,\bar{X}_i^n),\cdot)$$

により定義すると，$i=1,\dots,n-1$ に対してほぼ明らかに

$$H(\mu_1^n|\lambda) = \mathbb{E}\,[H(\bar{\mu}_1^n|\lambda)],\quad \int_{S^i} H(\mu_{i+1|i}^n|\lambda)\mu_i^n(dx_1\cdots dx_i) = \mathbb{E}\left[H(\bar{\mu}_{i+1}^n|\lambda)\right]$$

が成立する．この式と補題 5.1.1 により次の変分公式を得る．

$$-\log \mathbb{E}\left[e^{-f(X_1,\dots,X_n)}\right] = \inf_{\mu^n\in\mathcal{M}_1(S^n)} \mathbb{E}\left[f(\bar{X}_1^n,\dots,\bar{X}_n^n) + \sum_{i=1}^n H(\bar{\mu}_i^n|\lambda)\right].$$

以上の議論をまとめて次の補題を得る．

補題 5.1.4. $\{X_i\}_{i\in\mathbb{N}}$ を独立同分布な S 値確率変数の列とし，X_1 の分布を $\lambda\in\mathcal{M}_1(S)$ とする．また $n\in\mathbb{N}$ と有界可測関数 $f: S^n\to\mathbb{R}$ を任意に固定する．このとき各 $\mu^n\in\mathcal{M}_1(S^n)$ に対して，S 値確率変数 $\bar{X}_1^n,\dots,\bar{X}_n^n$ と $\mathcal{M}_1(S)$ 値確率変数 $\bar{\mu}_1^n,\dots,\bar{\mu}_n^n$ を上記の要領で構成すれば以下の４条件が成立する．

(1) S^n 値確率変数 $(\bar{X}_1^n,\dots,\bar{X}_n^n)$ の分布は μ^n である．

(2) 各 $i=1,\dots,n$ に対して $\bar{\mu}_i^n$ は $\sigma(\bar{X}_1^n,\dots,\bar{X}_{i-1}^n)$ 可測である．ただし $i=1$ のときは $\{\emptyset,\Omega\}$ 可測と見なす．

(3) 各 $i=1,\dots,n$ に対して $\bar{X}_i^n$ を $\bar{X}_1^n,\dots,\bar{X}_{i-1}^n$ で条件付けたときの分布は $\bar{\mu}_i^n$ である．

(4) $\displaystyle\int_{S^n} f\,d\mu^n + H(\mu^n|\lambda^{\otimes n}) = \mathbb{E}\left[f(\bar{X}_1^n,\dots,\bar{X}_n^n) + \sum_{i=1}^n H(\bar{\mu}_i^n|\lambda)\right]$ が成立する．

さらに次の変分公式も成立する．

$$-\log \mathbb{E}\left[e^{-f(X_1,\dots,X_n)}\right] = \inf_{\mu^n\in\mathcal{M}_1(S^n)} \mathbb{E}\left[f(\bar{X}_1^n,\dots,\bar{X}_n^n) + \sum_{i=1}^n H(\bar{\mu}_i^n|\lambda)\right].$$

各 $\mu^n\in\mathcal{M}_1(S)$ に対して，補題 5.1.4 により与えられる S 値確率変数 $\bar{X}_1^n,\dots,\bar{X}_n^n$ と $\mathcal{M}_1(S)$ 値確率変数 $\bar{\mu}_1^n,\dots,\bar{\mu}_n^n$ を μ^n から定まる制御と呼ぶことにし，単に $(\bar{X}^n,\bar{\mu}^n)$ と書くことにする．$(\bar{X}^n,\bar{\mu}^n)$ が μ^n に依存することは明記しない．

本節の最後にこれまで述べてきたことをまとめる．定理 1.4.4（Laplace 原理）により，定理 5.0.1 を示すためには任意の有界連続関数 $F:\mathcal{M}_1(S)\to\mathbb{R}$ に対して

$$\lim_{n\to\infty} -\frac{1}{n}\log\mathbb{E}\left[e^{-nF(L^n)}\right] = \inf_{\mu\in\mathcal{M}_1(S)}[F(\mu)+H(\mu|\lambda)]$$

を示せば十分である．補題 5.1.4 における $f : S^n \to \mathbb{R}$ として

$$f(x_1, \dots, x_n) := nF\left(\frac{1}{n}\sum_{i=1}^n \delta_{x_i}\right)$$

を選ぶと，$f(X_1, \dots, X_n) = nF(L^n)$ なので

$$-\frac{1}{n}\log \mathbb{E}\left[e^{-nF(L^n)}\right] = \inf_{\mu^n \in \mathcal{M}_1(S^n)} \mathbb{E}\left[F(\bar{L}^n) + \frac{1}{n}\sum_{i=1}^n H(\bar{\mu}_i^n|\lambda)\right] \tag{5.1.5}$$

を得る．ここで $\bar{L}^n$ は制御 $\bar{X}_1^n, \dots, \bar{X}_n^n$ に対する経験測度

$$\bar{L}_n := \frac{1}{n}\sum_{i=1}^n \delta_{\bar{X}_i^n}$$

である．これらの議論により定理 5.0.1 を示すためには

$$\lim_{n\to\infty} \inf_{\mu^n \in \mathcal{M}_1(S^n)} \mathbb{E}\left[F(\bar{L}^n) + \frac{1}{n}\sum_{i=1}^n H(\bar{\mu}_i^n|\lambda)\right] = \inf_{\mu \in \mathcal{M}_1(S)} [F(\mu) + H(\mu|\lambda)] \tag{5.1.6}$$

を示せば十分である．

5.2 測度の緊密性

弱収束法による証明で基本的な役割を果たす，測度の緊密性に関する性質を本節で示す．まず測度の緊密性の定義を復習する．A を添字集合とし，$\{\gamma_a\}_{a\in A}$ を A で添字付けられた $\mathcal{M}_1(S)$ の元の族とする．$\{\gamma_a\}_{a\in A}$ が緊密であるとは，任意の $\varepsilon > 0$ に対してあるコンパクト集合 $K_\varepsilon \subset S$ が存在して次が成立することである．

$$\inf_{a\in A} \gamma_a(K_\varepsilon) \geq 1 - \varepsilon.$$

また $\{X_a\}_{a\in A}$ を A で添字付けられた S 値確率変数の族とするとき，$\{X_a\}_{a\in A}$ が緊密であるとは，その法則の族が緊密であることをいう．Borel 可測関数 $g : S \to [0, \infty]$ が緊密性関数であるとは，任意の $M \in [0, \infty)$ に対して $\{x \in S \mid g(x) \leq M\}$ が相対コンパクトになることをいう．

補題 5.2.1. $\{\gamma_a\}_{a\in A}$ が緊密であるための必要十分条件は，次を満たす緊密性関数 $g : S \to [0, \infty]$ が存在することである．

$$\sup_{a\in A} \int_S g \, d\gamma_a < \infty. \tag{5.2.1}$$

証明. $\{\gamma_a\}_{a\in A}$ は緊密であると仮定する．$i\in\mathbb{N}$ に対してコンパクト集合 $\widetilde{K}_i\subset S$ を，任意の $a\in A$ に対して $\gamma_a(\widetilde{K}_i^c)\le 2^{-i}$ となるように取り，$K_i:=\cup_{j=1}^i\widetilde{K}_i$ とおく．このとき $\{K_i\}_{i\in\mathbb{N}}$ は増大列であり．任意の $a\in A$ に対して $\gamma_a(K_i^c)\le 2^{-i}$ である．ここで関数 $g:S\to\mathbb{N}_0\cup\{\infty\}$ を次で定義する．

$$g(x):=\sum_{i=1}^\infty \mathbf{1}_{K_i^c}(x),\quad x\in S.$$

このとき g は緊密性関数であることを見る．実際任意の $M\in\mathbb{N}$ に対して，$g(x)\le M\Leftrightarrow x\in K_{M+1}$ であることに注意すると，$\{x\in S\mid g(x)\le M\}=K_{M+1}$ を得るので $\{x\in S\mid g(x)\le M\}$ はコンパクト集合である．よって g は緊密性関数である．また任意の $a\in A$ に対して

$$\int_S g\,d\gamma_a=\sum_{i=1}^\infty\gamma_a(K_i^c)\le 1$$

が成立するので (5.2.1) が示された．

逆に緊密性関数 $g:S\to[0,\infty]$ で (5.2.1) を満たすものが存在したと仮定する．ここで $M\in[0,\infty)$ に対して $C_M\subset S$ を $\{x\in S\mid g(x)\le M\}$ の閉包とする．g は緊密性関数なので C_M はコンパクト集合である．また $C_M^c\subset\{x\in S\mid g(x)>M\}$ である．よって Chebyshev の不等式により任意の $a\in A$ に対して

$$\gamma_a(C_M^c)\le\frac{1}{M}\int_S g\,d\gamma_a$$

が成立する．(5.2.1) を仮定したので右辺の積分値は a に依らない定数で評価できて，$\{\gamma_a\}_{a\in A}$ が緊密であることがわかる． □

次の補題は補題 5.2.1 により直ちに従う．

補題 5.2.2. $g:S\to[0,\infty]$ を緊密性関数とする．関数 $G:\mathcal{M}_1(S)\to[0,\infty]$ を次で定義する．

$$G(\mu):=\int_S g\,d\mu,\quad \mu\in\mathcal{M}_1(S).$$

このとき，任意の $M\in[0,\infty)$ に対して $\{\mu\in\mathcal{M}_1(S)\mid G(\mu)\le M\}$ は緊密である．特に G は $\mathcal{M}_1(S)$ 上の緊密性関数である．

次の補題は測度値確率変数を扱う分野において有名な事実である．

補題 5.2.3. $\{\Lambda_a\}_{a\in A}$ を $\mathcal{M}_1(S)$ 値確率変数の族とし，各 $a\in A$ に対して $\gamma_a\in\mathcal{M}_1(S)$ を $\gamma_a:=\mathbb{E}[\Lambda_a]$ と定義する．[10] このとき，$\{\Lambda_a\}_{a\in A}$ が確率変数の族として緊密であることと，$\{\gamma_a\}_{a\in A}$ が確率測度の族として緊密であることは同値である．

[10] γ_a はランダムでないことに注意せよ．

証明. $\{\Lambda_a\}_{a\in A}$ は緊密であると仮定する．Λ_a の $\mathcal{M}_1(S)$ 上の分布を η_a と表す．$\varepsilon > 0$ を固定する．$\{\Lambda_a\}_{a\in A}$ は緊密なので，あるコンパクト集合 $\mathcal{K}_\varepsilon \subset \mathcal{M}_1(S)$ に対して

$$\sup_{a\in A} \eta_a(\mathcal{K}_\varepsilon^c) \le \varepsilon \tag{5.2.2}$$

を満たす．$\mathcal{K}_\varepsilon$ は $\mathcal{M}_1(S)$ におけるコンパクト集合なので，Prokhorov の定理により $\mathcal{K}_\varepsilon$ は緊密である.[11] よって，あるコンパクト集合 $K_\varepsilon \subset S$ が存在して

$$\sup_{\gamma\in\mathcal{K}_\varepsilon} \gamma(K_\varepsilon^c) \le \varepsilon \tag{5.2.3}$$

が成立する．ここで任意の $a \in A$ に対して

$$\begin{aligned}
\gamma_a(K_\varepsilon^c) &= \int_{\mathcal{M}_1(S)} \gamma(K_\varepsilon^c)\eta_a(d\gamma) \\
&= \int_{\mathcal{K}_\varepsilon} \gamma(K_\varepsilon^c)\eta_a(d\gamma) + \int_{\mathcal{K}_\varepsilon^c} \gamma(K_\varepsilon^c)\eta_a(d\gamma) \\
&\le \int_{\mathcal{K}_\varepsilon} \gamma(K_\varepsilon^c)\eta_a(d\gamma) + \eta_a(\mathcal{K}_\varepsilon^c)
\end{aligned}$$

となっているので，(5.2.2) と (5.2.3) により

$$\sup_{a\in A} \gamma_a(K_\varepsilon^c) \le 2\varepsilon$$

が成立する．よって $\{\gamma_a\}_{a\in A}$ は緊密である．

逆に $\{\gamma_a\}_{a\in A}$ は緊密であると仮定する．補題 5.2.1 により緊密性関数 $g : S \to [0, \infty]$ で

$$\sup_{a\in A} \int_S g\, d\gamma_a < \infty$$

となるものが存在する．ここで $G : \mathcal{M}_1(S) \to [0, \infty]$ を次で定義する．

$$G(\gamma) := \int_S g\, d\gamma, \quad \gamma \in \mathcal{M}_1(S).$$

補題 5.2.2 により G は緊密性関数である．また

$$\sup_{a\in A} \int_{\mathcal{M}_1(S)} G(\gamma)\eta_a(d\gamma) = \sup_{a\in A} \mathbb{E}\left[G(\Lambda_a)\right] = \sup_{a\in A} \mathbb{E}\left[\int_S g\, d\Lambda_a\right] = \sup_{a\in A} \int_S g\, d\gamma_a < \infty$$

が成立する.（最後の等号を得るために単調収束定理を用いて単関数による近似を行なった.）よって補題 5.2.1 により $\{\Lambda_a\}_{a\in A}$ は緊密である．　□

[11] Prokhorov の定理については例えば [27, Theorem 5.2] を参照せよ．

5.3　制御に対する評価

本節では制御に対する評価をいくつか示す．それらは定理 5.0.1 と定理 5.0.2 の証明中で用いる．$(\bar{X}^n, \bar{\mu}^n)$ を $\mu^n \in \mathcal{M}_1(S^n)$ から定まる制御とする．ここで $\mathcal{M}_1(S)$ 値確率変数 $\bar{L}^n$ と $\widehat{\mu}^n$ をそれぞれ次で定義する．

$$\bar{L}^n := \frac{1}{n}\sum_{i=1}^n \delta_{\bar{X}_i^n}, \quad \widehat{\mu}^n := \frac{1}{n}\sum_{i=1}^n \bar{\mu}_i^n.$$

制御 $(\bar{X}^n, \bar{\mu}^n)$ は確率空間 $(S^n, \mathcal{B}(S^n), \mu^n)$ 上で定められているため，$\bar{L}^n$ と $\widehat{\mu}^n$ もこの確率空間上で定義された確率変数である．

補題 5.3.1. $(\mu^n)_{n\in\mathbb{N}} \in \prod_{n=1}^\infty \mathcal{M}_1(S^n)$ として，各 $n \in \mathbb{N}$ に対して $(\bar{X}^n, \bar{\mu}^n)$ を μ^n から定まる制御とする．$\{\mathbb{E}\left[H(\widehat{\mu}^n|\lambda)\right]\}_{n\in\mathbb{N}}$ は有界だと仮定すると，$\mathcal{M}_1(S)^2$ 値確率変数の列 $\{(\bar{L}^n, \widehat{\mu}^n)\}_{n\in\mathbb{N}}$ は緊密である．[12]

証明. $\{\widehat{\mu}^n\}_{n\in\mathbb{N}}$ と $\{\bar{L}^n\}_{n\in\mathbb{N}}$ の緊密性をそれぞれ示せば十分である．任意の有界連続関数 $f : S \to \mathbb{R}$ に対して，定理 4.1.3 により

$$H(\widehat{\mu}^n|\lambda) \geq \int_S f\, d\widehat{\mu}^n - \log \int_S e^f\, d\lambda$$

が成立する．両辺で期待値を取って，f についての上限を取ることにより

$$H(\mathbb{E}\left[\widehat{\mu}^n\right]|\lambda) \leq \mathbb{E}\left[H(\widehat{\mu}^n|\lambda)\right] \tag{5.3.1}$$

がわかる．右辺は仮定により n について有界なので

$$\sup_{n\in\mathbb{N}} H(\mathbb{E}\left[\widehat{\mu}^n\right]|\lambda) < \infty$$

がわかる．命題 4.1.4 により $H(\cdot|\lambda)$ は良い速度関数，すなわち緊密性関数なので，補題 5.2.1 により $\{\mathbb{E}\left[\widehat{\mu}^n\right]\}_{n\in\mathbb{N}}$ は緊密である．よって補題 5.2.3 により $\{\widehat{\mu}^n\}_{n\in\mathbb{N}}$ も緊密である．

一方で任意の非負もしくは有界可測関数 $f : S \to \mathbb{R}$ に対して，制御の性質 (3) により

$$\mathbb{E}\left[\int_S f\, d\bar{L}^n\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n f(\bar{X}_i^n)\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \int_S f\, d\bar{\mu}_i^n\right] = \mathbb{E}\left[\int_S f\, d\widehat{\mu}^n\right]$$

が成立するので，$\mathbb{E}\left[\bar{L}^n\right] = \mathbb{E}\left[\widehat{\mu}^n\right]$ がわかる．よって再び補題 5.2.3 により，$\{\bar{L}^n\}_{\in\mathbb{N}}$ が緊密であることがわかる．　□

[12] n ごとに $(\bar{L}^n, \widehat{\mu}^n)$ が定義されている確率空間は違うが，今はその法則のみを扱っているので，以下で問題は生じない．

補題 5.3.2. $(\mu^n)_{n\in\mathbb{N}} \in \prod_{n=1}^{\infty} \mathcal{M}_1(S^n)$ として，各 $n\in\mathbb{N}$ に対して $(\bar{X}^n, \bar{\mu}^n)$ を μ^n から定まる制御とする．また $\mathcal{M}_1(S)^2$ 値確率変数の列 $\{(\bar{L}^n, \widehat{\mu}^n)\}_{n\in\mathbb{N}}$ が部分列に沿って $\mathcal{M}_1(S)^2$ 値確率変数 $(\bar{L}, \widehat{\mu})$ に弱収束すると仮定する．このとき確率 1 で $\bar{L} = \widehat{\mu}$ が成立する．

証明. $\mathcal{M}_1(S)$ を分離する有界連続関数の列 $\{f_m\}_{m\in\mathbb{N}} \subset C_{\mathrm{b}}(S)$ を取る．各 $m\in\mathbb{N}$ と $i=1,\ldots,n$ に対して

$$\Delta^n_{m,i} := f_m(\bar{X}^n_i) - \int_S f_m d\bar{\mu}^n_i$$

と定義する．このとき

$$\int_S f_m d\bar{L}^n - \int_S f_m d\widehat{\mu}^n = \frac{1}{n}\sum_{i=1}^n \Delta^n_{m,i}$$

と書けるので，Chebyshev の不等式により任意の $\varepsilon>0$ に対して

$$\mathbb{P}\left(\left|\int_S f_m d\bar{L}^n - \int_S f_m d\widehat{\mu}^n\right| > \varepsilon\right) \le \frac{1}{\varepsilon^2 n^2}\sum_{i,j=1}^n \mathbb{E}\left[\Delta^n_{m,i}\Delta^n_{m,j}\right]$$

が成立する．ここで $i\neq j$ に対して

$$\mathbb{E}\left[\Delta^n_{m,i}\Delta^n_{m,j}\right] = 0$$

となる．実際，$\bar{\mathcal{F}}^n_i := \sigma(\bar{X}^n_1,\ldots,\bar{X}^n_{i-1})$ とおくと，$i>j$ に対して制御の性質 (2) と (3) により $\Delta^n_{m,j}$ は $\bar{\mathcal{F}}^n_i$ 可測かつ $\mathbb{E}\left[\Delta^n_{m,i}|\bar{\mathcal{F}}^n_i\right]=0$ なので

$$\mathbb{E}\left[\Delta^n_{m,i}\Delta^n_{m,j}\right] = \mathbb{E}\left[\mathbb{E}\left[\Delta^n_{m,i}\Delta^n_{m,j}|\bar{\mathcal{F}}^n_i\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\Delta^n_{m,i}|\bar{\mathcal{F}}^n_i\right]\Delta^n_{m,j}\right] = 0$$

となっている．また $|\Delta^n_{m,i}| \le 2\|f_m\|_\infty$ なので，$n\to\infty$ のとき

$$\mathbb{P}\left(\left|\int_S f_m\, d\bar{L}^n - \int_S f_m\, d\widehat{\mu}^n\right| > \varepsilon\right) \le \frac{4\|f_m\|^2_\infty}{n\varepsilon^2} \longrightarrow 0$$

が成立する．ここで $\mathcal{M}_1(S)^2$ の部分集合

$$\left\{(\mu_1,\mu_2)\in\mathcal{M}_1(S)^2 \;\middle|\; \left|\int_S f_m\, d\mu_1 - \int_S f_m\, d\mu_2\right| > \varepsilon\right\}$$

は弱収束の定義により開集合なので，弱収束の特徴付けにより[13]

$$\mathbb{P}\left(\left|\int_S f_m\, d\bar{L} - \int_S f_m\, d\widehat{\mu}\right| > \varepsilon\right) \le \varliminf_{n\to\infty} \mathbb{P}\left(\left|\int_S f_m\, d\bar{L}^n - \int_S f_m\, d\widehat{\mu}^n\right| > \varepsilon\right) = 0$$

が成立する．$\varepsilon>0$ は任意なので，任意の $m\in\mathbb{N}$ に対して

$$\mathbb{P}\left(\int_S f_m\, d\bar{L} = \int_S f_m\, d\widehat{\mu}\right) = 1.$$

$\{f_m\}_{m\in\mathbb{N}}$ の分離性により，確率 1 で $\bar{L}=\widehat{\mu}$ が成立する． □

[13] Portmanteau の定理と呼ばれることが多い．

次の 2 つの命題は定理 5.0.2 を示すために必要である．そのため $S = \mathbb{R}^d$ の場合のみを考える．$\lambda \in \mathcal{M}_1(\mathbb{R}^d)$ は $\mathbb{R}^d$ 値確率変数 X_1 の分布であったことを思い出しておく．

命題 5.3.3. $S = \mathbb{R}^d$ とし，条件 **(FEM)** を仮定する．

(1) $\theta \in \mathcal{M}_1(\mathbb{R}^d)$ は $H(\theta|\lambda) < \infty$ を満たすとする．このとき任意の $\sigma \geq 1$ と $M \geq 0$ に対して次が成立する．

$$\int_{\mathbb{R}^d} |y| \mathbf{1}_{\{|y| \geq M\}} \theta(dy) \leq \int_{\mathbb{R}^d} e^{\sigma|y|} \mathbf{1}_{\{|y| \geq M\}} \lambda(dy) + \frac{1}{\sigma} H(\theta|\lambda).$$

特に条件 **(FEM)** により右辺第 1 項は有限なので，$\int_{\mathbb{R}^d} |y| \theta(dy) < \infty$ であり，積分 $\int_{\mathbb{R}^d} y\theta(dy)$ は問題なく定まる．

(2) $\mathcal{M}_1(S)$ 値確率変数 L は $H(\mathbb{E}[L]|\lambda) < \infty$ を満たすとする．このとき任意の $\sigma \geq 1$ と $M \geq 0$ に対して次が成立する．

$$\mathbb{E}\left[\int_{\mathbb{R}^d} |y| \mathbf{1}_{\{|y| \geq M\}} L(dy)\right] \leq \int_{\mathbb{R}^d} e^{\sigma|y|} \mathbf{1}_{\{|y| \geq M\}} \lambda(dy) + \frac{1}{\sigma} H(\mathbb{E}[L]|\lambda). \tag{5.3.2}$$

特に $\mathbb{E}\left[\int_{\mathbb{R}^d} |y| L(dy)\right] < \infty$ であり，積分 $\int_{\mathbb{R}^d} yL(dy)$ は確率 1 で問題なく定まる．

証明. まず (1) を示す．任意の $b \geq 0$ と $\sigma \geq 1$ に対して次が成立することに注意する．

$$\sup_{a \in \mathbb{R}} \{ab - e^{\sigma a}\} = \frac{b}{\sigma}\left(\log \frac{b}{\sigma} - 1\right) \leq \frac{1}{\sigma}(b \log b - b + 1) =: \frac{1}{\sigma}\ell(b).$$

実際，はじめの等号は微分を用いた簡単な計算により従い，$\sigma \geq 1$ なので不等号は自明である．よって任意の $a, b \geq 0$ と $\sigma \geq 1$ に対して

$$ab \leq e^{\sigma a} + \frac{1}{\sigma}\ell(b) \tag{5.3.3}$$

が成立する．一方で $H(\theta|\lambda) < \infty$ なので θ は λ に対して絶対連続である．よって (5.3.3) により

$$\begin{aligned}
\int_{\mathbb{R}^d} |y| \mathbf{1}_{\{|y| \geq M\}} \theta(dy) &= \int_{\mathbb{R}^d} |y| \mathbf{1}_{\{|y| \geq M\}} \frac{d\theta}{d\lambda}(y) \lambda(dy) \\
&\leq \int_{\mathbb{R}^d} \left(e^{\sigma|y|} \mathbf{1}_{\{|y| \geq M\}} + \frac{1}{\sigma}\ell\left(\frac{d\theta}{d\lambda}(y)\right)\right) \lambda(dy) \\
&= \int_{\mathbb{R}^d} e^{\sigma|y|} \mathbf{1}_{\{|y| \geq M\}} \lambda(dy) + \frac{1}{\sigma} H(\theta|\lambda)
\end{aligned}$$

を得る．ただし 2 行目の不等式を得るために $\ell(\cdot) \geq 0$ であることを用いた．これで (1) が示された．

(2) は (1) で $\theta = \mathbb{E}[L]$ とすれば直ちにわかる．実際

$$\begin{aligned}\mathbb{E}\left[\int_{\mathbb{R}^d} |y|\mathbf{1}_{\{|y|\geq M\}} L(dy)\right] &= \int_{\mathbb{R}^d} |y|\mathbf{1}_{\{|y|\geq M\}} \mathbb{E}[L](dy) \\ &\leq \int_{\mathbb{R}^d} e^{\sigma|y|}\mathbf{1}_{\{|y|\geq M\}} \lambda(dy) + \frac{1}{\sigma} H(\mathbb{E}[L]|\lambda)\end{aligned}$$

を得るので，証明が終わる． □

命題 5.3.4. $S = \mathbb{R}^d$ とし，条件 **(FEM)** を仮定する．$(\mu^n)_{n\in\mathbb{N}} \in \prod_{n=1}^{\infty} \mathcal{M}_1(S^n)$ として，各 $n \in \mathbb{N}$ に対して $(\bar{X}^n, \bar{\mu}^n)$ を μ^n から定まる制御とする．$\{\mathbb{E}[H(\widehat{\mu}^n|\lambda)]\}_{n\in\mathbb{N}}$ は有界であり，$\{\bar{L}^n\}_{n\in\mathbb{N}}$ は $n \to \infty$ のとき $\mathcal{M}_1(S)$ 値確率変数 $\bar{L}$ に弱収束すると仮定する．このとき任意の有界連続関数 $f : \mathbb{R}^d \to \mathbb{R}$ に対して

$$\lim_{n\to\infty} \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\bar{L}^n(dy)\right)\right] = \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\bar{L}(dy)\right)\right].$$

証明. まず命題の主張に現れる積分が確率 1 で問題なく定まることを示す．$\bar{L}^n$ は Dirac 測度の和なので，積分 $\int_{\mathbb{R}^d} y\bar{L}^n(dy)$ は常に問題なく定まる．$\mathcal{M}_1(S)$ は可分完備距離空間で $\{\bar{L}^n\}_{n\in\mathbb{N}}$ は $n\to\infty$ のとき $\bar{L}$ に弱収束するので，Skorokhod の表現定理を適用することができる．Skorokhod の表現定理により与えられる確率空間を $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ と表し，記号の単純化のため，概収束する $\mathcal{M}_1(S)$ 値確率変数を同じ記号 $\bar{L}^n$ $(n \in \mathbb{N})$ と $\bar{L}$ で表すことにする．また $\widehat{\mathbb{P}}$ に対する期待値は $\widehat{\mathbb{E}}[\cdot]$ で表すことにする．$C_0 := \sup_{n\in\mathbb{N}} \mathbb{E}[H(\widehat{\mu}^n|\lambda)] < \infty$ とおくと，(5.3.1) と仮定により $H(\widehat{\mathbb{E}}[\widehat{\mu}^n]|\lambda) \leq C_0$ が成立する．また弱収束の定義と有界収束定理により，$\{\widehat{\mathbb{E}}[\bar{L}^n]\}_{n\in\mathbb{N}}$ が $\widehat{\mathbb{E}}[\bar{L}]$ に弱収束することが簡単にわかる．さらに補題 5.3.1 の証明で見たように $\widehat{\mathbb{E}}[\bar{L}^n] = \mathbb{E}[\bar{L}^n] = \mathbb{E}[\widehat{\mu}^n]$ が成立するので，

$$H(\widehat{\mathbb{E}}[\bar{L}]|\lambda) \leq \varliminf_{n\to\infty} H(\widehat{\mathbb{E}}[\bar{L}^n]|\lambda) = \varliminf_{n\to\infty} H(\mathbb{E}[\widehat{\mu}^n]|\lambda) \leq C_0 \tag{5.3.4}$$

が成立する．ただし最初の不等号では $H(\cdot|\lambda)$ の下半連続性を用いた．よって命題 5.3.3(2) により $\int_{\mathbb{R}^d} y\bar{L}(dy)$ は確率 1 で問題なく定まる．

次に $\psi(y) = y, y \in \mathbb{R}^d$ とおき，

$$\lim_{n\to\infty} \widehat{\mathbb{E}}\left[\left|\int_{\mathbb{R}^d} \psi\, d\bar{L}^n - \int_{\mathbb{R}^d} \psi\, d\bar{L}\right|\right] = 0$$

を示す．これが示されれば，$\int_{\mathbb{R}^d} \psi\, d\bar{L}^n$ は $n \to \infty$ のとき $\int_{\mathbb{R}^d} \psi\, d\bar{L}$ に法則収束することがわかるので，結論を得る．$M > 0$ に対して

$$\psi_M(y) := ((-M) \vee (y_i \wedge M))_{i=1}^d \in \mathbb{R}^d, \quad y = (y_1, \ldots, y_d) \in \mathbb{R}^d$$

とおく．このとき任意の $y \in \mathbb{R}^d$ に対して $|\psi(y) - \psi_M(y)| \leq |y|\mathbf{1}_{\{|y|\geq M\}}$ である．ここで

$$\widehat{\mathbb{E}}\left[\left|\int_{\mathbb{R}^d} \psi\, d\bar{L}^n - \int_{\mathbb{R}^d} \psi\, d\bar{L}\right|\right] = \widehat{\mathbb{E}}\left[\left|\int_{\mathbb{R}^d} \psi_M\, d(\bar{L}^n - \bar{L}) + \int_{\mathbb{R}^d} (\psi - \psi_M)\, d(\bar{L}^n - \bar{L})\right|\right]$$

$$\leq \widehat{\mathbb{E}}\left[\left|\int_{\mathbb{R}^d} \psi_M \, d(\bar{L}^n - \bar{L})\right|\right] + \widehat{\mathbb{E}}\left[\int_{\mathbb{R}^d} |\psi - \psi_M| \, d(\bar{L}^n + \bar{L})\right]$$

が成立する．$\int_{\mathbb{R}^d} \psi_M \, d\bar{L}^n$ は $\int_{\mathbb{R}^d} \psi_M \, d\bar{L}$ に概収束するので，右辺第1項は $n \to \infty$ のとき有界収束定理により 0 に収束する．一方で $\sigma \geq 1$ と $M \geq 0$ に対して (5.3.2) により

$$\widehat{\mathbb{E}}\left[\int_{\mathbb{R}^d} |y| \mathbf{1}_{\{|y| \geq M\}} \bar{L}^n(dy)\right] \leq \int_{\mathbb{R}^d} e^{\sigma |y|} \mathbf{1}_{\{|y| \geq M\}} \lambda(dy) + \frac{C_0}{\sigma} \tag{5.3.5}$$

を得る．さらに (5.3.2) と (5.3.4) により，(5.3.5) において $\bar{L}^n$ を $\bar{L}$ に取り替えたものも成立するので，

$$\widehat{\mathbb{E}}\left[\int_{\mathbb{R}^d} |\psi - \psi_M| \, d(\bar{L}^n + \bar{L})\right] \leq 2 \int_{\mathbb{R}^d} e^{\sigma |y|} \mathbf{1}_{\{|y| \geq M\}} \lambda(dy) + \frac{2C_0}{\sigma}$$

が成立する．条件 **(FEM)** により $e^{\sigma|y|}$ は λ について可積分なので，右辺第1項は $M \to \infty$ のとき 0 に収束する．よって $n \to \infty, M \to \infty, \sigma \to \infty$ の順に極限を取ることにより

$$\lim_{n \to \infty} \widehat{\mathbb{E}}\left[\left|\int_{\mathbb{R}^d} \psi \, d\bar{L}^n - \int_{\mathbb{R}^d} \psi \, d\bar{L}\right|\right] = 0$$

を得るので，証明が終わる． □

5.4 弱収束法による Sanov の定理の証明

本節では定理 5.0.1 を弱収束法により証明する．

定理 5.0.1 の証明. 第 5.1 節で述べたように，定理 5.0.1 を示すためには (5.1.6) を任意の有界連続関数 $F : \mathcal{M}_1(S) \to \mathbb{R}$ に対して示せば十分である．本証明中では任意に選んだ F を固定して議論する．また (5.1.6) の左辺に現れる下限を J^n とおく，つまり

$$J^n := \inf_{\mu^n \in \mathcal{M}_1(S^n)} \mathbb{E}\left[F(\bar{L}^n) + \frac{1}{n} \sum_{i=1}^{n} H(\bar{\mu}_i^n | \lambda)\right]$$

である．ただし J^n が F と λ に依存することは明記しない．

まず (5.1.6) の下からの評価を示す．$\varepsilon > 0$ を任意に固定する．各 $n \in \mathbb{N}$ に対して，J^n の定義によりある $\mu^n \in \mathcal{M}_1(S^n)$ が存在して，$(\bar{X}^n, \bar{\mu}^n)$ を μ^n から定まる制御とすれば

$$J^n + \varepsilon \geq \mathbb{E}\left[F(\bar{L}^n) + \frac{1}{n} \sum_{i=1}^{n} H(\bar{\mu}_i^n | \lambda)\right]$$

が成立する．また定理 4.1.3 の後で述べたように $H(\cdot|\lambda)$ は凸なので

$$J^n + \varepsilon \geq \mathbb{E}\left[F(\bar{L}^n) + H(\widehat{\mu}^n | \lambda)\right]$$

を得る．F は有界なので $\{\mathbb{E}[H(\widehat{\mu}^n|\lambda)]\}_{n\in\mathbb{N}}$ は有界である．よって補題 5.3.1 により，左辺の下極限を実現する $\{(\bar{L}^n,\widehat{\mu}^n)\}_{n\in\mathbb{N}}$ の部分列であって，ある $\mathcal{M}_1(S)^2$ 値確率変数 $(\bar{L},\widehat{\mu})$ に弱収束するものが存在する．部分列を再び同じ記号で表すことにすると，F は有界連続なので $\lim_{n\to\infty}\mathbb{E}\left[F(\bar{L}^n)\right]=\mathbb{E}\left[F(\bar{L})\right]$ である．また $H(\cdot|\lambda)$ の下半連続性と補題 5.0.4（および注意 5.0.5）により

$$\varliminf_{n\to\infty}\mathbb{E}[H(\widehat{\mu}^n|\lambda)]\geq\mathbb{E}[H(\widehat{\mu}|\lambda)]$$

が成立する．よって補題 5.3.2 により $\bar{L}=\widehat{\mu}$ が確率 1 で成立するので

$$J^n+\varepsilon\geq\mathbb{E}\left[F(\bar{L})+H(\widehat{\mu}|\lambda)\right]\geq\inf_{\mu\in\mathcal{M}_1(S)}[F(\mu)+H(\mu|\lambda)]$$

が成立する．$\varepsilon>0$ は任意なので，(5.1.6) の下からの評価を得る．

　次に (5.1.6) の上からの評価を示す．$\varepsilon>0$ を任意に取る．$\mu^*\in\mathcal{M}_1(S)$ を次を満たすように取る．

$$F(\mu^*)+H(\mu^*|\lambda)\leq\inf_{\mu\in\mathcal{M}_1(S)}[F(\mu)+H(\mu|\lambda)]+\varepsilon.$$

ここで $(\bar{X}^n,\bar{\mu}^n)$ を $(\mu^*)^{\otimes n}$ から定まる制御とすれば，$\bar{X}_1^n,\dots,\bar{X}_n^n$ は独立同分布で $\bar{X}_1^n$ の分布は μ^* である．さらに任意の $n\in\mathbb{N}$ と $i=1,\dots,n$ に対して $\bar{\mu}_i^n=\mu^*$ が成立する．このとき補題 5.3.1 と補題 5.3.2 により，$\{\bar{L}^n\}_{n\in\mathbb{N}}$ は $n\to\infty$ のとき μ^* に弱収束する．よって (5.1.5) により

$$\begin{aligned}\varlimsup_{n\to\infty}J^n&\leq\varlimsup_{n\to\infty}\mathbb{E}\left[F(\bar{L}^n)+H(\mu^*|\lambda)\right]\\&=F(\mu^*)+H(\mu^*|\lambda)\\&\leq\inf_{\mu\in\mathcal{M}_1(S)}[F(\mu)+H(\mu|\lambda)]+\varepsilon\end{aligned}$$

が成立する．$\varepsilon>0$ は任意なので，(5.1.6) の上からの評価を得る．　□

5.5　弱収束法による Cramér の定理の証明

　本節では $S=\mathbb{R}^d$ として定理 5.0.2 を弱収束法により証明する．まず定理 5.0.2 に表れる関数 I が良い速度関数であることを示す．

命題 5.5.1. 条件 (FEM) を仮定する．次で定義される関数は $\mathbb{R}^d$ 上の良い速度関数である．

$$I(x):=\inf\left\{H(\mu|\lambda)\;\middle|\;\mu\in\mathcal{M}_1(\mathbb{R}^d),\int_{\mathbb{R}^d}|y|\mu(dy)<\infty,\int_{\mathbb{R}^d}y\mu(dy)=x\right\},\quad x\in\mathbb{R}^d.$$

証明. まず $I(x) < \infty$ のときに，ある $\mu_x \in \mathcal{M}_1(\mathbb{R}^d)$ が存在して

$$I(x) = H(\mu_x|\lambda), \quad \int_{\mathbb{R}^d} |y|\mu_x(dy) < \infty, \quad \int_{\mathbb{R}^d} y\mu_x(dy) = x \tag{5.5.1}$$

が成立することを示す．$I(x)$ の下限を実現する最小化列 $\{\mu_x^n\}_{n\in\mathbb{N}} \subset \mathcal{M}_1(\mathbb{R}^d)$ を取る．このとき命題 4.1.4 により $H(\cdot|\lambda)$ は良い速度関数であって，$\{H(\mu_x^n|\lambda)\}_{n\in\mathbb{N}}$ は有界なので収束部分列が存在する．収束部分列を再び同じ記号で表すことにして，極限を μ_x と表せば，$H(\cdot|\lambda)$ の下半連続性により

$$H(\mu_x|\lambda) \leq \varliminf_{n\to\infty} H(\mu_x^n|\lambda) \leq \sup_{n\in\mathbb{N}} H(\mu_x^n|\lambda) < \infty \tag{5.5.2}$$

が成立する．$H(\mu_x|\lambda) < \infty$ なので，命題 5.3.3(1) により $\int_{\mathbb{R}^d} |y|\mu_x(dy) < \infty$ である．また $\{\mu_x^n\}_{n\in\mathbb{N}}$ と μ_x の相対エントロピーが一様に評価されているので，命題 5.3.4 の証明の後半において $\bar{L}^n$ と $\bar{L}$ の替わりに μ_x^n と μ^n とすることにより

$$\int_{\mathbb{R}^d} y\mu_x(dy) = \lim_{n\to\infty} \int_{\mathbb{R}^d} y\mu_x^n(dy) = x \tag{5.5.3}$$

を得る.（この場合にはランダム測度ではないので期待値 $\widehat{\mathbb{E}}\,[\cdot]$ は必要ない.）よって $I(x)$ の定義により $I(x) \leq H(\mu_x|\lambda)$ である．一方で $\{\mu_x^n\}_{n\in\mathbb{N}}$ の選び方により $I(x) = \lim_{n\to\infty} H(\mu_x^n|\lambda)$ なので，(5.5.2) の左側の不等式により $I(x) \geq H(\mu_x|\lambda)$ がわかる．以上により $I(x) = H(\mu_x|\lambda)$ を得る．

前半の事実を用いて $K \geq 0$ に対して $C_K := \{x \in \mathbb{R}^d \mid I(x) \leq K\}$ がコンパクト集合であることを示す．任意に点列 $\{x_n\}_{n\in\mathbb{N}} \subset C_K$ を取り，各 $n \in \mathbb{N}$ に対して (5.5.1) で x を x_n に取り替えたものを満たす $\mu_n \in \mathcal{M}_1(\mathbb{R}^d)$ を取る．命題 4.1.4 により $H(\cdot|\lambda)$ は良い速度関数であって，$\{\mu_n\}_{n\in\mathbb{N}}$ は緊密である．収束部分列を再び同じ記号で表すことにして，μ_* を $\{\mu_n\}_{n\in\mathbb{N}}$ の極限とする．このとき $H(\cdot|\lambda)$ の下半連続性により

$$H(\mu_*|\lambda) \leq K < \infty$$

が成立するので，命題 5.3.3(1) により $\int_{\mathbb{R}^d} |y|\mu_*(dy) < \infty$ である．また $x_* := \int_{\mathbb{R}^d} y\mu_*(dy) \in \mathbb{R}^d$ とおくと，$I(x_*) \leq H(\mu_*|\lambda)$ なので $x_* \in C_K$ である．また (5.5.3) と同じ理由で $n \to \infty$ のとき

$$x_n = \int_{\mathbb{R}^d} y\mu_n(dy) \longrightarrow \int_{\mathbb{R}^d} y\mu_*(dy) = x_*$$

を得る．よって C_K は点列コンパクトなので，C_K はコンパクトである． □

定理 5.0.2 の証明. 定理 1.4.4（Laplace 原理）により，任意の有界連続関数 $f : \mathbb{R}^d \to \mathbb{R}$ に対して

$$\lim_{n\to\infty} -\frac{1}{n} \log \mathbb{E}\left[\exp\left\{-nf\left(\int_{\mathbb{R}^d} yL^n(dy)\right)\right\}\right] = \inf_{x\in\mathbb{R}^d}[f(x) + I(x)] \tag{5.5.4}$$

を示せばよい．本証明中では任意に選んだ f を固定して議論する．補題 5.1.4 における有界可測関数として

$$(x_1, \ldots, x_n) \in (\mathbb{R}^d)^n \longmapsto nf\left(\frac{x_1 + \cdots + x_n}{n}\right) \in \mathbb{R}$$

を選ぶことにより，

$$\begin{aligned} &-\frac{1}{n}\log \mathbb{E}\left[\exp\left\{-nf\left(\int_{\mathbb{R}^d} yL^n(dy)\right)\right\}\right] \\ &= \inf_{\mu^n \in \mathcal{M}_1((\mathbb{R}^d)^n)} \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\bar{L}^n(dy)\right) + \frac{1}{n}\sum_{i=1}^n H(\bar{\mu}_i^n|\lambda)\right] \end{aligned} \tag{5.5.5}$$

を得る．

まず (5.5.4) の下からの評価を示す．$\varepsilon > 0$ を任意に取る．(5.5.5) によりある $\mu^n \in \mathcal{M}_1((\mathbb{R}^d)^n)$ が存在して，$(\bar{X}^n, \bar{\mu}^n)$ を μ^n から定まる制御とすれば

$$-\frac{1}{n}\log \mathbb{E}\left[\exp\left\{-nf\left(\int_{\mathbb{R}^d} yL^n(dy)\right)\right\}\right] + \varepsilon \geq \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\bar{L}^n(dy)\right) + \frac{1}{n}\sum_{i=1}^n H(\bar{\mu}_i^n|\lambda)\right]$$

を満たす．また相対エントロピーの凸性により

$$-\frac{1}{n}\log \mathbb{E}\left[\exp\left\{-nf\left(\int_{\mathbb{R}^d} yL^n(dy)\right)\right\}\right] + \varepsilon \geq \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\bar{L}^n(dy)\right) + H(\widehat{\mu}^n|\lambda)\right]$$

を得る．f は有界なので $\{\mathbb{E}[H(\widehat{\mu}^n|\lambda)]\}_{n\in\mathbb{N}}$ は有界である．よって補題 5.3.1 により，左辺の下極限を実現する $\{(\bar{L}^n, \widehat{\mu}^n)\}_{n\in\mathbb{N}}$ の部分列であって，ある $\mathcal{M}_1(S)^2$ 値確率変数 $(\bar{L}, \widehat{\mu})$ に弱収束するものが存在する．部分列を再び同じ記号で表すことにすると，f は有界連続なので命題 5.3.4 により

$$\lim_{n\to\infty} \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\bar{L}^n(dy)\right)\right] = \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\bar{L}(dy)\right)\right].$$

が成立する．命題 5.3.4 の証明中で見たように，確率 1 で $\int_{\mathbb{R}^d} |y|\bar{L}(dy) < \infty$ であることに注意する．また $H(\cdot|\lambda)$ の下半連続性と補題 5.0.4 により

$$\varliminf_{n\to\infty} \mathbb{E}[H(\widehat{\mu}^n|\lambda)] \geq \mathbb{E}[H(\widehat{\mu}|\lambda)]$$

が成立する．補題 5.3.1 により $\bar{L} = \widehat{\mu}$ が確率 1 で成立するので

$$\varliminf_{n\to\infty} -\frac{1}{n}\log \mathbb{E}\left[\exp\left\{-nf\left(\int_{\mathbb{R}^d} yL^n(dy)\right)\right\}\right] + \varepsilon \geq \mathbb{E}\left[f\left(\int_{\mathbb{R}^d} y\widehat{\mu}(dy)\right) + H(\widehat{\mu}|\lambda)\right]$$

が成立する．先に注意したように確率 1 で $\int_{\mathbb{R}^d}|y|\widehat{\mu}(dy)<\infty$ なので，右辺は下から

$$\inf\left\{f\left(\int_{\mathbb{R}^d}y\mu(dy)\right)+H(\mu|\lambda)\ \middle|\ \mu\in\mathcal{M}_1(\mathbb{R}^d),\int_{\mathbb{R}^d}|y|\mu(dy)<\infty\right\}$$

と評価できる．さらに $\int_{\mathbb{R}^d}y\mu(dy)=x$ と場合わけをすれば，この下限は I の定義により

$$\begin{aligned}&\inf_{x\in\mathbb{R}^d}\inf\left\{f(x)+H(\mu|\lambda)\ \middle|\ \mu\in\mathcal{M}_1(\mathbb{R}^d),\int_{\mathbb{R}^d}|y|\mu(dy)<\infty,\int_{\mathbb{R}^d}y\mu(dy)=x\right\}\\&\quad=\inf_{x\in\mathbb{R}^d}[f(x)+I(x)]\end{aligned}$$

と書くことができる．$\varepsilon>0$ は任意なので (5.5.4) の下からの評価を得る．

次に (5.5.4) の上からの評価を示す．$\varepsilon>0$ を任意に取る．下からの評価の証明の最後に行った下限の書き換えに注意して，$\mu^*\in\mathcal{M}_1(\mathbb{R}^d)$ を次を満たすように取る．

$$\int_{\mathbb{R}^d}|y|\mu^*(dy)<\infty,\quad f\left(\int_{\mathbb{R}^d}y\mu^*(dy)\right)+H(\mu^*|\lambda)\le\inf_{x\in\mathbb{R}^d}[f(x)+I(x)]+\varepsilon.$$

ここで $(\bar{X}^n,\bar{\mu}^n)$ を $(\mu^*)^{\otimes n}$ から定まる制御とすれば，$\bar{X}_1^n,\dots,\bar{X}_n^n$ は独立同分布で $\bar{X}_1^n$ の分布は μ^* である．さらに任意の $n\in\mathbb{N}$ と $i=1,\dots,n$ に対して $\bar{\mu}_i^n=\mu^*$ が成立する．このとき補題 5.3.1 と補題 5.3.2 により $\{\bar{L}^n\}_{n\in\mathbb{N}}$ は $n\to\infty$ のとき μ^* に弱収束する．よって (5.5.5) と命題 5.3.4 により

$$\begin{aligned}\varlimsup_{n\to\infty}-\frac{1}{n}\log\mathbb{E}\left[\exp\left\{-nf\left(\int_{\mathbb{R}^d}yL^n(dy)\right)\right\}\right]&\le\varlimsup_{n\to\infty}\mathbb{E}\left[f\left(\int_{\mathbb{R}^d}y\bar{L}^n(dy)\right)+H(\mu^*|\lambda)\right]\\&=f\left(\int_{\mathbb{R}^d}y\mu^*(dy)\right)+H(\mu^*|\lambda)\\&\le\inf_{x\in\mathbb{R}^d}[f(x)+I(x)]+\varepsilon\end{aligned}$$

が成立する．$\varepsilon>0$ は任意なので (5.5.4) の上からの評価を得る．以上により定理 5.0.2 の証明が終了する．　□

第6章　一般化されたSanovの定理に対する情報理論的証明

本章では一般化された Sanov の大偏差原理に情報理論的な証明を与える.[1] この証明は第 4 章で紹介した標準的な証明とは全く異なる発想に基づいており，特に Donsker-Varadhan の変分公式（定理 4.1.2 および定理 4.1.3）を用いない．非常に一般的に定式化されているにもかかわらず，測度論的に深い結果は証明中では使われず，初等的と言ってよい議論だけで証明が完成してしまう点がこの方法の魅力である．

本章では $(S,\mathcal{F})$ を可測空間とする．基本的に S 上には位相を導入しない．$(S,\mathcal{F})$ 上の確率測度全体を $\mathcal{M}_1(S)$ と表し，$\lambda \in \mathcal{M}_1(S)$ を任意に選ぶ．また S 上で定義された実数値の有界 $\mathcal{F}$ 可測関数の全体を $M_{\mathrm{b}}(S)$ と表す．$f \in M_{\mathrm{b}}(S)$ に対して

$$\mathcal{I}_f(\mu) := \int_S f(x)\mu(dx), \qquad \mu \in \mathcal{M}_1(S)$$

と表し，$\mathcal{I}_f\colon \mathcal{M}_1(S) \to \mathbb{R}$ という写像だとみなす.

$\mathcal{M}_1(S)$ 上には τ 位相を入れる．これは端的に言えば全ての $\mathcal{I}_f\colon \mathcal{M}_1(S) \to \mathbb{R}$ $(f \in M_{\mathrm{b}}(S))$ を連続にする位相のうち最弱なものである．厳密には以下のように各 $\mu \in \mathcal{M}_1(S)$ の基本近傍系を定めるとよい．$m \in \mathbb{N},\ f_1,\dots,f_m \in M_{\mathrm{b}}(S)$ と $\varepsilon \in (0,\infty)$ に対して

$$V(\mu; f_1,\dots,f_m,\varepsilon) = \left\{ \nu \in \mathcal{M}_1(S) \;\middle|\; \max_{1\le i\le m} |\mathcal{I}_{f_i}(\nu) - \mathcal{I}_{f_i}(\mu)| < \varepsilon \right\} \tag{6.0.1}$$

とおくと，

$$\{V(\mu; f_1,\dots,f_m,\varepsilon) \mid m \in \mathbb{N},\ f_1,\dots,f_m \in M_{\mathrm{b}}(S),\ \varepsilon \in (0,\infty)\} \tag{6.0.2}$$

が τ 位相に関する μ の基本近傍系を定める.[2] （なお S が位相空間で $\mathcal{F}$ がその位相に関する Borel 加法族の場合は，τ 位相は明らかに確率測度の弱収束の位相より強い.）τ 位相に関する $\mathcal{M}_1(S)$ 上の Borel 加法族を $\mathcal{G}^\tau$ と書く．$\mathcal{G}^\tau$ は扱いづらいので，むしろその部分 σ 加法族である $\mathcal{G}^{\mathrm{cyl}} := \sigma(\{\mathcal{I}_f \mid f \in M_{\mathrm{b}}(S)\})$ を用いる．$\mathcal{G}^{\mathrm{cyl}}$ は全ての $\mathcal{I}_f$ $(f \in M_{\mathrm{b}}(S))$ を可測にする $\mathcal{M}_1(S)$ 上の σ 加法族のうち最小のものである.

---

[1] 本章の記述は [50] を参考にした.

[2] τ 位相は定義の仕方から Hausdorff であるが，一般には距離付け可能ではないため，点列を使ってこの位相を論ずることはできない.

ここで S の有限分割を導入する．以下の 3 条件を満たすときに $\mathcal{A}=(A_1,\ldots,A_k)$ が S の長さ $k\in\mathbb{N}$ の $\mathcal{F}$ 可測分割であるという．(i) 各 $i=1,\ldots,k$ に対して $A_i\in\mathcal{F}$，(ii) $\{A_i\}_{i=1}^k$ は互いに素，(iii) $\cup_{i=1}^k A_i=S$．なお $(A_1,\ldots,A_k)$ には順序を与えて考えていることに注意せよ.[3] $k=|\mathcal{A}|$ とも書く．$\mathcal{A}'$ が S の $\mathcal{F}$ 可測分割であるとは，ある $k\in\mathbb{N}$ に対して $\mathcal{A}'$ が S の長さ k の $\mathcal{F}$ 可測分割であることをいう．S の長さ k の $\mathcal{F}$ 可測分割全体を $\mathbf{Prt}_k(S)$ と書き，S の $\mathcal{F}$ 可測分割全体を $\mathbf{Prt}(S)$ と書くと，明らかに $\mathbf{Prt}(S)=\cup_{k\in\mathbb{N}}\mathbf{Prt}_k(S)$ である．(以下では簡単のために，$\mathbf{Prt}_k(S)$ の元と $\mathbf{Prt}(S)$ の元をそれぞれ単に「長さ k の分割」「分割」と書くことにする.）分割 $\mathcal{E}=(E_1,\ldots,E_l)$ が分割 $\mathcal{A}=(A_1,\ldots,A_k)$ の細分であるとは，各 j $(1\le j\le l)$ に対して $E_j\subset A_i$ となる i $(1\le i\le k)$ が存在することをいう．2 つの分割 $\mathcal{A}=(A_1,\ldots,A_k)$ と $\mathcal{A}'=(A'_1,\ldots,A'_m)$ に対して，$\{A_i\cap A'_j \mid 1\le i\le k, 1\le j\le m\}$ を任意の順番で並べたものは $\mathcal{A}$ と $\mathcal{A}'$ の共通の細分である．

$k\in\mathbb{N}$ に対して $[\![1,k]\!]:=\{1,\ldots,k\}$ と表し，任意の部分集合が可測であるような σ 加法族を $[\![1,k]\!]$ 上に導入する．通常どおり，$\overline{p}\in\mathcal{M}_1([\![1,k]\!])$ は k 次元確率ベクトルと見なす．すなわち

$$\mathcal{M}_1([\![1,k]\!])=\{\overline{p}=(p_1,\ldots,p_k)\in\mathbb{R}^k \mid p_1,\ldots,p_k\in[0,1],\ p_1+\cdots+p_k=1\}$$

である．この集合には $\mathbb{R}^k$ からの相対位相を与えるが，それは当然 τ 位相と一致する．

さて $\overline{p}=(p_1,\ldots,p_k),\ \overline{q}=(q_1,\ldots,q_k)\in\mathcal{M}_1([\![1,k]\!])$ に対して

$$D_k(\overline{p}|\overline{q})=\sum_{i=1}^k p_i\log\frac{p_i}{q_i}$$

および

$$D_k(\overline{p})=-\sum_{i=1}^k p_i\log p_i$$

と定める．なお本章を通じて $0\log 0=0,\ 0\log\frac{0}{0}=0$ および $a\log\frac{a}{0}=\infty$（$0<a<\infty$ のとき）と約束する．Jensen の不等式を用いて，常に $D_k(\overline{p}|\overline{q})\in[0,\infty]$ であることが示せる.（D_k は $[\![1,k]\!]$ 上の相対エントロピーと一致する.）

長さ k の分割 $\mathcal{A}=(A_1,\ldots,A_k)$ が与えられたとき，$\mu\in\mathcal{M}_1(S)$ に対して $\overline{\mu}_{\mathcal{A}}=(\mu(A_1),\ldots,\mu(A_k))$ とおくと $\mathcal{M}_1([\![1,k]\!])$ の元が得られるので，これを用いて情報ダイバージェンスを $\mu,\nu\in\mathcal{M}_1(S)$ に対して

$$D(\mu|\nu)=\sup_{\mathcal{A}\in\mathbf{Prt}(S)}D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}})=\sup_{\mathcal{A}\in\mathbf{Prt}(S)}\sum_{i=1}^{|\mathcal{A}|}\mu(A_i)\log\frac{\mu(A_i)}{\nu(A_i)}$$

[3] つまりこの $\mathcal{A}=(A_1,\ldots,A_k)$ は $\mathcal{F}^k$（$\mathcal{F}$ の k 重直積集合）に属する元だと見なしており，$\mathcal{F}$ の k 点部分集合だとは見なしていない．中括弧ではなく丸括弧を用いるのもこの理由による．

と定める．一方で第 4 章の冒頭で導入した相対エントロピー $H(\mu|\nu)$ は一般の可測空間 $(S,\mathcal{F})$ 上でも同様に定義できる.（注意 4.0.2 もそのまま成立する.）すなわち $\mu,\nu\in\mathcal{M}_1(S)$ に対して

$$H(\mu|\nu):=\begin{cases}\displaystyle\int_S \frac{d\mu}{d\nu}\log\left(\frac{d\mu}{d\nu}\right)d\nu, & \mu\text{ が }\nu\text{ に対して絶対連続なとき},\\ \infty, & \text{それ以外のとき},\end{cases}$$

と定める．実は D と H は一致するのだが，本章では H ではなく D が主役を務める．

上記の状況において，D と H に関して以下の 2 つの命題が成立する.（証明は後出の第 6.1 節と第 6.2 節で与えるが，どちらもそれほど難しくない.）

命題 6.0.1. 任意の $\mu,\nu\in\mathcal{M}_1(S)$ に対して，$D(\mu|\nu)=H(\mu|\nu)$ が成り立つ.[4]

命題 6.0.2. 任意の $\nu\in\mathcal{M}_1(S)$ に対して，$D(\,\cdot\,|\nu)\colon\mathcal{M}_1(S)\to[0,\infty]$ は τ 位相に関して良い速度関数である．

本章の主定理を述べる前に経験平均分布を導入する．$n\in\mathbb{N}$ のとき，$(S^n,\mathcal{F}^{\otimes n})$ を $(S,\mathcal{F})$ の可測空間としての n 直積とする．$\mu\in\mathcal{M}_1(S)$ の n 直積測度を $\mu^{\otimes n}$ と表すが，これは $(S^n,\mathcal{F}^{\otimes n})$ 上の確率測度である．任意の $n\in\mathbb{N}$ と $\mathbf{x}=(x_1,\ldots,x_n)\in S^n$ に対して

$$\ell_{\mathbf{x}}(A)=\frac{|\{j\in[\![1,n]\!]\mid x_j\in A\}|}{n},\qquad A\in\mathcal{F}$$

とおくと,[5] 明らかに $\ell_{\mathbf{x}}\in\mathcal{M}_1(S)$ であるので，$\mathbf{x}\mapsto\ell_{\mathbf{x}}$ は S^n から $\mathcal{M}_1(S)$ への写像である．非常に一般的な設定で議論しているために，残念ながらこれが $\mathcal{F}^{\otimes n}/\mathcal{G}^\tau$ 可測写像である保証はない.[6] しかし，悪くない可測性を持つ $\mathcal{M}_1(S)$ の部分集合 Γ に対しては，上記の写像が何らか意味で可測性を持ちそうなものである．これに関しては次の事実がある．$\Gamma\in\mathcal{G}^{\mathrm{cyl}}$ のとき，$\{\mathbf{x}\in S^n\mid\ell_{\mathbf{x}}\in\Gamma\}\in\mathcal{F}^{\otimes n}$ が成り立つ．これを実際に確認しよう．$\mathscr{Z}:=\{\Delta\subset\mathcal{M}_1(S)\mid\ell^{-1}(\Delta)\in\mathcal{F}^{\otimes n}\}$ とおくと $\mathscr{Z}$ は自動的に σ 加法族になるので，$\mathscr{Z}\supset\mathcal{G}^{\mathrm{cyl}}$ を示せば十分であるが，そのためには各 $f\in M_{\mathrm{b}}(S)$ に対して $\mathscr{Z}\supset\sigma(\mathcal{I}_f)$ を見ればよい．写像

$$\mathbf{x}\in S^n\mapsto\mathcal{I}_f(\ell_{\mathbf{x}})=\frac{f(x_1)+\cdots+f(x_n)}{n}\in\mathbb{R}$$

[4] [7, 第 3.7 節] ではこの命題は GKY(Gelfand-Kolmogorov-Yaglom) の定理，D は GKY 相対エントロピーと呼ばれている．しかし別の文献では同じ命題に Gelfand-Yaglom-Perez の名前を冠している．

[5] 右辺は要するに $n^{-1}(\delta_{x_1}+\cdots+\delta_{x_n})(A)$ である．1 点集合が可測でないときも Dirac 測度 δ_x $(x\in S)$ を定義することはできる．

[6] 通常の Sanov の定理の記述のように，経験分布を $n^{-1}(\delta_{X_1}+\cdots+\delta_{X_n})$ と独立同分布列 $\{X_i\}$ を用いて書くと，これが $\mathcal{M}_1(S)$ 値確率変数（すなわち確率空間からの可測写像）であるような誤解を与える可能性が高いため，本章ではこの書き方を避ける．

は明らかに $\mathcal{F}^{\otimes n}/\mathcal{B}(\mathbb{R})$ 可測なので，$C \in \mathcal{B}(\mathbb{R})$ であれば $\ell^{-1}(\mathcal{I}_f^{-1}(C)) \in \mathcal{F}^{\otimes n}$ である．これで $\mathcal{Z} \supset \sigma(\mathcal{I}_f) = \{\mathcal{I}_f^{-1}(C) \mid C \in \mathcal{B}(\mathbb{R})\}$ が得られた．

それでは本章の主題である一般化された Sanov の大偏差原理を述べる．この定理は定義 1.1.2 の意味での大偏差原理ではないことに注意せよ.[7]

定理 6.0.3. $\lambda \in \mathcal{M}_1(S)$ とする．このとき，任意の $\Gamma \in \mathcal{G}^{\mathrm{cyl}}$ に対して，

$$\begin{aligned} -\inf_{\mu \in \Gamma^\circ} D(\mu|\lambda) &\leq \varliminf_{n\to\infty} \frac{1}{n} \log \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_{\mathbf{x}} \in \Gamma\}) \\ &\leq \varlimsup_{n\to\infty} \frac{1}{n} \log \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_{\mathbf{x}} \in \Gamma\}) \leq -\inf_{\mu \in \overline{\Gamma}} D(\mu|\lambda) \end{aligned}$$

が成立する．ここで Γ° と $\overline{\Gamma}$ はそれぞれ Γ の τ 位相に関する内部と閉包である．

注意 6.0.4. S が可分完備距離空間で $\mathcal{F} = \mathcal{B}(S)$ のときは実は $\mathcal{G}^{\mathrm{cyl}} = \mathcal{B}(\mathcal{M}_1(S))$ となるため，上記の定理 6.0.3 は通常の Sanov の大偏差原理（定理 4.0.1）を導く．

以下では $\mathcal{H} := \sigma(\{\mathcal{I}_f \mid f \in C_{\mathrm{b}}(S)\})$ とおき，$\mathcal{G}^{\mathrm{cyl}} = \mathcal{H} = \mathcal{B}(\mathcal{M}_1(S))$ となることを手短かに確認する．まず $f \in C_{\mathrm{b}}(S)$ のときは，弱収束の定義により $\mathcal{I}_f$ は連続なので，$\mathcal{B}(\mathcal{M}_1(S))/\mathcal{B}(\mathbb{R})$ 可測である．よって $\mathcal{B}(\mathcal{M}_1(S)) \supset \mathcal{H}$ である．逆向きの包含関係を見るには，$\mathcal{M}_1$ 上の距離の定義式 (4.0.1) を思い出そう．この式の右辺は $C_{\mathrm{b}}(S)$ の元のみで書けているため，$\mathcal{M}_1(S)$ の任意の球は $\mathcal{H}$ 可測集合である．$\mathcal{M}_1(S)$ は可分距離空間なので，第 2 可算公理（または Lindelöf 性）により，任意の開集合は球の可算和の形で書けるため $\mathcal{H}$ 可測集合である．これで $\mathcal{B}(\mathcal{M}_1(S)) \subset \mathcal{H}$ が示せた．

次は $\mathcal{G}^{\mathrm{cyl}} = \mathcal{H}$ を示すが，左辺の方が大きいことは自明なので，$\mathcal{G}^{\mathrm{cyl}} \subset \mathcal{H}$ のみを見ればよい．$\mathcal{Z} = \{f \in M_{\mathrm{b}}(S) \mid \mathcal{I}_f \colon \mathcal{M}_1(S) \to \mathbb{R}$ は $\mathcal{H}$ 可測$\}$ とおくと，明らかに $\mathcal{Z} \supset C_{\mathrm{b}}(S)$ である.（特に定数関数は $\mathcal{Z}$ に属す.）さらに $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{Z}$ が一様有界かつ各点で広義単調増加ならば，単調収束定理または有界収束定理により，各 $\mu \in \mathcal{M}_1(S)$ に対して $\lim_{n\to\infty} \mathcal{I}_{f_n}(\mu) = \mathcal{I}_{f_\infty}(\mu)$ となる．ここで f_∞ は $\{f_n\}$ の各点収束先である．よって $\mathcal{I}_{f_\infty}$ は $\mathcal{H}$ 可測関数列 $\{\mathcal{I}_{f_n}\}$ の各点収束先なので $\mathcal{H}$ 可測関数である．これは $f_\infty \in \mathcal{Z}$ を意味する．$C_{\mathrm{b}}(S)$ が代数であるために単調族定理の関数版が使えて，$\mathcal{Z}$ は全ての $\sigma(C_{\mathrm{b}}(S))$ 可測な有界関数を含む．ここで全ての $f \in C_{\mathrm{b}}(S)$ を可測にするような S 上の σ 加法族のうち最小のものを $\sigma(C_{\mathrm{b}}(S))$ と書いた．S は距離空間なので $\sigma(C_{\mathrm{b}}(S)) = \mathcal{B}(S)$ であることを思い出すと，$\mathcal{G}^{\mathrm{cyl}} = \mathcal{H}$ の証明が終わる.[8]以上で注意 6.0.4 を終わる．

---

[7]本書では定義 1.1.2 を大偏差原理の定義としたが，定義を一般化して扱う流派も存在する．例えば [33, p. 5] における定義の意味では定理 6.0.3 も大偏差原理と呼ばれる．

[8]$\sigma(C_{\mathrm{b}}(S)) = \mathcal{B}(S)$ は「S の任意の閉集合の定義関数がある有界連続関数列の各点収束の極限になる」ことから直ちに従う．

6.1　命題 6.0.1 の証明

本節の目的は命題 6.0.1（すなわち情報ダイバージェンス D と相対エントロピー H の一致）を証明することである.[9] まず簡単な補題を 1 つ示した後，命題 6.0.1 を証明する.

補題 6.1.1. $\mu, \nu \in \mathcal{M}_1(S)$ とし，$A \in \mathcal{F}$ は ν 零集合でないとする．このとき，

$$\mu(A) \log \frac{\mu(A)}{\nu(A)} \leq \int_A \frac{d\mu}{d\nu} \log\left(\frac{d\mu}{d\nu}\right) d\nu$$

が成立する.

証明. 右辺が有限の場合のみ示せばよい．$F := d\mu/d\nu$ と略記する．$t \mapsto t\log t$ は凸関数で $F \log F$ は ν 可積分なので，確率測度 $\nu(A)^{-1}\mathbf{1}_A d\nu$ に対して Jensen の不等式を用いると

$$\left(\frac{1}{\nu(A)}\int_A F d\nu\right) \log\left(\frac{1}{\nu(A)}\int_A F d\nu\right) \leq \frac{1}{\nu(A)}\int_A F \log F d\nu$$

を得る．分母の $\nu(A)$ を払った後で左辺を整理すると，示すべき式の左辺を得るので，補題の証明が終わる．□

命題 6.0.1 の証明. まず μ が ν に対して絶対連続でない場合を考える．このとき $\mu(A) > 0$ かつ $\nu(A) = 0$ となる $A \in \mathcal{F}$ が存在する．このとき $\mu(A)\log[\mu(A)/\nu(A)] = \infty$ なので，分割 $\mathcal{A} = (A, A^c)$ を考えれば $D(\mu|\nu) = \infty$ が直ちにわかる.

ここからは μ が ν に対して絶対連続な場合を考える．このとき $F := d\mu/d\nu \in L^1(\nu)$ なので，F の代表元として各点で定義された $[0, \infty)$ 値の $\mathcal{F}$ 可測関数が取れる．以下ではこのような代表元を 1 つ選び，再び同じ記号で表す.

さて $\mathcal{A} = (A_1, \ldots, A_k)$ を任意の分割とする．$\nu(A_i) = 0$ のときは絶対連続性により $\mu(A_i) = 0$ であり，$\nu(A_i) > 0$ のときは補題 6.1.1 が使えるので，

$$\begin{aligned}\sum_{i=1}^k \mu(A_i) \log \frac{\mu(A_i)}{\nu(A_i)} &= \sideset{}{'}\sum_{1\leq i\leq k} \mu(A_i) \log \frac{\mu(A_i)}{\nu(A_i)} \\ &\leq \sideset{}{'}\sum_{1\leq i\leq k} \int_{A_i} F \log F d\nu \leq \int_S F \log F d\nu\end{aligned}$$

を得る．ここで $\sum_i'$ は条件 $\nu(A_i) > 0$ を満たす $i \in [\![1, k]\!]$ に関して和を取ることを意味する．ここで $0 \log \frac{0}{0} = 0$ という規約を用いた．$\mathcal{A}$ に関して上限を取ると $D(\mu|\nu) \leq H(\mu|\nu)$ を得る.

[9] 本節の記述は [7, 第 3.7 節] を参考にした.

逆向きの不等式 $H(\mu|\nu) \le D(\mu|\nu)$ を示す．各 $m \in \mathbb{N}$ に対して，値域を幅 2^{-m} の小区間で区切ることにより F を近似する．具体的には

$$\begin{aligned} A_{m,i} &:= \{x \in S \mid i2^{-m} \le F(x) < (i+1)2^{-m}\} \qquad (0 \le i \le m2^m - 1), \\ A_{m,m2^m} &:= \{x \in S \mid F(x) \ge m\} \end{aligned}$$

とおくと，各 m に対して $(A_{m,i})_{0 \le i \le m2^m}$ は S の分割である．$m \to \infty$ のとき，明らかに $\nu(A_{m,m2^m}) \searrow 0$ となる．

$0 \le i \le m2^m - 1$ のとき，定義により

$$i2^{-m}\nu(A_{m,i}) \le \int_{A_{m,i}} F\, d\nu = \mu(A_{m,i}) \le (i+1)2^{-m}\nu(A_{m,i})$$

となる．よって $\nu(A_{m,i}) > 0$ であれば

$$i2^{-m} \le \frac{\mu(A_{m,i})}{\nu(A_{m,i})} \le (i+1)2^{-m}$$

となる．したがって $0 \le i \le m2^m - 1$ かつ $\nu(A_{m,i}) > 0$ のとき

$$\left| F(x) - \frac{\mu(A_{m,i})}{\nu(A_{m,i})} \right| \le 2^{-m}, \qquad x \in A_{m,i}$$

が成り立つ．

近似関数列を以下のように定める．$\nu(A_{m,i}) > 0$ を満たす i $(0 \le i \le m2^m)$ に対しては，$A_{m,i}$ 上で $F_m(x) = \mu(A_{m,i})/\nu(A_{m,i})$ と定め，それ以外の i に対しては $A_{m,i}$ 上で $F_m(x) = i2^{-m}$ と定める．すると $\{x \in S \mid F(x) < m\}$ 上で $|F(x) - F_m(x)| \le 2^{-m}$ が満たされているため，各 $x \in S$ に対して $\lim_{m\to\infty} F_m(x) = F(x)$ と収束する．よって特に $-1/e \le F_m \log F_m \to F \log F$ と各点収束するため，Fatouの補題を用いて次の評価を得る．

$$\begin{aligned} \int_S F \log F d\nu &\le \varliminf_{m\to\infty} \int_S F_m \log F_m\, d\nu \\ &= \varliminf_{m\to\infty} \sum_{i=0}^{m2^m} \mu(A_{m,i}) \log \frac{\mu(A_{m,i})}{\nu(A_{m,i})} \le D(\mu|\nu). \end{aligned}$$

なお i に関する和において，$\nu(A_{m,i}) = 0$ となる i に対応する項は消えていることに注意せよ.（実際，絶対連続性により $\mu(A_{m,i}) = 0$ なので，$0 \log \frac{0}{0} = 0$ という約束からそれが従う.） □

6.2　命題6.0.2の証明

本節の目的は命題6.0.2（すなわち $D(\,\cdot\,|\nu)$ が τ 位相に関して良い速度関数であること）を証明することである．

まず τ 位相に関する基本近傍系として，(6.0.1)–(6.0.2) に与えたものとは別のものを与える．各 $\mu \in \mathcal{M}_1(S)$ の新しい基本近傍系を以下のように定める．$m \in \mathbb{N}$, $A_1, \dots, A_m \in \mathcal{F}$ と $\varepsilon \in (0, \infty)$ に対して

$$U(\mu; A_1, \dots, A_m, \varepsilon) = \left\{ \nu \in \mathcal{M}_1(S) \;\middle|\; \max_{1 \le i \le m} |\nu(A_i) - \mu(A_i)| < \varepsilon \right\} \tag{6.2.1}$$

とおくと，

$$\{U(\mu; A_1, \dots, A_m, \varepsilon) \mid m \in \mathbb{N},\ A_1, \dots, A_m \in \mathcal{F},\ \varepsilon \in (0, \infty)\} \tag{6.2.2}$$

は基本近傍系の公理を満たす．これで決まる位相は任意の $A \in \mathcal{F}$ に対して写像 $\mu \mapsto \mu(A)$ を連続にする位相のうち最弱なものである．

この新しい基本近傍系が結局は τ 位相を生成することを確認しよう．$A \in \mathcal{F}$ ならば $\mathbf{1}_A \in M_{\mathrm{b}}(S)$ なので，(6.0.1)–(6.0.2) の基本近傍系の方が細かいことは明らかである．したがって，逆に (6.2.1)–(6.2.2) の基本近傍系の方が細かいことを示せばよい．そのためには $\mu \in \mathcal{M}_1(S)$ のとき,「任意の $m \in \mathbb{N}$, $f_1, \dots, f_m \in M_{\mathrm{b}}(S)$ と $\varepsilon \in (0, \infty)$ に対して，ある分割 $\mathcal{A} = (A_1, \dots, A_n)$ と $\delta \in (0, \infty)$ が存在して，

$$U(\mu; A_1, \dots, A_n, \delta) \subset V(\mu; f_1, \dots, f_m, \varepsilon) \tag{6.2.3}$$

となる」ことを見れば十分である．次の段落でこれを示す．

任意の $\mu \in \mathcal{M}_1(S)$ と $f_1, \dots, f_m \in M_{\mathrm{b}}(S)$ と $\varepsilon \in (0, \infty)$ が与えられたとする．有界可測関数は単関数で一様近似できるので，ある分割 $\mathcal{A} = (A_1, \dots, A_n)$ と $c_1^k, \dots, c_n^k \in \mathbb{R}$ が存在して，

$$\max_{1 \le k \le m} \sup_{x \in S} \left| f_k(x) - \sum_{j=1}^{n} c_j^k \mathbf{1}_{A_j}(x) \right| < \frac{\varepsilon}{3}$$

を満たす.[10] もし $\max_{1 \le i \le n} |\nu(A_i) - \mu(A_i)| < \varepsilon/(3nc)$（ただし $c := \max_{j,k} |c_j^k| + 1$ とおく）であれば，標準的な「3ε 論法」を用いて $\max_{1 \le k \le m} |\mathcal{I}_{f_k}(\nu) - \mathcal{I}_{f_k}(\mu)| < \varepsilon$ が得られる．これは $U(\mu; A_1, \dots, A_n, \varepsilon/(3nc)) \subset V(\mu; f_1, \dots, f_m, \varepsilon)$ を意味するので，(6.2.3) が確認できた．以上で 2 つの基本近傍系が同じ位相を生成することがわかった．

確率測度の空間の入れ物として直積空間 $[0,1]^{\mathcal{F}}$ を考える．$B \in \mathcal{F}$ のとき，第 B 座標への射影を $\pi_B\colon [0,1]^{\mathcal{F}} \to [0,1]$ と書く．すなわち，$\pi_B((s_A)_{A \in \mathcal{F}}) = s_B$ である．この直積空間には直積位相を入れるが，この位相は全ての射影 π_B $(B \in \mathcal{F})$ を連続にする位相の中で最弱のものである．Tychonoff の定理により $[0,1]^{\mathcal{F}}$ はコンパクトである.[11] $[0,1]^{\mathcal{F}}$ は自然に $\mathcal{F}$ から $[0,1]$ への写像全体がなす集合 $\mathbf{Map}(\mathcal{F}, [0,1])$ と同一視できるが，その場合は射

[10] ここで分割 $\mathcal{A}$ は k に無関係なものが取れることに注意せよ．実際，各 f_k の単関数近似に現れる分割 $\mathcal{A}^{(k)}$ $(1 \le k \le m)$ の共通の細分を $\mathcal{A}$ とおけばよい．

[11] Tychonoff の定理については [21, 定理 2.28] や [6, 定理 23.3] などを見よ．

影 π_B は写像に点 B を代入する操作，すなわち $\mathbf{Map}(\mathcal{F},[0,1]) \ni \psi \mapsto \psi(B) \in [0,1]$ に相当し，直積位相は各点収束の位相に相当する．$\mathcal{M}_1(S) \subset [0,1]^{\mathcal{F}}$ であるが，(6.2.1)–(6.2.2) の基本近傍系による位相の定義の仕方から τ 位相と $[0,1]^{\mathcal{F}}$ の部分集合としての相対位相は一致する．

$(S,\mathcal{F})$ 上の有限加法的な確率測度全体がなす集合を $\mathcal{M}_1^{\mathrm{FA}}(S)$ と表すと明らかに

$$\mathcal{M}_1(S) \subset \mathcal{M}_1^{\mathrm{FA}}(S) \subset [0,1]^{\mathcal{F}}$$

である．この集合の位相的性質については次の事実が知られている．

補題 6.2.1. $\mathcal{M}_1^{\mathrm{FA}}(S)$ は $[0,1]^{\mathcal{F}}$ のコンパクト部分集合である．

証明. $[0,1]^{\mathcal{F}}$ がコンパクトなので，$\mathcal{M}_1^{\mathrm{FA}}(S)$ が閉集合であることを見ればよい．互いに素な $A, B \in \mathcal{F}$ を任意に選び，$\widetilde{\pi}_{A,B} = \pi_A + \pi_B - \pi_{A\cup B}$ とおくと，$\widetilde{\pi}_{A,B}$ は連続写像である．明らかに $\widetilde{\pi}_{A,B}^{-1}(\{0\})$ は $\mathcal{M}_1^{\mathrm{FA}}(S)$ を包含する閉集合であるので，$\mathcal{M}_1^{\mathrm{FA}}(S)$ の閉包も包含する．これは閉包に属する任意の元 m が $m(A)+m(B) = m(A \sqcup B)$ を満たすことを意味するので，m は有限加法的測度である．同様に $\pi_S^{-1}(\{1\})$ は $\mathcal{M}_1^{\mathrm{FA}}(S)$ を包含する閉集合であるので，$\mathcal{M}_1^{\mathrm{FA}}(S)$ の閉包も包含する．これは $m(S) = 1$ を意味するので，$m \in \mathcal{M}_1^{\mathrm{FA}}(S)$ がわかった． □

$D(\,\cdot\,|\nu)$ が τ 位相に関して良い速度関数であることを証明する．

命題 6.0.2 の証明. $q \in [0,1]$ のとき，$p \in [0,1] \mapsto p\log(p/q)$ は下半連続である．実際，$q \neq 0$ のときは連続関数で，$q = 0$ のときは 2 点集合 $\{0,\infty\}$ に値を取る下半連続関数である．$\mathbb{R}^k$ の各成分への射影が連続なので，$\overline{q} = (q_1,\ldots,q_k) \in \mathcal{M}_1([\![1,k]\!])$ のとき，

$$\mathcal{M}_1([\![1,k]\!]) \ni \overline{p} = (p_1,\ldots,p_k) \mapsto D_k(\overline{p}|\overline{q}) = \sum_{i=1}^{k} p_i \log\frac{p_i}{q_i} \in [0,\infty]$$

も下半連続である．

任意の $\nu \in \mathcal{M}_1(S)$ を取り，以下ではそれを固定して議論する．各分割 $\mathcal{A} = (A_1,\ldots,A_{|\mathcal{A}|})$ に対して，$\mu \mapsto (\mu(A_1),\ldots,\mu(A_{|\mathcal{A}|}))$ の連続性により，写像

$$\mathcal{M}_1^{\mathrm{FA}}(S) \ni \mu \mapsto D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}}) = \sum_{i=1}^{|\mathcal{A}|} \mu(A_i)\log\frac{\mu(A_i)}{\nu(A_i)} \in [0,\infty]$$

は下半連続である．($D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}})$ と $D(\mu|\nu)$ の定義が有限加法的確率測度の場合まで自然に延長することは自明であろう.）したがって，補題 1.2.5 により

$$\mathcal{M}_1^{\mathrm{FA}}(S) \ni \mu \mapsto D(\mu|\nu) = \sup_{\mathcal{A}\in\mathbf{Prt}(S)} D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}}) \in [0,\infty]$$

も下半連続であるため，各 $r \in [0,\infty)$ に対して $\left\{\mu \in \mathcal{M}_1^{\mathrm{FA}}(S) \mid D(\mu|\nu) \le r\right\}$ はコンパクト集合の閉部分集合，よってコンパクトであることがわかる．

$\mathcal{M}_1(S)$ 上の 2 種類の位相が一致することを思い出すと，次の事実を確認すれば証明が完成する．

$$\left\{\mu \in \mathcal{M}_1^{\mathrm{FA}}(S) \mid D(\mu|\nu) \le r\right\} = \{\mu \in \mathcal{M}_1(S) \mid D(\mu|\nu) \le r\}.$$

左辺が右辺を包含することは明らかなので，逆向きの包含のみが問題である．μ が可算加法的でないならば，互いに素な $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{F}$ で $\mu(\cup_{i\in\mathbb{N}} A_i) \neq \sum_{i\in\mathbb{N}} \mu(A_i)$ を満たすものがある．各 $n \in \mathbb{N}$ に対して $C_n := \cup_{i=n}^{\infty} A_i$ とおくと，$n \to \infty$ のとき $C_n \searrow \emptyset$ である．さらに有限加法性により $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{n-1} \mu(A_i) + \mu(C_n)$ なので，$\lim_{n\to\infty} \mu(C_n) > 0$ を満たす．分割 $\mathcal{A}_n := (C_n, C_n^c)$ を考えると，

$$D_2(\overline{\mu}_{\mathcal{A}_n}|\overline{\nu}_{\mathcal{A}_n}) = \mu(C_n)\log\frac{\mu(C_n)}{\nu(C_n)} + \mu(C_n^c)\log\frac{\mu(C_n^c)}{\nu(C_n^c)}$$

となる．ν は可算加法的なので $\lim_{n\to\infty} \nu(C_n) = 0$ となるため，上式の右辺は $n \to \infty$ のとき ∞ に発散する．よって $D(\mu|\nu) = \infty$ である． □

6.3 離散確率論からの準備事項

本節を通じて $n, k \in \mathbb{N}$ とする．$\mathcal{A} = (A_1, \cdots, A_k)$ が長さ k の分割で $\mathbf{x} \in S^n$ のとき，$\overline{(\ell_{\mathbf{x}})}_{\mathcal{A}} = (\ell_{\mathbf{x}}(A_1), \ldots, \ell_{\mathbf{x}}(A_k))$ は k 次元確率ベクトルであるが，その各成分は $n^{-1}\mathbb{N}_0$ に属する．したがって，次のような $\mathcal{M}_1([\![1,k]\!])$ の部分集合を考えるのは自然であろう．

$$\mathcal{P}_n([\![1,k]\!]) := \left\{\overline{p} = \left(\tfrac{n_1}{n}, \ldots, \tfrac{n_k}{n}\right) \;\middle|\; n_1, \ldots, n_k \in \mathbb{N}_0,\ n_1 + \cdots + n_k = n\right\}.$$

各 n_i $(1 \le i \le k)$ が取り得る値は $0, 1, \ldots, n$ の高々$(n+1)$ 個なので，次の不等式はほぼ自明である．

$$|\mathcal{P}_n([\![1,k]\!])| \le (n+1)^k. \tag{6.3.1}$$

さて $\mathcal{P}_n([\![1,k]\!])$ の任意の元 $\overline{p} = \left(\frac{n_1}{n}, \ldots, \frac{n_k}{n}\right)$ が与えられたとする．このとき

$$\mathcal{T}_n(\overline{p}) = \{(a_1, \ldots, a_n) \in [\![1,k]\!]^n \mid \text{各}\ i \in [\![1,k]\!]\ \text{に対して}\ a_l = i\ \text{となる}\ l\ \text{の数が}\ n_i\ \text{個}\}$$

と定める．表現を変えると，

$$(\overbrace{1, \ldots, 1}^{n_1}, \overbrace{2, \ldots, 2}^{n_2}, \ldots\ldots, \overbrace{k, \ldots, k}^{n_k})$$

という長さ n の数列を並べ変えてできる数列を全て集めて作る集合が $\mathcal{T}_n(\overline{p})$ である．したがって，$|\mathcal{T}_n(\overline{p})| = n!/(n_1! \cdots n_k!)$ である．本節の目標は $\mathcal{T}_n(\overline{p})$ の重みに関する次の命題である．(証明は本節の最後の方で与える．)

命題 6.3.1. $\overline{p} = \left(\frac{n_1}{n}, \dots, \frac{n_k}{n}\right) \in \mathcal{P}_n([\![1,k]\!])$ および $\overline{q} \in \mathcal{M}_1([\![1,k]\!])$ とすると，

$$(n+1)^{-k} e^{-nD_k(\overline{p}|\overline{q})} \leq \overline{q}^{\otimes n}(\mathcal{T}_n(\overline{p})) \leq e^{-nD_k(\overline{p}|\overline{q})}$$

が成立する．

上の命題を証明する前に補題をいくつか証明する．

補題 6.3.2. $l \in \mathbb{N}$ かつ $m \in \mathbb{N}_0$ のとき，$m!/l! \geq l^{m-l}$ が成り立つ．

証明. $m \geq l$ のとき，示すべき不等式は $m! \geq l^{m-l} l!$ と同値である．これを書き下すと，

$$\underbrace{m(m-1)\cdots(l+1)}_{m-l} \times l! \geq \underbrace{l \cdots l}_{m-l} \times l!$$

となるので，明らかに成り立つことがわかる．

$m < l$ のとき，示すべき不等式は $l^{l-m} m! \geq l!$ と同値である．これを書き下すと，

$$\underbrace{l \cdots l}_{l-m} \times m! \geq \underbrace{l(l-1)\cdots(m+1)}_{l-m} \times m!$$

となるので，この場合も明らかに成り立つことがわかる． □

補題 6.3.3. $\overline{p} = \left(\frac{n_1}{n}, \dots, \frac{n_k}{n}\right) \in \mathcal{P}_n([\![1,k]\!])$ および $\overline{q} \in \mathcal{M}_1([\![1,k]\!])$ とすると，任意の $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{T}_n(\overline{p})$ に対して

$$\overline{q}^{\otimes n}(\{\mathbf{a}\}) = e^{-n\{D_k(\overline{p}) + D_k(\overline{p}|\overline{q})\}} \tag{6.3.2}$$

が成立する．特に $\overline{q} = \overline{p}$ のときは $D_k(\overline{p}|\overline{q}) = 0$ なので，次が成立する．

$$\overline{p}^{\otimes n}(\{\mathbf{a}\}) = e^{-nD_k(\overline{p})}.$$

証明. $\overline{p} = (p_1, \dots, p_k)$ および $\overline{q} = (q_1, \dots, q_k)$ と表す．全ての q_i が正のときは，単純計算により

$$\begin{aligned}
\overline{q}^{\otimes n}(\{\mathbf{a}\}) &= \overline{q}^{\otimes n}\Big(\{(\overbrace{1,\dots,1}^{n_1}, \overbrace{2,\dots,2}^{n_2}, \dots\dots, \overbrace{k,\dots,k}^{n_k})\}\Big) \\
&= q_1^{n_1} \cdots q_k^{n_k} = q_1^{np_1} \cdots q_k^{np_k} \\
&= \exp\Big(n \sum_{i=1}^{k} p_i \log q_i\Big) \\
&= \exp\Big(-n \sum_{i=1}^{k} (-p_i \log p_i + p_i \log \tfrac{p_i}{q_i})\Big)
\end{aligned}$$

$$= \exp\bigl(-n\{D_k(\overline{p}) + D_k(\overline{p}|\overline{q})\}\bigr)$$

となり，(6.3.2) が得られた．

$q_i = 0$ かつ $p_i = n_i/n > 0$ を満たす i が存在するときは，(6.3.2) の左辺は明らかに 0 で，右辺も $e^{-\infty} = 0$ なので，この場合も (6.3.2) は成り立つ．

$q_i = 0$ ならば $p_i = n_i/n = 0$ となる場合は，$q_i = 0$ となる i に対応する情報は (6.3.2) の左辺には現れず，$0\log 0 = 0\log\frac{0}{0} = 0$ という約束により右辺にも現れない．よって，この場合も (6.3.2) は成り立つ．□

補題 6.3.4. $\overline{p} = \left(\frac{n_1}{n}, \ldots, \frac{n_k}{n}\right) \in \mathcal{P}_n([\![1,k]\!])$ とすると，

$$(n+1)^{-k} e^{nD_k(\overline{p})} \leq |\mathcal{T}_n(\overline{p})| \leq e^{nD_k(\overline{p})}$$

が成り立つ．

証明. $\mathcal{T}_n(\overline{p})$ に属する全ての元は $\overline{p}^{\otimes n}$ に関して同じ重みを持つので，補題 6.3.3 により

$$1 \geq \overline{p}^{\otimes n}(\mathcal{T}_n(\overline{p})) = |\mathcal{T}_n(\overline{p})| e^{-nD_k(\overline{p})}. \tag{6.3.3}$$

となり，上からの評価を得る．

下からの評価を示す．$\overline{p} = (p_1, \ldots, p_k)$ の台を $\Delta(\overline{p}) := \{i \in [\![1,k]\!] \mid p_i > 0\}$ と表す．$\Delta(\overline{r}) \subset \Delta(\overline{p})$ を満たす $\overline{r} = (r_1, \ldots, r_k) \in \mathcal{P}_n([\![1,k]\!])$ を任意に取る．このとき $p_i = 0$ ならば $r_i = 0$ なので，$\mathcal{T}_n(\overline{r})$ に属する数列は i という文字を含まない．したがって，

$$\overline{p}^{\otimes n}(\mathcal{T}_n(\overline{r})) = |\mathcal{T}_n(\overline{r})| \prod_{i\in\Delta(\overline{p})} p_i^{nr_i} = n! \prod_{i\in\Delta(\overline{p})} p_i^{nr_i}/(nr_i)!$$

と求まるので，補題 6.3.2 を用いて

$$\begin{aligned}
\frac{\overline{p}^{\otimes n}(\mathcal{T}_n(\overline{p}))}{\overline{p}^{\otimes n}(\mathcal{T}_n(\overline{r}))} &= \prod_{i\in\Delta(\overline{p})} \frac{(nr_i)!}{(np_i)!}(np_i)^{np_i - nr_i} \cdot n^{nr_i - np_i} \\
&\geq \prod_{i\in\Delta(\overline{p})} n^{nr_i - np_i} \\
&= n^{n\sum_{i\in\Delta(\overline{p})} r_i - n\sum_{i\in\Delta(\overline{p})} p_i} = n^{n(1-1)} = 1
\end{aligned}$$

を得る．よってこの場合は

$$\overline{p}^{\otimes n}(\mathcal{T}_n(\overline{p})) \geq \overline{p}^{\otimes n}(\mathcal{T}_n(\overline{r})) \tag{6.3.4}$$

が示せた．なお $\overline{r} \in \mathcal{P}_n([\![1,k]\!])$ が $\Delta(\overline{r}) \subset \Delta(\overline{p})$ を満たさないときは，$p_i = 0$ かつ $r_i > 0$ となる i が存在するので，(6.3.4) の右辺は 0 である．よって，いずれにせよ (6.3.4) は成り立つ．

$[\![1,k]\!]^n$ が $\mathcal{T}_n(\overline{r})$ たちの非交和に分解するため，

$$1 = \sum_{\overline{r}\in\mathcal{P}_n([\![1,k]\!])} \overline{p}^{\otimes n}(\mathcal{T}_n(\overline{r})) \le |\mathcal{P}_n([\![1,k]\!])|\, \overline{p}^{\otimes n}(\mathcal{T}_n(\overline{p})) \le (n+1)^k |\mathcal{T}_n(\overline{p})|\, e^{-nD_k(\overline{p})}$$

が成り立つ．ここで (6.3.1), (6.3.3), (6.3.4) を用いた．これで求める下からの評価を得る． □

命題 6.3.1 の証明． 補題 6.3.3 から直ちに

$$\overline{q}^{\otimes n}(\mathcal{T}_n(\overline{p})) = |\mathcal{T}_n(\overline{p})| e^{-n\{D_k(\overline{p})+D_k(\overline{p}|\overline{q})\}}$$

が従うので，この式の右辺に補題 6.3.4 を適用すればよい． □

本節の最後に経験平均分布 $\ell_{\mathbf{x}}$ と分割 $\mathcal{A}$ に関する簡単な事実を述べる．

補題 6.3.5. $\mathcal{A} = (A_1,\ldots,A_k)$ を分割，$\overline{p} = \left(\frac{n_1}{n},\ldots,\frac{n_k}{n}\right) \in \mathcal{P}_n([\![1,k]\!])$，$\nu \in \mathcal{M}_1(S)$ とする．このとき次が成り立つ．

$$\nu^{\otimes n}(\{\mathbf{x} = (x_1,\ldots,x_n) \in S^n \mid \overline{(\ell_{\mathbf{x}})}_{\mathcal{A}} = \overline{p}\}) = (\overline{\nu}_{\mathcal{A}})^{\otimes n}(\mathcal{T}_n(\overline{p})).$$

証明. $\mathbf{x}$ に対する条件 $\overline{(\ell_{\mathbf{x}})}_{\mathcal{A}} = \overline{p}$ は各 $i \in [\![1,k]\!]$ に対して $|\{j \in [\![1,n]\!] \mid x_j \in A_i\}| = n_i$ となることと同値である．

$$S^n = \bigcup \{A_{i_1} \times \cdots \times A_{i_n} \mid (i_1,\ldots,i_n) \in [\![1,k]\!]^n\}$$

という S^n の非交和分解により，これらの条件は「ある $(i_1,\ldots,i_n) \in \mathcal{T}_n(\overline{p})$ に対して $\mathbf{x} \in A_{i_1} \times \cdots \times A_{i_n}$」とも同値である．したがって，示すべき式の左辺は次に等しい．

$$\sum_{(i_1,\ldots,i_n)\in\mathcal{T}_n(\overline{p})} \nu^{\otimes n}(A_{i_1} \times \cdots \times A_{i_n}) = \sum_{(i_1,\ldots,i_n)\in\mathcal{T}_n(\overline{p})} \nu(A_{i_1}) \times \cdots \times \nu(A_{i_n}) = (\overline{\nu}_{\mathcal{A}})^{\otimes n}(\mathcal{T}_n(\overline{p})).$$

これで補題の証明が終わった． □

6.4　下からの評価の証明

本節では一般化された Sanov の大偏差原理（定理 6.0.3）の下からの評価を証明する．

定理 6.0.3 の下からの評価の証明． 任意の $\Gamma \in \mathcal{G}^{\mathrm{cyl}}$ に対して以下の不等式を示せば十分である．

$$\varliminf_{n\to\infty} \frac{1}{n} \log \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_{\mathbf{x}} \in \Gamma\}) \ge -D(\mu|\lambda), \qquad \mu \in \Gamma^\circ. \tag{6.4.1}$$

以下では λ と Γ を固定して，この評価を示す．

ここから (6.4.1) の証明を始める．τ 位相に関する近傍 $U(\mu; A_1, \ldots, A_k, \varepsilon)$ の定義に関しては (6.2.1) を見よ．$(A_1, \ldots, A_k) =: \mathcal{A}$ が S の分割をなす場合はこの近傍を単に $U(\mu; \mathcal{A}, \varepsilon)$ と表す．$\mu \in \Gamma^\circ$ を任意に取る．包含 (6.2.3) により，ある分割 $\mathcal{A} = (A_1, \ldots, A_k)$ と $\varepsilon \in (0, \infty)$ が存在して，$U(\mu; \mathcal{A}, \varepsilon) \subset \Gamma^\circ$ となる．この近傍の部分集合

$$W(\mu; \mathcal{A}, \varepsilon) := \{\gamma \in U(\mu; \mathcal{A}, \varepsilon) \mid \mu(A_i) = 0 \text{ となる } i \in [\![1, k]\!] \text{ に対して } \gamma(A_i) = 0\}$$

の重みを下から評価する．

各 $n \in \mathbb{N}$ に対して，$\overline{p(n)} = (p(n)_1, \ldots, p(n)_k) \in \mathcal{P}_n([\![1, k]\!])$ を次の２条件を満たすように選ぶ．

- $\mu(A_i) = 0$ となる $i \in [\![1, k]\!]$ に対して $p(n)_i = 0$.
- 各 $i \in [\![1, k]\!]$ に対して $\lim_{n\to\infty} p(n)_i = \mu(A_i)$.

以下では $\varepsilon_n := \max_{1\le i\le k} |p(n)_i - \mu(A_i)|$ をおく．明らかに $\lim_{n\to\infty} \varepsilon_n = 0$ である．

$\varepsilon_n \le \varepsilon$ のとき，定義から明らかに $\overline{(\ell_\mathbf{x})}_\mathcal{A} = \overline{p(n)}$ ならば $\ell_\mathbf{x} \in W(\mu; \mathcal{A}, \varepsilon)$ なので，

$$\begin{aligned}
\lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_\mathbf{x} \in \Gamma\}) &\ge \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_\mathbf{x} \in W(\mu; \mathcal{A}, \varepsilon)\}) \\
&\ge \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \overline{(\ell_\mathbf{x})}_\mathcal{A} = \overline{p(n)}\}) \\
&= (\overline{\lambda}_\mathcal{A})^{\otimes n}(\mathcal{T}_n(\overline{p(n)})) \\
&\ge (n+1)^{-k} \exp\left[-nD_k(\overline{p(n)}|\overline{\lambda}_\mathcal{A})\right]
\end{aligned}$$

上の３行目で補題 6.3.5 を，４行目で命題 6.3.1 を用いた．D_k の定義と $\overline{p(n)}$ の選び方により，$\lambda(A_i) = 0$ かどうかで場合わけをすると $\lim_{n\to\infty} D_k(\overline{p(n)}|\overline{\lambda}_\mathcal{A}) = D_k(\overline{\mu}_\mathcal{A}|\overline{\lambda}_\mathcal{A})$ が簡単に得られる．したがって

$$\varliminf_{n\to\infty} \frac{1}{n} \log \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_\mathbf{x} \in \Gamma\}) \ge -D_k(\overline{\mu}_\mathcal{A}|\overline{\lambda}_\mathcal{A}) \ge -D(\mu|\lambda)$$

となり，目標とする (6.4.1) が証明できた．以上で下からの評価の証明を終わる．　□

6.5　上からの評価の証明

本節では一般化された Sanov の大偏差原理（定理 6.0.3）の上からの評価を証明する．証明の鍵になるのは次の不等式である．

補題 6.5.1. 任意の $\nu \in \mathcal{M}_1(S)$ と $\Gamma \in \mathcal{G}^{\mathrm{cyl}}$ に対して，次の不等式が成立する．

$$\sup_{\mathcal{A} \in \mathbf{Prt}(S)} \inf\{D_{|\mathcal{A}|}(\overline{\mu}_\mathcal{A}|\overline{\nu}_\mathcal{A}) \mid \mu \in \Gamma\} \ge \inf\{D(\mu|\nu) \mid \mu \in \overline{\Gamma}\}. \tag{6.5.1}$$

この補題の証明は後回しにして，先にこの補題を用いて上からの評価を証明しよう．

定理 6.0.3 の上からの評価の証明． 分割 $\mathcal{A} = (A_1, \ldots, A_k)$ に対して，写像 $\mu \in \mathcal{M}_1(S) \mapsto \overline{\mu}_{\mathcal{A}} \in \mathcal{M}_1([\![1,k]\!])$ を $\rho_{\mathcal{A}}$ と表す．なお $\mu \in \rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))$ であることは，ある $\widetilde{\mu} \in \Gamma$ が存在して各 $i \in [\![1,k]\!]$ に対して $\mu(A_i) = \widetilde{\mu}(A_i)$ を満たすことと同値である．当然 $\rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma)) \supset \Gamma$ である．これを用いて次の評価を得る．

$$
\begin{aligned}
\lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_{\mathbf{x}} \in \Gamma\}) &\leq \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_{\mathbf{x}} \in \rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))\}) \\
&= \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \overline{(\ell_{\mathbf{x}})}_{\mathcal{A}} \in \rho_{\mathcal{A}}(\Gamma) \cap \mathcal{P}_n([\![1,k]\!])\}) \\
&\leq (n+1)^k \max\{(\overline{\lambda}_{\mathcal{A}})^{\otimes n}(\mathcal{T}_n(\overline{p})) \mid \overline{p} \in \rho_{\mathcal{A}}(\Gamma) \cap \mathcal{P}_n([\![1,k]\!])\} \\
&\leq (n+1)^k \exp\left[-n \inf\{D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}} | \overline{\lambda}_{\mathcal{A}}) \mid \mu \in \Gamma\}\right].
\end{aligned}
$$

下から2つ目の不等号では (6.3.1) と補題 6.3.5 を，最後の不等号では命題 6.3.1 を用いた．したがって，

$$
\begin{aligned}
\varlimsup_{n\to\infty} \frac{1}{n} \log \lambda^{\otimes n}(\{\mathbf{x} \in S^n \mid \ell_{\mathbf{x}} \in \Gamma\}) &\leq \inf_{\mathcal{A} \in \mathbf{Prt}(S)} \left[-\inf\{D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}} | \overline{\lambda}_{\mathcal{A}}) \mid \mu \in \Gamma\}\right] \\
&= -\sup_{\mathcal{A} \in \mathbf{Prt}(S)} \inf\{D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}} | \overline{\lambda}_{\mathcal{A}}) \mid \mu \in \Gamma\} \\
&\leq -\inf\{D(\mu|\lambda) \mid \mu \in \overline{\Gamma}\}
\end{aligned}
$$

が成り立つ．最後の不等号で補題 6.5.1 を用いた．これで目標とする定理 6.0.3 の上からの評価が証明された． □

補題 6.5.1 の証明． 本証明中では $\alpha \in [0,\infty)$ とし，ダイバージェンス球を $K(\nu,\alpha) := \{\mu \in \mathcal{M}_1(S) \mid D(\mu|\nu) \leq \alpha\}$ と表す．命題 6.0.2 で見たように，$K(\nu,\alpha)$ は τ 位相に関してコンパクトである．

不等式 (6.5.1) を示すためには，

$$
\alpha > \sup_{\mathcal{A} \in \mathbf{Prt}(S)} \inf\{D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}} | \overline{\nu}_{\mathcal{A}}) \mid \mu \in \Gamma\} \quad \Longrightarrow \quad \overline{\Gamma} \cap K(\nu,\alpha) \neq \emptyset \tag{6.5.2}
$$

を確認すれば十分である．以下でこのことを示す．

【第1段】まず本段では

$$
\overline{\Gamma} = \bigcap_{\mathcal{A} \in \mathbf{Prt}(S)} \overline{\rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))}
$$

を示す．$\Gamma \subset \rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))$ なので，左辺が右辺に包含されることは明らかである．逆向きの包含が問題だが，それは次の条件と同値である．「任意の分割 $\mathcal{A}$ に対して $\mu \in \overline{\rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))}$ ならば，任意の分割 $\mathcal{A}$ と $\varepsilon \in (0,\infty)$ に対して $U(\mu;\mathcal{A},\varepsilon) \cap \Gamma \neq \emptyset$ となる．」（ここで (6.2.3) を暗に用いている．）

上記の鉤括弧内の条件を示す．$\varepsilon \in (0,\infty)$ を任意とする．任意の分割 $\mathcal{A}=(A_1,\dots,A_k)$ に対して，$U(\mu;\mathcal{A},\varepsilon)\cap\rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))\neq\emptyset$ なので，ある μ' がこの集合に属する．このとき，ある $\widetilde{\mu}\in\Gamma$ が存在して各 $i\in[\![1,k]\!]$ に対して $\mu'(A_i)=\widetilde{\mu}(A_i)$ を満たす．$U(\mu;\mathcal{A},\varepsilon)$ の定義により，$\widetilde{\mu}\in U(\mu;\mathcal{A},\varepsilon)$ とわかるので，$U(\mu;\mathcal{A},\varepsilon)\cap\Gamma\neq\emptyset$ が得られた．以上で第１段を終わる．

【第２段】続いて本段では

$$\begin{aligned}\alpha &> \sup_{\mathcal{A}\in\mathbf{Prt}(S)}\inf\{D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}})\mid \mu\in\Gamma\}\\ &\Longrightarrow\quad \text{任意の分割}\,\mathcal{A}\,\text{に対して}\ \rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))\cap K(\nu,\alpha)\neq\emptyset\end{aligned}$$

を示す．

左側の条件を仮定すると，任意の分割 $\mathcal{A}=(A_1,\dots,A_k)$ に対して，ある $\mu\in\Gamma$ が存在して，$D_k(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}})<\alpha$ を満たす．特に $D_k(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}})$ は有限なので，$\nu(A_i)=0$ のときは $\mu(A_i)=0$ であることに注意せよ．この μ から別の測度 $\gamma\in\mathcal{M}_1(S)$ を以下のように定める．

$$\gamma(C)=\sum_{1\le i\le k}{}^{\prime}\frac{\mu(A_i)}{\nu(A_i)}\nu(A_i\cap C),\qquad C\in\mathcal{F}.$$

ここで $\sum_i'$ は条件 $\nu(A_i)>0$ を満たす $i\in[\![1,k]\!]$ に関して和を取ることを意味する．したがって，各 i に対して $\gamma(A_i)=\mu(A_i)$，すなわち $\overline{\gamma}_{\mathcal{A}}=\overline{\mu}_{\mathcal{A}}$ が成り立つ．特に $\gamma\in\rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))$ である．

$\mathcal{E}=(E_1,\dots,E_m)$ を任意の分割とする．作り方から明らかに γ は ν に対して絶対連続なので，$\nu(A_i)=0$ のときは $\gamma(A_i)=0$ となる．したがって，

$$\begin{aligned}D_{|\mathcal{E}|}(\overline{\gamma}_{\mathcal{E}}|\overline{\nu}_{\mathcal{E}}) &= \sum_{1\le j\le m}{}^{\prime\prime}\gamma(E_j)\log\frac{\gamma(E_j)}{\nu(E_j)}\\ &= \sum_{1\le j\le m}{}^{\prime\prime}\nu(E_j)\left(\sum_{1\le i\le k}{}^{\prime}\frac{\mu(A_i)}{\nu(A_i)}\frac{\nu(A_i\cap E_j)}{\nu(E_j)}\right)\log\left(\sum_{1\le i\le k}{}^{\prime}\frac{\mu(A_i)}{\nu(A_i)}\frac{\nu(A_i\cap E_j)}{\nu(E_j)}\right)\\ &\le \sum_{1\le j\le m}{}^{\prime\prime}\nu(E_j)\sum_{1\le i\le k}{}^{\prime}\frac{\mu(A_i)}{\nu(A_i)}\log\left(\frac{\mu(A_i)}{\nu(A_i)}\right)\frac{\nu(A_i\cap E_j)}{\nu(E_j)}\\ &= \sum_{1\le i\le k}{}^{\prime}\left(\sum_{1\le j\le m}{}^{\prime\prime}\nu(A_i\cap E_j)\right)\frac{\mu(A_i)}{\nu(A_i)}\log\frac{\mu(A_i)}{\nu(A_i)}\\ &= \sum_{1\le i\le k}{}^{\prime}\mu(A_i)\log\frac{\mu(A_i)}{\nu(A_i)}=D_{|\mathcal{A}|}(\overline{\mu}_{\mathcal{A}}|\overline{\nu}_{\mathcal{A}})<\alpha\end{aligned}$$

が成り立つ．ここで $\sum_j''$ は条件 $\nu(E_j)>0$ を満たす $j\in[\![1,m]\!]$ に関して和を取ることを意味する．なお上の不等号では，$\sum_i'\nu(A_i\cap E_j)/\nu(E_j)=1$ であることから $\sum_i'$ に関す

る Jensen の不等式を用いた．$\mathcal{E}$ に関する上限を取ると，$D(\gamma|\nu) \leq \alpha$ を得る．[12] これで $\gamma \in \rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma)) \cap K(\nu, \alpha)$ がわかった．以上で第２段を終わる．

【第 3 段】本段では (6.5.2) の証明を仕上げる．$m \in \mathbb{N}$ と分割 $\mathcal{A}(1), \ldots, \mathcal{A}(m)$ を任意に取る．分割 $\mathcal{E}$ が $\mathcal{A}(j)$ の細分のとき，$\rho_{\mathcal{A}(j)}^{-1}(\rho_{\mathcal{A}(j)}(\Gamma)) \supset \rho_{\mathcal{E}}^{-1}(\rho_{\mathcal{E}}(\Gamma))$ となるので，特に $\mathcal{E}$ を $\mathcal{A}(1), \ldots, \mathcal{A}(m)$ の共通の細分とすれば，

$$\bigcap_{j=1}^{m} \left[\rho_{\mathcal{A}(j)}^{-1}(\rho_{\mathcal{A}(j)}(\Gamma)) \cap K(\nu, \alpha)\right] \supset \rho_{\mathcal{E}}^{-1}(\rho_{\mathcal{E}}(\Gamma)) \cap K(\nu, \alpha)$$

である．第２段の結果により右辺は空でないので左辺も空でない．よって

$$\left\{\overline{\rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))} \cap K(\nu, \alpha) \;\middle|\; \mathcal{A} \in \mathbf{Prt}(S)\right\}$$

はコンパクト集合 $K(\nu, \alpha)$ 内の閉集合の族で，有限交叉性を持つ．したがって，

$$\emptyset \neq \bigcap_{\mathcal{A} \in \mathbf{Prt}(S)} \left\{\overline{\rho_{\mathcal{A}}^{-1}(\rho_{\mathcal{A}}(\Gamma))} \cap K(\nu, \alpha)\right\} = \overline{\Gamma} \cap K(\nu, \alpha)$$

を得る．ここで第１段の結果を用いた．以上で (6.5.2) が得られたので，第３段が終わり，補題 6.5.1 の証明が全て終わった． □

[12] 命題 6.0.1 で $D(\gamma|\nu) = H(\gamma|\nu)$ を得たので，$H(\gamma|\nu) \leq \alpha$ を示すという方法もある．これも難しくない．

第7章　ランダム行列に由来する大偏差原理

本章ではランダム行列理論において非常に有名な大偏差原理を解説する．$\mathbb{R}^N$ 上において下記の (7.0.1) で定まる確率測度を導入し，それから（Sanov の定理の場合と類似の方法で）経験測度を定める．この確率模型は対数気体と呼ばれている．このように定めた経験測度の列が $N \to \infty$ のときに $\mathbb{R}$ 上の確率測度全体がなす空間 $\mathcal{M}_1(\mathbb{R})$ 上で大偏差原理を満たすことを証明する．[1]

まず記号を設定する．本章では $N \in \mathbb{N}$ とし，$\mathbb{R}^N$ の一般的な元を $\lambda = (\lambda_1, \ldots, \lambda_N)$ と書き，N 次元 Lebesgue 測度を $d\lambda = \prod_{i=1}^N d\lambda_i$ と書く．また $\Delta(\lambda) = \prod_{1 \le i < j \le N} (\lambda_i - \lambda_j)$ と定める．これは差積（または Vandermonde 行列式）と呼ばれる．

正定数 β と連続関数 $V \colon \mathbb{R} \to \mathbb{R}$ から決まる $\mathbb{R}^N$ 上の確率測度を考察するが，本章を通じて β と V に対して以下の条件を課す．

(H) ある $\beta' > 1$ が存在して，$\beta' \ge \beta$ および以下の不等式を満たす．

$$\varliminf_{|x| \to \infty} \frac{V(x)}{\beta' \log |x|} > 1.$$

この条件 **(H)** を満たす V は下から有界で，任意の $c \ge 1$ に対して $\int_{\mathbb{R}} e^{-cV(x)} dx < \infty$ となる．この状況の下で $\mathbb{R}^N$ 上の Borel 確率測度 $\mathbb{Q}_{V,\beta}^N$ を

$$\mathbb{Q}_{V,\beta}^N(d\lambda) = (Z_{V,\beta}^N)^{-1} |\Delta(\lambda)|^\beta e^{-N \sum_{i=1}^N V(\lambda_i)} d\lambda \tag{7.0.1}$$

と定義する．ここで

$$Z_{V,\beta}^N := \int_{\mathbb{R}^N} |\Delta(\lambda)|^\beta e^{-N \sum_{i=1}^N V(\lambda_i)} d\lambda \ \in (0, \infty) \tag{7.0.2}$$

は正規化定数である．（後で確認するが，実は $Z_{V,\beta}^N$ の値は有限である．）また正規化する前の有限測度を $\widetilde{\mathbb{Q}}_{V,\beta}^N := Z_{V,\beta}^N \mathbb{Q}_{V,\beta}^N$ と書く．

---

[1] 本章の記述は [26] および [38, 51] を参考にした．

さて連続写像 $L_N\colon \mathbb{R}^N \to \mathcal{M}_1(\mathbb{R})$ を

$$L_N(\lambda) = \frac{\delta_{\lambda_1} + \cdots + \delta_{\lambda_N}}{N}$$

と定める.[2] ここで $\mathcal{M}_1(\mathbb{R})$ は $\mathbb{R}$ 上の Borel 確率測度全体がなす集合に弱収束の位相を入れたものである．$\mathbb{Q}^N_{V,\beta}$ の下での L_N の法則の漸近挙動は興味深い研究対象である．本章では $N \to \infty$ のときに $\{\mathbb{Q}^N_{V,\beta} \circ L_N^{-1}\}_{N\in\mathbb{N}}$ が速度 N^2 である良い速度関数に対して大偏差原理を満たすことを証明する．ここからこの大偏差原理を正確に述べるための準備を始める．

まず自由エントロピー $\Sigma\colon \mathcal{M}_1(\mathbb{R}) \to [-\infty, \infty)$ を

$$\Sigma(\mu) = \begin{cases} \int_{\mathbb{R}^2} \log|x-y|\mu(dx)\mu(dy), & \int_{\mathbb{R}} \log(|x|+1)\mu(dx) < \infty \text{ のとき}, \\ -\infty, & \text{それ以外のとき}, \end{cases} \tag{7.0.3}$$

と定める．$\Sigma(\mu)$ が有限であれば，$\mathbb{R}^2$ の対角線は $\mu \otimes \mu$ に関して零集合である（特に μ に関して任意の 1 点集合は零集合である）．さらに拡大非負値関数 $I^V_\beta\colon \mathcal{M}_1(\mathbb{R}) \to [0, \infty]$ を

$$I^V_\beta(\mu) = \begin{cases} \int_{\mathbb{R}} V(x)\mu(dx) - \frac{\beta}{2}\Sigma(\mu) - c^V_\beta, & \int_{\mathbb{R}} V(x)\mu(dx) < \infty \text{ のとき}, \\ \infty, & \text{それ以外のとき}, \end{cases} \tag{7.0.4}$$

と定める．ここで

$$c^V_\beta := \inf_{\nu \in \mathcal{M}_1(\mathbb{R})} \left\{ \int_{\mathbb{R}} V(x)\nu(dx) - \frac{\beta}{2}\Sigma(\nu) \right\}$$

とおいた．なお V は下から有界なので，$\int_{\mathbb{R}} V(x)\mu(dx)$ は積分確定であることに注意せよ．後述する補題 7.1.1 において，Σ と I^V_β が矛盾なく定義されていることと，c^V_β が有限であることを示す．また後述する補題 7.1.3 において，I^V_β が良い速度関数であることも示す．条件 **(H)** により $\int_{\mathbb{R}} V(x)\mu(dx) < \infty$ ならば $\int_{\mathbb{R}} \log(|x|+1)\mu(dx) < \infty$ であることに注意せよ．

それでは本章の主定理を述べる．Sanov の定理の場合と同じく，これは確率測度の空間上の確率測度の列に対する大偏差原理であることに注意せよ．

定理 7.0.1. 条件 **(H)** を仮定する．このとき確率測度の列 $\{\mathbb{Q}^N_{V,\beta} \circ L_N^{-1}\}_{N\in\mathbb{N}}$ は，$N \to \infty$ とするとき速度 N^2 で良い速度関数 I^V_β に対して $\mathcal{M}_1(\mathbb{R})$ 上で大偏差原理を満たす．

注意 7.0.2. $\mathbb{Q}^N_{V,\beta}$ の例として最も重要なのは Gauss 型 β 統計集団（を尺度変換したもの）であるので,[3] 本注意でこれを解説する．$\beta > 0$ に対して $V_\beta(x) = \beta x^2/4$ と定めると，明らかに **(H)** は満たされている．このとき $\mathbb{R}^N$ 上の確率測度

$$\mathbb{P}^N_{V_\beta,\beta}(d\lambda) = (W^N_{V_\beta,\beta})^{-1}|\Delta(\lambda)|^\beta e^{-\beta\sum_{i=1}^N \lambda_i^2/4} d\lambda$$

---

[2] この確率模型はポテンシャル V と逆温度 β を持つ対数気体（log-gas）と呼ばれる．

[3] 数学的には,「統計集団（ensemble)」とは要するに統計力学的な解釈を持つ確率測度という意味である．

を Gauss 型 β 統計集団と呼ぶ．ここで $W^N_{V_\beta,\beta} \in (0,\infty)$ は正規化定数である．このとき，スカラー倍写像 $(\lambda_1,\dots,\lambda_N) \mapsto (\lambda_1/\sqrt{N},\dots,\lambda_N/\sqrt{N})$ による $\mathbb{P}^N_{V_\beta,\beta}$ の像測度が $\mathbb{Q}^N_{V_\beta,\beta}$ である．言葉を換えると，$(\Lambda_1,\dots,\Lambda_N)$ がある確率空間の上で定義された $\mathbb{R}^N$ 値確率変数でその法則が $\mathbb{P}^N_{V_\beta,\beta}$ であれば，$(\Lambda_1/\sqrt{N},\dots,\Lambda_N/\sqrt{N})$ の法則が $\mathbb{Q}^N_{V_\beta,\beta}$ になる．

$\beta = 1,2,4$ のときは $\mathbb{P}^N_{V_\beta,\beta}$ は有名な $N \times N$ 型ランダム行列（すなわち行列値確率変数）の固有値の分布であることが古くから知られている．$\beta = 1,2,4$ に応じて，それぞれ Gauss 型直交，ユニタリ，シンプレクティック統計集団の固有値の分布である.[4] ここでは直交統計集団のみを正確に紹介する.（他の 2 つに関してもほぼ同様の事実が知られている.）

$\{H_{jk} \mid 1 \le j \le k < \infty\}$ を独立な平均 0 の実数値 Gauss 確率変数列とし，H_{jj} $(j \ge 1)$ の分散は 1，H_{jk} $(j<k)$ の分散は 1/2 とする．$j>k$ に対しては $H_{jk} = H_{kj}$ とおく．このとき $[H_{jk}]_{1\le j,k\le N}$ は N 次対称行列に値を取る確率変数になるが，これ（この法則）を Gauss 型直交統計集団と呼ぶ．$[H_{jk}]_{1\le j,k\le N}$ の N 個の固有値 $\Lambda^{(N)}_1,\dots,\Lambda^{(N)}_N$ は全て実数であり，これらを並べたものの法則が $\mathbb{P}^N_{V_1,1}$ であることが知られている.（本章では固有値たちに順序をつけずに,「対称化」したものを考える.）よって，$(\Lambda^{(N)}_1/\sqrt{N},\dots,\Lambda^{(N)}_N/\sqrt{N})$ の法則が $\mathbb{Q}^N_{V_1,1}$ に等しい．さて高名な Wigner の半円則によれば，$\mathcal{M}_1(\mathbb{R})$ 値確率変数

$$L_N(\Lambda^{(N)}_1/\sqrt{N},\dots,\Lambda^{(N)}_N/\sqrt{N}) = N^{-1}\left\{\delta_{\Lambda^{(N)}_1/\sqrt{N}} + \cdots + \delta_{\Lambda^{(N)}_N/\sqrt{N}}\right\}$$

は $N\to\infty$ のときに半円分布 $\mu_{\mathrm{sc}}(dx) := (2\pi)^{-1}\sqrt{(4-x^2)\vee 0}\,dx$ に確率収束する.[5] 特に $\mathcal{M}_1(\mathbb{R})$ 上の確率測度列の弱収束の意味で $\lim_{N\to\infty}\mathbb{Q}^N_{V_1,1}\circ L_N^{-1} = \delta_{\mu_{\mathrm{sc}}}$ となる．ここで右辺は μ_{sc} に集中する点測度である．したがって $\{\mathbb{Q}^N_{V_1,1}\circ L_N^{-1}\}_{N\in\mathbb{N}}$ に関する大偏差原理を問うことは自然な問題であろう．

ちなみに一般の $\beta > 0$ の場合も $\mathbb{P}^N_{V_\beta,\beta}$ がランダム行列の固有値の分布として実現できることが比較的最近になって証明された.[6]

注意 7.0.3. 定理 7.0.1 において現れる速度関数 I^V_β は，以下のように現れることを形式的に確かめておこう.（本注意の内容は厳密ではない.）まず $\Delta(\lambda)$ と $L_N = L_N(\lambda)$ の定義により

$$\log|\Delta(\lambda)|^\beta = \beta\sum_{1\le i<j\le N}\log|\lambda_i-\lambda_j| \approx \frac{\beta N^2}{2}\int_{\mathbb{R}^2}\log|x-y|L_N(dx)L_N(dy) = \frac{\beta N^2}{2}\Sigma(L_N)$$

と書けることに注意する．ただし近似 $\approx$ をする際に対角線からの寄与を無視した．また $\sum_{i=1}^N V(\lambda_i) = N\int_{\mathbb{R}} V(x)L_N(dx)$ であるので，$N\to\infty$ とするときに Laplace 原理を直感

---

[4] 例えば [26, 第 2.5 節と第 4.1 節] を見よ．

[5] 例えば [26, 定理 2.1.1] を見よ．なお本章で示す結果の応用として，この収束は確率収束から概収束に強められる．これについては後述の注意 7.4.1 を見よ．

[6] [26, 第 4.5 節] を見よ．

的に用いると

$$\frac{1}{N^2}\log Z^N_{V,\beta} = \frac{1}{N^2}\log\int_{\mathbb{R}^N}\exp\left\{N^2\left(\frac{\beta}{2}\Sigma(L_N) - \int_{\mathbb{R}} V(x)L_N(dx)\right)\right\}d\lambda$$
$$\sim \sup_{\nu\in\mathcal{M}_1(\mathbb{R})}\left\{\frac{\beta}{2}\Sigma(\nu) - \int_{\mathbb{R}} V(x)\nu(dx)\right\} = -c^V_\beta$$

となることが期待される.（もちろん $\mathbb{R}^N$ 上の Lebesgue 測度は確率測度ではないので，これは大偏差原理の一般論からは従わないことに注意せよ.）これらのことから一般の $\mu\in\mathcal{M}_1(\mathbb{R})$ に対して $N\to\infty$ とするとき

$$\frac{1}{N^2}\log\mathbb{Q}^N_{V,\beta}(L_N\approx\mu) \sim c^V_\beta + \frac{\beta}{2}\Sigma(\mu) - \int_{\mathbb{R}} V(x)\mu(dx) = -I^V_\beta(\mu)$$

となり，速度関数が I^V_β であることが形式的に確かめられた.

7.1　速度関数の性質

本節では定理 7.0.1 の証明に必要な事実のうち，速度関数の性質に関連するものを列挙する．以下では $f\colon\mathbb{R}^2\to(-\infty,\infty]$ を

$$f(x,y) = \frac{1}{2}V(x) + \frac{1}{2}V(y) - \frac{\beta}{2}\log|x-y|, \qquad (x,y)\in\mathbb{R}^2 \tag{7.1.1}$$

と定める.

補題 7.1.1. $\Sigma\colon\mathcal{M}_1(\mathbb{R})\to[-\infty,\infty)$ と $I^V_\beta\colon\mathcal{M}_1(\mathbb{R})\to[0,\infty]$ は矛盾なく定義されており，さらに $c^V_\beta\in\mathbb{R}$ である.

証明. 明らかに

$$\log(|x|+1)+\log(|y|+1)-\log(|x|+|y|) = \log\frac{(|x|+1)(|y|+1)}{|x|+|y|} \geq 0$$

となるので，

$$\log|x-y| \leq \log(|x|+|y|) \leq \log(|x|+1)+\log(|y|+1), \qquad (x,y)\in\mathbb{R}^2 \tag{7.1.2}$$

という不等式が成り立つ．よって $\int_{\mathbb{R}}\log(|x|+1)\mu(dx)<\infty$ のとき，

$$\int_{\mathbb{R}^2} 0\vee\log|x-y|\mu(dx)\mu(dy) \leq \int_{\mathbb{R}^2}\{\log(|x|+1)+\log(|y|+1)\}\mu(dx)\mu(dy)$$
$$= 2\int_{\mathbb{R}}\log(|x|+1)\mu(dx) < \infty \tag{7.1.3}$$

となる．これで $\Sigma(\mu)\in[-\infty,\infty)$ とわかる．

次は $c_\beta^V<\infty$ を見よう．$\nu\in\mathcal{M}_1(\mathbb{R})$ として特に $\nu(dx)=\mathbf{1}_{[0,1]}(x)dx$（区間 $[0,1]$ 上の Lebesgue 測度）を取る．このとき

$$\begin{aligned}\Sigma(\nu)&=\iint_{[0,1]^2}\log|x-y|dxdy\\&=2\iint_{\{0<y<x<1\}}\log(x-y)dxdy\\&=2\int_0^1\left(\Big[(x-y)\log(x-y)-(x-y)\Big]_{x=y}^{x=1}\right)dy\\&=2\int_0^1\{(1-y)\log(1-y)-(1-y)\}dy\ \in\mathbb{R}\end{aligned}$$

である．最右辺の dy 積分の被積分関数は連続なので，最右辺は有限である．$\int_{\mathbb{R}}V(x)\nu(dx)$ が有限なのは明らかなので，これで $c_\beta^V<\infty$ が確認できた．

ここからしばらく (7.1.1) で定義された関数 f を調べる．(7.1.2) により，明らかに

$$f(x,y)\geq\frac{1}{2}\{V(x)-\beta\log(|x|+1)\}+\frac{1}{2}\{V(y)-\beta\log(|y|+1)\}$$

が成り立つ．$|x|\to\infty$ のときに

$$V(x)-\beta\log(|x|+1)=\left\{\frac{V(x)}{\beta'\log|x|}-\frac{\beta}{\beta'}\frac{\log(|x|+1)}{\log|x|}\right\}\beta'\log|x|$$

と変形すると，**(H)** により中括弧内の第１項の下極限は１より大であり，中括弧内の第２項の極限は１以下である．よって $\lim_{|x|\to\infty}\{V(x)-\beta\log(|x|+1)\}=\infty$ である．これから

$$\lim_{|(x,y)|\to\infty}f(x,y)=\infty,\qquad\lim_{|x-y|\searrow0}f(x,y)=\infty$$

とわかる．なお２つ目の極限はほぼ自明である．より正確に述べると，$\lim_{m\to\infty}a_m=\infty$ となる正数列 $\{a_m\}_{m\in\mathbb{N}}$ が存在して

$$|x|>m\quad\text{または}\quad|y|>m\quad\text{または}\quad|x-y|<1/m\quad\Longrightarrow\quad f(x,y)\geq a_m\qquad(7.1.4)$$

となる．コンパクト集合 $\{(x,y)\in\mathbb{R}^2\mid|x|\leq m,|y|\leq m,|x-y|\geq1/m\}$ 上では f は連続なので，f は下から有界である．以下では $\underline{f}:=\inf\{f(x,y)\mid(x,y)\in\mathbb{R}^2\}$ と表す．

定義の仕方から明らかに

$$c_\beta^V=\inf\left\{\int_{\mathbb{R}}V(x)\nu(dx)-\frac{\beta}{2}\Sigma(\nu)\,\middle|\,\nu\in\mathcal{M}_1(\mathbb{R}),\ \int_{\mathbb{R}}\log(|x|+1)\nu(dx)<\infty\right\}$$

である．条件 $\int_{\mathbb{R}}\log(|x|+1)\nu(dx)<\infty$ の下で，

$$\int_{\mathbb{R}}V(x)\nu(dx)-\frac{\beta}{2}\Sigma(\nu)$$

$$= \frac{1}{2}\int_{\mathbb{R}} V(x)\nu(dx) + \frac{1}{2}\int_{\mathbb{R}} V(y)\nu(dy) + \int_{\mathbb{R}^2} \left(-\frac{\beta}{2}\right)\log|x-y|\nu(dx)\nu(dy)$$
$$= \int_{\mathbb{R}^2} f(x,y)\nu(dx)\nu(dy) \geq \underline{f} > -\infty$$

が成り立つ．(なお上記の積分計算の中で “$\infty - \infty$” という不定形は生じていないことに注意せよ.）これで $c_\beta^V > -\infty$ とわかったので，結局 $c_\beta^V \in \mathbb{R}$ が示せた．

$\int_{\mathbb{R}} V(x)\mu(dx) < \infty$ のときは $I_\beta^V(\mu) = \int_{\mathbb{R}} V(x)\mu(dx) - \frac{\beta}{2}\Sigma(\mu) - c_\beta^V$ であるが，右辺の第 1 項と第 3 項は有限なので，$I_\beta^V(\mu) \in [0,\infty]$ である．これで I_β^V が拡大非負値関数として矛盾なく定義されていることが示せた． □

補題 7.1.2. 直積測度を作る以下の対応は連続写像である．

$$\mathcal{M}_1(\mathbb{R}) \times \mathcal{M}_1(\mathbb{R}) \ni (\mu,\nu) \mapsto \mu \otimes \nu \in \mathcal{M}_1(\mathbb{R}^2).$$

ここで確率測度の空間の位相は弱収束の位相であり，左辺の直積集合の位相は直積位相である．[7]

証明. $\mathcal{M}_1(\mathbb{R})$ の列 $\{\mu_n\}$ と $\{\nu_n\}$ がそれぞれ $\lim_{n\to\infty}\mu_n = \mu_\infty$ および $\lim_{n\to\infty}\nu_n = \nu_\infty$ と $\mathcal{M}_1(\mathbb{R})$ において弱収束していると仮定し，$\lim_{n\to\infty}\mu_n\otimes\nu_n = \mu_\infty\otimes\nu_\infty$ と $\mathcal{M}_1(\mathbb{R}^2)$ において弱収束することを示せばよい．

仮定により $\{\mu_n\}$ と $\{\nu_n\}$ は緊密なので，任意の $\varepsilon \in (0,1)$ に対して $\mathbb{R}$ のコンパクト集合 A_ε と B_ε が存在して，任意の n に対して $\mu_n(A_\varepsilon^c) \vee \nu_n(B_\varepsilon^c) < \varepsilon$ を満たす．このとき $A_\varepsilon \times B_\varepsilon$ もコンパクトで

$$\mu_n \otimes \nu_n((A_\varepsilon \times B_\varepsilon)^c) \leq \mu_n\otimes\nu_n(\mathbb{R}\times B_\varepsilon^c) + \mu_n\otimes\nu_n(A_\varepsilon^c\times\mathbb{R}) < 2\varepsilon$$

となるため，$\{\mu_n\otimes\nu_n\}$ も $\mathcal{M}_1(\mathbb{R}^2)$ において緊密（よって相対コンパクト）である．よって，この列の集積点が $\mu_\infty\otimes\nu_\infty$ のみであることを示せば証明が終わる．

η を上記の列の任意の集積点とし，$\{\mu_{n(k)}\otimes\nu_{n(k)}\}_{k\in\mathbb{N}}$ を η に弱収束する部分列とすると，$\mathbb{R}$ 上の任意の有界連続関数 f と g に対して

$$\begin{aligned}\int_{\mathbb{R}^2} f(x)g(y)\eta(dxdy) &= \lim_{k\to\infty}\int_{\mathbb{R}^2} f(x)g(y)\mu_{n(k)}(dx)\nu_{n(k)}(dy) \\ &= \lim_{k\to\infty}\int_{\mathbb{R}} f(x)\mu_{n(k)}(dx)\int_{\mathbb{R}} g(y)\nu_{n(k)}(dy) \\ &= \int_{\mathbb{R}} f(x)\mu_\infty(dx)\int_{\mathbb{R}} g(y)\nu_\infty(dy) \qquad (7.1.5)\end{aligned}$$

[7] ちなみに直積測度を作る操作の弱収束に関する連続性は，より一般に 2 つの可分距離空間の直積の場合でも同様に成立する．詳しくは [27, 定理 2.8] を見よ．そこでの証明法はここで与えるものとははっきり異なる．

を満たす．直感的には自明であろうが，この条件を満たす $\eta \in \mathcal{M}_1(\mathbb{R}^2)$ は $\mu_\infty \otimes \nu_\infty$ に限ることを以下で厳密に示すので，それで本補題の証明が完成する．

A と B を $\mathbb{R}$ の任意の閉集合とする．このとき $\mathbf{1}_A$ と $\mathbf{1}_B$ にそれぞれ各点収束する $[0,1]$ 値連続関数の列 $\{f_n\}$ と $\{g_n\}$ が存在するので，これらを (7.1.5) に代入して有界収束定理を用いると，

$$\eta(A \times B) = \mu_\infty(A)\nu_\infty(B) \tag{7.1.6}$$

を得る.（例えば，$f_n(x) = \{1 + nd_{\mathbb{R}}(x, A)|\}^{-1}$ および $g_n(x) = \{1 + nd_{\mathbb{R}}(x, B)|\}^{-1}$ とおけばよい.）

$\mathcal{B}(\mathbb{R})$ を $\mathbb{R}$ の Borel 加法族とし，

$$\mathcal{L}_1 := \{A \in \mathcal{B}(\mathbb{R}) \mid \text{任意の閉集合 } B \subset \mathbb{R} \text{ に対して (7.1.6) が成立}\}$$

と定めると，これは $\mathbb{R}$ の閉集合全体を含む $\mathbb{R}$ 上の λ 族である．$\mathbb{R}$ の閉集合全体は π 族であるから π-λ 定理（Dynkin 族定理）が使えて，$\mathcal{L}_1$ は $\mathbb{R}$ の任意の Borel 集合を含むことがわかるが，これは (7.1.6) が任意の Borel 集合 A と任意の閉集合 B に対して成立することを意味する．そこで次は役割を交換して

$$\mathcal{L}_2 := \{B \in \mathcal{B}(\mathbb{R}) \mid \text{任意の } A \in \mathcal{B}(\mathbb{R}) \text{ に対して (7.1.6) が成立}\}$$

と定めて同様に議論を進めると，(7.1.6) が任意の $A, B \in \mathcal{B}(\mathbb{R})$ に対して成立することが示せる．直積測度の定義により，これは $\eta = \mu_\infty \otimes \nu_\infty$ と同値である．　□

次の補題で I_β^V が良い速度関数であることを証明する．この補題中で I_β^V を別の形に書き換えるが，この新しい形は煩わしい場合分けから解放されているために扱いやすく便利である．なお補題 7.1.1 の証明中で見たように，f は下から有界なので，(7.1.7) の右辺第 1 項の積分は $\mathbb{R} \cup \{\infty\}$ に値を取る．

補題 7.1.3. I_β^V は $\mathcal{M}_1(\mathbb{R})$ 上の良い速度関数である．また

$$I_\beta^V(\mu) = \int_{\mathbb{R}^2} f(x,y)\mu(dx)\mu(dy) - c_\beta^V, \qquad \mu \in \mathcal{M}_1(\mathbb{R}) \tag{7.1.7}$$

が成り立つ．

証明. **(H)** の下では $\int_{\mathbb{R}} V(x)\mu(dx) < \infty$ ならば $\int_{\mathbb{R}} \log(|x|+1)\mu(dx) < \infty$ であることをまず最初に注意しておく．よって，このとき $\Sigma(\mu) = \int_{\mathbb{R}^2} \log|x-y|\mu(dx)\mu(dy)$ である．

$I_\beta^V(\mu) < \infty$ とする．上の注意と (7.1.3) を合わせて考えると，このとき $\int_{\mathbb{R}} V(x)\mu(dx) \in \mathbb{R}$ かつ $\Sigma(\mu) \in \mathbb{R}$ である．すると次が成り立つ．

$$I_\beta^V(\mu) = \int_{\mathbb{R}} V(x)\mu(dx) - \frac{\beta}{2}\Sigma(\mu) - c_\beta^V$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^2} f(x,y)\mu(dx)\mu(dy) - c_\beta^V \\
&= \sup_{M>0} \int_{\mathbb{R}^2} f(x,y) \wedge M\, \mu(dx)\mu(dy) - c_\beta^V. \qquad (7.1.8)
\end{aligned}
$$

(7.1.8) の最後の等号では，f が下から有界であることと単調収束定理を用いた．各 M に対して $f \wedge M$ は有界連続なので，補題 7.1.2 により $\mu \mapsto \int_{\mathbb{R}^2} f(x,y) \wedge M\, \mu(dx)\mu(dy)$ は連続である．したがって，各 $r < \infty$ に対して

$$
\left\{\mu \in \mathcal{M}_1(\mathbb{R}) \mid I_\beta^V(\mu) \le r\right\} = \bigcap_{M>0} \left\{\mu \in \mathcal{M}_1(\mathbb{R}) \,\middle|\, \int_{\mathbb{R}^2} f(x,y) \wedge M \mu(dx)\mu(dy) \le r + c_\beta^V\right\}
$$

は閉集合であるから，I_β^V は下半連続である．これで I_β^V が速度関数であることがわかった．

(7.1.7) の証明を完成させるためには，後は $\int_{\mathbb{R}^2} f(x,y)\mu(dx)\mu(dy) < \infty$ ならば $I_\beta^V(\mu) < \infty$ となることを見れば十分である．$\delta > 0$ が十分小さいときには，$V^\delta := (1-\delta)V$ が再び **(H)** を満たすことに注意せよ．(7.1.1) における V をこの V^δ に取り替えて定まる関数を f^δ と書くことにする．すると

$$
\begin{aligned}
\infty &> \int_{\mathbb{R}^2} \left\{\frac{1}{2}V(x) + \frac{1}{2}V(y) - \frac{\beta}{2}\log|x-y|\right\} \mu(dx)\mu(dy) \\
&= \int_{\mathbb{R}^2} \left\{\frac{\delta}{2}V(x) + \frac{\delta}{2}V(y) + f^\delta(x,y)\right\} \mu(dx)\mu(dy) \qquad (7.1.9)
\end{aligned}
$$

となるが，最右辺の 3 個の被積分関数 $V(x)$, $V(y)$, $f^\delta(x,y)$ はどれも下から有界なので，結局どれも可積分である．特に $\int_{\mathbb{R}} V(x)\mu(dx) \in \mathbb{R}$ が示せた．この可積分性により，(7.1.9) の 1 行目の積分をバラして移項していいので，$\infty > -(\beta/2)\int_{\mathbb{R}^2} \log|x-y|\mu(dx)\mu(dy)$ を得る．以上をまとめて $\Sigma(\mu) \in \mathbb{R}$ とわかる．これで $I_\beta^V(\mu) < \infty$ が証明できた．

最後に $I_\beta^V(\mu)$ が良いことを示す．(7.1.4) により

$$
\{(x,y) \in \mathbb{R}^2 \mid |x| > m,\, |y| > m\} \subset \{(x,y) \in \mathbb{R}^2 \mid f(x,y) \ge a_m\}
$$

となるので，$I_\beta^V(\mu) \le r < \infty$ であれば，

$$
\begin{aligned}
\mu(|x| > m)^2 &= \mu \otimes \mu(|x| > m,\, |y| > m) \\
&\le \mu \otimes \mu(f(x,y) \ge a_m) \\
&\le \frac{1}{a_m - \underline{f}} \int_{\mathbb{R}^2} \{f(x,y) - \underline{f}\}\mu(dx)\mu(dy) \\
&= \frac{I_\beta^V(\mu) + c_\beta^V - \underline{f}}{a_m - \underline{f}} \le \frac{r + c_\beta^V - \underline{f}}{a_m - \underline{f}} \qquad (7.1.10)
\end{aligned}
$$

が成立する．ここで (7.1.7) を使った．$m \to \infty$ のとき (7.1.10) の最右辺は 0 に収束するので，閉集合 $\{\mu \in \mathcal{M}_1(\mathbb{R}) \mid I_\beta^V(\mu) \le r\}$ は緊密（すなわち相対コンパクト）なので，結局コンパクトである．これで $I_\beta^V(\mu)$ が良いことが示せたので，補題の証明が完成した．$\square$

次の補題は自由エントロピー Σ がある意味で狭義凹であることを主張する．この事実は速度関数 I_β^V の零点が唯一であることの証明に用いられる．

補題 7.1.4. $\mu_1, \mu_2 \in \mathcal{M}_1(\mathbb{R})$ をコンパクトな台を持つ相異なる確率測度とし，$\Sigma(\mu_1), \Sigma(\mu_2) \in \mathbb{R}$ とする．このとき

$$\Sigma(\theta\mu_1 + (1-\theta)\mu_2) > \theta\Sigma(\mu_1) + (1-\theta)\Sigma(\mu_2), \qquad \theta \in (0,1) \tag{7.1.11}$$

が成立する．

証明. $\nu \in \mathcal{M}_1(\mathbb{R})$ と $r \in (0,\infty)$ とする．任意の Borel 集合 $A \subset \mathbb{R}$ に対して $\nu^r(A) = \nu(rA)$ とおいて $\nu^r \in \mathcal{M}_1(\mathbb{R})$ を定義する．ここで $rA := \{rx \mid x \in A\}$ である．ν がコンパクトな台を持つとき，$\Sigma(\nu)$ を定義する 2 重積分は積分確定なので，

$$\Sigma(\nu^r) = \int_{\mathbb{R}^2} \log|x-y|\nu^r(dx)\nu^r(dy) = \int_{\mathbb{R}^2} \log|rx-ry|\nu(dx)\nu(dy) = \Sigma(\nu) + \log r$$

と計算してよい．これから

$$\Sigma(\theta\mu_1^r + (1-\theta)\mu_2^r) - \theta\Sigma(\mu_1^r) - (1-\theta)\Sigma(\mu_2^r) = \Sigma(\theta\mu_1 + (1-\theta)\mu_2) - \theta\Sigma(\mu_1) - (1-\theta)\Sigma(\mu_2)$$

を得る．したがって (7.1.11) において (μ_1,μ_2) を (μ_1^r,μ_2^r) に取り替えた不等式も元の (7.1.11) と同値である．また r を十分大きく取れば，μ_1^r の台も μ_2^r の台も $[-1/2,1/2]$ に含まれる．以上により，台が $[-1/2,1/2]$ に含まれるような (μ_1,μ_2) に対してのみ (7.1.11) を証明すれば十分だとわかる．

ここでいくつか定義する．$i,j=1,2$ に対して

$$\Sigma(\mu_i,\mu_j) := \int_{\mathbb{R}^2} \log|x-y|\mu_i(dx)\mu_j(dy) \ \in [-\infty, 0]$$

と定め，

$$\begin{aligned}\Sigma(\mu_1-\mu_2,\mu_1-\mu_2) := \Sigma(\mu_1,\mu_1) - \Sigma(\mu_1,\mu_2)& \\ - \Sigma(\mu_2,\mu_1) + \Sigma(\mu_2,\mu_2) \ &\in (-\infty,\infty]\end{aligned} \tag{7.1.12}$$

と定める．台に対する仮定により $\Sigma(\mu_i,\mu_j) \le 0$ である．定義により $\Sigma(\mu_i,\mu_i) = \Sigma(\mu_i) \in \mathbb{R}$ $(i=1,2)$ なので，(7.1.12) の右辺で “$\infty - \infty$” という不定形は生じていない．

ここから $\Sigma(\mu_i,\mu_j)$ を計算し，有限であることを見る．

$$\frac{1}{z} = \frac{1}{2z}\int_0^\infty e^{-u/2}du, \qquad z > 0$$

というほぼ自明の等式において $u = z^2/t$ とおき変数を u から t に変換すると，

$$\frac{1}{z} = \frac{1}{2}\int_0^\infty \frac{z}{t^2}e^{-z^2/(2t)}dt, \qquad z > 0$$

を得る．次に z を 1 から $|x-y|$ まで積分して，Fubini の定理を用いると

$$\log|x-y| = \int_0^\infty \frac{1}{2t}\Big\{\exp\Big(-\frac{1}{2t}\Big) - \exp\Big(-\frac{|x-y|^2}{2t}\Big)\Big\}dt$$

という対数関数の表示式を得る．この式の両辺を $\mu_i \otimes \mu_j$ で積分する．このとき $\mu_i \otimes \mu_j$ に関してほとんど全ての (x,y) に対して $|x-y| \le 1$，よって被積分関数は非正であるために，Fubini の定理が再び使えて

$$\Sigma(\mu_i,\mu_i) = \int_0^\infty \frac{1}{2t}\Big[\int_{\mathbb{R}^2}\Big\{\exp\Big(-\frac{1}{2t}\Big) - \exp\Big(-\frac{|x-y|^2}{2t}\Big)\Big\}\mu_i(dx)\mu_j(dy)\Big]dt \qquad (7.1.13)$$

を得る．この (7.1.13) を (7.1.12) に代入する．(7.1.12) の右辺の各項を t に関する積分と見た場合，$\Sigma(\mu_1,\mu_1)+\Sigma(\mu_2,\mu_2)$ は可積分，$-\Sigma(\mu_1,\mu_2)-\Sigma(\mu_2,\mu_1)$ は積分確定で $[0,\infty]$ 値である．特に "$\infty-\infty$" という不定形は生じていないので，t に関する 4 つの積分を一つにまとめてよい．すなわち

$$\begin{aligned}
&\Sigma(\mu_1-\mu_2,\mu_1-\mu_2)\\
&= \int_0^\infty \frac{1}{2t}\Big[\int_{\mathbb{R}^2}\Big\{\exp\Big(-\frac{1}{2t}\Big) - \exp\Big(-\frac{|x-y|^2}{2t}\Big)\Big\}\\
&\qquad\cdot\{\mu_1(dx)\mu_1(dy) - \mu_1(dx)\mu_2(dy) - \mu_2(dx)\mu_1(dy) + \mu_2(dx)\mu_2(dy)\}\Big]dt\\
&= -\int_0^\infty \frac{1}{2t}\Big[\int_{\mathbb{R}^2}\exp\Big(-\frac{|x-y|^2}{2t}\Big)(\mu_1-\mu_2)(dx)(\mu_1-\mu_2)(dy)\Big]dt \qquad (7.1.14)
\end{aligned}$$

である．ここで (x,y) に関する定数項 $\exp(-1/(2t))$ が消えた点が重要である．

次は (7.1.14) の最右辺の被積分関数を計算する．Gauss 関数の Fourier 変換に関する有名な公式

$$\exp\Big(-\frac{|x-y|^2}{2t}\Big) = \sqrt{\frac{t}{2\pi}}\int_{\mathbb{R}}\exp\Big(-\frac{t\xi^2}{2}\Big)e^{\sqrt{-1}(x-y)\xi}d\xi, \qquad t>0$$

を用いると，任意の $t>0$ と $(x,y)\in\mathbb{R}^2$ に対して

$$\begin{aligned}
&\int_{\mathbb{R}^2}\exp\Big(-\frac{|x-y|^2}{2t}\Big)(\mu_1-\mu_2)(dx)(\mu_1-\mu_2)(dy)\\
&= \sqrt{\frac{t}{2\pi}}\int_{\mathbb{R}}\Big|\int_{\mathbb{R}} e^{\sqrt{-1}x\xi}(\mu_1-\mu_2)(dx)\Big|^2\exp\Big(-\frac{t\xi^2}{2}\Big)d\xi \ge 0 \qquad (7.1.15)
\end{aligned}$$

となることがわかる．(7.1.14) と (7.1.15) により $\Sigma(\mu_1-\mu_2,\mu_1-\mu_2) \le 0$ であるため，$\Sigma(\mu_1,\mu_2)+\Sigma(\mu_2,\mu_1) \ge \Sigma(\mu_1)+\Sigma(\mu_2) > -\infty$ を得るので，特に $\Sigma(\mu_1,\mu_2)=\Sigma(\mu_2,\mu_1)$ も有限である．これで (7.1.12) の右辺の 4 つの項は全て有限であることがわかった．

仮に $\Sigma(\mu_1-\mu_2,\mu_1-\mu_2)=0$ だとすると，(7.1.14) と (7.1.15) により，$d\xi$ に関してほとんど全ての ξ で，

$$\int_{\mathbb{R}} e^{\sqrt{-1}x\xi}\mu_1(dx)=\int_{\mathbb{R}} e^{\sqrt{-1}x\xi}\mu_2(dx)$$

となる．確率測度の特性関数（Fourier 変換）は ξ の連続関数なので，実は上の等式は全ての ξ で成立している．特性関数が一致するため $\mu_1=\mu_2$ を得るが，これは本補題の仮定に反する．よって $\Sigma(\mu_1-\mu_2,\mu_1-\mu_2)<0$ である．

(7.1.11) の左辺を展開して計算すると，

$$\begin{aligned}
&\Sigma(\theta\mu_1+(1-\theta)\mu_2)-\theta\Sigma(\mu_1)-(1-\theta)\Sigma(\mu_2)\\
&=\theta^2\Sigma(\mu_1)+\theta(1-\theta)\{\Sigma(\mu_1,\mu_2)+\Sigma(\mu_2,\mu_1)\}+(1-\theta)^2\Sigma(\mu_2)\\
&\qquad-\theta\Sigma(\mu_1)-(1-\theta)\Sigma(\mu_2)\\
&=(\theta^2-\theta)\Sigma(\mu_1-\mu_2,\mu_1-\mu_2)>0
\end{aligned}$$

となる．すでに見たように上の計算で現れる項は全て有限である．以上で本補題が証明できた．　□

補題 7.1.3 で見たように I_β^V は良い速度関数であるため，補題 1.2.6 により I_β^V は最小値 0 を取る．この最小値を達成する確率測度がどのような性質を持つのかを次の補題で見る．確率測度 ν の台を $\mathrm{supp}(\nu)$ と書く．

補題 7.1.5. $I_\beta^V(\sigma_\beta^V)=0$ を満たす $\sigma_\beta^V\in\mathcal{M}_1(\mathbb{R})$ が唯一存在し，$\mathrm{supp}(\sigma_\beta^V)$ はコンパクトである．さらに σ_β^V は次の性質を持つ．ある定数 $A_\beta^V\in\mathbb{R}$ に対して，

$$V(x)-\beta\int_{\mathbb{R}}\log|x-y|\sigma_\beta^V(dy)=A_\beta^V,\qquad \sigma_\beta^V\text{ に関してほとんど全ての }x\in\mathbb{R},\tag{7.1.16}$$

$$V(x)-\beta\int_{\mathbb{R}}\log|x-y|\sigma_\beta^V(dy)\ge A_\beta^V,\qquad \text{全ての }x\notin\mathrm{supp}(\sigma_\beta^V),\tag{7.1.17}$$

という 2 条件が成り立つ.（実は必然的に $A_\beta^V=2c_\beta^V-\int_{\mathbb{R}}V(y)\sigma_\beta^V(dy)$ である.）

証明. 順番が前後するが，まず零点の一意性を見る．本証明の後半で示すが，I_β^V の零点である確率測度は必然的にコンパクトな台を持つ．$I_\beta^V(\sigma_\beta^V)=0=I_\beta^V(\tau_\beta^V)$ となる $\sigma_\beta^V,\tau_\beta^V$ $(\sigma_\beta^V\neq\tau_\beta^V)$ が存在したとする．このとき σ_β^V と τ_β^V は共に補題 7.1.4 の仮定を満たす．したがって $\theta\in(0,1)$ であれば，$\theta\sigma_\beta^V+(1-\theta)\tau_\beta^V$ もコンパクトな台を持ち，さらに

$$\begin{aligned}
I_\beta^V(\theta\sigma_\beta^V+(1-\theta)\tau_\beta^V)&=I_\beta^V(\theta\sigma_\beta^V+(1-\theta)\tau_\beta^V)-\theta I_\beta^V(\sigma_\beta^V)-(1-\theta)I_\beta^V(\tau_\beta^V)\\
&=-\frac{\beta}{2}\{\Sigma(\theta\sigma_\beta^V+(1-\theta)\tau_\beta^V)-\theta\Sigma(\sigma_\beta^V)-(1-\theta)\Sigma(\tau_\beta^V)\}<0
\end{aligned}$$

となり，I_β^V の最小値を下回ってしまう．これは矛盾である．

ここからは $I^V_\beta(\sigma^V_\beta)=0$ を満たす σ^V_β がどのような性質を持つかを調べる．補題 7.1.3 の証明中で見たように，このとき $\int_{\mathbb{R}} V(x)\sigma^V_\beta(dx)\in\mathbb{R}$ かつ $\Sigma(\sigma^V_\beta)\in\mathbb{R}$ である．特に $(x,y)\mapsto\log|x-y|$ は $\sigma^V_\beta\otimes\sigma^V_\beta$ に関して可積分である．

φ と ψ を $\mathbb{R}$ 上のコンパクトな台を持つ有界 Borel 可測関数とする．さらに ψ は非負値であり，$\int_{\mathbb{R}}\varphi(x)\sigma^V_\beta(dx)+\int_{\mathbb{R}}\psi(x)dx=0$ となることも仮定する．このとき

$$m(dx):=\varphi(x)\sigma^V_\beta(dx)+\psi(x)dx$$

とおくと，m は $m(\mathbb{R})=0$ を満たす $\mathbb{R}$ 上の符号付き測度である．$\varepsilon\in(0,1)$ が十分小さいときは,

$$(\sigma^V_\beta+\varepsilon m)(dx)=\{1+\varepsilon\varphi(x)\}\sigma^V_\beta(dx)+\varepsilon\psi(x)dx$$

は（非負値）確率測度であることに注意すると，

$$I^V_\beta(\sigma^V_\beta+\varepsilon m)\geq I^V_\beta(\sigma^V_\beta)=0 \tag{7.1.18}$$

である．

m の台はコンパクトなので $\int_{\mathbb{R}} V(x)\{\sigma^V_\beta+\varepsilon m\}(dx)\in\mathbb{R}$ である．したがって $I^V_\beta(\sigma^V_\beta+\varepsilon m)$ の定義式 (7.0.4) と $\Sigma(\sigma^V_\beta+\varepsilon m)$ の定義式 (7.0.3) ではどちらも場合分けの１つ目が生じている．このとき

$$A_1:=\int_{\mathbb{R}^2}\log|x-y|\sigma^V_\beta(dx)m(dy),\quad A_2:=\int_{\mathbb{R}^2}\log|x-y|m(dx)m(dy)$$

と定めると

$$\Sigma(\sigma^V_\beta+\varepsilon m)=\Sigma(\sigma^V_\beta)+2\varepsilon A_1+\varepsilon^2 A_2\ \in[-\infty,\infty) \tag{7.1.19}$$

である．上で述べたように $\Sigma(\sigma^V_\beta)\in\mathbb{R}$ である．以下で $A_1,A_2\in\mathbb{R}$ であることを見る．

$R:=\sup_{x\in\mathbb{R}}|\varphi(x)|\vee|\psi(x)|$ とおく．十分大きな $L>1$ を取れば，コンパクト集合 $K_L:=\{(x,y)\in\mathbb{R}^2\mid|x+y|\leq L,|x-y|\leq L\}$ の外で $\varphi(x)\varphi(y)$, $\varphi(x)\psi(y)$, $\psi(x)\varphi(y)$, $\psi(x)\psi(y)$ は消える．A_2 を展開すると

$$\begin{aligned}A_2=&\int_{\mathbb{R}^2}\log|x-y|\varphi(x)\varphi(y)\sigma^V_\beta(dx)\sigma^V_\beta(dy)\\&+2\int_{K_L}\log|x-y|\varphi(x)\psi(y)\sigma^V_\beta(dx)dy+\int_{K_L}\log|x-y|\psi(x)\psi(y)dxdy\end{aligned}$$

となる．右辺第 1 項は $\sigma^V_\beta\otimes\sigma^V_\beta$ に関して可積分であるので有限である．第 3 項の正定数倍が

$$\int_0^L dv\{\int_0^1(-\log u)du+\int_1^L\log u du\}=L(L\log L-L+2)<\infty$$

と評価できる．(変数変換 $(u,v)=(x-y,x+y)$ を行った．）第 2 項を評価すると

$$\begin{aligned}&\Big|\int_{K_L}\log|x-y|\varphi(x)\psi(y)\sigma_\beta^V(dx)dy\Big|\\&\le R^2\Big\{\int_{\{|x-y|<1\}}-\log|x-y|dy\sigma_\beta^V(dx)+\int_{\{|x-y|\ge1\}\cap K_L}\log|x-y|dy\sigma_\beta^V(dx)\Big\}\\&\le R^2(2+2L\log L)<\infty\end{aligned}$$

となり，これも有限である．以上で A_2 が有限であることが示せた．

A_1 を展開すると

$$\begin{aligned}A_1=&\int_{\mathbb{R}^2}\log|x-y|\varphi(y)\sigma_\beta^V(dx)\sigma_\beta^V(dy)\\&+\int_{\{|x-y|<1\}}\log|x-y|\psi(y)dy\sigma_\beta^V(dx)+\int_{\{|x-y|\ge1\}}\log|x-y|\psi(y)dy\sigma_\beta^V(dx)\end{aligned}$$

この右辺第 1 項と第 2 項は A_2 の計算中に出てきたものと同じ理由で有限である．第 3 項は非積分関数が非負なので拡大非負値である．しかし $A_1<\infty$ なので第 3 項も有限になる．以上で A_1 も有限であることが示せた．

(7.1.18) と (7.1.19) により，

$$I_\beta^V(\sigma_\beta^V+\varepsilon m)-I_\beta^V(\sigma_\beta^V)=\varepsilon\int_{\mathbb{R}}V(x)m(dx)-\varepsilon\beta A_1-\varepsilon^2\frac{\beta}{2}A_2\ge0$$

となる．上式の両辺を ε で割った後に $\varepsilon\searrow0$ として，次の不等式を得る．

$$\int_{\mathbb{R}}\Big\{\underbrace{V(x)-\beta\int_{\mathbb{R}}\log|x-y|\sigma_\beta^V(dy)}_{=:G(x)}\Big\}m(dx)\ge0. \tag{7.1.20}$$

以下では簡単のために $G(x):=V(x)-\beta\int_{\mathbb{R}}\log|x-y|\sigma_\beta^V(dy)$ と表す．すでに述べたように G は σ_β^V に関して可積分である．特に σ_β^V に関してほとんど確実に $|G|<\infty$ である．(7.1.20) において特に $m(dx)=\pm\varphi(x)\sigma_\beta^V(dx)+0\,dx$ となる特殊な場合を考えると，任意のコンパクトな台を持つ有界 Borel 可測関数 φ で，$\int_{\mathbb{R}}\varphi(x)\sigma_\beta^V(dx)=0$ を満たすものに対して

$$\int_{\mathbb{R}}G(x)\varphi(x)\sigma_\beta^V(dx)=0 \tag{7.1.21}$$

という等式を得る．

本段落では (7.1.21) から出発して，ある定数 $A_\beta^V\in\mathbb{R}$ に対して (7.1.16) が成立することを証明する．各 $n\in\mathbb{N}$ に対して $B_n:=\{x\in\mathbb{R}\mid|G(x)|\le n\}\cap[-n,n]$ とおくと，$\{B_n\}$ は広義単調増大であり，$\sigma_\beta^V(\cup_{n\in\mathbb{N}}B_n)=1$ である．よって n が十分大きいとき $\sigma_\beta^{V,n}:=\sigma_\beta^V(\cdot\cap B_n)$ は非自明な有限測度である．(7.1.21) において B_n の外では消え

る φ を考えると，B_n 上の有界 Borel 可測関数 φ で，$\int_{B_n}\varphi(x)\sigma_\beta^{V,n}(dx)=0$ を満たすものに対して $\int_{B_n}G(x)\varphi(x)\sigma_\beta^V(dx)=0$ が成り立つことがわかる．一般の $\eta\in L^2(B_n,\sigma_\beta^{V,n})$ で $\int_{B_n}\eta(x)\sigma_\beta^{V,n}(dx)=0$ を満たすものを考えよう．このとき各 $N\in\mathbb{N}$ に対して，$\eta^N:=(-N)\vee(\eta\wedge N)$ および $\varphi^N:=\eta^N-\sigma_\beta^{V,n}(B_n)^{-1}\int_{B_n}\eta^N d\sigma_\beta^{V,n}$ とおいた関数は有界で積分値が 0 なので，

$$\begin{aligned}
0&=\lim_{N\to\infty}\int_{B_n}G(x)\varphi^N(x)\sigma_\beta^{V,n}(dx)\\
&=\lim_{N\to\infty}\left\{\int_{B_n}G(x)\eta^N(x)\sigma_\beta^{V,n}(dx)-\sigma_\beta^{V,n}(B_n)^{-1}\int_{B_n}G(x)\sigma_\beta^{V,n}(dx)\int_{B_n}\eta^N(x)\sigma_\beta^{V,n}(dx)\right\}\\
&=\int_{B_n}G(x)\eta(x)\sigma_\beta^{V,n}(dx)
\end{aligned}$$

となる．ここで G が B_n 上で有界であることと Lebesgue の収束定理を用いた．したがって $G\restriction_{B_n}$ は積分値が 0 である $L^2(B_n,\sigma_\beta^{V,n})$ の任意の元と直交することがわかった．これは $G\restriction_{B_n}$ が $L^2(B_n,\sigma_\beta^{V,n})$ の元として定数関数であることを意味する．n の任意性と $\{B_n\}$ の単調性により (7.1.16) が証明できた．

この段落では (7.1.16) から σ_β^V の台がコンパクトであることを導く．(7.1.2) により

$$G(x)\geq\{V(x)-\beta\log(|x|+1)\}-\beta\int_{\mathbb{R}}\log(|y|+1)\sigma_\beta^V(dy)$$

と下から評価できる．右辺の積分は有限である．また補題 7.1.1 の証明中で示したように，$|x|\to\infty$ のとき右辺の中括弧の中は ∞ に発散する．よって G はほとんどいたる所定数に等しいにもかかわらず $\lim_{|x|\to\infty}G(x)=\infty$ となる．これは σ_β^V の台がコンパクトでない限り成立しない．

本段落では G が $\mathrm{supp}(\sigma_\beta^V)^c$ 上連続であることを示す．そのためには

$$H(x):=\int_{\mathrm{supp}(\sigma_\beta^V)}\log|x-y|\sigma_\beta^V(dy)$$

とおいて，H が開集合 $\mathrm{supp}(\sigma_\beta^V)^c$ 上連続であることを示せばよい．$x\in\mathrm{supp}(\sigma_\beta^V)^c$ を任意に取る．このとき $\delta>0$ を十分小さく選べば $(x-2\delta,x+2\delta)\cap\mathrm{supp}(\sigma_\beta^V)=\emptyset$ となる．$\{x_n\}_{n\in\mathbb{N}}\subset(x-\delta,x+\delta)$ を x に収束する任意の点列とすると，(7.1.2) により

$$\log\delta\leq\log|x_n-y|\leq\log(|x_n|+1)+\log(|y|+1)\leq\log(|x|+\delta+1)+\log(|y|+1)$$

という不等式が任意の $y\in\mathrm{supp}(\sigma_\beta^V)$ と n に対して成り立つ．この式の最右辺と最左辺は n に無関係で σ_β^V 可積分な関数なので，Lebesgue の収束定理により $\lim_{n\to\infty}H(x_n)=H(x)$ を得る．これで所望の連続性が示せた．

さて (7.1.20) における m として

$$m(dx)=-\Big(\int_{\mathbb{R}}\psi(x')dx'\Big)\mathbf{1}_{\mathrm{supp}(\sigma_\beta^V)}(x)\sigma_\beta^V(dx)+\psi(x)dx$$

を代入しよう．これと (7.1.16) と合わせて計算すると，

$$-\Big(\int_{\mathbb{R}} \psi(x)dx\Big)A_\beta^V + \int_{\mathbb{R}} G(x)\psi(x)dx = \int_{\mathbb{R}} \big\{-A_\beta^V + G(x)\big\}\psi(x)dx \geq 0$$

が任意のコンパクト台を持つ非負有界関数 ψ に対して成立することがわかる．これから Lebesgue 測度 dx に関してほとんど全ての x について $G(x) \geq A_\beta^V$ が成り立つことがわかる．G は台の外側では連続なので (7.1.17) が成り立つ．

定数 A_β^V の表示を求める．まず (7.1.16) の両辺を σ_β^V で積分すると

$$\int_{\mathbb{R}} V(x)\sigma_\beta^V(dx) - \beta\Sigma(\sigma_\beta^V) = A_\beta^V$$

となる．これと

$$I_\beta^V(\sigma_\beta^V) = \int_{\mathbb{R}} V(x)\sigma_\beta^V(dx) - \frac{\beta}{2}\Sigma(\sigma_\beta^V) - c_\beta^V = 0$$

から，簡単な計算で $A_\beta^V = 2c_\beta^V - \int_{\mathbb{R}} V(y)\sigma_\beta^V(dy)$ と求まる． □

7.2 上からの評価

本節では本章の主定理（定理 7.0.1）の上からの評価について議論する．$\mathcal{M}_1(\mathbb{R})$ 上の距離としては計算の便宜のために Prokhorov 距離を採用する．[8] 本節では記号を簡単にするために，必要に応じて $L_N(\lambda) = L_N^\lambda$ と書く．また $f\colon \mathbb{R}^2 \to (-\infty, \infty]$ は (7.1.1) で定義された関数で，補題 7.1.1 の証明中で見たように下から有界である．f の下限を $\underline{f}$ と表す．

まず最初に有限測度 $\widetilde{\mathbb{Q}}_{V,\beta}^N \circ L_N^{-1} := Z_{V,\beta}^N \mathbb{Q}_{V,\beta}^N \circ L_N^{-1}$ を上から評価する．$A \subset \mathcal{M}_1(\mathbb{R})$ が Borel 可測のとき，任意の $M \in (0, \infty)$ に対して

$$\begin{aligned}
&\widetilde{\mathbb{Q}}_{V,\beta}^N(L_N \in A) \\
&= \int_{\{L_N^\lambda \in A\}} |\Delta(\lambda)|^\beta e^{-N\sum_{i=1}^N V(\lambda_i)} d\lambda \\
&= \int_{\{L_N^\lambda \in A\}} \prod_{i<j} |\lambda_i - \lambda_j|^\beta e^{-\sum_{i=1}^N V(\lambda_i)} \exp\Big(-\sum_{i<j}\{V(\lambda_i) + V(\lambda_j)\}\Big) d\lambda \\
&= \int_{\{L_N^\lambda \in A\}} e^{-\sum_{i=1}^N V(\lambda_i)} \exp\big(-2\sum_{i<j} f(\lambda_i, \lambda_j)\big) d\lambda
\end{aligned}$$

[8] Prokhorov 距離を復習する．（例えば [27, Theorem 6.8] を参照せよ．）$P, Q \in \mathcal{M}_1(\mathbb{R})$ に対して，

$$d(P,Q) := \inf\{\varepsilon > 0 \mid \text{任意の } A \in \mathcal{B}(\mathbb{R}) \text{ に対して，} P(A) \leq Q(A^{(\varepsilon)}) + \varepsilon \text{ かつ } Q(A) \leq P(A^{(\varepsilon)}) + \varepsilon\}$$

と定義する．ただし $A^{(\varepsilon)} := \{x \in \mathbb{R} \mid d(x, A) < \varepsilon\}$ である．このとき d は $\mathcal{M}_1(\mathbb{R})$ 上の距離であり，d が誘導する位相は弱収束の位相と一致する．さらに $(\mathcal{M}_1(\mathbb{R}), d)$ は完備かつ可分である．

$$
\begin{aligned}
&= \int_{\{L_N^\lambda \in A\}} e^{-\sum_{i=1}^N V(\lambda_i)} \exp\Big(-N^2 \int_{\{x\neq y\}} f(x,y) L_N^\lambda(dx) L_N^\lambda(dy)\Big) d\lambda \\
&\le \int_{\{L_N^\lambda \in A\}} e^{-\sum_{i=1}^N V(\lambda_i)} \exp\Big(-N^2 \int_{\{x\neq y\}} f(x,y)\wedge M\, L_N^\lambda(dx) L_N^\lambda(dy)\Big) d\lambda \\
&= \int_{\{L_N^\lambda \in A\}} e^{-\sum_{i=1}^N V(\lambda_i)} \exp\Big(-N^2 \int_{\mathbb{R}^2} f(x,y)\wedge M\, L_N^\lambda(dx) L_N^\lambda(dy)\Big) e^{MN} d\lambda \\
&\le \int_{\{L_N^\lambda \in A\}} e^{-\sum_{i=1}^N V(\lambda_i)} \exp\Big(-N^2 \inf_{\mu\in A} \int_{\mathbb{R}^2} f(x,y)\wedge M\, \mu(dx)\mu(dy)\Big) e^{MN} d\lambda \\
&\le \Big(e^M \int_{\mathbb{R}} e^{-V(x)} dx\Big)^N \exp\Big(-N^2 \inf_{\mu\in A} \int_{\mathbb{R}^2} f(x,y)\wedge M\, \mu(dx)\mu(dy)\Big) \qquad (7.2.1)
\end{aligned}
$$

と上から評価できる．Lebesgue 測度 $d\lambda$ に関して $\{\lambda \in \mathbb{R}^N \mid$ ある $i \neq j$ に対して $\lambda_i = \lambda_j\}$ は零集合であることを使っている．$\int_{\mathbb{R}} e^{-V(x)}dx < \infty$ かつ f は下から有界なので，(7.2.1) の最右辺は有限である．

特に $A = \mathcal{M}_1(\mathbb{R})$ のとき (7.2.1) の左辺の値は $Z_{V,\beta}^N$ に等しいため，$Z_{V,\beta}^N \in (0,\infty)$ とわかる．より詳しくは

$$
Z_{V,\beta}^N \le \Big(e^M \int_{\mathbb{R}} e^{-V(x)} dx\Big)^N \exp\Big(-N^2 \inf_{\mu\in\mathcal{M}_1(\mathbb{R})} \int_{\mathbb{R}^2} f(x,y)\wedge M\, \mu(dx)\mu(dy)\Big)
$$

と評価できるため，任意の $M \in (0,\infty)$ に対して

$$
\varlimsup_{N\to\infty} \frac{1}{N^2} \log Z_{V,\beta}^N \le -\inf_{\mu\in\mathcal{M}_1(\mathbb{R})} \int_{\mathbb{R}^2} f(x,y)\wedge M\, \mu(dx)\mu(dy) \le -\underline{f} < \infty \qquad (7.2.2)
$$

と評価できる．なお (7.2.1) で $M \to \infty$ とする誘惑に駆られるが，実はこの極限操作は μ に関する下限と簡単には交換しないことを注意しておく．

大偏差原理の局所版を示すことにより主定理を証明する予定である．

補題 7.2.1. 任意の $\mu \in \mathcal{M}_1(\mathbb{R})$ に対して，次の不等式が成立する．

$$
\lim_{\kappa\searrow 0} \varlimsup_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}_{V,\beta}^N(L_N \in B(\mu,\kappa)) \le -I_\beta^V(\mu) - c_\beta^V.
$$

証明. 任意の $\mu \in \mathcal{M}_1(\mathbb{R})$ と $\kappa \in (0,\infty)$ に対して，(7.2.1) において $A = B(\mu,\kappa)$ と取ると，

$$
\varlimsup_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}_{V,\beta}^N(L_N \in B(\mu,\kappa)) \le -\inf_{\nu\in B(\mu,\kappa)} \int_{\mathbb{R}^2} f(x,y)\wedge M\, \nu(dx)\nu(dy)
$$

となる．$f\wedge M$ は $\mathbb{R}^2$ 上で有界連続なので，補題 7.1.2 により $\nu \mapsto \int_{\mathbb{R}^2} f(x,y)\wedge M\, \nu(dx)\nu(dy)$ は連続である．したがって，上式で $\kappa \searrow 0$ として

$$
\lim_{\kappa\searrow 0} \varlimsup_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}_{V,\beta}^N(L_N \in B(\mu,\kappa)) \le -\int_{\mathbb{R}^2} f(x,y)\wedge M\, \mu(dx)\mu(dy)
$$

を得る．f は下から有界なので，$M \to \infty$ のときに単調収束定理を用いて

$$\lim_{\kappa \searrow 0} \overline{\lim_{N\to\infty}} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in B(\mu,\kappa)) \le -\int_{\mathbb{R}^2} f(x,y)\mu(dx)\mu(dy) = -I^V_\beta(\mu) - c^V_\beta$$

を得る．右側の等式は補題 7.1.3 による．　□

補題 7.2.2. 有限測度の列 $\{\widetilde{\mathbb{Q}}^N_{V,\beta} \circ L_N^{-1}\}_{N\in\mathbb{N}}$ は $N \to \infty$ のとき指数的に緊密である．より詳しくは，任意の $R \in (1,\infty)$ に対してコンパクト集合 $K_R \subset \mathcal{M}_1(\mathbb{R})$ が存在して，以下の不等式を満たす.

$$\overline{\lim_{N\to\infty}} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in (K_R)^c) \le -R.$$

証明. $\delta \in (0,1)$ が十分小さいときには，$(1-\delta)V$ は再び **(H)** を満たすことを思い出そう．以下ではこのような δ を１つ固定して使う．$r \in (1,\infty)$ に対して $J_r := \{\mu \in \mathcal{M}_1(\mathbb{R}) \mid \int_{\mathbb{R}} V(x)\mu(dx) \le r\}$ とおく．$\lim_{|x|\to\infty} V(x) = \infty$ なので，J_r は緊密すなわち相対コンパクトである．ここで緊密性関数に関する補題 5.2.1 を用いた．この補集合の測度を評価すると，

$$\begin{aligned}
\widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in (J_r)^c) &= \int_{\{L_N^\lambda \notin J_r\}} |\Delta(\lambda)|^\beta e^{-N\sum_{i=1}^N V(\lambda_i)} d\lambda \\
&= \int_{\{N^{-1}\sum_{i=1}^N V(\lambda_i) > r\}} |\Delta(\lambda)|^\beta e^{-\delta N \sum_{i=1}^N V(\lambda_i)} e^{-N\sum_{i=1}^N (1-\delta)V(\lambda_i)} d\lambda \\
&\le e^{-N^2\delta r} \int_{\mathbb{R}^N} |\Delta(\lambda)|^\beta e^{-N\sum_{i=1}^N (1-\delta)V(\lambda_i)} d\lambda \\
&= e^{-N^2\delta r} Z^N_{(1-\delta)V,\beta}
\end{aligned}$$

および

$$\overline{\lim_{N\to\infty}} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in (J_r)^c) \le -\delta r + \overline{\lim_{N\to\infty}} \frac{1}{N^2} \log Z^N_{(1-\delta)V,\beta} \le -\delta r - \underline{f^\delta}$$

を得る．ここで (7.2.2) を用いた．よって $r \to \infty$ のとき，右辺は $-\infty$ に発散する．$K_r = \overline{J_r}$ とすると，明らかに K_r はコンパクトで $(K_r)^c \subset (J_r)^c$ を満たすので，指数的緊密性が証明できた.　□

補題 7.2.3. 次の不等式が成立する.

$$\overline{\lim_{N\to\infty}} \frac{1}{N^2} \log Z^N_{V,\beta} \le -c^V_\beta.$$

証明. $r \in (c^V_\beta, \infty)$ を任意に取る．補題 7.2.2 により，このとき十分大きなコンパクト集合 $K = K_r \subset \mathcal{M}_1(\mathbb{R})$ が存在して，$\overline{\lim}_{N\to\infty} N^{-2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in K^c) \le -r$ を満たす.

次は $\varepsilon \in (0,1)$ を任意に取る．補題 7.2.1 により，任意の $\mu \in K$ に対して $\kappa_\mu \in (0,1)$ が存在して，次の不等式を満たす．

$$\varlimsup_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in B(\mu,\kappa_\mu)) \le -\min\{I^V_\beta(\mu) + c^V_\beta - \varepsilon,\, \varepsilon^{-1}\}.$$

K のコンパクト性により，有限個の $\mu_1,\ldots,\mu_m \in K$ が存在して $K \subset \cup_{i=1}^m B(\mu_i,\kappa_{\mu_i})$ と被覆できる．

上の評価と補題 1.2.2 を組み合わせると

$$\begin{aligned}
&\varlimsup_{N\to\infty} \frac{1}{N^2} \log Z^N_{V,\beta} \\
&= \varlimsup_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in \mathcal{M}_1(\mathbb{R})) \\
&\le \varlimsup_{N\to\infty} \frac{1}{N^2} \log \left(\widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in K^c) + \sum_{i=1}^m \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in B(\mu_i,\kappa_{\mu_i})) \right) \\
&\le \left(\varlimsup_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in K^c) \right) \vee \max_{1\le i\le m} \varlimsup_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in B(\mu_i,\kappa_{\mu_i})) \\
&\le (-r) \vee \max_{1\le i\le m} \left\{ -\min\{I^V_\beta(\mu_i) + c^V_\beta - \varepsilon,\, \varepsilon^{-1}\} \right\} \\
&\le (-r) \vee \left[-\min\{I^V_\beta(\mu_1) + c^V_\beta - \varepsilon, \ldots, I^V_\beta(\mu_m) + c^V_\beta - \varepsilon,\, \varepsilon^{-1}\} \right]
\end{aligned}$$

を得る．全ての i に対して $I^V_\beta(\mu_i) = \infty$ ならば右辺は $(-r) \vee (-\varepsilon^{-1})$ 以下である．ある i に対して $I^V_\beta(\mu_i) < \infty$ ならば，ε が十分小さいとき右辺は $(-r) \vee (-c^V_\beta + \varepsilon)$ 以下である．いずれにせよ $\varepsilon \searrow 0$ とすると，$-r \le -c^V_\beta$ であるため左辺が $-c^V_\beta$ 以下だとわかる． □

7.3 下からの評価

本節では本章の主定理（定理 7.0.1）の下からの評価について議論する．

補題 7.3.1. 任意の $\mu \in \mathcal{M}_1(\mathbb{R})$ に対して，次の不等式が成立する．

$$\lim_{\kappa\searrow 0} \varliminf_{N\to\infty} \frac{1}{N^2} \log \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in B(\mu,\kappa)) \ge -I^V_\beta(\mu) - c^V_\beta. \tag{7.3.1}$$

特に

$$\varliminf_{N\to\infty} \frac{1}{N^2} \log Z^N_{V,\beta} \ge -c^V_\beta. \tag{7.3.2}$$

証明. $Z^N_{V,\beta} = \widetilde{\mathbb{Q}}^N_{V,\beta}(L_N \in \mathcal{M}_1(\mathbb{R}))$ なので，(7.3.2) は (7.3.1) から直ちに従う．よって (7.3.1) を示すことに集中する．

明らかに $I^V_\beta(\mu) < \infty$ となる μ に関してのみ (7.3.1) を示せば十分である．このとき $\mu \otimes \mu$ は対角線に重みを与えない．特に μ はアトムを持たない（すなわち任意の 1 点集合は μ 零集合である）ので，特に分布関数 $F(x) := \mu((-\infty,x])$ は連続関数である．

さらに $\mathrm{supp}(\mu)$ がコンパクトであるという仮定を追加して (7.3.1) を示せば十分であることを見る．$M \in \mathbb{N}$ を十分大として $\mu_M(dx) := \mu([-M,M])^{-1}\mathbf{1}_{[-M,M]}(x)\mu(dx)$ と定める．明らかに $\mathrm{supp}(\mu_M)$ はコンパクトであり，Lebesgue の収束定理を用いて $\lim_{M\to\infty}\mu_M = \mu$ と弱収束することを簡単に証明できる．f が下から有界なので単調収束定理を（$f - \underline{f}$ に対して）用いると，

$$\begin{aligned}\lim_{M\to\infty}\{I_\beta^V(\mu_M) + c_\beta^V\} &= \lim_{M\to\infty}\int_{\mathbb{R}^2} f(x,y)\mu_M(dx)\mu_M(dy)\\ &= \int_{\mathbb{R}^2} f(x,y)\mu(dx)\mu(dy) = I_\beta^V(\mu) + c_\beta^V\end{aligned}$$

とわかる．なお補題 7.1.3 も用いた．特に M が十分大きいとき $I_\beta^V(\mu_M) < \infty$ である．仮に (7.3.1) が任意の μ_M について成り立つとする．このとき任意の $\kappa, \varepsilon \in (0,1)$ に対して，M が十分大きいければ $\mu_M \in B(\mu,\kappa)$ かつ $-I_\beta^V(\mu_M) - c_\beta^V \geq -I_\beta^V(\mu) - c_\beta^V - \varepsilon$ となる．よって

$$\begin{aligned}\varliminf_{N\to\infty}\frac{1}{N^2}\log\widetilde{\mathbb{Q}}_{V,\beta}^N(L_N \in B(\mu,\kappa)) &\geq \lim_{\kappa'\searrow 0}\varliminf_{N\to\infty}\frac{1}{N^2}\log\widetilde{\mathbb{Q}}_{V,\beta}^N(L_N \in B(\mu_M,\kappa'))\\ &\geq -I_\beta^V(\mu_M) - c_\beta^V \geq -I_\beta^V(\mu) - c_\beta^V - \varepsilon\end{aligned}$$

が成り立つ．$\varepsilon \searrow 0$ とすると，(7.3.1) が μ に対しても成り立つことがわかる．

ここからはコンパクトな台を持ち $I_\beta^V(\mu) < \infty$ となる μ のみを考える．$N \in \mathbb{N}$ に対して，$-\infty < x_{0,N} < x_{1,N} < \cdots < x_{N,N} < \infty$ を $x_{0,N} = \min\mathrm{supp}(\mu)$, $x_{N,N} = \max\mathrm{supp}(\mu)$ および $x_{i,N} = \min\{x \in \mathbb{R} \mid F(x) = i/N\}$ $(1 \leq i \leq N-1)$ と定める．確率測度 $N^{-1}\sum_{i=1}^N \delta_{x_{i,N}}$ の分布関数 F_N は階段状の関数であり，$\lim_{N\to\infty} F_N(x) = F(x)$ と各点収束することが簡単に示せる．よって $\lim_{N\to\infty} N^{-1}\sum_{i=1}^N \delta_{x_{i,N}} = \mu$ と弱収束する．したがって，任意の $\kappa \in (0,1)$ に対してある $N_{\kappa/2} \in \mathbb{N}$ が存在して，$N \geq N_{\kappa/2}$ であれば $d(\mu, N^{-1}\sum_{i=1}^N \delta_{x_{i,N}}) < \kappa/2$ を満たす．Prokhorov 距離の定義から直ちに，

$$|\lambda_i - x_{i,N}| < \frac{\kappa}{2} \quad (1 \leq i \leq N) \quad \Longrightarrow \quad d\Big(N^{-1}\sum_{i=1}^N \delta_{x_{i,N}},\ N^{-1}\sum_{i=1}^N \delta_{\lambda_i}\Big) < \frac{\kappa}{2}$$

となることに注意すると，$N \geq N_{\kappa/2}$ のとき

$$\{\lambda \in \mathbb{R}^N \mid |\lambda_i - x_{i,N}| < \kappa/2\ (1 \leq i \leq N)\} \subset \{\lambda \in \mathbb{R}^N \mid d(\mu, L_N^\lambda) < \kappa\}$$

となる．

この包含関係と $d\lambda$ の平行移動不変性により

$$\begin{aligned}&\widetilde{\mathbb{Q}}_{V,\beta}^N(L_N \in B(\mu,\kappa))\\ &\geq \int_{\cap_{1\leq i\leq N}\{|\lambda_i - x_{i,N}| < \kappa/2\}} |\Delta(\lambda)|^\beta e^{-N\sum_{i=1}^N V(\lambda_i)} d\lambda\end{aligned}$$

$$
\begin{aligned}
&= \int_{\cap_{1\le i\le N}\{|\lambda_i|<\kappa/2\}} \prod_{i<j} |x_{i,N}-x_{j,N}+\lambda_i-\lambda_j|^\beta e^{-N\sum_{i=1}^N V(x_{i,N}+\lambda_i)} d\lambda \\
&\ge \int_{\cap_{1\le i\le N}\{|\lambda_i|<\kappa/2\}\cap\{\lambda_1<\cdots<\lambda_N\}} \\
&\qquad\qquad \prod_{i<j} |x_{i,N}-x_{j,N}+\lambda_i-\lambda_j|^\beta e^{-N\sum_{i=1}^N V(x_{i,N}+\lambda_i)} d\lambda \\
&\ge X_N Y_N^\kappa
\end{aligned}
\tag{7.3.3}
$$

と評価される．ただしここで以下のようにおいた．

$$
\begin{aligned}
X_N &:= \prod_{i+1<j} |x_{i,N}-x_{j,N}|^\beta \prod_i |x_{i,N}-x_{i+1,N}|^{\beta/2}\, e^{-N\sum_{i=1}^N V(x_{i,N})}, \\
Y_N^\kappa &:= \int_{\cap_{1\le i\le N}\{|\lambda_i|<\kappa/2\}\cap\{\lambda_1<\cdots<\lambda_N\}} \prod_i |\lambda_i-\lambda_{i+1}|^{\beta/2} e^{-N\sum_{i=1}^N\{V(x_{i,N}+\lambda_i)-V(x_{i,N})\}} d\lambda.
\end{aligned}
$$

(7.3.3) の最後の不等号に関する説明を補足する．まず $a,b\in[0,\infty)$ のとき $|a+b|\ge|a|\vee|b|$ となるという自明な事実を，$a=x_{j,N}-x_{i,N}$ と $b=\lambda_j-\lambda_i$ の場合に用いた．また $\prod_{i<j}$ を「番号が１つ違うもの」と「番号が２つ以上違うもの」 に分割し，さらに前者を２分割した．

まず先に Y_N の漸近挙動を評価する．まず $\varepsilon>0$ に対して

$$
\begin{aligned}
M_V(\varepsilon) := \sup\{|V(x)-V(y)| \mid |x-y|\le\varepsilon,& \\
\min\operatorname{supp}(\mu)-1\le x\le y\le \max\operatorname{supp}(\mu)+1&\}
\end{aligned}
$$

とおく．V は有界閉区間上で一様連続なので $\lim_{\varepsilon\searrow 0} M_V(\varepsilon)=0$ となる．この記号を用いると，Y_N^κ を定義する積分の中にある指数関数の部分は

$$
e^{-N\sum_{i=1}^N\{V(x_{i,N}+\lambda_i)-V(x_{i,N})\}} \ge e^{-N^2 M_V(\kappa/2)}
$$

と下から評価できる．積分変数を $\lambda=(\lambda_1,\ldots,\lambda_N)$ から $u=(u_1,\ldots,u_N)$ に以下のように変換する．$u_1=\lambda_1,\ u_i=\lambda_i-\lambda_{i-1}\ (2\le i\le N)$．この変数変換の Jacobi 行列式は 1 である．ほぼ明らかに

$$
0<u_i<\frac{\kappa}{2N}\ \ (1\le i\le N) \quad\Longrightarrow\quad \text{“}|\lambda_i|<\frac{\kappa}{2}\ \ (1\le i\le N)\text{”}\ \text{かつ}\ \text{“}\lambda_1<\cdots<\lambda_N\text{”}
$$

となることに注意すると，

$$
\begin{aligned}
Y_N^\kappa &\ge e^{-N^2M_V(\kappa/2)} \int_{\cap_{1\le i\le N}\{|\lambda_i|<\kappa/2\}\cap\{\lambda_1<\cdots<\lambda_N\}} \prod_i |\lambda_i-\lambda_{i+1}|^{\beta/2} d\lambda \\
&\ge e^{-N^2M_V(\kappa/2)} \int_{\cap_{1\le i\le N}\{0<u_i<\kappa/(2N)\}} \prod_{2\le i\le N} |u_i|^{\beta/2} du
\end{aligned}
$$

$$= e^{-N^2 M_V(\kappa/2)} \frac{\kappa}{2N} \left[\left(\frac{\beta}{2} + 1 \right)^{-1} \left(\frac{\kappa}{2N} \right)^{(\frac{\beta}{2}+1)} \right]^{N-1}$$

と下から評価できる．$N \to \infty$ とした後に $\kappa \searrow 0$ とすると

$$\varliminf_{\kappa \searrow 0} \varliminf_{N \to \infty} \frac{1}{N^2} \log Y_N^\kappa \geq - \lim_{\kappa \searrow 0} M_V(\kappa/2) = 0 \tag{7.3.4}$$

を得る．Y_N^κ は κ に関して広義単調なので，左辺の極限は存在する．

次は X_N の漸近挙動を評価する．$\{x_{i,N}\}_{i=1}^N \subset \mathrm{supp}(\mu)$ かつ $N^{-1} \sum_{i=1}^N \delta_{x_{i,N}}$ は μ に弱収束するので，$\lim_{N\to\infty} N^{-1} \sum_{i=1}^N V(x_{i,N}) = \int_{\mathbb{R}} V(x)\mu(dx)$ と収束する.（これを見るには，μ の台の外で V の値を修正し，$\mathbb{R}$ 上の有界連続関数になるようにすればよい.）

関数 $(0,\infty) \ni x \mapsto \log x$ は単調増大で，各 i に対して $\mu([x_{i,N}, x_{i+1,N}]) = 1/N$ であるから，

$$\begin{aligned}
&\int_{\{x_{1,N} < x < y \leq x_{N,N}\}} \log|y-x| \mu(dx)\mu(dy) \\
&\leq \sum_{1 \leq i \leq j \leq N-1} \log(x_{j+1,N} - x_{i,N}) \int_{[x_{i,N}, x_{i+1,N}] \times [x_{j,N}, x_{j+1,N}]} \mathbf{1}_{\{x<y\}} \mu(dx)\mu(dy) \\
&= \frac{1}{N^2} \sum_{1 \leq i < j \leq N-1} \log(x_{j+1,N} - x_{i,N}) + \frac{1}{2N^2} \sum_{i=1}^{N-1} \log(x_{i+1,N} - x_{i,N})
\end{aligned} \tag{7.3.5}$$

という不等式を得る．ここで $\mu \otimes \mu$ は対角線に重みを与えないことを用いた．特に任意の１点集合は μ 零集合なので，小区間の端点について気を使う必要はない．また $N \to \infty$ のとき $x_{1,N} \to x_{0,N}$ となるので，単調収束定理により (7.3.5) の左辺は $(1/2)\Sigma(\mu)$ に収束する．ここでも１点集合が零集合であることを使った.

X_N の定義と (7.3.5) により，

$$\begin{aligned}
\frac{1}{N^2} \log X_N &= \frac{\beta}{N^2} \sum_{1 \leq i < j \leq N-1} \log(x_{j+1,N} - x_{i,N}) \\
&\qquad + \frac{\beta}{2N^2} \sum_{i=1}^{N-1} \log(x_{i+1,N} - x_{i,N}) - N^{-1} \sum_{i=1}^N V(x_{i,N}) \\
&\geq \beta \int_{\{x_{1,N} < x < y \leq x_{N,N}\}} \log|y-x| \mu(dx)\mu(dy) - N^{-1} \sum_{i=1}^N V(x_{i,N})
\end{aligned}$$

を得るが，この両辺の下極限を取ると

$$\begin{aligned}
\varliminf_{N \to \infty} \frac{1}{N^2} \log X_N &\geq \frac{\beta}{2} \Sigma(\mu) - \int_{\mathbb{R}} V(x)\mu(dx) \\
&= - \int_{\mathbb{R}^2} f(x,y)\mu(dx)\mu(dy) = -I_\beta^V(\mu) - c_\beta^V
\end{aligned} \tag{7.3.6}$$

を得る．(7.3.3), (7.3.4), (7.3.6) から直ちに (7.3.1) が従う．これで補題 7.3.1 の証明が終わった. □

7.4　定理 7.0.1 の証明

本節では本章の主定理（定理 7.0.1）を証明を完成させる．

定理 7.0.1 の証明. 補題 7.1.1，補題 7.2.3, 補題 7.3.1 により

$$\lim_{N\to\infty}\frac{1}{N^2}\log Z_{V,\beta}^N = -c_\beta^V \ \in \mathbb{R}$$

が成り立つ．これを用いて補題 7.2.1 と補題 7.3.1 中の測度を正規化すると，各 $\mu \in \mathcal{M}_1(\mathbb{R})$ に対して

$$\lim_{\kappa\searrow 0}\varlimsup_{N\to\infty}\frac{1}{N^2}\log \mathbb{Q}_{V,\beta}^N(L_N \in B(\mu,\kappa)) \le -I_\beta^V(\mu),$$
$$\lim_{\kappa\searrow 0}\varliminf_{N\to\infty}\frac{1}{N^2}\log \mathbb{Q}_{V,\beta}^N(L_N \in B(\mu,\kappa)) \ge -I_\beta^V(\mu)$$

を得る．命題 1.3.2 により，これから $\{\mathbb{Q}_{V,\beta}^N \circ L_N^{-1}\}$ の速度関数 I_β^V に対する弱大偏差原理が導かれる．同様に補題 7.2.2 中の測度を正規化して，$\{\mathbb{Q}_{V,\beta}^N \circ L_N^{-1}\}$ が指数的に緊密であることがわかる．よって命題 1.3.1 により，目標とする定理 7.0.1 の証明が終わる．　□

注意 7.4.1. 各 $N \in \mathbb{N}$ に対して $\Lambda^{(N)} = (\Lambda_1^{(N)}, \ldots, \Lambda_N^{(N)})$ は（N によらない共通の）確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上で定義された $\mathbb{R}^N$ 値確率変数の列で，その法則が $\mathbb{Q}_{V,\beta}^N$ に等しいとする．このとき

$$L_N(\Lambda^{(N)}) = N^{-1}\left\{\delta_{\Lambda_1^{(N)}} + \cdots + \delta_{\Lambda_N^{(N)}}\right\}$$

とおくと，$N \to \infty$ のとき $\mathcal{M}_1(\mathbb{R})$ 値の確率変数列 $\{L_N(\Lambda^{(N)})\}_{N\in\mathbb{N}}$ はほとんど確実に σ_β^V に収束する．ここで σ_β^V は速度関数 I_β^V の唯一の零点で，もちろん非ランダムである（補題 7.1.5 および命題 1.2.15 を見よ）．この事実は Wigner の半円則の「概収束版」を一般化したものだと解釈できる．

第8章　Mogulskiiの大偏差原理

$\{X_i\}_{i\in\mathbb{N}}$ を確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上で定義された独立同分布な $\mathbb{R}^d$ 値確率変数列とする．本章では次の条件を常に仮定する．[1]

(FEM)　任意の $\lambda \in \mathbb{R}^d$ に対して，$\mathbb{E}\left[e^{\langle \lambda, X_1\rangle}\right] < \infty$ となる．

したがって，$\{X_i\}$ の経験平均について Cramér の大偏差原理（定理 2.0.1）が成立する状況である．特に X_1 は 2 乗可積分であるので，Kolmogorov の不等式により任意の $\delta > 0$ に対して

$$\mathbb{P}\left(\sup_{t\in[0,T]}\left|\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor}(X_i - \mathbb{E}[X_1])\right| \geq \delta\right) \leq \frac{\lfloor nT\rfloor}{n^2\delta^2}\mathbb{E}\left[(X_1 - \mathbb{E}[X_1])^2\right] \xrightarrow{n\to\infty} 0$$

が成立する．よって経験平均を適切な経路空間値の確率変数列 $\{n^{-1}\sum_{i=1}^{\lfloor n\cdot\rfloor} X_i\}_{n\in\mathbb{N}}$ と見ると，$n \to \infty$ のときに線形関数 $t \mapsto t\mathbb{E}[X_1]$ に確率収束していることがわかる（大数の法則の成立）．本章の主題である Mogulskii の定理はこの大数の法則に対応する大偏差原理を述べるものである．[2]

まず経路空間の記号を設定する．本章では $d \in \mathbb{N}$ および $T > 0$ とし，関数空間の記号にはこれらを明示しないことにする．まず $\mathcal{M} := \{w \mid w\colon [0,T] \to \mathbb{R}^d\}$ とし，各点収束の位相を入れる．要するに $\mathcal{M}$ は $[0,T]$ から $\mathbb{R}^d$ への写像全体がなす空間である．$w_0 = 0$ を満たす $\mathcal{M}$ の元全体がなす部分集合を $\mathcal{M}_0$ とおき，相対位相（部分集合としての位相）を入れる．$\mathcal{M}$ と $\mathcal{M}_0$ は Hausdorff 空間ではあるが，距離（付け可能な）空間ではないことに注意せよ．次に連続関数の空間を $\mathcal{C} = \{w \in \mathcal{M} \mid w \text{ は連続}\}$ および $\mathcal{C}_0 = \mathcal{C} \cap \mathcal{M}_0$ とおく．これら 2 つの空間は通常の一様ノルム $\|w\|_\infty = \sup_{0\leq t\leq T}|w_t|$ の下で可分 Banach 空間になる．さらに $\mathcal{L}^\infty := L^\infty([0,T], \mathbb{R}^d)$ と書く．これは $[0,T]$ 上の Lebesgue 測度に関する L^∞ 空間なので，本質的上限ノルム $\|w\|_\infty = \operatorname{ess.sup}_{0\leq t\leq T}|w_t|$ の下で非可分な Banach 空間になる．($\|\cdot\|_\infty$ を僅かに異なる 2 種類の意味で使っているが，w が連続のときは結局一致するので問題は生じない.）任意の空でない開集合の Lebesgue 測度が正であるため，ごく自然に $\mathcal{C} \hookrightarrow \mathcal{L}^\infty$ と等長埋め込みされる．これは $\mathcal{C}$ および $\mathcal{C}_0$ の位相は，$\mathcal{L}^\infty$ の部分集合だとみな

[1] この条件はほぼ明らかに「任意の $r > 0$ に対して $\mathbb{E}\left[e^{r|X_1|}\right] < \infty$」と同値である．同値であることを見るには，**(FEM)** において $\lambda = \pm r e_i$ $(r > 0,\ i = 1, \ldots, d)$ と取るとよい．

[2] 本章の記述は [33, 第 5.1 節] を参考にした．

した場合の相対位相と一致することを意味する．最後に $\mathcal{AC} = \{w \in \mathcal{C} \mid w \text{ は絶対連続}\}$ および $\mathcal{AC}_0 = \mathcal{AC} \cap \mathcal{C}_0$ とおく．なお w が絶対連続であるとは，任意の $\kappa > 0$ に対してある $\delta > 0$ が存在して以下の条件を満たすことである.[3]

$$k \in \mathbb{N},\ 0 \le s_1 < t_1 \le s_2 < t_2 \le \cdots \le s_k < t_k \le T,\ \sum_{i=1}^{k}(t_i - s_i) < \delta \\ \Longrightarrow \sum_{i=1}^{k} |w_{t_i} - w_{s_i}| < \kappa.$$

上記の独立確率変数列 $\{X_i\}$ から，各 $n \in \mathbb{N}$ ごとに確率過程を 2 つ構成しよう．まず確率過程 $Z^n = (Z_t^n)$ を

$$Z_t^n := \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor} X_i, \qquad t \ge 0 \tag{8.0.1}$$

により定める．ちなみに $t \in [k/n, (k+1)/n)$ のとき（$k \in \mathbb{N}_0$），$Z_t^n = n^{-1}\sum_{i=1}^{k} X_i$ であるため Z^n はこの小区間上で定数関数である．特に Z^n はランダムな階段関数であるので，$(Z_t^n)_{0 \le t \le T}$ を $\mathcal{L}^\infty$ 値確率変数と見なせる．$\mathcal{L}^\infty$ 上に誘導される $(Z_t^n)_{0 \le t \le T}$ の法則を μ^n と書く．

次に確率過程 $\widetilde{Z}^n = (\widetilde{Z}_t^n)$ を

$$\widetilde{Z}_t^n := Z_t^n + \left(t - \frac{\lfloor nt \rfloor}{n}\right) X_{\lfloor nt \rfloor + 1}, \qquad t \ge 0$$

により定める．これは明らかにランダムな折れ線の式であり，$t = k/n$（$k \in \mathbb{N}_0$）のとき $\widetilde{Z}_t^n = Z_t^n$ となるので，いわば Z^n の折れ線版だと思える．$(\widetilde{Z}_t^n)_{0 \le t \le T}$ は $\mathcal{C}_0$ 値確率変数である．$\mathcal{C}_0$ 上に誘導される $(\widetilde{Z}_t^n)_{0 \le t \le T}$ の法則を $\widetilde{\mu}^n$ と書く．埋め込み $\mathcal{C}_0 \hookrightarrow \mathcal{L}^\infty$ を通じて $(\widetilde{Z}_t^n)_{0 \le t \le T}$ は $\mathcal{L}^\infty$ 値，したがって $\widetilde{\mu}^n$ も $\mathcal{L}^\infty$ 上の確率測度だとも思える．

Mogulskii の大偏差原理を述べる前に，まず速度関数になる関数を導入しよう．第 2 章と同様に X_1 の対数積率母関数とその Fenchel-Legendre 変換をそれぞれ

$$\Lambda(\lambda) = \log \mathbb{E}\left[e^{\langle \lambda, X_1 \rangle}\right], \qquad \lambda \in \mathbb{R}^d,$$
$$\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, x \rangle - \Lambda(\lambda)\}, \qquad x \in \mathbb{R}^d$$

と表す．次に $\mathcal{L}^\infty$ 上の関数 I を次で定義する.[4]

$$I(w) = \begin{cases} \int_0^T \Lambda^*(w_t')dt, & w \in \mathcal{AC}_0 \text{ のとき}, \\ +\infty, & \text{それ以外のとき}. \end{cases} \tag{8.0.2}$$

---

[3] 区間上の絶対連続関数に関する基本事項については例えば [3, 第 19 章] を参照せよ．

[4] この関数 I は自然に $\mathcal{M}_0$ 上の関数だとも思える．

絶対連続な w に対しては，Lebesgue 測度 dt に関してほとんど全ての t で微分 w'_t が存在し，$|w'|$ は $[0,T]$ 上可積分であることを思い出そう.[5] Λ^* は非負値なので，I も非負値である.

ここで本章の主題である Mogulskii の大偏差原理の主張を正確に述べる．上記の I が良い速度関数であることももちろん主張に含まれている.

定理 8.0.1. 条件 **(FEM)** を仮定する．$n \to \infty$ のとき，上記の確率測度の列 $\{\mu^n\}_{n\in\mathbb{N}}$ は $\mathcal{L}^\infty$ 上で (8.0.2) で定義された良い速度関数 I に対して速度 n で大偏差原理を満たす.

8.1 Mogulskii の大偏差原理の証明

上記の Mogulskii の大偏差原理（定理 8.0.1）を証明するためには，**(FEM)** の下で成立する以下の 3 つの補題が必要である.

補題 8.1.1. $n \to \infty$ のとき，確率測度の列 $\{\mu^n\}_{n\in\mathbb{N}}$ と $\{\widetilde{\mu}^n\}_{n\in\mathbb{N}}$ は $\mathcal{L}^\infty$ 上で指数的に同値である.

証明. 本証明中では時間区間 J 上の（本質的）上限ノルムを $\|\cdot\|_{\infty,J}$ と表すことにする．$J(n,k) = [k/n,(k+1)/n) \cap [0,T]$ と表す．当然 $\|w\|_{\infty,[0,T]} = \max_{0\le k\le \lfloor nT\rfloor} \|w\|_{\infty,J(n,k)}$ が成り立つ．$|\widetilde{Z}^n_t - Z^n_t| \le n^{-1}|X_{\lfloor nt\rfloor+1}|$ が任意の $t \in [0,T)$ に対して成り立つので，$\{X_i\}$ が同分布であることと合わせて，任意の $\delta,\lambda \in (0,\infty)$ に対して

$$
\begin{aligned}
\mathbb{P}\left(\sup_{0\le t\le T}|\widetilde{Z}^n_t - Z^n_t| > \delta\right) &= \mathbb{P}\left(\bigcup_{k=0}^{\lfloor nT\rfloor}\{\|\widetilde{Z}^n - Z^n\|_{\infty,J(n,k)} > \delta\}\right) \\
&\le \sum_{k=0}^{\lfloor nT\rfloor}\mathbb{P}\left(n^{-1}|X_{k+1}| > \delta\right) \\
&\le (\lfloor nT\rfloor + 1)\mathbb{P}\left(|X_1| > n\delta\right) \\
&\le (nT+2)e^{-n\lambda\delta}\mathbb{E}\left[e^{\lambda|X_1|}\right]
\end{aligned}
$$

となる．**(FEM)** により最右辺の期待値は有限であることに注意せよ．この両辺の $(1/n)\log$ を取った式で，まず $n \to \infty$ とした後に $\lambda \to \infty$ とすると

$$
\overline{\lim_{n\to\infty}}\frac{1}{n}\log\mathbb{P}\left(\sup_{0\le t\le T}|\widetilde{Z}^n_t - Z^n_t| > \delta\right) = -\infty \tag{8.1.1}
$$

を得る．$\delta > 0$ は任意なので，これで $\{\mu^n\}$ と $\{\widetilde{\mu}^n\}$ が指数的に同値であることが示せた. □

[5] [3, 第 19 章] を参照せよ.

上の補題と命題 1.6.2 により，$\{\mu^n\}$ に対する大偏差原理と $\{\widetilde{\mu}^n\}$ に対する大偏差原理は同値である．したがって，問題は $\{\widetilde{\mu}^n\}$ の大偏差原理に帰着されるのだが，いきなりそれを示すのは難しいので，まず $\{\widetilde{\mu}^n\}$ が弱い位相に関して大偏差原理を満たすことを示す．この補題の証明の鍵は射影極限法（Dawson-Gärtner の定理．定理 1.9.3）である．連続な埋め込み $\mathcal{C}_0 \hookrightarrow \mathcal{M}_0$ を通じて，$\widetilde{\mu}^n$ は $\mathcal{M}_0$ 上の測度だとみなせることに注意せよ．

補題 8.1.2. $n \to \infty$ のとき，確率測度の列 $\{\widetilde{\mu}^n\}_{n\in\mathbb{N}}$ は $\mathcal{M}_0$ 上で良い速度関数 I に対して速度 n で大偏差原理を満たす．

$\{\widetilde{\mu}^n\}$ に対する大偏差原理の位相を強めるためには，定石通りに指数的緊密性を示す．

補題 8.1.3. $n \to \infty$ のとき，確率測度の列 $\{\widetilde{\mu}^n\}_{n\in\mathbb{N}}$ は $\mathcal{C}_0$ 上で指数的に緊密である．

補題 8.1.2 と補題 8.1.3 の証明は後回しにする．上記の 3 つの補題を用いると，目的とする Mogulskii の大偏差原理が比較的簡単に証明できるので，先にそれを終わらせておく．

定理 8.0.1 の証明. 明らかに $\{\widetilde{\mu}^n\}$ は $\mathcal{C}_0$ 上の確率測度の族であり，埋め込み $\mathcal{C}_0 \hookrightarrow \mathcal{M}_0$ は連続である．この族は補題 8.1.2 と補題 8.1.3 にあるように $\mathcal{M}_0$ 上で大偏差原理を満たし，$\mathcal{C}_0$ 上で指数的に緊密である．したがって逆縮小原理（命題 1.5.2）により，この大偏差原理は実は $\mathcal{C}_0$ 上で良い速度関数 I に対して成立することがわかる．[6]

$\mathcal{C}_0$ が $\mathcal{L}^\infty$ の閉部分空間であるため命題 1.2.12(1) が使えて，この大偏差原理は $\mathcal{L}^\infty$ 上で良い速度関数 I に対して成立することがわかる．最後に補題 8.1.1 で示した $\mathcal{L}^\infty$ 上での指数的な同値性と補題 1.6.2 を用いると，所望の $\{\mu^n\}$ に対する大偏差原理が得られる．　□

8.2　補題 8.1.2 の証明

本節の目的は補題 8.1.2 を証明することである．この補題の証明を以下のようにさらに 2 つの補題に分解して与える．これらの補題中では

$$\mathcal{Q} := \{t_1 < t_2 < \cdots < t_{|\mathcal{Q}|}\} \subset (0,T] \tag{8.2.1}$$

という形の有限集合を $(0,T]$ の分割と呼ぶことにする（分点の数 $|\mathcal{Q}|$ に制限はない）．この形の分割 $\mathcal{Q}$ に対して，自然な射影 $\pi_{\mathcal{Q}}\colon \mathcal{M}_0 \to (\mathbb{R}^d)^{|\mathcal{Q}|}$ を

$$\pi_{\mathcal{Q}}(w) = (w_{t_1}, \ldots, w_{t_{|\mathcal{Q}|}}) \ \in (\mathbb{R}^d)^{|\mathcal{Q}|}, \qquad w \in \mathcal{M}_0 \tag{8.2.2}$$

と定める．以下では $(\mathbb{R}^d)^{|\mathcal{Q}|}$ の一般的な元を $\mathbf{z} = (z_1, \ldots, z_{|\mathcal{Q}|})$ と表す．

[6] $\mathcal{C}_0$ は $\mathcal{M}_0$ において Borel 可測でないため（例えば [20] の定理 7.1.1 直後にある説明と問題 7.1.4 を見よ），この部分の証明に命題 1.2.12 は使えない．

補題 8.2.1. $\mathcal{Q}$ を (8.2.1) で与えられる $(0,T]$ の分割とする．$n\to\infty$ とするとき，射影 $\pi_{\mathcal{Q}}$ が誘導する法則の列 $\{\mu^n\circ\pi_{\mathcal{Q}}^{-1}\}_{n\in\mathbb{N}}$ および $\{\widetilde{\mu}^n\circ\pi_{\mathcal{Q}}^{-1}\}_{n\in\mathbb{N}}$ はともに $(\mathbb{R}^d)^{|\mathcal{Q}|}$ 上で良い速度関数 $I_{\mathcal{Q}}\colon(\mathbb{R}^d)^{|\mathcal{Q}|}\to[0,\infty]$ に対して速度 n で大偏差原理を満たす．ここで

$$I_{\mathcal{Q}}(\mathbf{z}):=\sum_{l=1}^{|\mathcal{Q}|}(t_l-t_{l-1})\,\Lambda^*\left(\frac{z_l-z_{l-1}}{t_l-t_{l-1}}\right),\qquad \mathbf{z}\in(\mathbb{R}^d)^{|\mathcal{Q}|}$$

と定義される．なお簡単のため $t_0=0,\ z_0=0$ とおいた．

証明. $\mathcal{Q}$ は固定して議論するので，本証明中では $N:=|\mathcal{Q}|$ と表す．定義により $\mu^n\circ\pi_{\mathcal{Q}}^{-1}$ と $\widetilde{\mu}^n\circ\pi_{\mathcal{Q}}^{-1}$ はそれぞれ

$$Z^n_{\mathcal{Q}}:=(Z^n_{t_1},Z^n_{t_2},\dots,Z^n_{t_N}),\qquad \widetilde{Z}^n_{\mathcal{Q}}:=(\widetilde{Z}^n_{t_1},\widetilde{Z}^n_{t_2},\dots,\widetilde{Z}^n_{t_N}),$$

の法則と等しい．また

$$Y^n_{\mathcal{Q}}:=(Z^n_{t_1},\,Z^n_{t_2}-Z^n_{t_1},\dots,\,Z^n_{t_N}-Z^n_{t_{N-1}})\tag{8.2.3}$$

と定める.（n に依存しない）$(\mathbb{R}^d)^N$ 内の線形同型写像

$$(z_1,z_2,\dots,z_N)\mapsto(z_1,z_2-z_1,\dots,z_N-z_{N-1})$$

により，$Y^n_{\mathcal{Q}}$ は $Z^n_{\mathcal{Q}}$ に写される．縮小原理（命題 1.5.1）により，$\{\mu^n\circ\pi_{\mathcal{Q}}^{-1}\}$ に対する大偏差原理を示すには，$\{Y^n_{\mathcal{Q}}\}$ が下記の $\Lambda^*_{\mathcal{Q}}$ を良い速度関数として大偏差原理を満たすことを示せば十分である．

$$\Lambda^*_{\mathcal{Q}}(\mathbf{y}):=\sum_{l=1}^{N}(t_l-t_{l-1})\Lambda^*\left(\frac{y_l}{t_l-t_{l-1}}\right),\qquad \mathbf{y}=(y_1,\dots,y_N)\in(\mathbb{R}^d)^N.$$

まず $Y^n_{\mathcal{Q}}$ の各成分は独立なので，$Y^n_{\mathcal{Q}}$ の法則は各成分の法則の積測度に等しい．また各成分の法則は以下で与えられる.（ただし $Z^n_{t_0}=0$ とおく.）

$$Z^n_{t_l}-Z^n_{t_{l-1}}\overset{\text{Law}}{=}\frac{X_1+\cdots+X_{\lfloor nt_l\rfloor-\lfloor nt_{l-1}\rfloor}}{n},\qquad 1\le l\le N.$$

次で示す補題 8.2.2 により，$n\to\infty$ のときに $\{Z^n_{t_l}-Z^n_{t_{l-1}}\}$ は指数的に緊密であり，かつ良い速度関数 $(t_l-t_{l-1})\Lambda^*(\cdot/(t_l-t_{l-1}))$ に対して大偏差原理を満たす．よって積測度の大偏差原理に関する命題 1.8.1 により，$\Lambda^*_{\mathcal{Q}}$ が良い測度関数であることと $\{Y^n_{\mathcal{Q}}\}$ が所望の大偏差原理を満たすことが従う．

最後に $\{\widetilde{\mu}^n\circ\pi_{\mathcal{Q}}^{-1}\}$ について述べる．補題 8.1.1 の証明中にある (8.1.1) により，任意の $\delta\in(0,\infty)$ に対して

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\max_{1\le l\le N}|\widetilde{Z}^n_{t_l}-Z^n_{t_l}|>\delta\right)=-\infty$$

がわかるので，$\{\mu^n\circ\pi_{\mathcal{Q}}^{-1}\}$ と $\{\widetilde{\mu}^n\circ\pi_{\mathcal{Q}}^{-1}\}$ は指数的に同値である.（補題 8.1.1 の主張そのものはここでは使っていない.）命題 1.6.2 により，$\{\widetilde{\mu}^n\circ\pi_{\mathcal{Q}}^{-1}\}$ も所望の大偏差原理を満たす．　□

各 $t>0$ に対して $\{Z_t^n\}_{n\in\mathbb{N}}$ に対する大偏差原理を示したいが，その速度関数は $t\Lambda^*(\cdot/t)$ であることが形式的には次のようにわかる．$x\in\mathbb{R}^d$ に対して

$$\begin{aligned}\frac{1}{n}\log\mathbb{P}\left(Z_t^n\approx x\right)&=\frac{\lfloor nt\rfloor}{n}\frac{1}{\lfloor nt\rfloor}\log\mathbb{P}\left(\frac{\lfloor nt\rfloor}{n}\frac{X_1+\cdots+X_{\lfloor nt\rfloor}}{\lfloor nt\rfloor}\approx x\right)\\&\sim t\frac{1}{\lfloor nt\rfloor}\log\mathbb{P}\left(\frac{X_1+\cdots+X_{\lfloor nt\rfloor}}{\lfloor nt\rfloor}\approx\frac{x}{t}\right)\sim -t\Lambda^*(x/t).\end{aligned}$$

最後の近似は Cramér の大偏差原理（定理 2.0.1）による．（ここで近似記号 $\approx$ と $\sim$ は直感的な意味で使った.）この議論を厳密にすると以下のようになる．

補題 8.2.2. $\gamma\in(0,\infty)$ を正定数とする．また $\{\beta(n)\}_{n\in\mathbb{N}}$ を $\lim_{n\to\infty}\beta(n)/n=\gamma$ を満たす自然数列とする．この仮定の下で，$n\to\infty$ のときに $\{Z_{\beta(n)}^n\}_{n\in\mathbb{N}}$ は指数的に緊密であり，かつ良い速度関数 $\gamma\Lambda^*(\cdot/\gamma)$ に対して速度 n で大偏差原理を満たす.

証明. 基本的に Cramér の定理の変種である．Cramér の定理の証明に似ている部分については，足早に議論を進める．

まず指数的緊密性を確認する．$R\in(0,\infty)$ を任意とする．Chebyshev の不等式と $\{X_n\}$ が独立同分布であることにより，

$$\begin{aligned}\mathbb{P}\left(|Z_{\beta(n)}^n|>R\right)&\le\mathbb{P}\left(e^{|X_1|+\cdots+|X_{\beta(n)}|}>e^{nR}\right)\\&\le e^{-nR}\mathbb{E}\left[e^{|X_1|+\cdots+|X_{\beta(n)}|}\right]\\&\le e^{-nR}\mathbb{E}\left[e^{|X_1|}\right]^{\beta(n)}\end{aligned}\tag{8.2.4}$$

となる．**(FEM)** を仮定したので $\mathbb{E}\left[e^{|X_1|}\right]\in(0,\infty)$ であることに注意せよ．したがって，

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(|Z_{\beta(n)}^n|>R\right)\le -R+\gamma\log\mathbb{E}\left[e^{|X_1|}\right]$$

と評価できる．右辺は $R\nearrow\infty$ のとき明らかに $-\infty$ に発散するので，これで指数的緊密性が示せた.

大偏差原理の証明に入る．もちろん **(FEM)** の下で Cramér の大偏差原理が成り立つことを用いる．まず各 $n\in\mathbb{N}$ に対して，新しい確率変数を次のように定める．

$$V^n:=\frac{X_1+\cdots+X_{\beta(n)}}{\beta(n)}\gamma.$$

このとき Cramér の大偏差原理により，任意の閉集合 $F\subset\mathbb{R}^d$ に対して

$$\begin{aligned}\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(V^n\in F\right)&\le\lim_{n\to\infty}\frac{\beta(n)}{n}\cdot\varlimsup_{n\to\infty}\frac{1}{\beta(n)}\log\mathbb{P}\left(\frac{X_1+\cdots+X_{\beta(n)}}{\beta(n)}\in\gamma^{-1}F\right)\\&\le\gamma\varlimsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{X_1+\cdots+X_n}{n}\in\gamma^{-1}F\right)\end{aligned}$$

$$\leq -\gamma \inf_{x \in \gamma^{-1}F} \Lambda^*(x) = -\gamma \inf_{x \in F} \Lambda^*(x/\gamma)$$

となる．ここで $\lim_{n\to\infty} \beta(n) = \infty$ を使っている．同様に任意の開集合 $O \subset \mathbb{R}^d$ に対して

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(V^n \in O\right) \geq -\gamma \inf_{x \in O} \Lambda^*(x/\gamma)$$

も得る．これで $\{V^n\}$ は良い速度関数 $\gamma\Lambda^*(\cdot/\gamma)$ に対して（速度 $\beta(n)$ でなく）速度 n の大偏差原理を満たすことがわかった．

最後に $\{Z^n_{\beta(n)}\}$ と $\{V^n\}$ が指数的に同値であることを見れば十分である．命題 1.6.2 を参照せよ．仮定により任意の $\kappa \in (0,1)$ に対して n が十分大きければ $|1-(\gamma n)/\beta(n)| \leq \kappa$ となる．$\delta \in (0,\infty)$ のとき，次のように (8.2.4) とほぼ同じ評価ができる．

$$\begin{aligned}
\mathbb{P}\left(|Z^n_{\beta(n)} - V^n| > \delta\right) &= \mathbb{P}\left(\left|1 - \frac{\gamma n}{\beta(n)}\right| |Z^n_{\beta(n)}| > \delta\right) \\
&\leq \mathbb{P}\left(|Z^n_{\beta(n)}| > \delta/\kappa\right) \\
&\leq e^{-n\delta/\kappa} \mathbb{E}\left[e^{|X_1|}\right]^{\beta(n)}.
\end{aligned}$$

したがって，

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(|Z^n_{\beta(n)} - V^n| > \delta\right) \leq -\frac{\delta}{\kappa} + \gamma \log \mathbb{E}\left[e^{|X_1|}\right]$$

となる．この式で $\kappa \searrow 0$ として指数的同値性を得る．　□

注意 8.2.3. 上の議論では (8.2.3) で定めた $\{Y^n_{\mathcal{Q}}\}$ が大偏差原理を満たすことを，Cramér の大偏差原理（定理 2.0.1）と積測度に対する大偏差原理（命題 1.8.1）を用いて証明した．証明が初等的であるのは長所だが，やや長くなったきらいもある．実は Gärtner-Ellis の定理（定理 3.0.1）を用いると，この部分を簡潔に証明できるので，以下で紹介する．（ただし，Gärtner-Ellis の定理の証明自体がかなり難しいので，全体として見れば簡単になったとは言えないと思う．）

まず速度関数の良さに関してであるが，Λ^* が良い速度関数であることを用いると，命題 1.8.1 の証明の第 1 段落と同じ議論で $\Lambda^*_{\mathcal{Q}}$ も良い速度関数であることが直接示せる．[7] Λ^* の定義と簡単な計算により，

$$\begin{aligned}
\Lambda^*_{\mathcal{Q}}(\mathbf{y}) &= \sum_{l=1}^{N} (t_l - t_{l-1}) \sup_{\lambda_l \in \mathbb{R}^d} \left\{ \left\langle \lambda_l, \frac{y_l}{t_l - t_{l-1}} \right\rangle - \Lambda(\lambda_l) \right\} \\
&= \sum_{l=1}^{N} \sup_{\lambda_l \in \mathbb{R}^d} \left\{ \langle \lambda_l, y_l \rangle - (t_l - t_{l-1})\Lambda(\lambda_l) \right\}
\end{aligned}$$

[7] 直接示すのではなく，Gärtner-Ellis の定理にある一般論を引用する方法もある．

$$= \sup_{\boldsymbol{\lambda}\in(\mathbb{R}^d)^N}\left[\sum_{l=1}^{N}\{\langle\lambda_l, y_l\rangle - (t_l - t_{l-1})\Lambda(\lambda_l)\}\right]$$
$$= \sup_{\boldsymbol{\lambda}\in(\mathbb{R}^d)^N}\{\langle\boldsymbol{\lambda},\mathbf{y}\rangle - \Lambda_{\mathcal{Q}}(\boldsymbol{\lambda})\}, \qquad \mathbf{y}\in(\mathbb{R}^d)^N$$

となる．ただし $N = |\mathcal{Q}|$，$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_N)$ および $\Lambda_{\mathcal{Q}}(\boldsymbol{\lambda}) := \sum_{l=1}^{N}(t_l - t_{l-1})\Lambda(\lambda_l)$ とおいた.（ちなみに 3 つ目の等号で sup と $\sum_l$ の順番を交換したが，これは定義に従って議論すれば簡単に示せる.）特に $\Lambda_{\mathcal{Q}}^*$ は微分可能な有限値関数の Fenchel-Legendre 変換であることがわかった．$Y_{\mathcal{Q}}^n$ の各成分は独立であることを注意すると

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left[e^{n\langle\boldsymbol{\lambda}, Y_{\mathcal{Q}}^n\rangle}\right] = \lim_{n\to\infty}\sum_{l=1}^{N}\frac{1}{n}\log\mathbb{E}\left[e^{n\langle\lambda_l, Z_{t_l}^n - Z_{t_{l-1}}^n\rangle}\right]$$
$$= \lim_{n\to\infty}\sum_{l=1}^{N}\frac{1}{n}\left(\lfloor nt_l\rfloor - \lfloor nt_{l-1}\rfloor\right)\Lambda(\lambda_l)$$
$$= \sum_{l=1}^{N}(t_l - t_{l-1})\Lambda(\lambda_l) = \Lambda_{\mathcal{Q}}(\boldsymbol{\lambda})$$

となり，$\boldsymbol{\lambda}$ の関数として至る所微分可能である．ただし $Z_{t_0}^n = 0$ とおいた．また 2 つ目の等号で $\{X_i\}$ が独立列であることを用いた．これで Gärtner-Ellis の定理を用いて，$\{Y_{\mathcal{Q}}^n\}$ が所望の大偏差原理を満たすことが示せる．以上で注意 8.2.3 を終わる．

補題 8.1.2 の証明. (8.2.1) にある $(0, T]$ の分割の全体は集合としての包含関係を順序とみなすと自然に有向集合になる．そのような $\mathcal{Q} := \{t_1 < t_2 < \cdots < t_{|\mathcal{Q}|}\}$ に対して，$\mathcal{Y}_{\mathcal{Q}} := (\mathbb{R}^d)^{|\mathcal{Q}|}$ と定め，座標番号を $(1, \dots, |\mathcal{Q}|)$ ではなく $(t_1, \cdots, t_{|\mathcal{Q}|})$ と名付ける．すると $\mathcal{Q} \supset \mathcal{Q}'$ のとき射影 $\pi_{\mathcal{Q}',\mathcal{Q}}\colon \mathcal{Y}_{\mathcal{Q}} \to \mathcal{Y}_{\mathcal{Q}'}$ が自然に定まる．例 1.9.1 で見たように，この射影系の射影極限は位相空間として $\mathcal{M}_0$ と自然に同一視できる．$\mathcal{Q}$ に対して定まる射影 $\pi_{\mathcal{Q}}\colon \mathcal{M}_0 \to \mathcal{Y}_{\mathcal{Q}}$ は (8.2.2) と一致する．

補題 8.2.1 と射影極限に対する Dawson-Gärtner の定理（定理 1.9.3）により，$\{\widetilde{\mu}^n\}_{n\in\mathbb{N}}$ は $\mathcal{M}_0$ 上で次の式で与えられる良い速度関数 $\widetilde{I}\colon \mathcal{M}_0 \to [0, \infty]$ に対して速度 n で大偏差原理を満たす．$w \in \mathcal{M}_0$ のとき，

$$\widetilde{I}(w) := \sup\left\{\sum_{l=1}^{|\mathcal{Q}|}(t_l - t_{l-1})\Lambda^*\left(\frac{w_{t_l} - w_{t_{l-1}}}{t_l - t_{l-1}}\right)\,\middle|\, \mathcal{Q} = \{t_1 < \cdots < t_{|\mathcal{Q}|}\} \subset (0, T]\right\}$$
$$= \sup\left\{\sum_{l=1}^{|\mathcal{Q}|}(t_l - t_{l-1})\Lambda^*\left(\frac{w_{t_l} - w_{t_{l-1}}}{t_l - t_{l-1}}\right)\,\middle|\, \mathcal{Q} = \{t_1 < \cdots < t_{|\mathcal{Q}|} = T\} \subset (0, T]\right\}$$

と定める．2 つ目の等号は Λ^* が非負値であることによる．この $\widetilde{I}$ が (8.0.2) で定めた I

と等しいことを示せば証明が終わる.[8] $w \in \mathcal{AC}_0$ かつ $t_{|\mathcal{Q}|} = T$ のとき，Λ^* の凸性により Jensen の不等式が使えるので

$$\begin{aligned}\sum_{l=1}^{|\mathcal{Q}|}(t_l - t_{l-1})\Lambda^*\left(\frac{w_{t_l} - w_{t_{l-1}}}{t_l - t_{l-1}}\right) &= \sum_{l=1}^{|\mathcal{Q}|}(t_l - t_{l-1})\Lambda^*\left(\frac{1}{t_l - t_{l-1}}\int_{t_{l-1}}^{t_l} w'_s ds\right) \\ &\le \sum_{l=1}^{|\mathcal{Q}|}\int_{t_{l-1}}^{t_l}\Lambda^*(w'_s)ds = I(w)\end{aligned}$$

となる．これで $\widetilde{I} \le I$ が示された．

逆向きの不等式 $\widetilde{I} \ge I$ を示す．まず $w \in \mathcal{AC}_0$ の場合を考える．記号を簡単にするため $t > T$ に対して $w_t = w_T$ と定め，w の定義域を $[0,\infty)$ に拡張しておく．まず $g \in L^1([0,\infty),\mathbb{R}^d)$ を $g = w'$ とおく．すなわち $g_t = w'_t$ $(0 \le t \le T)$ および $g_t = 0$ $(t > T)$ とおく．さらに $k \in \mathbb{N}$ に対して

$$g^k_t := k\int_{\lfloor kt\rfloor/k}^{(\lfloor kt\rfloor+1)/k} g_s ds, \qquad t \ge 0$$

とおく．これは g を長さ $1/k$ の小区間上で平均して作った階段関数である．すると

$$\begin{aligned}\widetilde{I}(w) &\ge \varliminf_{k\to\infty}\sum_{l=1}^{\infty}\frac{1}{k}\Lambda^*\left(k\{w_{l/k} - w_{(l-1)/k}\}\right) \\ &= \varliminf_{k\to\infty}\sum_{l=1}^{\infty}\frac{1}{k}\Lambda^*\left(g^k_{(l-1)/k}\right) = \varliminf_{k\to\infty}\int_0^T\Lambda^*(g^k_s)ds.\end{aligned}$$

上の l に関する和は実は有限和である.（ここでは Λ^* の凸性は使っていない.）Lebesgue の微分定理により,[9]ほとんど全ての s に対して $\lim_{k\to\infty} g^k_s = g_s$ となる．Fatou の補題と Λ^* の下半連続性（および非負性）を用いると

$$\widetilde{I}(w) \ge \varliminf_{k\to\infty}\int_0^T\Lambda^*(g^k_s)ds \ge \int_0^T\varliminf_{k\to\infty}\Lambda^*(g^k_s)ds \ge \int_0^T\Lambda^*(g_s)ds = I(w)$$

となり，求める評価が得られた．

次は $w \in \mathcal{M}_0 \setminus \mathcal{AC}_0$ の場合を考える．絶対連続性の条件を否定すると，定数 $\kappa > 0$ と自然数列 $\{k_n\}_{n\in\mathbb{N}}$ と分割の列 $\{\{0 \le s^n_1 < t^n_1 \le \cdots \le s^n_{k_n} < t^n_{k_n} \le T\}\}_{n\in\mathbb{N}}$ で

$$\lim_{n\to\infty}\sum_{i=1}^{k_n}(t^n_i - s^n_i) = 0, \qquad \sum_{i=1}^{k_n}|w_{t^n_i} - w_{s^n_i}| \ge \kappa$$

---

[8] より正確には「(8.0.2) で定めた I を $\mathcal{M}_0$ 上の関数だと思い直したもの」と等しいことを示せば証明が終わる.

[9] [3, 定理 19.3] を参照せよ.

を満たすものが存在することがわかる．$\Lambda^* \geq 0$ なので，

$$\begin{aligned}\widetilde{I}(w) &\geq \sup\left\{ \sum_{l=1}^{k} (t_l - s_l)\Lambda^*\left(\frac{w_{t_l} - w_{s_l}}{t_l - s_l}\right) \,\middle|\, \{0 \leq s_1 < t_1 \leq \cdots \leq s_k < t_k \leq T\} \right\} \\ &= \sup\Big\{ \sum_{l=1}^{k} \sup_{\lambda_l \in \mathbb{R}^d} [\langle \lambda_l, w_{t_l} - w_{s_l}\rangle - (t_l - s_l)\Lambda(\lambda_l)] \ \Big| \\ &\qquad\qquad\qquad \{0 \leq s_1 < t_1 \leq \cdots \leq s_k < t_k \leq T\} \Big\}\end{aligned}$$

となる．$r > 0$ を任意に取る．上の不等式で特に $k = k_n,\ t_l = t_l^n,\ s_l = s_l^n$ の場合に $\lambda_l := r(w_{t_l^n} - w_{s_l^n})/|w_{t_l^n} - w_{s_l^n}|$（ただし $w_{t_l^n} - w_{s_l^n} \neq 0$ のとき）$\lambda_l := 0$（ただし $w_{t_l^n} - w_{s_l^n} = 0$ のとき）と取ると，

$$\widetilde{I}(w) \geq \varlimsup_{n\to\infty} \Big\{ r\sum_{l=1}^{k_n} |w_{t_l^n} - w_{s_l^n}| - [\max_{|\lambda|\leq r} \Lambda(\lambda)] \sum_{l=1}^{k_n} (t_l^n - s_l^n) \Big\} \geq r\kappa$$

となる．Λ は連続なので，$\max_{|\lambda|\leq r} \Lambda(\lambda)$ は存在し有限である．$r \nearrow \infty$ とすると $\widetilde{I}(w) = \infty$ を得る．これで $\widetilde{I} \geq I$ が示されたので，補題 8.1.2 の証明が終わった． □

8.3 補題 8.1.3 の証明

本節の目的は補題 8.1.3 の証明であるが，その前にまず $\mathbb{R}$ 値確率変数に関する補題を１つ証明しよう．W を $\mathbb{R}$ 値確率変数とし，W から決まる対数積率母関数とその Fenchel-Legendre 変換をそれぞれ Λ_W と Λ_W^* と表す．

補題 8.3.1. W に対して条件 **(FEM)**，すなわち任意の $\lambda \in \mathbb{R}$ に対して $\Lambda_W(\lambda) < \infty$ となることを仮定する．このとき任意の $\delta \in (0, 1)$ に対して，

$$\mathbb{E}\left[\exp(\delta\Lambda_W^*(W))\right] \leq \frac{1+\delta}{1-\delta} < \infty$$

が成立する．

証明. 仮定により W は可積分なので，本証明中では $m := \mathbb{E}[W] \in \mathbb{R}$ と表す．Λ_W^* は下半連続な非負凸関数で，Jensen の不等式により $\Lambda_W^*(m) = 0$ を満たす．凸性により Λ_W^* は $[m, \infty)$ 上で広義単調増大，$(-\infty, m]$ 上で広義単調減少である（系 A.1.4 を見よ）.[10]

さらに $\lim_{x\to\pm\infty} \Lambda_W^*(x) = \infty$ となることを確認しよう．これを否定すると，ある $c \in (0, \infty)$ と $\lim_{n\to\infty} |x_n| = \infty$ を満たす実数列 $\{x_n\}_{n\in\mathbb{N}}$ が存在して，任意の $\lambda \in \mathbb{R}$ に対して

[10]区間上の凸関数の常識を付録（第 A.1 節）にまとめた．

$\lambda x_n - \Lambda_W(\lambda) \le c$ を満たす．この不等式で特に $\lambda = 2c/x_n$ と選ぶと $2c - \Lambda_W(2c/x_n) \le c$ を得る．最後に $n \to \infty$ として，**(FEM)** の下で Λ_W が原点で連続であることを用いると $2c \le c$ を得るが，これは明らかに矛盾である．

さて $x \in \mathbb{R}$ のとき，次が成り立つ．

$$e^{\Lambda_W(\lambda)} = \mathbb{E}\left[e^{\lambda W}\right] \ge \begin{cases} e^{\lambda x}\mathbb{P}(W > x), & \lambda \ge 0 \text{ のとき}, \\ e^{\lambda x}\mathbb{P}(W < x), & \lambda \le 0 \text{ のとき}. \end{cases}$$

Jensen の不等式から導かれる不等式 $\lambda x - \Lambda_W(\lambda) \le \lambda(x - m)$ をこれと組み合わせると，

$$\Lambda_W^*(x) = \begin{cases} \sup_{\lambda \ge 0}\{\lambda x - \Lambda_W(\lambda)\} \le -\log \mathbb{P}(W > x), & x \ge m \text{ のとき}, \\ \sup_{\lambda \le 0}\{\lambda x - \Lambda_W(\lambda)\} \le -\log \mathbb{P}(W < x), & x \le m \text{ のとき}, \end{cases}$$

が比較的簡単に示せる．整理すると，$x \ge m$ のときは $\mathbb{P}(W > x) \le \exp(-\Lambda_W^*(x))$ であり，$x \le m$ のときは $\mathbb{P}(W < x) \le \exp(-\Lambda_W^*(x))$ である．

$s \in (0, \infty)$ を任意とする．最初に注意した Λ_W^* の性質により，$\{x \in \mathbb{R} \mid \Lambda_W^*(x) \le s\}$ は有界閉区間（１点集合に退化する場合も含む）であるので，この区間を $[a(s), b(s)]$ と表すと $-\infty < a(s) \le m \le b(s) < \infty$ である．s に関して $b(s)$ は広義単調増大なので，$b := \lim_{s\to\infty} b(s) \in [m, \infty]$ が存在する．同様に $a(s)$ は広義単調減少なので，$a := \lim_{s\to\infty} a(s) \in [-\infty, m]$ が存在する．（ほぼ明らかだが，$b = \inf\{x > m \mid \Lambda_W^*(x) = \infty\}$ かつ $a = \sup\{x < m \mid \Lambda_W^*(x) = \infty\}$ である．）作り方により

$$\mathbb{P}(\Lambda_W^*(W) > s) = \mathbb{P}(W < a(s)) + \mathbb{P}(W > b(s)) \tag{8.3.1}$$

が明らかに成り立つ．

もし $b < \infty$ とすると，(b, ∞) 上で $\Lambda_W^* \equiv \infty$ であるから

$$\mathbb{P}(W > b) = \lim_{k\to\infty} \mathbb{P}(W > b + 1/k) \le \lim_{k\to\infty} \exp(-\Lambda_W^*(b + 1/k)) = 0$$

である．同様に $a > -\infty$ とすると，$\mathbb{P}(W < a) = 0$ である．よって，仮に $a = b$ であればほとんど確実に $W = m$（定値写像）であるため，補題は明らかに成立する．よって以下では $a \ne b$ と仮定するが，そのとき必然的に $a < m < b$ となる．

開区間上で定義された有限値凸関数は自動的に連続である（命題 A.1.3 を見よ）．$b(s) < b$ とすると，Λ_W^* は $b(s)$ を含むある開区間上で有限値であるため，連続でもある．したがって $s = \Lambda_W^*(b(s))$，特に $\mathbb{P}(W > b(s)) \le e^{-s}$ を得る．$b(s) = b$ の場合もこの評価は成立していたので，いずれにせよ $\mathbb{P}(W > b(s)) \le e^{-s}$ が成り立つ．同様に $\mathbb{P}(W < a(s)) \le e^{-s}$ も成り立つ．これと (8.3.1) により，尾確率の評価 $\mathbb{P}(\Lambda_W^*(W) > s) \le 2e^{-s}$ が得られた．

尾確率の評価から可積分性を導く標準的な手法により

$$\mathbb{E}\left[\exp(\delta\Lambda_W^*(W))\right] = 1 + \mathbb{E}\left[\int_0^\infty \mathbf{1}_{\{s < \delta\Lambda_W^*(W)\}} e^s ds\right]$$

$$= 1 + \int_0^\infty \mathbb{P}\left(s < \delta\Lambda_W^*(W)\right)e^s ds$$
$$\leq 1 + \int_0^\infty 2e^{(1-1/\delta)s}ds = \frac{1+\delta}{1-\delta}$$

とわかる．なお２つ目の等号で Fubini の定理を用いた． □

補題 8.1.3 の証明. $\{\widetilde{\mu}^n\}$ が $\mathcal{C}_0$ において指数的に緊密であることを示す．$X_1 = (X_1^1, \ldots, X_1^d)$ と成分表示し，各 j $(1 \leq j \leq d)$ に対して

$$\Lambda_j(\lambda) := \log \mathbb{E}\left[e^{\lambda X_1^j}\right], \qquad \lambda \in \mathbb{R}$$

と定め，Λ_j^* を Λ_j の Fenchel-Legendre 変換とする．また各 $r \in (0, \infty)$ と j に対して，

$$K_r^j := \left\{ w = (w^1, \ldots, w^d) \in \mathcal{AC}_0 \,\middle|\, \int_0^T \Lambda_j^*((w^j)'_u)du \leq r \right\}, \qquad K_r := \bigcap_{1\leq j\leq d} K_r^j$$

と定める．Lebesgue 測度に関してほとんど全ての t に対して $(\widetilde{Z}^n)'_t = X_{\lfloor nt\rfloor+1}$ なので，

$$\begin{aligned}
\widetilde{\mu}^n(K_r^c) = \widetilde{\mu}^n\left(\bigcup_{1\leq j\leq d}(K_r^j)^c\right) &\leq \sum_{j=1}^d \mathbb{P}\left(\int_0^T \Lambda_j^*(X^j_{\lfloor ns\rfloor+1})ds > r\right) \\
&\leq \sum_{j=1}^d \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor+1} \Lambda_j^*(X_i^j) > r\right) \\
&\leq d \max_{1\leq j\leq d} \mathbb{P}\left(\exp(\sum_{i=1}^{\lfloor nT\rfloor+1} \Lambda_j^*(X_i^j)/2) > e^{nr/2}\right) \\
&\leq de^{-nr/2} \max_{1\leq j\leq d} \mathbb{E}\left[\prod_{i=1}^{\lfloor nT\rfloor+1} \exp(\Lambda_j^*(X_i^j)/2)\right] \\
&\leq de^{-nr/2} \max_{1\leq j\leq d} \mathbb{E}\left[\exp(\Lambda_j^*(X_1^j)/2)\right]^{\lfloor nT\rfloor+1} < \infty
\end{aligned}$$

となる．最後から２つ目の不等式で Chebyshev の不等式を使い，最後の不等式で各 j について $\{X_i^j\}_{i\in\mathbb{N}}$ が独立同分布であることを使った．なお最右辺の期待値は補題 8.3.1 により有限である．これから直ちに

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{n} \log \widetilde{\mu}^n(K_r^c) \leq \lim_{r\to\infty}\left(-\frac{r}{2} + T \max_{1\leq j\leq d} \log \mathbb{E}\left[\exp(\Lambda_j^*(X_1^j)/2)\right]\right) = -\infty$$

を得る．

後は K_r が相対コンパクトであることを見ると指数的に緊密であることがわかる．Ascoli-Arzelà の定理により，これは K_r が一様有界かつ同程度連続な関数族であることと同値である．$w \in K_r$ を任意に取る．まず Λ_j^* の凸性と Jensen の不等式により次を得る．

$$\Lambda_j^*\left(\frac{w_t^j - w_s^j}{t-s}\right) \le \frac{1}{t-s}\int_s^t \Lambda_j^*((w^j)'_u)du \le \frac{r}{t-s}, \qquad 0 \le s < t \le T,\, 1 \le j \le d.$$

これをほぼ自明の不等式

$$\Lambda_j^*(x) \ge M|x| - \{\Lambda_j(M) \vee \Lambda_j(-M)\}, \qquad M > 0,\, x \in \mathbb{R}$$

と組み合わせると，任意の $M > 0$, $0 \le s < t \le T$ および $1 \le j \le d$ に対して

$$|w_t^j - w_s^j| \le \frac{1}{M}\big[r + (t-s)\{\Lambda_j(M) \vee \Lambda_j(-M)\}\big] \tag{8.3.2}$$

が得られる．右辺に現れる量はどれも w には依存しないことに注意せよ．この式で $s=0$ および $M=1$ とすると，$\sup_{w\in K_r}\|w^j\|_\infty < \infty$ がわかる．これで一様有界性が示せた．

次は同程度連続性を示そう．以下では $\delta \in (0,T]$ とする．Λ_j は $\mathbb{R}$ 上の連続関数で $\Lambda_j(0) = 0$ なので，

$$M_j(\delta) := \delta^{-1} \wedge \sup\{\tau > 0 \mid \sup_{|\lambda|\le\tau} \Lambda_j(\lambda) \le \delta^{-1}\}$$

と定めると，$\Lambda_j(M_j(\delta)) \vee \Lambda_j(-M_j(\delta)) \le \delta^{-1}$ かつ $\lim_{\delta\searrow 0} M_j(\delta) = \infty$ となる．これを (8.3.2) に戻すと

$$\sup_{0\le t-s\le\delta} |w_t^j - w_s^j| \le (r+1)/M_j(\delta)$$

を得る．右辺は $w \in K_r$ には依存せず，$\delta \searrow 0$ のとき 0 に収束するので，これで同程度連続性が示せた．以上で補題 8.1.3 の証明を終える．□

8.4　連続時間パラメータの場合への拡張

本章の最後に Mogulskii の大偏差原理を離散パラメータ $n \in \mathbb{N}$ $(n \to \infty)$ の場合から，連続パラメータ $\varepsilon \in (0,1]$ $(\varepsilon \searrow 0)$ の場合に拡張しておく．$\varepsilon \in (0,1]$ に対して，新しい確率過程 $Y^\varepsilon = (Y_t^\varepsilon)$ を

$$Y_t^\varepsilon := \varepsilon \sum_{i=1}^{\lfloor t/\varepsilon \rfloor} X_i, \qquad t \ge 0 \tag{8.4.1}$$

により定める．($\varepsilon = 1/n$ のときは $Y^\varepsilon = Z^n$ である．) $(Y_t^\varepsilon)_{0\le t\le T}$ が $\mathcal{L}^\infty$ 上に誘導する法則を ν^ε と書く．

定理 8.4.1. 条件 **(FEM)** を仮定する．$\varepsilon \searrow 0$ とするとき，上記の確率測度の族 $\{\nu^\varepsilon\}_{\varepsilon\in(0,1]}$ は $\mathcal{L}^\infty$ 上で (8.0.2) で定義された良い速度関数 I に対して速度 $1/\varepsilon$ で大偏差原理を満たす．

証明. 定理の主張は $\lim_{m\to\infty}\varepsilon_m = 0$ となる任意の狭義単調減少列 $\{\varepsilon_m\}_{m\in\mathbb{N}}$ に対して，$\{\nu^{\varepsilon_m}\}_{m\in\mathbb{N}}$ は $m\to\infty$ のときに $\mathcal{L}^\infty$ 上で I に対して速度 $1/\varepsilon_m$ で大偏差原理を満たすことと同値である．以下ではこれを示す．

もし任意の m に対して $\varepsilon_m^{-1}\in\mathbb{N}$ であるなら，単に Mogulskii の大偏差原理の離散版（定理 8.0.1）の部分列を取っただけなので，明らかにこの大偏差原理は成立する．次は一般の場合を考える．$n_m := \lfloor \varepsilon_m^{-1}\rfloor$ とおくと，定理 8.0.1 により $\{\mu_{n_m}\}_{m\in\mathbb{N}}$ は $m\to\infty$ のときに I に対して速度 n_m で大偏差原理を満たす．したがって，任意の $\delta>0$ に対して

$$\varlimsup_{m\to\infty}\frac{1}{n_m}\log\mathbb{P}\left(\|Y^{\varepsilon_m}-Z^{n_m}\|_\infty>\delta\right)=-\infty \tag{8.4.2}$$

を示せば十分である．実際 (8.4.2) は $\{Y^{\varepsilon_m}\}$ と $\{Z^{n_m}\}$ が速度 n_m に対して指数的に同値であることを意味するため，命題 1.6.2 および $\lim_{m\to\infty} n_m\varepsilon_m=1$ により，$\{Y^{\varepsilon_m}\}$ に対する所望の大偏差原理が従う．

以下で (8.4.2) を示す．まず $1-\varepsilon_m\le\varepsilon_m n_m\le 1$ と $\lfloor n_m t\rfloor\le\lfloor t/\varepsilon_m\rfloor\le\lfloor n_m t\rfloor+\lfloor T\rfloor+1$ に注意せよ．[11] Z^n と Y^ε の定義式 (8.0.1) と (8.4.1) により

$$\begin{aligned}
|Y_t^{\varepsilon_m}-Z_t^{n_m}| &\le (1-\varepsilon_m n_m)|Z_t^{n_m}|+\varepsilon_m\sum_{k=1}^{\lfloor T\rfloor+1}|X_{\lfloor n_m t\rfloor+k}|\\
&\le \frac{\varepsilon_m}{n_m}\sum_{i=1}^{\lfloor n_m t\rfloor}|X_i|+\varepsilon_m\sum_{k=1}^{\lfloor T\rfloor+1}|X_{\lfloor n_m t\rfloor+k}|\\
&\le 2\varepsilon_m(\lfloor T\rfloor+1)\max_{1\le i\le\lfloor n_m T\rfloor+\lfloor T\rfloor+1}|X_i|
\end{aligned}$$

と評価できる．この右辺は t に依存しない．すると $\{X_i\}$ が同分布なので，

$$\begin{aligned}
&\frac{1}{n_m}\log\mathbb{P}\left(\|Y^{\varepsilon_m}-Z^{n_m}\|_\infty>\delta\right)\\
&\le\frac{1}{n_m}\log\mathbb{P}\left(\cup_{1\le i\le\lfloor n_m T\rfloor+\lfloor T\rfloor+1}\{2\varepsilon_m(\lfloor T\rfloor+1)|X_i|>\delta\}\right)\\
&\le\frac{1}{n_m}\log(\lfloor n_m T\rfloor+\lfloor T\rfloor+1)+\frac{1}{n_m}\log\mathbb{P}\left(2\varepsilon_m(\lfloor T\rfloor+1)|X_1|>\delta\right)
\end{aligned}\tag{8.4.3}$$

を得る．明らかに (8.4.3) の右辺第 1 項は $m\to\infty$ のとき 0 に収束する．よって (8.4.3) の右辺第 2 項が $-\infty$ に発散することを確認すると (8.4.2) が得られて，本定理の証明が終わる．

簡単のために $\widetilde{\varepsilon}_m := 2\varepsilon_m(\lfloor T\rfloor+1)\sqrt{d}$ とおく．仮定 **(FEM)** により X_1 の各成分は，任意の $r>0$ に対して $\mathbb{E}\left[\exp(r|X_i^j|)\right]<\infty$ を満たす．$\kappa>0$ のとき，$|X_1|\ge\kappa$ ならばある j $(1\le j\le d)$ に対して $|X_1^j|\ge\kappa/\sqrt{d}$ が成り立つため，

$$\frac{1}{n_m}\log\mathbb{P}\left(2\varepsilon_m(\lfloor T\rfloor+1)|X_1|>\delta\right)$$

[11] 不等式 $\lfloor a+b\rfloor\le\lfloor a\rfloor+\lfloor b\rfloor+1$ $(a,b\in\mathbb{R})$ を用いた．

$$
\begin{aligned}
&= \frac{1}{n_m} \log \mathbb{P}\left(\cup_{1\le j\le d}\{|X_1^j| > \delta/\widetilde{\varepsilon}_m\}\right) \\
&\le \frac{\log d}{n_m} + \max_{1\le j\le d} \frac{1}{n_m} \log \mathbb{P}\left(\exp(r|X_1^j|) > \exp(r\delta/\widetilde{\varepsilon}_m)\right) \\
&\le \frac{\log d}{n_m} + \max_{1\le j\le d} \frac{1}{n_m} \log\Big(\exp(-r\delta/\widetilde{\varepsilon}_m)\mathbb{E}\left[\exp(r|X_1^j|)\right]\Big) \\
&\le \frac{\log d}{n_m} - \frac{r\delta}{n_m\widetilde{\varepsilon}_m} + \max_{1\le j\le d} \frac{1}{n_m} \log \mathbb{E}\left[\exp(r|X_1^j|)\right] \qquad (8.4.4)
\end{aligned}
$$

と評価できる．上で注意した指数可積分性により，(8.4.4) の右辺第１項と第３項は $m \to \infty$ のとき 0 に収束する．よって $\lim_{m\to\infty} n_m\varepsilon_m = 1$ を思い出すと，

$$
\overline{\lim_{m\to\infty}} \frac{1}{n_m} \log \mathbb{P}\left(2\varepsilon_m(\lfloor T\rfloor + 1)|X_1| > \delta\right) \le -\frac{r\delta}{2(\lfloor T\rfloor + 1)\sqrt{d}}
$$

となる．ここで r は任意なので，$r \nearrow \infty$ として求める評価を得る．これで (8.4.2) が証明できた． □

第9章　Schilderの大偏差原理

本章ではSchilderの大偏差原理を論ずる.[1] これは $\varepsilon > 0$ を十分小さな正定数とするときに，$\sqrt{\varepsilon}$ 倍されたBrown運動の法則が満たす大偏差原理のことである.

$T > 0$ と $d \in \mathbb{N}$ を任意とし，原点から出発する d 次元の連続な経路全体がなす空間を

$$\mathcal{C}_0 = \{w = (w^1, \dots, w^d)\colon [0,T] \to \mathbb{R}^d \mid w \text{ は連続かつ } w_0 = 0\} \tag{9.0.1}$$

と書く．通常の一様ノルム $\|w\|_\infty = \sup_{0\le t\le T} |w_t|$ の下で $\mathcal{C}_0$ は可分なBanach空間になるので，$\mathcal{C}_0$ 上にはBorel加法族 $\mathcal{B}(\mathcal{C}_0)$ を入れる．よく知られているように，この空間上にWiener測度という非常に重要な確率測度があるが，本章ではこれを $\mathbb{P}$ と書く.（$\mathbb{P}$ は d 次元標準Brown運動の法則だとも言える.）当然ながら $\mathbb{P}$ の下では座標過程 $(w_t)_{0\le t\le T}$ は d 次元標準Brown運動である．本章では基本的に確率空間 $(\mathcal{C}_0, \mathcal{B}(\mathcal{C}_0), \mathbb{P})$ 上で作業する.

任意の $\varepsilon > 0$ に対して，$\mathbb{P}$ の下での定数倍写像 $w \in \mathcal{C}_0 \mapsto \sqrt{\varepsilon} w \in \mathcal{C}_0$ の法則を $\mathbb{P}^\varepsilon$ と書く．すなわち任意の $A \in \mathcal{B}(\mathcal{C}_0)$ に対して，$\mathbb{P}^\varepsilon(A) = \mathbb{P}(\sqrt{\varepsilon} w \in A)$ として $\mathbb{P}^\varepsilon$ を定める．$\mathbb{P}^\varepsilon$ は分散が ε 倍された d 次元Brown運動の法則である．$\varepsilon \searrow 0$ の時に $\mathbb{P}^\varepsilon$ は $\mathcal{C}_0$ の零元におけるDirac測度 δ_0 に弱収束することに注意すると，これは大偏差原理を問う典型的な状況であることがわかる．この問いに精密な答えを与えるのが本章の主題であるSchilderの大偏差原理である.

次にBrown運動に関する解析において非常に重要な役割を果たすCameron-Martin空間

$$\mathcal{H} = \left\{h = (h^1, \dots, h^d) \in \mathcal{C}_0 \mid h \text{ は絶対連続かつ } \|h\|_{\mathcal{H}}^2 := \textstyle\int_0^T |h'_t|^2 dt < \infty\right\} \tag{9.0.2}$$

を思い出そう．$h, k \in \mathcal{H}$ に対して，内積は通常通り

$$\langle h, k\rangle_{\mathcal{H}} := \int_0^T h'_t \cdot k'_t \, dt$$

と定義される．定義の仕方により，時間微分および不定積分という簡単な操作を通じて，Cameron-Martin空間と $L^2([0,T], \mathbb{R}^d)$ はユニタリ同型である．特に $\mathcal{H}$ は可分なHilbert空間である.

[1] 本章ではEuclid空間の標準内積を $\langle v, v'\rangle_{\mathbb{R}^d}$ ではなく $v \cdot v'$ と書く（$v, v' \in \mathbb{R}^d$）．ノルムは $|v| := \sqrt{v \cdot v}$ と書く．一方，2本線のノルム $\|\cdot\|$ は無限次元空間のノルムに対してのみ使う.

Schilder の大偏差原理に現れる速度関数 $I\colon \mathcal{C}_0 \to [0,\infty]$ は以下のように定義される．

$$I(w) = \begin{cases} \frac{1}{2}\|w\|_{\mathcal{H}}^2, & w \in \mathcal{H}, \\ \infty, & w \notin \mathcal{H}. \end{cases} \tag{9.0.3}$$

後で示すが，これは良い速度関数である．(証明は簡単ではないが，実は $\mathbb{P}^{\varepsilon}(\mathcal{H}) = 0$ である．ところが Schilder の大偏差原理に「効いている」のがこの零集合 $\mathcal{H}$ だけだというのは，もし予備知識がなければ少し不思議に思える現象である.)

Schilder の大偏差原理の主張を正確に述べる．本章の残りでこの大偏差原理に 3 種類の証明を与える．

定理 9.0.1. $\varepsilon \searrow 0$ とするとき，$\mathcal{C}_0$ 上の Borel 確率測度の族 $\{\mathbb{P}^{\varepsilon}\}_{\varepsilon\in(0,1]}$ は良い速度関数 I に対して速度 $1/\varepsilon$ で大偏差原理を満たす.

後述する定理 9.5.3 で見るように，実はこの大偏差原理の位相は α-Hölder 位相（$0 < \alpha < 1/2$）に強めることができる．

9.1 経路空間に関する基本事項

$\mathcal{C}_0$ と $\mathcal{H}$ に関する基本的な事実のうち，定理 9.0.1 の証明中で使うものを本節で簡潔にまとめておく．

補題 9.1.1. 埋め込み写像 $\mathcal{H} \hookrightarrow \mathcal{C}_0$ はコンパクトである．すなわち $\mathcal{H}$ の任意の有界集合は，$\mathcal{C}_0$ において相対コンパクトである．またこの埋め込みは稠密である．

証明. まず $\mathcal{H}$ 内の任意の有界点列 $\{h(n)\}_{n\in\mathbb{N}}$ から，$\mathcal{C}_0$ の位相で収束する部分列を抜き出せることを示す．$R := \sup_{n\in\mathbb{N}} \|h(n)\|_{\mathcal{H}} < \infty$ とおくと，$0 \le s \le t \le T$ のときに Schwarz の不等式により

$$|h(n)_t - h(n)_s| \le \int_s^t |h(n)'_u| du \le \sqrt{t-s}\Big\{\int_s^t |h(n)'_u|^2 du\Big\}^{1/2} \le R\sqrt{t-s}$$

が成り立つ．右辺は n に依存しないので，$\{h(n)\}$ は同程度連続である．また $h(n)_0 = 0$ であるので，この式から $\|h(n)\|_\infty \le R\sqrt{T}$ も従うため，$\{h(n)\}$ は一様有界でもある．よって Ascoli-Arzelà の定理により，$\{h(n)\}$ は $\mathcal{C}_0$ の位相で収束する部分列を持つことがわかる．これで埋め込みがコンパクトであることが示せた．

折れ線近似を用いて稠密性を示す．$w \in \mathcal{C}_0$ の連続率を $M_w(\delta) := \sup_{0\le t-s\le\delta} |w_t - w_s|$ とおく $(0 < \delta \le T)$．w の一様連続性から各 w に対して $\lim_{\delta\searrow 0} M_w(\delta) = 0$ である．w に対して n 等分割 $\{jT/n \mid 0 \le j \le n\}$ に対応する折れ線近似を $w(n)$ と書く．このとき

$$\|w - w(n)\|_\infty = \sup_{0\le t\le T} |w_t - w(n)_t|$$

$$\leq \sup_{0\leq t\leq T}\{|w_t - w_{[t]_n}| + |w_{[t]_n} - w(n)_t|\} \leq 2M_w(1/n).$$

よって $\lim_{n\to\infty}\|w-w(n)\|_\infty = 0$ である．ここで $[t]_n := \lfloor tn/T\rfloor\cdot(T/n)$ は t を含む小区間の左端である．折れ線は明らかに $\mathcal{H}$ に属するので，これは $\mathcal{H}$ の稠密性を意味する． □

ここで $\mathcal{H}$ の弱収束や弱位相について復習しよう．$\mathcal{H}$ 内の点列 $\{h(n)\}_{n\in\mathbb{N}}$ が $h\in\mathcal{H}$ に弱収束するとは，各 $k\in\mathcal{H}$ に対して $\lim_{n\to\infty}\langle h(n),k\rangle_{\mathcal{H}} = \langle h,k\rangle_{\mathcal{H}}$ が成立することである．もちろん強収束（すなわち $\lim_{n\to\infty}\|h(n)-h\|_{\mathcal{H}} = 0$）すれば弱収束もするが，逆は一般には成立しない．

$\mathcal{H}$ の弱位相は弱すぎるため点列の情報だけからでは位相を特定できない．弱位相を正確に記述するには，各 $g\in\mathcal{H}$ における基本近傍系を指定するとよい．

$$\widetilde{V}_\varepsilon^{l_1,\dots,l_m}(g) := \{f\in\mathcal{H} \mid \max_{1\leq i\leq m}|\langle f,l_i\rangle_{\mathcal{H}} - \langle g,l_i\rangle_{\mathcal{H}}| < \varepsilon\}$$

とおき，$\{\widetilde{V}_\varepsilon^{l_1,\dots,l_m}(g) \mid \varepsilon>0,\, m\in\mathbb{N},\, l_1,\dots,l_m\in\mathcal{H}\}$ を弱位相に関する g における基本近傍系と定める．与えられた $\widetilde{V}_\varepsilon^{l_1,\dots,l_m}(g)$ に対して十分小さな $\delta>0$ を取れば，明らかに $B_{\mathcal{H}}(g,\delta)\subset\widetilde{V}_\varepsilon^{l_1,\dots,l_m}(g)$ となるため，$\mathcal{H}$ の強位相（ノルムから定まる位相）は弱位相より強い．

補題 9.1.2. $r\in(0,\infty)$ を任意とする．$\mathcal{H}$ の閉球 $\overline{B}_{\mathcal{H}}(0,r)$ は $\mathcal{H}$ の弱位相の下でコンパクトかつ距離付け可能である．

証明. $\mathcal{H}$ は可分なので強位相に関して $\mathcal{H}$ で稠密な可算集合 $\{a_j \mid j\in\mathbb{N}\}$ が存在する．以下ではこれを固定して用いる．

まず対角線論法を用いて，弱位相の下での点列コンパクト性を証明する．(これは有名な Banach-Alaoglu の定理の特殊な場合である．) 任意の点列 $\{h(n)\}_{n\in\mathbb{N}}\subset\overline{B}_{\mathcal{H}}(0,r)$ から弱位相に関して収束する部分列を抜き出そう．$|\langle h(n),a_j\rangle_{\mathcal{H}}|\leq r\|a_j\|_{\mathcal{H}}$ となるので，各 j に対して $\{\langle h(n),a_j\rangle_{\mathcal{H}}\}_{n\in\mathbb{N}}\subset\mathbb{R}$ は有界列である．したがって，ある部分列 $\{n_{1,k}\}_k$ に対して $\{\langle h(n_{1,k}),a_1\rangle_{\mathcal{H}}\}_k$ は $k\to\infty$ のときに収束する．するとさらなる部分列 $\{n_{2,k}\}_k$ が存在して $\{\langle h(n_{2,k}),a_2\rangle_{\mathcal{H}}\}_k$ は収束する．この要領で部分列を抜き出すことを繰り返すと次を得る．各 j に対して部分列 $\{n_{j,k}\}_k$ が存在して，(i) $\{n_{j,k}\}_k$ は $\{n_{j-1,k}\}_k$ の部分列であり，(ii) $\{\langle h(n_{j,k}),a_j\rangle_{\mathcal{H}}\}_k$ は収束する．ここで対角線論法を使うと，各 j に対して $\{\langle h(n_{k,k}),a_j\rangle_{\mathcal{H}}\}_k$ は収束することがわかる．

次は一般の $l\in\mathcal{H}$ に対して $\{\langle h(n_{k,k}),l\rangle_{\mathcal{H}}\}_k$ が収束することを示す．稠密性により，任意の $\delta>0$ に対してある $j\in\mathbb{N}$ が存在して $\|l-a_j\|_{\mathcal{H}}<\delta$ となるため，

$$\begin{aligned}|\langle h(n_{k,k}),l\rangle_{\mathcal{H}} - \langle h(n_{k',k'}),l\rangle_{\mathcal{H}}| &\leq |\langle h(n_{k,k}),l\rangle_{\mathcal{H}} - \langle h(n_{k,k}),a_j\rangle_{\mathcal{H}}| \\ &\quad + |\langle h(n_{k,k}),a_j\rangle_{\mathcal{H}} - \langle h(n_{k',k'}),a_j\rangle_{\mathcal{H}}| \\ &\quad + |\langle h(n_{k',k'}),a_j\rangle_{\mathcal{H}} - \langle h(n_{k',k'}),l\rangle_{\mathcal{H}}|\end{aligned}$$

$$\leq 2r\delta + |\langle h(n_{k,k}), a_j\rangle_{\mathcal{H}} - \langle h(n_{k',k'}), a_j\rangle_{\mathcal{H}}|$$

と評価できる．上極限を取った後に $\delta \searrow 0$ とすると

$$\overline{\lim_{k,k'\to\infty}} |\langle h(n_{k,k}), l\rangle_{\mathcal{H}} - \langle h(n_{k',k'}), l\rangle_{\mathcal{H}}| = 0$$

が得られるので，$\{\langle h(n_{k,k}), l\rangle_{\mathcal{H}}\}_{k\in\mathbb{N}}$ が Cauchy 列であることがわかる．以下では $\xi\langle l\rangle := \lim_{k\to\infty}\langle h(n_{k,k}), l\rangle_{\mathcal{H}}$ と表す．極限を取る前の線形性が取った後にも遺伝するため，$l \mapsto \xi\langle l\rangle$ の線形性はほぼ明らかである．また $|\langle h(n_{k,k}), l\rangle_{\mathcal{H}}| \leq r\|l\|_{\mathcal{H}}$ の極限を取ると，$|\xi\langle l\rangle| \leq r\|l\|_{\mathcal{H}}$ を得るが，これは $\xi \in \mathcal{H}^*$ かつ $\|\xi\|_{\mathcal{H}^*} \leq r$ を意味する．よって，Riesz の表現定理で ξ に対応する元を $h \in \mathcal{H}$ と書けば，$h \in \overline{B}_{\mathcal{H}}(0,r)$ かつ $\xi\langle l\rangle = \langle h, l\rangle_{\mathcal{H}}$ である．以上により，$\{h(n_{k,k})\}_{k\in\mathbb{N}}$ がこの h に弱収束することがわかり，点列コンパクト性の証明が終わる．

最後に距離付け可能であることを示す．

$$\rho(f,g) := \sum_{j=1}^{\infty} 2^{-j}\{1 \wedge |\langle f, a_j\rangle_{\mathcal{H}} - \langle g, a_j\rangle_{\mathcal{H}}|\}, \qquad f, g \in \overline{B}_{\mathcal{H}}(0,r)$$

とおくと，ρ が $\overline{B}_{\mathcal{H}}(0,r)$ 上の距離関数になることは容易に示せる．以下で ρ が誘導する位相と弱位相が等しいことを示す．そのためには任意の $g \in \overline{B}_{\mathcal{H}}(0,r)$ における 2 つの位相の基本近傍系が同値であることを示せばよい．まず

$$U_\varepsilon := \{f \in \overline{B}_{\mathcal{H}}(0,r) \mid \rho(f,g) < \varepsilon\}$$

とおくと，$\{U_\varepsilon \mid \varepsilon > 0\}$ は ρ が誘導する位相に関して g における基本近傍系をなす．次は

$$V_\varepsilon^{l_1,\ldots,l_m} := \{f \in \overline{B}_{\mathcal{H}}(0,r) \mid \max_{1\leq i\leq m} |\langle f, l_i\rangle_{\mathcal{H}} - \langle g, l_i\rangle_{\mathcal{H}}| < \varepsilon\}$$

とおくと，$\{V_\varepsilon^{l_1,\ldots,l_m} \mid \varepsilon > 0,\, m \in \mathbb{N},\, l_1,\ldots,l_m \in \mathcal{H}\}$ は弱位相に関して g における基本近傍系をなす．

まず弱位相の方が強いことを見る．$\varepsilon > 0$ を任意とする．このとき $m \in \mathbb{N}$ を $\sum_{j=m+1}^{\infty} 2^{-j} < \varepsilon/2$ となるように取ると，$V_{\varepsilon/2}^{a_1,\ldots,a_m} \subset U_\varepsilon$ となる．実際，$f \in V_{\varepsilon/2}^{a_1,\ldots,a_m}$ のとき

$$\rho(f,g) \leq \sum_{j=1}^{m} 2^{-j}\{1 \wedge (\varepsilon/2)\} + \sum_{j=m+1}^{\infty} 2^{-j} < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

となるからである．次は ρ が誘導する位相の方が強いことを見る．$V_\varepsilon^{l_1,\ldots,l_m} = \cap_{i=1}^m V_\varepsilon^{l_i}$ に注意すると，任意の $\varepsilon > 0$ と $l \in \mathcal{H}$ に対してある $\delta > 0$ が $U_\delta \subset V_\varepsilon^l$ を満たすことを示せば十分である．これらの ε と l に対して，$\|l - a_j\|_{\mathcal{H}} < \varepsilon/(3r)$ となる a_j を取り，$\delta := 2^{-j}\{1 \wedge (\varepsilon/3)\}$ とおく．$f \in U_\delta$ とすると，$1 \wedge |\langle f, a_j\rangle_{\mathcal{H}} - \langle g, a_j\rangle_{\mathcal{H}}| < 1 \wedge (\varepsilon/3)$ であり，特に $|\langle f, a_j\rangle_{\mathcal{H}} - \langle g, a_j\rangle_{\mathcal{H}}| < \varepsilon/3$ である．すると

$$|\langle f, l\rangle_{\mathcal{H}} - \langle g, l\rangle_{\mathcal{H}}| \leq |\langle f, l\rangle_{\mathcal{H}} - \langle f, a_j\rangle_{\mathcal{H}}| + |\langle f, a_j\rangle_{\mathcal{H}} - \langle g, a_j\rangle_{\mathcal{H}}| + |\langle g, a_j\rangle_{\mathcal{H}} - \langle g, l\rangle_{\mathcal{H}}|$$

$$\begin{aligned}&\le \|f\|_{\mathcal{H}}\|l-a_j\|_{\mathcal{H}} + |\langle f, a_j\rangle_{\mathcal{H}} - \langle g, a_j\rangle_{\mathcal{H}}| + \|g\|_{\mathcal{H}}\|l-a_j\|_{\mathcal{H}} \\ &< r\varepsilon/(3r) + \varepsilon/3 + r\varepsilon/(3r) = \varepsilon\end{aligned}$$

となるため，$U_\delta \subset V^l_\varepsilon$ が得られる．これで 2 種類の位相が等しいことが証明できた． □

区分的 C^2 級関数についてここでまとめておく．$g \in \mathcal{C}_0$ が区分的 C^2 級であるとは，ある $[0,T]$ の分割 $\{0 = t_0 < t_1 < \cdots < t_N = T\}$ が存在して，各小区間 $[t_{i-1}, t_i]$ 上では g は C^2 級であることとする．区分的 C^2 級の関数は当然 $\mathcal{H}$ に属する．このような区分的 C^2 級の関数の全体を $\mathcal{P}C^2_0 := \{g \in \mathcal{H} \mid g$ は区分的 C^2 級 $\}$ と書く．C^1 級関数の全体が $L^2([0,T], \mathbb{R}^d)$ で稠密なことから，$\mathcal{P}C^2_0$ が $\mathcal{H}$ で稠密なことがわかる．

上記のような $g \in \mathcal{P}C^2_0$ と一般の $h \in \mathcal{H}$ に対して，

$$\begin{aligned}\langle g, h\rangle_{\mathcal{H}} &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} g'_t \cdot h'_t \, dt \\ &= \sum_{i=1}^N \left\{ g'_{t_i - 0} \cdot h_{t_i} - g'_{t_{i-1}+0} \cdot h_{t_{i-1}} - \int_{t_{i-1}}^{t_i} g''_t \cdot h_t \, dt \right\} \qquad (9.1.1)\end{aligned}$$

と部分積分できる．この右辺の形を見ると，$\langle g, \cdot\rangle_{\mathcal{H}} \in \mathcal{H}^*$ は $\mathcal{C}^*_0$ の元（すなわち $\mathcal{C}_0$ 上の連続汎関数）になるように拡張できる．それを下付き添字を消して $\langle g, \cdot\rangle$ と書く．（要するに $\langle g, w\rangle$ は (9.1.1) の右辺において h を w で置き換えたものである．）

$\tau \in [0,T]$ かつ $1 \le i \le d$ のときに，g として特に $g^{\tau,i}_t := (\tau \wedge t)e_i$ を選ぶと，任意の $w \in \mathcal{C}_0$ に対して $\langle g^{\tau,i}, w\rangle = w^i_\tau$ が成り立つことが簡単にわかる．（なお $\{e_i\}_{i=1}^d$ は $\mathbb{R}^d$ の標準基底である．）したがって，全ての $g \in \mathcal{P}C^2_0$ に対して $\langle g, w\rangle = \langle g, \widehat{w}\rangle$ であれば，$w = \widehat{w} \in \mathcal{C}_0$ である．

補題 9.1.3. $w \in \mathcal{C}_0$ のとき，

$$\|w\|_{\mathcal{H}} = \sup\{|\langle g, w\rangle| \mid g \in \mathcal{P}C^2_0,\ \|g\|_{\mathcal{H}} = 1\}. \qquad (9.1.2)$$

ただし，通常どおりに $w \in \mathcal{C}_0 \setminus \mathcal{H}$ のときは $\|w\|_{\mathcal{H}} = \infty$ と定めた．

証明. 簡単のため，(9.1.2) の右辺を $M_w \in [0,\infty]$ と書く．$\mathcal{P}C^2_0$ は $\mathcal{H}$ において稠密なので，$w \in \mathcal{H}$ のときは $M_w = \sup\{|\langle g, w\rangle| \mid g \in \mathcal{H}, \|g\|_{\mathcal{H}} = 1\} = \|w\|_{\mathcal{H}} < \infty$ が成立する．

逆に $M_w < \infty$ を仮定する．このとき $|\langle g, w\rangle| \le M_w \|g\|_{\mathcal{H}}$ が全ての $g \in \mathcal{P}C^2_0$ に対して成り立つが，稠密性と合わせて考えるとこれは $\langle \cdot, w\rangle$ が $\mathcal{H}^*$ の元に一意的に拡張できることを意味する．よって Riesz の表現定理により，$k \in \mathcal{H}$ が一意的に存在して $\langle g, w\rangle = \langle g, k\rangle_{\mathcal{H}}$ が全ての $g \in \mathcal{P}C^2_0$ に対して成り立つ．ここで g として上記の $g^{\tau,i}_t := (\tau \wedge t)e_i$ を取ると，$w = k$ が従う．特に $w \in \mathcal{H}$ である．これで証明が終わった． □

本節の最後に $h \in \mathcal{H}$ に対する Wiener 積分を思い出そう.[2] 通常どおりに

$$\langle h, w\rangle := \int_0^T h'_t \cdot dw_t$$

と定め，これを確率空間 $(\mathcal{C}_0, \mathcal{B}(\mathcal{C}_0), \mathbb{P})$ 上で定義された $\mathbb{R}$ 値確率変数とみなす．一般の $h \in \mathcal{H}$ に対しては，$\langle h, \cdot\rangle$ は $\mathcal{C}_0$ 上の連続線形汎関数にならないことを念のため注意しておく．ただし $g \in \mathcal{P}C_0^2$ のときは Wiener 積分に関しても部分積分が使えて

$$\langle g, w\rangle = \sum_{i=1}^N \left\{ g'_{t_i-0} \cdot w_{t_i} - g'_{t_{i-1}+0} \cdot w_{t_{i-1}} - \int_{t_{i-1}}^{t_i} g''_t \cdot w_t \, dt \right\}$$

となり (9.1.1) の右辺の h を w で置き換えたものと一致する.（つまりこの場合は $\langle g, w\rangle$ という記号が 2 種類の意味で定義されているが結局は一致する.）$\{\langle h, \cdot\rangle \mid h \in \mathcal{H}\}$ はよく知られているように $\mathcal{H}$ を添字集合にもつ平均 0 の Gauss 系である.[3] すなわち，$\langle h, \cdot\rangle$ の分布は平均 0 の正規分布で共分散は $\mathbb{E}\left[\langle h_1, \cdot\rangle\langle h_2, \cdot\rangle\right] = \langle h_1, h_2\rangle_{\mathcal{H}}$ である．

9.2　速度関数

本節では Schilder の大偏差原理に現れる速度関数 I が良いことを示す．

命題 9.2.1. 定義式 (9.0.3) で定めた関数 $I\colon \mathcal{C}_0 \to [0, \infty]$ は良い速度関数である．

証明. 任意の $r \in [0, \infty)$ に対して，$I^{-1}([0, r]) = \overline{B}_{\mathcal{H}}(0, \sqrt{2r})$ が $\mathcal{C}_0$ 内のコンパクト集合であることを示せばよい．補題 9.1.1 により，この集合は $\mathcal{C}_0$ 内で相対コンパクトなので，$\mathcal{C}_0$ の位相に関して閉であることを示せばよい．

$\{h(n)\}_{n\in\mathbb{N}}$ を $I^{-1}([0, r])$ 内の点列とし，$\mathcal{C}_0$ の位相に関してある $w \in \mathcal{C}_0$ に収束するとする．このとき w も再び $I^{-1}([0, r])$ に属すことを示す．補題 9.1.2 により Hilbert 空間の閉球は弱位相に関してコンパクトなので，適当な部分列 $\{h(n_j)\}_{j\in\mathbb{N}}$ と $I^{-1}([0, r])$ 内の点 $h(\infty)$ が存在して，$\lim_{j\to\infty} h(n_j) = h(\infty)$ と弱収束する.（すなわち任意の $k \in \mathcal{H}$ に対して $\lim_{j\to\infty}\langle h(n_j), k\rangle_{\mathcal{H}} = \langle h(\infty), k\rangle_{\mathcal{H}}$ となる.）

各 $\tau \in [0, T]$ と $1 \leq i \leq d$ に対して k として特に $g_t^{\tau,i} := (\tau \wedge t)e_i$ を選ぶと，任意の $h \in \mathcal{H}$ に対して $\langle h, g^{\tau,i}\rangle_{\mathcal{H}} = h_\tau^i$ が成り立つことから，$\lim_{j\to\infty} h(n_j)_\tau^i = h(\infty)_\tau^i$ となる．一方で $\{h(n_j)\}_{j\in\mathbb{N}}$ は一様ノルムに関して w に収束するので，当然 $\lim_{j\to\infty} h(n_j)_\tau^i = w_\tau^i$ となるが，これは $w = h(\infty)$ を意味する．これで $w \in I^{-1}([0, r])$ がわかった．　□

[2] Wiener 積分とは端的に言えば被積分関数が非ランダムな場合の（伊藤流）確率積分のことである．詳しくは [14, 第 1.7 節] を参照せよ．

[3] Gauss 系については例えば [2, 第 5.7 節] を参照せよ．

9.3 下からの評価

本節ではSchilderの大偏差原理（定理9.0.1）の下からの評価を示す．証明の中で鍵となるのはWiener測度の平行移動に関するCameron-Martinの定理である.[4]

命題 9.3.1. 任意の開集合 $O \subset \mathcal{C}_0$ に対して，次が成立する．

$$\varliminf_{\varepsilon \searrow 0} \varepsilon \log \mathbb{P}^{\varepsilon}(O) \geq -\inf_{w \in O} I(w).$$

証明. $O \neq \emptyset$ の場合のみを示せば十分である．補題9.1.1で示した稠密性により，このとき $O \cap \mathcal{H} \neq \emptyset$ であるため，証明すべき不等式の右辺は有限である．

$\mathcal{P}C_0^2$ が $\mathcal{H}$ において稠密であることと $\mathcal{H}$ の位相が $\mathcal{C}_0$ の位相よりも強いことから，任意の $h \in O \cap \mathcal{H}$ に対して関数列 $\{g_n\}_{n\in\mathbb{N}} \subset O \cap \mathcal{P}C_0^2$ で $\lim_{n\to\infty} \|g_n - h\|_{\mathcal{H}} = 0$ となるものが存在する．特に $O \cap \mathcal{P}C_0^2 \neq \emptyset$ であり，さらに次が成り立つ．

$$\inf_{w \in O} I(w) = \inf_{h \in O \cap \mathcal{H}} I(h) = \inf_{g \in O \cap \mathcal{P}C_0^2} I(g).$$

さて $g \in O \cap \mathcal{P}C_0^2$ としよう．このとき $\langle g, \cdot \rangle \in \mathcal{C}_0^*$ だったので，$M_g := \|\langle g, \cdot\rangle\|_{\mathcal{C}_0^*}$ と書こう．O は開なので，$\delta > 0$ が十分小さければ $B_{\mathcal{C}_0}(g, \delta) \subset O$ となる．したがって

$$\begin{aligned}
\mathbb{P}^{\varepsilon}(O) &\geq \mathbb{P}^{\varepsilon}(B_{\mathcal{C}_0}(g, \delta)) \\
&= \mathbb{P}\left(\sqrt{\varepsilon} w \in B_{\mathcal{C}_0}(g, \delta)\right) \\
&= \mathbb{E}\left[\mathbf{1}_{\{\sqrt{\varepsilon} w \in B_{\mathcal{C}_0}(0,\delta)\}} \exp\Big(-\frac{1}{\sqrt{\varepsilon}}\langle g, w\rangle - \frac{1}{2\varepsilon}\int_0^T |g_t'|^2 dt\Big)\right] \\
&\geq e^{-I(g)/\varepsilon}\mathbb{E}\left[\mathbf{1}_{\{\sqrt{\varepsilon} w \in B_{\mathcal{C}_0}(0,\delta)\}} \exp\big(-M_g\|w\|_\infty/\sqrt{\varepsilon}\big)\right] \\
&\geq e^{-\{I(g)+\delta M_g\}/\varepsilon}\mathbb{P}\left(\sqrt{\varepsilon} w \in B_{\mathcal{C}_0}(0, \delta)\right) \qquad (9.3.1)
\end{aligned}$$

となる．二つ目の等号において w を $g/\sqrt{\varepsilon}$ だけ平行移動する際に $\mathbb{P}$ に対するCameron-Martinの定理を用いた．

$\varepsilon \searrow 0$ のときに $\mathbb{P}^{\varepsilon}$ が $\mathcal{C}_0$ の零元における点測度 δ_0 に収束するので，開球の重みは $\mathbb{P}(\sqrt{\varepsilon} w \in B_{\mathcal{C}_0}(0,\delta)) = \mathbb{P}^{\varepsilon}(B_{\mathcal{C}_0}(0, \delta)) \to 1$ と収束する．したがって (9.3.1) の両辺の対数を取ると，

$$\varliminf_{\varepsilon \searrow 0} \varepsilon \log \mathbb{P}^{\varepsilon}(O) \geq -I(g) - \delta M_g$$

を得る．$\delta \searrow 0$ とした後で，g を $O \cap \mathcal{P}C_0^2$ を渡らせて上限を取ると

$$\varliminf_{\varepsilon \searrow 0} \varepsilon \log \mathbb{P}^{\varepsilon}(O) \geq -\inf_{g \in O \cap \mathcal{P}C_0^2} I(g) = -\inf_{w \in O} I(w)$$

となり，求める下からの評価が得られた． □

---

[4] Cameron-Martinの定理については例えば [14, 第1.7節] を見よ．（この定理はGirsanovの定理の特殊な場合である．）

9.4　上からの評価（その１）

本節では Schilder の大偏差原理（定理 9.0.1）の上からの評価に１つ目の証明を与える.[5] この証明方法は Brown 運動の初等的な性質のみを使うのでかなり平易である．その反面，Brown 運動に特有の性質を使うために，多くの確率過程に適用できるほどの一般性はない．

上からの評価を始める前にまず１変数の広義積分に関する補題を証明しよう．

補題 9.4.1. (1) $x > 0$ のとき，次が成り立つ．

$$\int_x^\infty e^{-y^2/2} dy \le \frac{1}{x} e^{-x^2/2}.$$

(2) $\alpha \ge 0$ とする．α のみに依存する正定数 C が存在して次の不等式を満たす．

$$\int_x^\infty e^{-y} y^\alpha dy \le C e^{-x}(x^\alpha + 1), \qquad x \ge 0.$$

証明. (1) は有名な誤差関数の評価である．

$$\int_x^\infty e^{-y^2/2} dy \le \int_x^\infty \frac{y}{x} e^{-y^2/2} dy = \frac{-1}{x} \left[e^{-y^2/2} \right]_{y=x}^\infty = \frac{1}{x} e^{-x^2/2}.$$

(2) を示す．$z = y - x$ とおいて変数を y から z に変換すると

$$\begin{aligned}\int_x^\infty e^{-y} y^\alpha dy &= e^{-x} \int_0^\infty e^{-z} (x+z)^\alpha dz \\ &\le e^{-x} \int_0^\infty e^{-z} 2^{(\alpha-1)\vee 0} (x^\alpha + z^\alpha) dz \\ &\le 2^{(\alpha-1)\vee 0} e^{-x} \{ x^\alpha + \Gamma(\alpha+1) \}\end{aligned}$$

と評価できる．ここで Γ はガンマ関数である． □

次に Brown 運動 $(w_t)_{t\in[0,T]} = (w_t^1, \ldots, w_t^d)_{t\in[0,T]}$ がある時刻までに原点の δ 近傍から出ていく確率に対する指数的な評価を思い出そう.[6]（なお次の補題では最終時刻 T も任意であることに注意せよ.）[7]

補題 9.4.2. d 次元 Brown 運動 $(w_t)_{t\in[0,T]}$ に対して，次の評価が成立する．

$$\mathbb{P}\left(\sup_{0\le s\le T} |w_s| > R \right) \le 2d \exp\left(-\frac{R^2}{2dT} \right), \qquad R > 0.$$

---

[5] 本節の記述は [46, 第 5 章] を参考にした．

[6] 反射原理 $\mathbb{P}(M_t > u) = 2\mathbb{P}(B_t > u)$ を使ってもこの種の評価は証明できる．反射原理については例えば [8, 第 2.6 節] を参照せよ．

[7] この補題だけを独立して読むことが可能である．

証明. まず $d=1$ の場合を示す．$\lambda>0$ に対して $M_t^\lambda := \exp(\lambda w_t - \lambda^2 t/2)$ と定めると，よく知られているように $(M_t^\lambda)_{t\in[0,T]}$ は真のマルチンゲールになる．特に任意の λ と t に対して $\mathbb{E}\left[M_t^\lambda\right]=1$ である．非負の劣マルチンゲールに対する Doob の不等式により[8]

$$\mathbb{P}\left(\sup_{0\le t\le T} M_t^\lambda \ge r\right) \le \frac{1}{r}\mathbb{E}\left[M_T^\lambda\right] = \frac{1}{r}, \qquad r>0$$

が成り立つ．もし $\sup_{0\le t\le T} w_t > R$ だとすると，

$$\sup_{0\le t\le T} M_t^\lambda \ge \exp\left(\lambda R - \lambda^2 T/2\right)$$

が成り立つため，

$$\mathbb{P}\left(\sup_{0\le t\le T} w_t > R\right) \le \mathbb{P}\left(\sup_{0\le t\le T} M_t^\lambda \ge \exp\left(\lambda R - \lambda^2 T/2\right)\right) \le \exp\left(-\lambda R + \lambda^2 T/2\right)$$

を得るが，特に $\lambda = R/T$ を代入すると結局 $\mathbb{P}\left(\sup_{0\le t\le T} w_t > R\right) \le e^{-R^2/(2T)}$ を得る．対称性から $-w$ も Brown 運動なので，$\mathbb{P}\left(\inf_{0\le t\le T} w_t < -R\right) \le e^{-R^2/(2T)}$ も明らかに成り立つ．よって

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le T}|w_t|>R\right) &= \mathbb{P}\left(\left\{\sup_{0\le t\le T} w_t > R\right\}\cup\left\{\inf_{0\le t\le T} w_t < -R\right\}\right)\\ &\le \mathbb{P}\left(\sup_{0\le t\le T} w_t > R\right) + \mathbb{P}\left(\inf_{0\le t\le T} w_t < -R\right) \le 2e^{-R^2/(2T)}\end{aligned}$$

を得る．これで $d=1$ の場合が証明できた．

一般の $d\in\mathbb{N}$ の場合を示す．$R>0$ のとき，

$$\begin{aligned}\|w\|_\infty > R \quad &\Longrightarrow \quad \sum_{i=1}^d \|w^i\|_\infty^2 \ge \sup_{t\in[0,T]}\sum_{i=1}^d |w_t^i|^2 > R^2\\ &\Longrightarrow \quad \text{ある } i\in\{1,\dots,d\} \text{ に対して } \ \|w^i\|_\infty > R/\sqrt{d}\end{aligned} \tag{9.4.1}$$

となる．よって，$\{\|w\|_\infty > R\} \subset \cup_{i=1}^d \{\|w^i\|_\infty > R/\sqrt{d}\}$ であるので，

$$\mathbb{P}\left(\|w\|_\infty > R\right) \le \sum_{i=1}^d \mathbb{P}\left(\|w^i\|_\infty > R/\sqrt{d}\right) \le 2de^{-R^2/(2dT)}$$

となる．ここで $d=1$ の場合の評価を使った．これで補題の証明が終わった．　□

Schilder の大偏差原理（定理 9.0.1）の上からの評価を示す．

[8]確率過程の径数が連続時間である場合の Doob の不等式については [11, 13, 16] などを参照せよ．

命題 9.4.3. 任意の閉集合 $F \subset \mathcal{C}_0$ に対して，次が成立する．

$$\varlimsup_{\varepsilon \searrow 0} \varepsilon \log \mathbb{P}^\varepsilon(F) \leq -\inf_{w \in F} I(w). \tag{9.4.2}$$

証明. F が空集合でないと仮定してよい．$\delta > 0$ に対して $F^\delta = \{w \in \mathcal{C}_0 \mid d_{\mathcal{C}_0}(w, F) \leq \delta\}$ および $\ell_\delta = \inf_{w \in F^\delta} I(w)$ とおく．F が閉であることと I が良いことに注意して補題 1.2.8 を使うと $\lim_{\delta \searrow 0} \ell_\delta = \inf_{w \in F} I(w)$ とわかる．F^δ は半径 δ の球を含むので，$\mathcal{H}$ の $\mathcal{C}_0$ 内での稠密性から F^δ 内に $\mathcal{H}$ の元は無限個存在する．特に $\ell_\delta < \infty$ である．

$[0,T]$ の n 等分点に対応した折れ線近似を与える写像 $\pi_n \colon \mathcal{C}_0 \to \mathcal{C}_0$ を考える（$n \geq 2$）．すなわち $\pi_n w$ は各 j $(1 \leq j \leq n)$ に対して $[(j-1)T/n, jT/n]$ 上で線分であり，かつ $(\pi_n w)_{jT/n} = w_{jT/n}$ を満たすものである．このとき，

$$\begin{aligned} w \in F &\Longleftrightarrow \text{“}w \in F \text{ かつ } \|w - \pi_n w\|_\infty \leq \delta\text{” または “}w \in F \text{ かつ } \|w - \pi_n w\|_\infty > \delta\text{”} \\ &\Longrightarrow \text{“}\pi_n w \in F^\delta\text{”} \quad \text{または} \quad \text{“}\|w - \pi_n w\|_\infty > \delta\text{”} \end{aligned}$$

である．これから直ちに

$$\begin{aligned} \mathbb{P}^\varepsilon(F) &\leq \mathbb{P}^\varepsilon\left(\pi_n w \in F^\delta\right) + \mathbb{P}^\varepsilon\left(\|w - \pi_n w\|_\infty > \delta\right) \\ &\leq \mathbb{P}^\varepsilon\left(I(\pi_n w) \geq \ell_\delta\right) + \mathbb{P}^\varepsilon\left(\|w - \pi_n w\|_\infty > \delta\right) \\ &= \mathbb{P}\left(\varepsilon I(\pi_n w) \geq \ell_\delta\right) + \mathbb{P}\left(\sqrt{\varepsilon}\|w - \pi_n w\|_\infty > \delta\right) \end{aligned} \tag{9.4.3}$$

が従う．２つ目の不等式は ℓ_δ の定義から明らかである．

まず不等式 (9.4.3) の右辺第１項を評価しよう．直接の計算で

$$2I(\pi_n w) = \int_0^T |(\pi_n w)'_u|^2 du = \sum_{i=1}^d \sum_{j=1}^n \left(\frac{w^i_{jT/n} - w^i_{(j-1)T/n}}{\sqrt{T/n}} \right)^2$$

となる．Brown 運動の独立増分性により，右辺は $\mathbb{P}$ の下では nd 個の独立な標準正規分布に従う確率変数の 2 乗和であるので，自由度 dn のカイ 2 乗分布に従う．その確率密度関数は

$$\mathbf{1}_{(0,\infty)}(y) 2^{-dn/2} \Gamma(dn/2)^{-1} y^{dn/2-1} e^{-y/2} dy$$

であることを思い出そう.[9]よって，

$$\begin{aligned} \mathbb{P}\left(\varepsilon I(\pi_n w) \geq \ell_\delta\right) &= \mathbb{P}\left(2I(\pi_n w) \geq 2\ell_\delta/\varepsilon\right) \\ &= C_1 \int_{2\ell_\delta/\varepsilon}^\infty y^{dn/2-1} e^{-y/2} dy \\ &= C_2 \int_{\ell_\delta/\varepsilon}^\infty y^{dn/2-1} e^{-y} dy \end{aligned}$$

[9]例えば [4, 第 6.3 節] や [25, 332L] を見よ．

$$\leq C_3 e^{-\ell_\delta/\varepsilon}\{(\ell_\delta/\varepsilon)^{dn/2-1}+1\}$$

と評価できる．ここで補題 9.4.1 (2) を使った．なお C_1, C_2, C_3 は d, n のみによる適当な正定数である．$\ell_\delta > 0$ のときは両辺の $\varepsilon \log$ の上極限を取って

$$\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\varepsilon I(\pi_n w) \geq \ell_\delta\right) \leq -\ell_\delta \tag{9.4.4}$$

を得る．この不等式は $\ell_\delta = 0$ のときも明らかに成り立つ．

不等式 (9.4.3) の右辺第２項を評価する前に簡単な事実をいくつか列挙する．$[0,T]$ の部分区間 J に対して，$\|w\|_{\infty,J} := \sup_{t\in J}|w_t|$ とおく．以下では $J(n,j) = [(j-1)T/n, jT/n]$ と書く．

- $r > 0$ のとき，$\{\|w\|_{\infty,J} > r\} \subset \cup_{i=1}^d \{\|w^i\|_{\infty,J} > r/\sqrt{d}\}$ である．これは補題 9.4.2 の証明中にある (9.4.1) 式ですでに事実上示した．

- $n \geq 2$ および $R > 0$ のとき，

$$\begin{aligned}\|w\|_\infty > R \quad &\Longleftrightarrow \quad \max_{1\leq j\leq n}\|w\|_{\infty,J(n,j)} > R \\ &\Longleftrightarrow \quad \text{ある } j \in \{1,\ldots,d\} \text{ に対して } \|w\|_{\infty,J(n,j)} > R\end{aligned}$$

  であるから，$\{\|w\|_\infty > R\} = \cup_{j=1}^d \{\|w\|_{\infty,J(n,j)} > R\}$ である．

- $n \geq 2$ および $R > 0$ のとき，

$$\|w\|_{\infty,J(n,1)} \leq R/2 \quad \Longrightarrow \quad \|w - \pi_n w\|_{\infty,J(n,1)} = \sup_{0\leq t\leq T/n}|w_t - nT^{-1}tw_{T/n}| \leq R$$

  である．この対偶をとると，$\{\|w-\pi_n w\|_{\infty,J(n,1)} > R\} \subset \{\|w\|_{\infty,J(n,1)} > R/2\}$ とわかる．

- w の時間に関する平行移動不変性，すなわち任意の $s \geq 0$ に対し $(w_{t+s} - w_s)_{t\geq 0}$ が再び Brown 運動であることを思い出そう．このとき任意の j に対して

$$\begin{aligned}&\|w - \pi_n w\|_{\infty,J(n,j)} \\ &= \sup_{(j-1)T/n\leq t\leq jT/n}\left|w_t - w_{(j-1)T/n} - nT^{-1}\big(t-(j-1)T/n\big)\big(w_{jT/n} - w_{(j-1)T/n}\big)\right| \\ &= \sup_{0\leq u\leq T/n}|\widetilde{w}_u - nT^{-1}u\widetilde{w}_{T/n}| \overset{\text{Law}}{=} \|w - \pi_n w\|_{\infty,J(n,1)}\end{aligned}$$

  となるので，$\{\|w - \pi_n w\|_{\infty,J(n,j)}\}_{j=1}^n$ は同分布である．(ここで $u := t - (j-1)T/n$ および $\widetilde{w}_u := w_{u+(j-1)T/n} - w_{(j-1)T/n}$ と書き，上記の平行移動不変性を用いた.）

上記の事実を使って，不等式 (9.4.3) の右辺第 2 項を評価すると

$$\begin{aligned}\mathbb{P}\left(\sqrt{\varepsilon}\|w-\pi_n w\|_\infty > \delta\right) &\le \mathbb{P}\left(\bigcup_{j=1}^{n}\{\|w-\pi_n w\|_{\infty,J(n,j)} > \delta/\sqrt{\varepsilon}\}\right)\\ &\le \sum_{j=1}^{n}\mathbb{P}\left(\|w-\pi_n w\|_{\infty,J(n,j)} > \delta/\sqrt{\varepsilon}\right)\\ &= n\mathbb{P}\left(\|w-\pi_n w\|_{\infty,J(n,1)} > \delta/\sqrt{\varepsilon}\right)\\ &\le n\mathbb{P}\left(\|w\|_{\infty,J(n,1)} > \delta/\sqrt{4\varepsilon}\right)\\ &\le 2dn\exp\left(-\frac{n\delta^2}{8Td\varepsilon}\right) \end{aligned} \tag{9.4.5}$$

となる．ここで補題 9.4.2 を使った．これから直ちに

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}\left(\sqrt{\varepsilon}\|w-\pi_n w\|_\infty > \delta\right) \le -\frac{n\delta^2}{8Td} \tag{9.4.6}$$

と求まる．(9.4.3)，(9.4.4)，(9.4.6) および補題 1.2.2 により

$$\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}^\varepsilon(F) \le -\min\left\{\ell_\delta, \frac{n\delta^2}{8Td}\right\}$$

を得るが，$\ell_\delta < \infty$ なので $n\to\infty$ として $\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}^\varepsilon(F) \le -\ell_\delta$ を得る．最後に $\delta\searrow 0$ として，この証明の冒頭に注意した事実を用いると証明が終わる．□

9.5　上からの評価（その 2）

本節では Schilder の大偏差原理（定理 9.0.1）の上からの評価（命題 9.4.3）に 2 つ目の証明を与える.[10] この証明法は Fernique の定理に代表される Brown 運動の Gauss 的な性質を用いるものの，Brown 運動に特有な性質は用いない．

まず α-Hölder ノルムを通常どおり以下のように定める（$0<\alpha\le 1$）．

$$\|w\|_\alpha := \sup_{0\le s<t\le T}\frac{|w_t-w_s|}{(t-s)^\alpha} \in [0,\infty], \qquad w\in\mathcal{C}_0$$

また α-Hölder 空間を $\mathcal{C}_0^\alpha = \{w\in\mathcal{C}_0 \mid \|w\|_\alpha<\infty\}$ とおく．補題 1.2.5 により，$w\mapsto\|w\|_\alpha$ は $\mathcal{C}_0$ 上で下半連続であり，特に Borel 可測である．よく知られているように，$0<\alpha<1/2$ のときに $\mathbb{E}[\|w\|_\alpha]<\infty$ となるので，[11] 特に $\mathbb{P}(\|w\|_\alpha<\infty)=1$ であり，Wiener 測度 $\mathbb{P}$ は実は $\mathcal{C}_0^\alpha$ 上の測度である．

[10] 本節の記述は主に [45, 第 4.1.5 節] を参考にした．

[11] Kolmogorov(-Čentsov) の連続変形定理による．[2, 定理 5.24]，[8, 定理 2.8]，[16, 定理 1.3.2] などを参照せよ．

まず Wiener 測度の下で α-Hölder ノルムは２乗指数可積分性を持つことを主張する Fernique の定理を思い出そう．

命題 9.5.1. $0<\alpha<1/2$ とする．このとき適当な正定数 $c=c_\alpha$ が存在して

$$\mathbb{E}\left[\exp(c\|w\|_\alpha^2)\right]<\infty$$

を満たす．

証明. 本証明中では次元を明記して $\mathbb{P}=\mathbb{P}_d$ および $\mathcal{C}_0^\alpha=\mathcal{C}_{0,d}^\alpha$ と書く．直積空間

$$\mathcal{C}_{0,2d}^\alpha=\mathcal{C}_{0,d}^\alpha\times\mathcal{C}_{0,d}^\alpha=\{(w,z)\mid w,z\in\mathcal{C}_{0,d}^\alpha\}$$

上に直積測度 $\mathbb{P}_{2d}=\mathbb{P}_d\otimes\mathbb{P}_d$ を入れる．$\mathbb{P}_{2d}$ は $2d$ 次元の Wiener 測度である．すなわちこの測度の下で座標過程 $(w_t,z_t)_{0\le t\le T}$ は $2d$ 次元の標準 Brown 運動である．Brown 運動を直交行列で回転させても法則は不変なので，この測度の下で

$$\left(\frac{w+z}{\sqrt{2}},\frac{w-z}{\sqrt{2}}\right)=\left(\frac{w_t+z_t}{\sqrt{2}},\frac{w_t-z_t}{\sqrt{2}}\right)_{0\le t\le T}$$

も再び $2d$ 次元の標準 Brown 運動になることに注意せよ．特に $\mathbb{P}_{2d}$ の下で，$(w+z)/\sqrt{2}$ と $(w-z)/\sqrt{2}$ は独立同分布で，共に $\mathbb{P}_d$ に従う．

以上により，$0<s<t$ のとき

$$\begin{aligned}
&\mathbb{P}_d\left(\|w\|_\alpha\le s\right)\mathbb{P}_d\left(\|w\|_\alpha>t\right)\\
&=\mathbb{P}_{2d}\left(\frac{\|w-z\|_\alpha}{\sqrt{2}}\le s\right)\mathbb{P}_{2d}\left(\frac{\|w+z\|_\alpha}{\sqrt{2}}>t\right)\\
&=\mathbb{P}_{2d}\left(\frac{\|w-z\|_\alpha}{\sqrt{2}}\le s,\quad\frac{\|w+z\|_\alpha}{\sqrt{2}}>t\right)\\
&\le\mathbb{P}_{2d}\left(\big|\|w\|_\alpha-\|z\|_\alpha\big|\le\sqrt{2}s,\quad\|w\|_\alpha+\|z\|_\alpha>\sqrt{2}t\right)\\
&\le\mathbb{P}_{2d}\left(\|w\|_\alpha>\frac{t-s}{\sqrt{2}},\quad\|z\|_\alpha>\frac{t-s}{\sqrt{2}}\right)\\
&=\mathbb{P}_d\left(\|w\|_\alpha>\frac{t-s}{\sqrt{2}}\right)^2.
\end{aligned}\tag{9.5.1}$$

なお下から２行目の不等号は，$\|w\|_\alpha$ と $\|z\|_\alpha$ の大小を場合分けすればごく簡単に示せる．

さて $\lim_{s\to\infty}\mathbb{P}_d\left(\|w\|_\alpha\le s\right)=1$ なので，s を十分大きく取れば $\mathbb{P}_d\left(\|w\|_\alpha\le s\right)\ge\frac{9}{10}$ となる．そのような s を一つ固定しよう．単調増加列 $\{t_n\}_{n=0}^\infty$ を $t_0:=s$ および $t_n:=s+\sqrt{2}t_{n-1}$（$n\in\mathbb{N}$）によって定める．すると (9.5.1) により

$$\mathbb{P}_d\left(\|w\|_\alpha\le s\right)\mathbb{P}_d\left(\|w\|_\alpha>t_n\right)\le\mathbb{P}_d\left(\|w\|_\alpha>t_{n-1}\right)^2$$

を得る．さらにこれから直ちに

$$\frac{\mathbb{P}_d(\|w\|_\alpha > t_n)}{\mathbb{P}_d(\|w\|_\alpha \le s)} \le \left(\frac{\mathbb{P}_d(\|w\|_\alpha > t_{n-1})}{\mathbb{P}_d(\|w\|_\alpha \le s)}\right)^2$$

という簡単な漸化式を得る．したがって帰納法により

$$\frac{\mathbb{P}_d(\|w\|_\alpha > t_n)}{\mathbb{P}_d(\|w\|_\alpha \le s)} \le \left(\frac{\mathbb{P}_d(\|w\|_\alpha > s)}{\mathbb{P}_d(\|w\|_\alpha \le s)}\right)^{2^n}$$

となる．直接計算により

$$t_n = (1 + 2^{1/2} + \cdots + 2^{n/2})s = \frac{2^{(n+1)/2} - 1}{\sqrt{2} - 1} s \le 4 \cdot 2^{n/2} s$$

と求まるので，任意の n に対して

$$\begin{aligned}
\mathbb{P}_d\left(\|w\|_\alpha > 4 \cdot 2^{n/2} s\right) &\le \mathbb{P}_d(\|w\|_\alpha > t_n) \\
&\le \mathbb{P}_d(\|w\|_\alpha \le s) \left(\frac{\mathbb{P}_d(\|w\|_\alpha > s)}{\mathbb{P}_d(\|w\|_\alpha \le s)}\right)^{2^n} \\
&\le \left(\frac{1/10}{9/10}\right)^{2^n} \le \left(\frac{1}{9}\right)^{2^n}
\end{aligned}$$

が成立する．

これを用いて所望の可積分性を示す．

$$\begin{aligned}
\mathbb{E}_d\left[\exp\left(\|w\|_\alpha^2/(32s^2)\right)\right] &\le e^{\frac{1}{2}} \mathbb{P}_d(\|w\|_\alpha \le 4s) \\
&\quad + \sum_{n=0}^{\infty} e^{2^n} \mathbb{P}_d\left(4 \cdot 2^{n/2} s < \|w\|_\alpha \le 4 \cdot 2^{(n+1)/2} s\right) \\
&\le e^{\frac{1}{2}} + \sum_{n=0}^{\infty} \left(\frac{e}{9}\right)^{2^n} < \infty.
\end{aligned}$$

したがって，正定数 $c = (32s^2)^{-1}$ に対して本命題の主張が成り立つ．　□

上記の Fernique の定理を用いると，$\{\mathbb{P}^\varepsilon\}_{0<\varepsilon\le 1}$ の指数的緊密性が簡単に証明できる．

命題 9.5.2. $\varepsilon \searrow 0$ のとき，$\{\mathbb{P}^\varepsilon\}_{0<\varepsilon\le 1}$ は $\mathcal{C}_0$ 上で指数的に緊密である．

証明. まず $\alpha \in (0, 1/2)$ を一つ取る．$r > 0$ のとき，$A_r := \{w \in \mathcal{C}_0 \mid \|w\|_\alpha \le r\}$ とおくと，A_r は一様有界かつ同程度連続であるため，Ascoli-Arzelà の定理により $\mathcal{C}_0$ において相対コンパクトである．（これは補題 9.1.1 の証明と事実上同じ議論である．）また $w \mapsto \|w\|_\alpha$ の下半連続性により A_r は閉であるため，結局 A_r はコンパクトである．

正定数 c を Fernique の定理（命題 9.5.1）に現れるものとすると，

$$\begin{aligned}\mathbb{P}^{\varepsilon}\left((A_r)^c\right) &= \mathbb{P}^{\varepsilon}\left(\|w\|_{\alpha} > r\right) \\ &= \mathbb{P}\left(\|w\|_{\alpha} > r/\sqrt{\varepsilon}\right) \\ &= \mathbb{P}\left(\exp(c\|w\|_{\alpha}^2) > \exp(cr^2/\varepsilon)\right) \\ &\leq \exp(-cr^2/\varepsilon)\mathbb{E}\left[\exp(c\|w\|_{\alpha}^2)\right]\end{aligned}$$

となる．最後の式変形で Chebyshev の不等式を使った．Fernique の定理により最右辺の期待値は有限なので，

$$\overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mathbb{P}^{\varepsilon}\left((A_r)^c\right) \leq -cr^2$$

を得る．r は任意なので，r を十分大きく取ると $\{\mathbb{P}^{\varepsilon}\}$ が指数的に緊密であることがわかる．　□

それでは Schilder の大偏差原理（定理 9.0.1）の上からの評価に２つ目の証明を与える．

命題 9.4.3 の証明（その２）．まず次の事実を示す．

$$\overline{\lim_{r\searrow 0}}\,\overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mathbb{P}^{\varepsilon}\left(B_{\mathcal{C}_0}(w,r)\right) \leq -I(w), \qquad w\in\mathcal{C}_0. \tag{9.5.2}$$

$w\in\mathcal{C}_0$ と $g\in\mathcal{P}C_0^2$ を任意とする．(9.1.1) の直後の注意により，$\langle g,\cdot\rangle\in\mathcal{C}_0^*$ なので

$$\begin{aligned}\mathbb{P}^{\varepsilon}\left(B_{\mathcal{C}_0}(w,r)\right) &= \mathbb{P}\left(B_{\mathcal{C}_0}(w/\sqrt{\varepsilon}, r/\sqrt{\varepsilon})\right) \\ &= \mathbb{E}\left[e^{-\langle g,\cdot\rangle/\sqrt{\varepsilon}}e^{\langle g,\cdot\rangle/\sqrt{\varepsilon}}, B_{\mathcal{C}_0}(w/\sqrt{\varepsilon}, r/\sqrt{\varepsilon})\right] \\ &\leq \exp\left(-\varepsilon^{-1}\langle g,w\rangle + \varepsilon^{-1}r\|\langle g,\cdot\rangle\|_{\mathcal{C}_0^*}\right)\mathbb{E}\left[e^{\langle g,\cdot\rangle/\sqrt{\varepsilon}}\right] \\ &= \exp\left(-\varepsilon^{-1}\langle g,w\rangle + \varepsilon^{-1}r\|\langle g,\cdot\rangle\|_{\mathcal{C}_0^*} + \varepsilon^{-1}\|g\|_{\mathcal{H}}^2/2\right)\end{aligned}$$

となる．最後の等号で $\langle g,\cdot\rangle/\sqrt{\varepsilon}$ が平均 0 で分散 $\|g\|_{\mathcal{H}}^2/\varepsilon$ の正規分布に従うことと積分公式

$$\frac{1}{\sqrt{2\pi v}}\int_{\mathbb{R}} e^x e^{-x^2/(2v)}dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{v/2}e^{-(x-\sqrt{v})^2/2}dx = e^{v/2}, \qquad v>0$$

を用いた．

上の球の重みに関する等式において g は任意なので，直ちに次の評価を得る．

$$\overline{\lim_{r\searrow 0}}\,\overline{\lim_{\varepsilon\searrow 0}}\,\varepsilon\log\mathbb{P}^{\varepsilon}\left(B_{\mathcal{C}_0}(w,r)\right) \leq -\sup_{g\in\mathcal{P}C_0^2}\left(\langle g,w\rangle - \frac{1}{2}\|g\|_{\mathcal{H}}^2\right).$$

したがって，次の等式を示せば (9.5.2) が証明できる．

$$\sup_{g\in\mathcal{P}C_0^2}\left(\langle g,w\rangle - \frac{1}{2}\|g\|_{\mathcal{H}}^2\right) = \sup_{g\in\mathcal{P}C_0^2:\|g\|_{\mathcal{H}}=1}\frac{1}{2}\langle g,w\rangle^2 = I(w). \tag{9.5.3}$$

(9.5.3) の右側の等号は補題 9.1.3 で示されたので，これから左側の等号を示す．以下では w を固定して $\Phi(g) = \langle g, w\rangle - \frac{1}{2}\|g\|_{\mathcal{H}}^2$ と定める．さらに $g \neq 0$ も固定して，$t \in \mathbb{R}$ の 2 次関数 φ を $\varphi(t) = \Phi(tg)$ と定めると，

$$\varphi(t) = t\langle g, w\rangle - \frac{t^2}{2}\|g\|_{\mathcal{H}}^2 = -\frac{1}{2}\|g\|_{\mathcal{H}}^2\left(\frac{\langle g, w\rangle}{\|g\|_{\mathcal{H}}^2} - t\right)^2 + \frac{1}{2}\left\langle \frac{g}{\|g\|_{\mathcal{H}}}, w\right\rangle^2$$

と平方完成できるので，$t = \langle g, w\rangle / \|g\|_{\mathcal{H}}^2$ のときに φ は最大値 $\langle g/\|g\|_{\mathcal{H}}, w\rangle^2/2 \geq 0$ を取ることがわかる．よって

$$\begin{aligned}\sup_{g\in\mathcal{P}C_0^2}\left(\langle g, w\rangle - \frac{1}{2}\|g\|_{\mathcal{H}}^2\right) &= 0 \vee \sup_{g\in\mathcal{P}C_0^2, g\neq 0}\left(\langle g, w\rangle - \frac{1}{2}\|g\|_{\mathcal{H}}^2\right)\\ &= 0 \vee \sup_{g\in\mathcal{P}C_0^2, g\neq 0}\frac{1}{2}\left\langle\frac{g}{\|g\|_{\mathcal{H}}}, w\right\rangle^2 = \sup_{g\in\mathcal{P}C_0^2 : \|g\|_{\mathcal{H}}=1}\frac{1}{2}\langle g, w\rangle^2\end{aligned}$$

と求まる．これで (9.5.3) が証明できた．

次は (9.5.2) から弱大偏差原理の上からの評価（すなわち任意のコンパクト集合 $F \subset \mathcal{C}_0$ に対して不等式 (9.4.2)）が従うことを見る．$F \neq \emptyset$ としてよい．$\kappa \in (0,1)$ とする．(9.5.2) により，任意の $w \in F$ に対して，ある $r(w) > 0$ と $\varepsilon(w) \in (0,1]$ が存在して，$0 < \varepsilon < \varepsilon(w)$ のときに

$$\varepsilon \log \mathbb{P}^{\varepsilon}\left(B_{\mathcal{C}_0}(w, r(w))\right) \leq \begin{cases} -\frac{1-\kappa}{2}\|w\|_{\mathcal{H}}^2, & w \in \mathcal{H} \text{ のとき}, \\ -\frac{1}{\kappa}, & w \notin \mathcal{H} \text{ のとき} \end{cases}$$

となる．コンパクト性により有限個の $w_1, \ldots, w_N \in F$ に対し，$F \subset \cup_{i=1}^N B_{\mathcal{C}_0}(w_i, r(w_i))$ と被覆できる．したがって

$$\begin{aligned}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}^{\varepsilon}(F) &\leq \varlimsup_{\varepsilon\searrow 0}\varepsilon\log\left(\sum_{i=1}^N \mathbb{P}^{\varepsilon}\big(B_{\mathcal{C}_0}(w_i, r(w_i))\big)\right)\\ &\leq \max_{1\leq i\leq N}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}^{\varepsilon}\big(B_{\mathcal{C}_0}(w_i, r(w_i))\big)\\ &\leq -\min\left\{\frac{1-\kappa}{2}\inf_{w\in F}\|w\|_{\mathcal{H}}^2,\ \frac{1}{\kappa}\right\}\end{aligned}$$

という評価を得る．ここで補題 1.2.2 も用いた．最後に $\kappa \searrow 0$ として，弱大偏差原理の上からの評価を得る．これを命題 9.5.2 で示した $\{\mathbb{P}^{\varepsilon}\}$ の指数的緊密性と合わせると，命題 1.3.1 により求める大偏差原理の上からの評価が従うので証明が終わる． □

この証明方法の長所として，ごく簡単に Schilder の大偏差原理の位相が Hölder 位相に強められる点がある．そうできる最大の理由は，Fernique の定理（命題 9.5.1）を証明する労力が一様収束位相の場合と Hölder 位相の場合で違わないからである．

定理 9.5.3. $0 < \alpha < 1/2$ とする．Schilder の大偏差原理（定理 9.0.1）は Hölder 空間 $\mathcal{C}_0^{\alpha}$ 上で成り立つ．

証明．$\alpha' \in (\alpha, 1/2)$ を取る．このとき埋め込み写像 $\mathcal{C}_0^{\alpha'} \hookrightarrow \mathcal{C}_0^{\alpha}$ はコンパクトである．これは Ascoli-Arzelà の定理とほぼ自明の不等式 $\|w\|_\alpha \le \|w\|_{\alpha'}^{\alpha/\alpha'}(2\|w\|_\infty)^{1-(\alpha/\alpha')}$ により簡単に確かめられる．α'-Hölder 位相に関する閉球の重さを評価することにより，命題 9.5.2 と同様の議論で $\{\mathbb{P}^\varepsilon\}_{0<\varepsilon\le 1}$ は $\mathcal{C}_0^\alpha$ 上で指数的に緊密であることが示せる．これで埋め込み $\mathcal{C}_0^\alpha \hookrightarrow \mathcal{C}_0$ に対して逆縮小原理（命題 1.5.2）を使えるので，証明が終わる． □

9.6 Mogulskii の大偏差原理（または射影極限法）を用いる証明

本節では第 8 章で証明した Mogulskii の大偏差原理を用いて，Schilder の大偏差原理（定理 9.0.1）に３つ目の証明を与える．Mogulskii の大偏差原理は射影極限法を用いて示すので，射影極限法を使う証明法だとも言える．なお本節では基本的に 9.1 節–9.5 節で示した諸事実をほとんど使っておらず，唯一の例外は補題 9.4.2 で示した１次元 Brown 運動が原点の近傍から脱出する確率に対する評価のみである．

定理 9.0.1 の証明（その３）．記述の便宜上，$(w_t)_{t\ge 0}$ を時間区間を $[0,\infty)$ に延長した d 次元 Brown 運動とする．$t \ge 0$ に対して $Y_t^\varepsilon := \sqrt{\varepsilon}w_{\varepsilon\lfloor t/\varepsilon\rfloor}$ （すなわち $k \in \mathbb{N}$, $t \in [(k-1)\varepsilon, k\varepsilon)$ のときに $Y_t^\varepsilon := \sqrt{\varepsilon}w_{(k-1)\varepsilon}$）と定める．Brown 運動の基本性質からすぐに見えるが，$\{(w_{k\varepsilon} - w_{(k-1)\varepsilon})/\sqrt{\varepsilon}\}_{k\in\mathbb{N}}$ は $\mathbb{R}^d$ 値の独立同分布列で，$(w_{k\varepsilon} - w_{(k-1)\varepsilon})/\sqrt{\varepsilon}$ の分布は d 次元の標準正規分布である．この Y^ε は (8.4.1) で定義した確率過程の特殊例であることに注意せよ．実際，$t \in [(k-1)\varepsilon, k\varepsilon)$ のときに $Y_t^\varepsilon = \varepsilon\sum_{i=1}^{k-1}(w_{i\varepsilon} - w_{(i-1)\varepsilon})/\sqrt{\varepsilon}$ となるので，結局 (8.4.1) の定義と一致している．よって Mogulskii の大偏差原理の連続パラメータ版（定理 8.4.1）を $(Y_t^\varepsilon)_{t\in[0,T]}$ に適用できる．

簡単な平方完成により，この場合の Λ と Λ^* はそれぞれ

$$\Lambda(\lambda) = \log\int_{\mathbb{R}^d} e^{\lambda\cdot x}\frac{e^{-|x|^2/2}}{(2\pi)^{d/2}}dx = \frac{1}{2}|\lambda|^2, \qquad \lambda \in \mathbb{R}^d,$$
$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}^d}\left\{\lambda\cdot x - \frac{1}{2}|\lambda|^2\right\} = \frac{1}{2}|x|^2, \qquad x \in \mathbb{R}^d$$

と求まる．Λ^* の形から直ちに Mogulskii の大偏差原理中の速度関数は (9.0.3) で定めたものと一致することがわかる．以上により $\{Y^\varepsilon\}_{\varepsilon\in(0,1]}$ は $L^\infty([0,T],\mathbb{R}^d)$ 上で $\varepsilon \searrow 0$ のときに，良い速度関数 I に対して大偏差原理を満たすことがわかった．

以下 $\delta > 0$ とし，区間 J 上の L^∞ ノルムを $\|\cdot\|_{\infty,J}$ と表す．命題 9.4.3 の証明中の不等式 (9.4.5) とほぼ同じ議論により

$$\mathbb{P}\left(\|\sqrt{\varepsilon}w - Y^\varepsilon\|_{\infty,[0,T]} > \delta\right) \le \mathbb{P}\left(\bigcup_{1\le k\le\lfloor T/\varepsilon\rfloor+1}\{\|\sqrt{\varepsilon}w - Y^\varepsilon\|_{\infty,[(k-1)\varepsilon,k\varepsilon)} > \delta\}\right)$$

$$
\begin{aligned}
&\le \sum_{k=1}^{\lfloor T/\varepsilon \rfloor+1} \mathbb{P}\left(\|\sqrt{\varepsilon}w - Y^\varepsilon\|_{\infty,[(k-1)\varepsilon,k\varepsilon)} > \delta\right) \\
&= \sum_{k=1}^{\lfloor T/\varepsilon \rfloor+1} \mathbb{P}\left(\sup_{0\le s\le\varepsilon} |\sqrt{\varepsilon}(w_{s+(k-1)\varepsilon} - w_{(k-1)\varepsilon})| > \delta\right) \\
&= (\lfloor T/\varepsilon \rfloor + 1)\mathbb{P}\left(\sup_{0\le s\le\varepsilon} |w_s| > \delta/\sqrt{\varepsilon}\right) \\
&\le 2d(\lfloor T/\varepsilon \rfloor + 1)e^{-\delta^2/(2d\varepsilon^2)}
\end{aligned}
$$

と評価できる．ここで $u \ge 0$ のとき $(w_{t+u} - w_u)_{t\ge 0}$ が再び Brown 運動になることと補題 9.4.2 を用いた．これから直ちに

$$
\overline{\lim_{\varepsilon\searrow 0}}\, \varepsilon \log \mathbb{P}\left(\|\sqrt{\varepsilon}w - Y^\varepsilon\|_{\infty,[0,T]} > \delta\right) = -\infty, \qquad \delta > 0
$$

が従うので，$\sqrt{\varepsilon}w$ の法則と Y^ε の法則は指数的に同値である．したがって命題 1.6.2 により $\sqrt{\varepsilon}w$ の法則も $L^\infty([0,T],\mathbb{R}^d)$ 上で $\varepsilon \searrow 0$ のときに，同じ良い速度関数 I に対して大偏差原理を満たす．

連続な経路空間 $\mathcal{C}_0$ は $L^\infty([0,T],\mathbb{R}^d)$ の閉部分集合である．ノルムの定義から明らかだが，前者の位相は後者の位相の相対位相であることに注意せよ．また明らかに $\sqrt{\varepsilon}w$ の法則は $\mathcal{C}_0$ 上の測度であり，I の実効定義域は $\mathcal{C}_0$ に含まれる．したがって命題 1.2.12(2) によりこの大偏差原理は実は $\mathcal{C}_0$ 上で成立している．以上で Schilder の大偏差原理（定理 9.0.1）の上からの評価に対する 3 つ目の証明が終わった． □

9.7 抽象 Wiener 空間上の Schilder 型大偏差原理に関する補足

端的に言うと，可分 Banach 空間に平均 0 の Gauss 測度を入れたものを抽象 Wiener 空間と呼ぶ.[12] Schilder の大偏差原理は抽象 Wiener 空間上の大偏差原理として見事に一般化されており，現代ではこの一般化された形を学ぶのが普通である．本節ではこれを紹介する.（抽象 Wiener 空間の一般論を説明するのにかなりの時間がかかるため，本節では事実を紹介するだけにとどめ，基本的に証明は与えない.）

まず抽象 Wiener 空間を導入する．$\mathcal{X}$ を可分 Banach 空間，$\mathcal{H}$ を $\mathcal{X}$ に連続かつ稠密に埋め込まれた可分 Hilbert 空間，μ を $\mathcal{X}$ 上の Borel 確率測度で

$$
\int_{\mathcal{X}} \exp(\sqrt{-1}\langle\varphi, w\rangle)\mu(dw) = \exp\left(-\frac{1}{2}\|\varphi\|_{\mathcal{H}^*}^2\right), \qquad \varphi \in \mathcal{X}^* \tag{9.7.1}
$$

[12]抽象 Wiener 空間については，例えば [12, 28] を参照せよ．

を満たすときに，$(\mathcal{X}, \mathcal{H}, \mu)$ を抽象 Wiener 空間と呼ぶ．ここで連続埋め込み $\mathcal{H} \hookrightarrow \mathcal{X}$ の双対を取ることで，$\mathcal{X}^* \hookrightarrow \mathcal{H}^*$ と連続かつ稠密に埋め込まれている．ちなみに $\mathcal{H}$ を Cameron-Martin 空間と呼ぶ．

実は以下のように直感的にわかりやすい方法で μ を構成できる．説明の便宜上，$\dim \mathcal{H} = \infty$ とする．($\dim \mathcal{H} < \infty$ の場合は簡単なので省略する．) $\{h_i\}_{i\in\mathbb{N}}$ を $\mathcal{H}$ の任意の完全正規直交系とし，各 Z_i が標準正規分布に従う独立同分布な $\mathbb{R}$ 値確率変数列 $\{Z_i\}_{i\in\mathbb{N}}$ とする．このとき測度の等式

$$\mu \overset{\mathrm{Law}}{=} \sum_{i=1}^{\infty} Z_i h_i \tag{9.7.2}$$

が成り立つことが知られている．右辺は $\mathcal{X}$ 値確率変数列としての法則収束を意味する．

(9.7.1) において φ を $t\varphi$ ($t \in \mathbb{R}$) で置き換えると，実数値確率変数 φ の法則の特性関数が $\exp(-t^2\|\varphi\|_{\mathcal{H}^*}^2/2)$ だとわかる．よって φ は平均 0 で分散 $\|\varphi\|_{\mathcal{H}^*}^2$ の Gauss 分布に従う．分極化を用いて，φ と ψ の共分散が $\langle \varphi, \psi \rangle_{\mathcal{H}^*}$ であることも簡単に示せる．すなわち $\mathcal{X}^*$ は μ の下で平均 0 で共分散が $\mathcal{H}^*$ の内積であるような Gauss 系をなす．

$\varepsilon > 0$ のとき，μ の下で定数倍写像 $w \in \mathcal{X} \mapsto \sqrt{\varepsilon} w \in \mathcal{X}$ の法則を μ^ε と書く．すなわち任意の Borel 集合 $A \subset \mathcal{X}$ に対して，$\mu^\varepsilon(A) = \mu\left(\sqrt{\varepsilon} w \in A\right)$ として μ^ε を定める．μ^ε は元の Gauss 測度 μ の分散を ε 倍した Gauss 測度である．また $J\colon \mathcal{X} \to [0, \infty]$ を以下のように定めると，実は J は良い速度関数である．

$$J(w) = \begin{cases} \frac{1}{2}\|w\|_{\mathcal{H}}^2, & w \in \mathcal{H}, \\ \infty, & w \in \mathcal{X} \setminus \mathcal{H}. \end{cases}$$

抽象 Wiener 空間上の Schilder 型大偏差原理を紹介する．本書では証明を与えないが，第 9.5 節で紹介した Fernique の定理を用いる 2 つ目の証明法とほぼ同じ論法で証明できる.[13] それ以外に抽象 Wiener 空間上の等周不等式を用いる証明法も知られている.[14]

定理 9.7.1. $\varepsilon \searrow 0$ とするとき，$\mathcal{X}$ 上の Borel 確率測度の族 $\{\mu^\varepsilon\}_{\varepsilon\in(0,1]}$ は良い速度関数 J に対して速度 $1/\varepsilon$ で大偏差原理を満たす.

本章で議論した古典的 Wiener 空間 $(\mathcal{C}_0, \mathcal{H}, \mathbb{P})$ がもちろん抽象 Wiener 空間の最重要例である．また自明な例として，Euclid 空間 $\mathbb{R}^d$ に標準正規分布を入れたものがある．それ以外の典型的な例をいくつか紹介しよう．

次の例を見ると，定理 9.5.3 は定理 9.7.1 の特殊例であることがわかる．

例 9.7.2. 第 9.5 節の冒頭にあるように，$0 < \alpha < 1/2$ とし，$\mathcal{C}_0^\alpha$ を α-Hölder 空間とする．また $\mathcal{H}$ を標準 Brown 運動に関する Cameron-Martin 空間とする．d 次元 Wiener 測度 $\mathbb{P}$

---

[13] 例えば [45, 第 4.1.5 節] または [44, 第 8.4 節] を参照せよ．
[14] この方法については [28, 第 4.9 節] または [24, 第 8.3 節] などを参照せよ．

は $\mathcal{C}_0^\alpha$ 上の測度で，さらに $\mathcal{H} \subset \mathcal{C}_0^\alpha$ であるが，実は $\mathcal{C}_0^\alpha$ は可分でない．そこで次で定義される部分空間を考える.[15]

$$\widetilde{\mathcal{C}}_0^\alpha = \left\{ w \in \mathcal{C}_0^\alpha \,\middle|\, \lim_{\kappa \searrow 0} \sup_{0<t-s\le\kappa} \frac{|w_t - w_s|}{(t-s)^\alpha} = 0 \right\}.$$

ここで上限は $0 < t-s \le \kappa$ を満たす $s, t \in [0,T]$ を渡って取る．この $\widetilde{\mathcal{C}}_0^\alpha$ は $\mathcal{C}_0^\alpha$ の可分な部分 Banach 空間であることが知られている．さらに $0 < \alpha < \beta < 1/2$ のとき $\mathcal{C}_0^\beta \subset \widetilde{\mathcal{C}}_0^\alpha$ と包含されるため，$\mathcal{H} \subset \widetilde{\mathcal{C}}_0^\alpha$ かつ $\mathbb{P}(\widetilde{\mathcal{C}}_0^\alpha) = 1$ である．これから予想できるように $(\widetilde{\mathcal{C}}_0^\alpha, \mathcal{H}, \mathbb{P})$ は抽象 Wiener 空間になる．

時間区間を無限区間 $[0,\infty)$ に取り替えた場合でも，d 次元標準 Brown 運動 $(w_t)_{t\ge0}$ に対する Schilder の大偏差原理が成り立つことが次の例からわかる．この場合も $(w_t)_{t\ge0}$ の法則（すなわち Wiener 測度）を引き続き $\mathbb{P}$ と表すことにする．

例 9.7.3. $d \in \mathbb{N}$ とし，$\mathcal{C}_0([0,\infty),\mathbb{R}^d) := \{w\colon [0,\infty) \to \mathbb{R}^d \mid w \text{ は連続かつ } w_0 = 0\}$ に広義一様収束の位相を入れる．この空間は Fréchet 空間ではあるが Banach 空間ではない．そこで

$$\mathcal{X} := \{w \in \mathcal{C}_0([0,\infty),\mathbb{R}^d) \mid \lim_{t\to\infty} |w_t|/(1+t) = 0\}, \qquad \|w\|_{\mathcal{X}} := \sup_{t\ge0} |w_t|/(1+t)$$

と定めると，$\mathcal{X}$ は可分な Banach 空間になる．この例における Cameron-Martin 空間は

$$\mathcal{H} := \left\{ h \in \mathcal{C}_0([0,\infty),\mathbb{R}^d) \mid h \text{ は絶対連続かつ } \|h\|_{\mathcal{H}}^2 := \textstyle\int_0^\infty |h_t'|^2 dt < \infty \right\}$$

である．実は $\mathbb{P}$ は $\mathcal{X}$ 上の確率測度であり，$(\mathcal{X}, \mathcal{H}, \mathbb{P})$ は抽象 Wiener 空間であることが知られている.[16]（なお $\mathcal{X}$ は $\mathcal{C}_0([0,\infty),\mathbb{R}^d)$ に連続に埋め込まれているので，広義一様位相に関しても Schilder の大偏差原理は成り立つ.）

Brown 運動でない Gauss 過程の典型例として，非整数 Brown 運動が挙げられる．この過程は Hurst 指数と呼ばれるパラメーター $\gamma \in (0,1)$ で添字付けられている．定義により $\gamma = 1/2$ の場合はこの過程は Brown 運動と一致する．一方 $\gamma \neq 1/2$ の場合はこの過程は半マルチンゲールでも Markov 過程でもないことが知られている．次の例により非整数 Brown 運動も Schilder 型大偏差原理を満たすことがわかる．

例 9.7.4. $d \in \mathbb{N}, T > 0, \gamma \in (0,1)$ とする．$\mathbb{R}^d$ 値確率過程 $(w_t^\gamma)_{t\in[0,T]} = (w_t^{\gamma,1}, \dots, w_t^{\gamma,d})_{t\in[0,T]}$ が平均 0 の連続 Gauss 過程であり，さらに共分散が

$$\mathbb{E}\left[w_s^{\gamma,i} w_t^{\gamma,j}\right] = \frac{\delta_{ij}}{2}(s^{2\gamma} + t^{2\gamma} - |t-s|^{2\gamma}), \qquad s,t \in [0,T],\ i,j \in \{1,\dots,d\}$$

[15]これを「小 Hölder 空間」という名前で呼ぶ文献もある．

[16]例えば [44, 第 8.1 節] を参照せよ．

で与えられるとき，(w_t^γ) を Hurst 指数 γ を持つ非整数 Brown 運動であるという．ここで δ_{ij} は Kronecker のデルタ記号である．この確率過程の法則を μ^γ と表す．対応する Cameron-Martin 空間 $\mathcal{H}^\gamma$ は Volterra 型積分核を用いて具体的に表示できることが知られているが，煩雑なのでここでは省略する.[17] $(\mathcal{C}_0, \mathcal{H}^\gamma, \mu^\gamma)$ は抽象 Wiener 空間であることが知られている．さらに $0 < \alpha < \gamma$ ならば，実は $(\widetilde{\mathcal{C}}_0^\alpha, \mathcal{H}^\gamma, \mu^\gamma)$ も抽象 Wiener 空間であることが知られている.（ここで $\widetilde{\mathcal{C}}_0^\alpha$ は例 9.7.2 で定義された空間である.）

最後の例として無限次元 Brown 運動を挙げる．抽象 Wiener 空間の理論ではよく知られているのだが，抽象 Wiener 空間 $(\mathcal{X}, \mathcal{H}, \mu)$ に値を取る Brown 運動で時刻 $t = 1$ における分布が μ であるものが存在する．無限次元 Brown 運動（あるいは可算無限個の 1 次元標準 Brown 運動）を扱う枠組みとしては，これが便利である．

例 9.7.5. $(\mathcal{X}, \mathcal{H}, \mu)$ を無限次元の抽象 Wiener 空間とし，$\{h_i\}_{i\in\mathbb{N}}$ を $\mathcal{H}$ の任意の完全正規直交系とする．また $\mathcal{Y} := \{w \in \mathcal{C}_0([0,\infty), \mathcal{X}) \mid \lim_{t\to\infty} \|w_t\|_{\mathcal{X}}/(1+t) = 0\}$ および $\|w\|_{\mathcal{Y}} := \sup_{t\geq 0} \|w_t\|_{\mathcal{X}}/(1+t)$ と定めると，$\mathcal{Y}$ は可分な Banach 空間になる．

$\{(w_t^i)_{t\geq 0} \mid i \in \mathbb{N}\}$ を適当な確率空間の上で定義された独立同分布な 1 次元の標準 Brown 運動の列だとする．このとき

$$\sum_{i=1}^{\infty} w_\cdot^i h_i := \lim_{N\to\infty} \sum_{i=1}^{N} w_\cdot^i h_i \tag{9.7.3}$$

は $\mathcal{Y}$ 値確率変数列として，完全正規直交系 $\{h_i\}$ に依存しないある確率測度 $\mathbb{P}_{\mathcal{X}}$ に法則収束することが知られている.[18] 以下 $(W_t)_{t\geq 0}$ を $\mathbb{P}_{\mathcal{X}}$ を法則とする確率過程だとする．(9.7.2) と (9.7.3) を比較すればすぐわかるが，$s \geq 0$ のとき W_s が $\mathcal{X}$ 上に誘導する法則は μ^s（μ の分散を s 倍した Gauss 測度）である．また 1 次元 Brown 運動の基本性質により，$0 \leq s \leq t$ のとき，$W_t - W_s$ は $\sigma\{W_u \mid 0 \leq u \leq s\}$ に独立で，その法則は W_{t-s} の法則と等しい．これらの性質により，$\mathbb{P}_{\mathcal{X}}$ を法則とする確率過程を $\mathcal{X}$ 値 Brown 運動 $(W_t)_{t\geq 0}$ と呼んでおかしくないことがわかる．

この場合の Cameron-Martin 空間 $\mathcal{K}$ は以下で与えられる．

$$\mathcal{K} = \left\{k \colon [0,\infty) \to \mathcal{H} \mid k_0 = 0,\ k \text{ は絶対連続かつ } \|k\|_{\mathcal{K}}^2 := \textstyle\int_0^\infty \|k_t'\|_{\mathcal{H}}^2 dt < \infty\right\}.$$

上記の 3 つ組 $(\mathcal{Y}, \mathcal{K}, \mathbb{P}_{\mathcal{X}})$ は抽象 Wiener 空間であることが知られている.[19]

[17] 例えば [30, 第 4.2 節] を参照せよ．

[18] これは伊藤・西尾の定理 [52] の特殊例だと思える．ちなみにその定理によれば，実はこの無限和は概収束している．

[19] 例えば [44, 第 8.6 節] を見よ．

第10章　Freidlin-Wentzellの大偏差原理

$T>0$ と $d\in\mathbb{N}$ を任意とし，第 9 章と同じく時間区間 $[0,T]$ 上の $\mathbb{R}^d$ 値標準 Brown 運動に関する Wiener 空間と Cameron-Martin 空間を $(\mathcal{C}_0,\mathcal{B}(\mathcal{C}_0),\mathbb{P})$ および $\mathcal{H}$ と書く．すなわち $\mathcal{C}_0$ と $\mathcal{H}$ はそれぞれ (9.0.1) と (9.0.2) で定義される経路空間で，$\mathbb{P}$ は d 次元の Wiener 測度，$\mathcal{B}(\mathcal{C}_0)$ は $\mathcal{C}_0$ の Borel 加法族である．$\mathbb{P}$ の下では座標過程 $(w_t)_{0\le t\le T}$ は d 次元標準 Brown 運動である．$\mathcal{N}$ を $\mathbb{P}$ 零集合の全体とし，$\mathcal{F}=\mathcal{B}(\mathcal{C}_0)\vee\mathcal{N}$ を Borel 加法族の $\mathbb{P}$ 完備化とする．また $0\le t\le T$ のとき，$\mathcal{F}_t=\sigma\{w_s\mid 0\le s\le t\}\vee\mathcal{N}$ とおく．これは時刻 t までの Brown 運動の全情報を $\mathbb{P}$ に関して完備化したものだと思える．よく知られているように情報系 $\{\mathcal{F}_t\}_{t\in[0,T]}$ は右連続なので,[1] ４つ組 $(\mathcal{C}_0,\mathcal{F},\mathbb{P};\{\mathcal{F}_t\}_{t\in[0,T]})$ はいわゆる「通常の条件」を満たす．本章で確率微分方程式（SDE）の計算をする際には，基本的にこの４つ組の上で作業する.[2]

さて伊藤型の SDE を導入しよう．$e\in\mathbb{N}$ を任意とし,[3] $\sigma\colon\mathbb{R}^e\to\mathbb{R}^{e\times d}$ と $b\colon\mathbb{R}^e\to\mathbb{R}^e$ を十分良い関数だとする．$\mathbb{R}^{e\times d}$ は $e\times d$ 型実行列全体のなす集合である．本章では $\varepsilon\in(0,1]$ に対して，ε を添字に持つ次の SDE を考える.[4]

$$X_t^\varepsilon=a+\int_0^t b(X_s^\varepsilon)ds+\sqrt{\varepsilon}\int_0^t\sigma(X_s^\varepsilon)dw_s,\qquad t\in[0,T]. \tag{10.0.1}$$

右辺の第 3 項は伊藤積分である．初期値 $a\in\mathbb{R}^e$ は任意だが非ランダムとし，以下では固定して議論する．X^ε は e 次元の経路空間 $\widetilde{\mathcal{C}}:=\{x\colon[0,T]\to\mathbb{R}^e\mid x\text{ は連続}\}$ に値を取る確率変数と見なすのが自然である．$\widetilde{\mathcal{C}}$ は通常の一様ノルム $\|x\|_\infty:=\sup_{t\in[0,T]}|x_t|$ の下で可分 Banach 空間になる．$\widetilde{\mathcal{C}}$ にはこの Banach 位相から決まる Borel 集合体 $\mathcal{B}(\widetilde{\mathcal{C}})$ を入れる．

上記の SDE(10.0.1) において，$\varepsilon\searrow 0$ のときに，X^ε に関する量がどのような極限挙動を示すかを調べる問題を小雑音問題と称する.（そう称する場合には $\sqrt{\varepsilon}$ は σ の係数では

---

[1]例えば [16, 補題 1.4.10]，[13, 定理 3.17]，[11, 命題 3.7.5] などを見よ．

[2]本章でも Euclid 空間の標準内積を $\langle v,v'\rangle_{\mathbb{R}^d}$ ではなく $v\cdot v'$ と書く（$v,v'\in\mathbb{R}^d$）．ノルムは $|v|=\sqrt{v\cdot v}$ と書く．一方，２本線のノルム $\|\cdot\|$ は無限次元空間のノルムに対してのみ使う．なお “SDE” は stochastic differential equation の略語である．同様に “ODE” は ordinary differential equation の略語である．

[3]本章では e は自然対数の底ではなく次元を表す．指数関数は exp と表す．

[4]SDE に関する予備知識のうち本章で仮定しているのは，標準的な SDE の教科書において巻頭から「大域的 Lipschitz 条件下での（強い）解の存在と一意性」までの部分に書いてある事項である．そのような教科書の例として [8, 11, 13, 16, 18, 23] などを挙げておく．

なく，むしろ雑音 dw_t の係数だと見なしている.）このとき，少なくとも形式的には X^ε は次の常微分方程式（ODE）の解 X^0 に収束する.

$$X^0_t = a + \int_0^t b(X^0_s)ds, \qquad t \in [0,T]. \tag{10.0.2}$$

この X^0 は非ランダムなので，その法則は Dirac 測度 δ_{X^0} である．したがって，$\widetilde{\mathcal{C}}$ に値を取る確率変数の族 $\{X^\varepsilon\}_{\varepsilon\in(0,1]}$ が $\varepsilon \searrow 0$ のときに大偏差原理を満たすかどうかを問うのは自然であろう．これは Freidlin-Wentzell の大偏差原理と呼ばれ，小雑音問題の代表例である.

次は SDE(10.0.1) に付随する骨格 ODE を導入する．$h \in \mathcal{H}$ を任意とし，(10.0.1) における $\sqrt{\varepsilon}dw_s$ を $dh_s = h'_s ds$ で形式的に置き換えると，h により制御される次の骨格 ODE を得る.

$$\varphi^h_t = a + \int_0^t b(\varphi^h_s)ds + \int_0^t \sigma(\varphi^h_s)h'_s ds, \qquad t \in [0,T]. \tag{10.0.3}$$

係数 σ と b が十分良い時には，任意の h に対して（時間大域的な）一意解 φ^h があるので，$h \mapsto \varphi^h$ は $\mathcal{H}$ から $\widetilde{\mathcal{C}}$ への写像である.

ここで係数行列に関する条件を導入しよう．本章では常に以下の大域的 Lipschitz 条件を仮定する．なお $e \times d$ 型行列 $A = (a_{ij})_{1\le i\le e, 1\le j\le d} \in \mathbb{R}^{e\times d}$ に対しては，そのノルムを $|A| := \{\sum_{i=1}^e \sum_{j=1}^d |a_{ij}|^2\}^{1/2}$ と定める.[5]

(LIP) 定数 $L > 0$ が存在して，以下の不等式を満たす.

$$|\sigma(x_1) - \sigma(x_2)| + |b(x_1) - b(x_2)| \le L|x_1 - x_2|, \qquad x_1, x_2 \in \mathbb{R}^e.$$

条件 **(LIP)** の下で，SDE (10.0.1) は任意の ε と a に対して，一意的な強い解 X^ε を持つことを思い出そう．その解は爆発せず，Picard 流の逐次近似法を用いて得られる.[6] 特に X^ε は $\widetilde{\mathcal{C}}$ 値の確率変数である.[7]また ODE (10.0.2) も任意の a に対して爆発しない一意解を持つ．同様に骨格 ODE (10.0.3) も任意の h と a に対して爆発しない一意解を持つ.

骨格 ODE の情報を用いて，以下のように関数 $J\colon \widetilde{\mathcal{C}} \to [0,\infty]$ を定義する．まず Schilder の大偏差原理に現れる良い速度関数 $I\colon \mathcal{C}_0 \to [0,\infty]$ を思い出そう.（これは (9.0.3) において導入された．$h \in \mathcal{H}$ のときは $I(h) = \|h\|^2_{\mathcal{H}}/2$ であり，$h \in \mathcal{C}_0 \setminus \mathcal{H}$ のときは $I(h) = \infty$ である.）この I を用いて

$$J(\xi) := \inf\{I(h) \mid h \in \mathcal{H},\ \ \xi = \varphi^h\}, \qquad \xi \in \widetilde{\mathcal{C}}. \tag{10.0.4}$$

---

[5]すなわち Frobenius ノルムである．Hilbert-Schmidt ノルムとも呼ばれる.

[6]例えば [16, 第 3.2 節，定理 3.2.1] や [13, 第 5.1 節，定理 5.3] や [18, 第 5.1 節，定理 5.2] を参照せよ.

[7]特にこの SDE は法則の意味での一意性も持つので，４つ組とブラウン運動 (w_t) の実現の仕方を取り替えても，解 X^ε の法則は変わらない.

と定める．なお通常どおり $\inf \emptyset = +\infty$ と約束する．実は J は良い速度関数であるが，そのことは後述する命題 10.1.3 で示す.

本章の主題である Freidlin-Wentzell の大偏差原理を正確に述べる．これは SDE の解 $\{X^\varepsilon\}_{\varepsilon\in(0,1]}$ の法則に対する大偏差原理である.

定理 10.0.1. 条件 **(LIP)** を仮定する．$\varepsilon \searrow 0$ とするとき，$\{X^\varepsilon\}_{\varepsilon\in(0,1]}$ は $\widetilde{\mathcal{C}}$ 上で良い速度関数 J に対して速度 $1/\varepsilon$ で大偏差原理を満たす.

注意 10.0.2. 本注意では上記の大偏差原理（定理 10.0.1）を形式的かつ直感的に観察する．X^1 を $\mathcal{C}_0$ から $\widetilde{\mathcal{C}}$ への写像だと見なすと，少なくとも形式的には $X^\varepsilon(w) = X^1(\sqrt{\varepsilon}w)$ と思える．仮に X^1 が連続写像であれば，縮小原理（命題 1.5.1）と Schilder の大偏差原理（定理 9.0.1）から直ちに定理 10.0.1 が従う．特に速度関数 J の形はこのように形式的に縮小原理を用いて得られるものと一致している．しかし（よほど例外的な場合を除いては）X^1 は連続ではないため，この種の論法は使えず，定理 10.0.1 の証明は簡単ではない.

本章の残りでは定理 10.0.1 に 3 種類の証明を与える．なお「$a = 0, b \equiv 0$ かつ $\sigma \equiv \mathrm{Id}_e$」という非常に特殊な場合には，この定理は前章で議論した Schilder の大偏差原理になることに注意せよ.（なお Id_e は e 次単位行列である.）

注意 10.0.3. 著者の知る限り，現時点では Freidlin-Wentzell の大偏差原理に対して 5 種類の有力な証明法が知られている．本書で触れない 2 種類は，粘性解と非線形半群を用いる偏微分方程式論的な方法とラフパス理論を用いる方法である.[8]

10.1 骨格 ODE に関する基本事項

骨格 ODE について丁寧に書いてある日本語文献は少ないので，この節で基本事項をまとめておく．まずは Lipschitz 条件の下での解の存在と一意性を確認しよう.

命題 10.1.1. 条件 **(LIP)** を仮定する．このとき任意の $h \in \mathcal{H}$ に対して，骨格 ODE (10.0.3) は時間大域的な一意解 $\varphi^h \in \widetilde{\mathcal{C}}$ を持つ．さらに任意の $r \in (0,\infty)$ に対して，次の評価が成り立つ.

$$\sup\{\|\varphi^h\|_\infty \mid h \in \mathcal{H},\, \|h\|_{\mathcal{H}} \le r\} < \infty.$$

証明. 本証明では，$K = \max\{L, |\sigma(0)|, |b(0)|\}$ とおく．すると $\max\{|\sigma(x)|, |b(x)|\} \le K(|x|+1)$ が任意の $x \in \mathbb{R}^e$ に対して成り立つ.

---

[8] 前者については [36] を見よ．後者については [5] または [37] を見よ.

まず逐次近似法を用いて，任意の h と a に対して時間大域解を構成しよう．写像 $G\colon \widetilde{\mathcal{C}} \to \widetilde{\mathcal{C}}$ を

$$G(z)_t = a + \int_0^t b(z_s)ds + \int_0^t \sigma(z_s)h'_s ds, \qquad t \in [0,T]$$

と定める．Schwarz の不等式により

$$\begin{aligned} |G(z)_t - G(\widehat{z})_t|^2 &\le 2\Big|\int_0^t \{b(z_s) - b(\widehat{z}_s)\}ds\Big|^2 + 2\Big|\int_0^t \{\sigma(z_s) - \sigma(\widehat{z}_s)\}h'_s ds\Big|^2 \\ &\le 2(T + \|h\|_{\mathcal{H}}^2)K^2 \int_0^t |z_s - \widehat{z}_s|^2 ds \end{aligned} \tag{10.1.1}$$

と評価できる．これから特に G の Lipschitz 連続性もわかる．

定石どおりに，$\varphi(0) \equiv a$ および

$$\varphi(n+1) = G(\varphi(n)), \qquad n \in \mathbb{N} \tag{10.1.2}$$

と定める．また簡単のため $C = 2(T + \|h\|_{\mathcal{H}}^2)K^2$ と書く．すると再帰的に次の不等式を得る．

$$\begin{aligned} |\varphi(n+1)_t - \varphi(n)_t|^2 &\le C\int_0^t |\varphi(n)_{t_1} - \varphi(n-1)_{t_1}|^2 dt_1 \\ &\le C^2 \int_0^t \int_0^{t_1} |\varphi(n-1)_{t_2} - \varphi(n-2)_{t_2}|^2 dt_2 dt_1 \\ &\le C^n \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} |\varphi(1)_{t_n} - \varphi(0)_{t_n}|^2 dt_n \cdots dt_2 dt_1 \\ &\le \frac{C^n T^n}{n!} \|\varphi(1) - \varphi(0)\|_\infty^2. \end{aligned}$$

したがって $\sum_{n=1}^\infty \|\varphi(n+1) - \varphi(n)\|_\infty < \infty$ となるので，$\{\varphi(n)\}_{n=1}^\infty$ は $\widetilde{\mathcal{C}}$ における Cauchy 列である．この極限を $\varphi(\infty)$ と書く．(10.1.2) において $n \to \infty$ とし，G の連続性を使うと $\varphi(\infty) = G(\varphi(\infty))$ を得るが，これは $\varphi(\infty)$ が時間大域解であることを意味する．

解の一意性は時間局所的な問題であるため，$t = 0$ の右近傍で確認すれば十分である．$0 < S \le T$ として，φ と $\widehat{\varphi}$ はともに $[0,S]$ 区間上の (10.0.3) の解だとする．$z = \varphi$ と $\widehat{z} = \widehat{\varphi}$ に対して (10.1.1) を用いると，$G(\varphi) = \varphi$ などが成立するので，

$$|\varphi_t - \widehat{\varphi}_t|^2 \le C\int_0^t |\varphi_s - \widehat{\varphi}_s|^2 ds, \qquad t \in [0,S]$$

を得る．Gronwall の不等式を用いると $[0,S]$ 上で φ と $\widehat{\varphi}$ が一致することがわかるので，これで一意性の証明が終わる．

最後に一意解 φ^h の評価を示す．骨格 ODE(10.0.3) に対して，(10.1.1) の計算と同じ要領で Schwarz の不等式を用いると，

$$1 + |\varphi_t^h|^2 \le 1 + 3|a|^2 + 6(T + \|h\|_{\mathcal{H}}^2)K^2 \int_0^t (1 + |\varphi_s^h|^2)ds, \qquad t \in [0,T]$$

を得る．再び Gronwall の不等式を用いて

$$1 + |\varphi_t^h|^2 \le (1 + 3|a|^2) \exp\bigl(6(T + \|h\|_{\mathcal{H}}^2) K^2 t\bigr), \qquad t \in [0, T]$$

と評価できるので，$t \in [0, T]$ に関する上限を取ると本命題の証明が終わる．　□

命題 10.1.2. 条件 **(LIP)** を仮定する．$\{h(n)\}_{n\in\mathbb{N}}$ は $\mathcal{H}$ の点列で，$n \to \infty$ のときに $h(\infty) \in \mathcal{H}$ に弱収束すると仮定する．このとき $\lim_{n\to\infty} \|\varphi^{h(n)} - \varphi^{h(\infty)}\|_\infty = 0$ となる．ここで，$\varphi^{h(n)}$ と $\varphi^{h(\infty)}$ はそれぞれ $h(n)$ と $h(\infty)$ で駆動される骨格 ODE (10.0.3) の解である．

証明. 簡単のため $\varphi^{h(n)} = \varphi(n)$ などと略記する．よく知られているように，一様有界性の原理により弱収束列は有界列なので，$M := \sup\{\|h(n)\|_{\mathcal{H}} \mid n \in \mathbb{N} \cup \{\infty\}\} < \infty$ である．すると命題 10.1.1 により，十分大きな定数 $R > 0$ を取れば

$$\{\varphi(n)_t \mid n \in \mathbb{N} \cup \{\infty\}, t \in [0, T]\} \subset B_{\mathbb{R}^e}(0, R)$$

となる．以下では $c_R := \sup\{|\sigma(x)| \vee |b(x)| \mid x \in B_{\mathbb{R}^e}(0, R)\}$ と書く．特に $\{\varphi(n)\}_{n\in\mathbb{N}}$ は一様有界である．

$0 \le s \le t \le T$ のとき，

$$\begin{aligned} |\varphi(n)_t - \varphi(n)_s| &\le \int_s^t |b(\varphi(n)_u)| du + \int_s^t |\sigma(\varphi(n)_u)|\,|h(n)'_u| du \\ &\le c_R(\sqrt{T} + M)\sqrt{t - s} \end{aligned}$$

と評価できる．この右辺は n に依存しないので，この評価は $\{\varphi(n)\}_{n\in\mathbb{N}}$ が同程度連続であることを意味する．Ascoli-Arzelà の定理により，点列 $\{\varphi(n)\}_{n\in\mathbb{N}}$ は $\widetilde{\mathcal{C}}$ において相対コンパクトである．

したがって，$\{\varphi(n)\}_{n\in\mathbb{N}}$ の任意の集積点 ψ が $\varphi(\infty)$ と等しいことを示せば命題の証明が終わる．$\{\varphi(n_k)\}_{k\in\mathbb{N}}$ を ψ に収束する部分列とする．(すなわち $\lim_{k\to\infty} \|\varphi(n_k) - \psi\|_\infty = 0$ である．）骨格 ODE

$$\varphi(n_k)_t = a + \int_0^t b(\varphi(n_k)_u) du + \int_0^t \sigma(\varphi(n_k)_u) h(n_k)'_u du$$

において，各 t を固定するごとに $k \to \infty$ とする．左辺が ψ_t に収束することは自明である．b の Lipschitz 連続性から，右辺第 2 項は $\int_0^t b(\psi_u) du$ に収束する．もし右辺第 3 項が

$$\lim_{k\to\infty} \int_0^t \sigma(\varphi(n_k)_u) h(n_k)'_u du = \int_0^t \sigma(\psi_u) h(\infty)'_u du \tag{10.1.3}$$

と収束するならば，ψ は任意の t に対して

$$\psi_t = a + \int_0^t b(\psi_u) du + \int_0^t \sigma(\psi_u) h(\infty)'_u du$$

を満たすので，骨格 ODE の解の一意性から $\psi = \varphi(\infty)$ が示せる．

したがって (10.1.3) を示すと命題の証明が終わる．$\sigma = (\sigma_{ij})_{1\le i\le e, 1\le j\le d}$ と成分表示する．σ の第 i 行ベクトルを $\sigma_{i\star} = (\sigma_{i1}, \dots, \sigma_{id})$ と書く（$1 \le i \le e$）．すると (10.1.3) の左辺の第 i 成分は

$$\int_0^t \sigma_{i\star}(\varphi(n_k)_u) h(n_k)'_u du = \left\langle \int_0^\cdot \sigma_{i\star}(\varphi(n_k)_u)\mathbf{1}_{\{u\le t\}} du,\ h(n_k) \right\rangle_{\mathcal{H}}$$

と $\mathcal{H}$ の内積の形で書ける．同様に右辺の第 i 成分は

$$\int_0^t \sigma_{i\star}(\psi_u) h(\infty)'_u du = \left\langle \int_0^\cdot \sigma_{i\star}(\psi_u)\mathbf{1}_{\{u\le t\}} du,\ h(\infty) \right\rangle_{\mathcal{H}}$$

である．差を評価すると，

$$\begin{aligned}
&\left| \left\langle \int_0^\cdot \sigma_{i\star}(\varphi(n_k)_u)\mathbf{1}_{\{u\le t\}} du,\ h(n_k) \right\rangle_{\mathcal{H}} - \left\langle \int_0^\cdot \sigma_{i\star}(\psi_u)\mathbf{1}_{\{u\le t\}} du,\ h(\infty) \right\rangle_{\mathcal{H}} \right| \\
&\le \left| \left\langle \int_0^\cdot \{\sigma_{i\star}(\varphi(n_k)_u) - \sigma_{i\star}(\psi_u)\}\mathbf{1}_{\{u\le t\}} du,\ h(n_k) \right\rangle_{\mathcal{H}} \right| \\
&\qquad\qquad + \left| \left\langle \int_0^\cdot \sigma_{i\star}(\psi_u)\mathbf{1}_{\{u\le t\}} du,\ h(n_k) - h(\infty) \right\rangle_{\mathcal{H}} \right| \\
&\le \|h(n_k)\|_{\mathcal{H}} \left\{ \int_0^t |\sigma_{i\star}(\varphi(n_k)_u) - \sigma_{i\star}(\psi_u)|^2 du \right\}^{1/2} \\
&\qquad\qquad + \left| \left\langle \int_0^\cdot \sigma_{i\star}(\psi_u)\mathbf{1}_{\{u\le t\}} du,\ h(n_k) - h(\infty) \right\rangle_{\mathcal{H}} \right| \\
&\le M\sqrt{T}L\|\varphi(n_k) - \psi\|_\infty + \left| \left\langle \int_0^\cdot \sigma_{i\star}(\psi_u)\mathbf{1}_{\{u\le t\}} du,\ h(n_k) - h(\infty) \right\rangle_{\mathcal{H}} \right|
\end{aligned}$$

を得る．$k \to \infty$ のとき，右辺第 1 項は明らかに 0 に収束し，右辺第 2 項も $\{h(n)\}_{n\in\mathbb{N}}$ が弱位相での収束することを仮定したために 0 に収束する．これで (10.1.3) が証明できた． □

ここで (10.0.4) で定義された J が良い速度関数であることを確認する．

命題 10.1.3. 条件 **(LIP)** の下で，$J\colon \widetilde{\mathcal{C}} \to [0,\infty]$ は良い速度関数である．

証明. 任意の $r \in [0,\infty)$ に対して $J^{-1}([0,r]) = \{\xi \in \widetilde{\mathcal{C}} \mid J(\xi) \le r\}$ がコンパクトであることを見ればよい．$\{\xi(n)\}_{n\in\mathbb{N}}$ を $J^{-1}([0,r])$ の任意の点列とする．J の定義により，任意の $n \in \mathbb{N}$ に対して $I(h(n)) \le r + 1/n$ かつ $\xi(n) = \varphi^{h(n)}$ となる $h(n) \in \mathcal{H}$ が存在する．$\{h(n)\}_{n\in\mathbb{N}}$ は $\mathcal{H}$ の有界列であるので，補題 9.1.2 により弱収束する部分列 $\{h(n_k)\}_{k\in\mathbb{N}}$ を持つ．この部分列の弱極限を $h(\infty)$ と書き，$\xi(\infty) := \varphi^{h(\infty)}$ とおく．命題 10.1.2 により

$\lim_{k\to\infty}\|\xi(n_k)-\xi(\infty)\|_\infty = 0$ である．よって $\{\xi(n)\}$ は $\widetilde{\mathcal{C}}$ における収束部分列も持つことがわかったので，最後は収束先 $\xi(\infty)$ が $J^{-1}([0,r])$ に属することを見ればよい．

まず

$$I(h) = \frac{1}{2}\|h\|_{\mathcal{H}}^2 = \frac{1}{2}\sup\{|\langle h,k\rangle_{\mathcal{H}}|^2 \mid k\in\mathcal{H},\ \|k\|_{\mathcal{H}}=1\}$$

と書けることを思い出そう．各 k に対して $\mathcal{H}\ni h\mapsto |\langle h,k\rangle_{\mathcal{H}}|$ は弱位相に関して明らかに連続なので，補題 1.2.5 によりそれらの上限の形で書ける I は弱位相に関して下半連続である．考えている点列は閉球 $\overline{B}_{\mathcal{H}}(0,\sqrt{2(r+1)})$ に値を取るが，補題 9.1.2 により閉球は弱位相の下でコンパクト距離空間である．I の下半連続性と補題 1.2.7 から

$$I(h(\infty)) \le \varliminf_{k\to\infty} I(h(n_k)) \le \varliminf_{k\to\infty}(r+1/n_k^{-1}) = r$$

が従う．これは $J(\xi(\infty))\le r$ も意味する． □

10.2 離散化と指数的に良い近似を用いる定理 10.0.1 の証明

本節では Freidlin-Wentzell の大偏差原理（定理 10.0.1）に１つ目の証明を与える．まず SDE の解に対して Euler・丸山近似を考える.[9] この近似は Brown 運動の連続写像として書けているので，Schilder の大偏差原理（定理 9.0.1）と縮小原理を組み合わせると近似過程が大偏差原理を満たすことはすぐわかる．極限に対しても大偏差原理が成り立つことを指数的に良い近似定理（命題 1.6.5）を用いて証明する．

$\varepsilon \in (0,1]$ とする．$m \in \mathbb{N}$ に対して，$X^{\varepsilon,m} = (X_t^{\varepsilon,m})_{t\in[0,T]}$ を SDE (10.0.1) の Euler・丸山近似とする．すなわち

$$X_t^{\varepsilon,m} = a + \int_0^t b(X_{\lfloor ms\rfloor/m}^{\varepsilon,m})ds + \sqrt{\varepsilon}\int_0^t \sigma(X_{\lfloor ms\rfloor/m}^{\varepsilon,m})dw_s, \qquad t \in [0,T].$$

簡単に言うと，幅 $1/m$ の小区間において dw_s と ds の係数を時間に関して定値なもので置き換えたわけである．

$m \in \mathbb{N}$ と $w \in \mathcal{C}_0$ に対して，$\gamma \in \widetilde{\mathcal{C}}$ を以下のように定める．$\gamma_0 = a$ および

$$\gamma_t = \gamma_{k/m} + b(\gamma_{k/m})\Big(t - \frac{k}{m}\Big) + \sigma(\gamma_{k/m})(w_t - w_{k/m}), \tag{10.2.1}$$
$$t \in \Big[\frac{k}{m}, \Big(\frac{k+1}{m}\Big)\wedge T\Big], \quad k = 0,1,\ldots,\lfloor mT\rfloor.$$

（$\lfloor mT\rfloor/m = T$ となる場合は最後の小区間は１点集合に退化するが，これは無視することにする.）写像 $\Phi^m\colon \mathcal{C}_0 \to \widetilde{\mathcal{C}}$ を $\Phi^m(w) := \gamma$ により定めると，定め方から Φ^m は連続であり，さらに $X^{\varepsilon,m} = \Phi^m(\sqrt{\varepsilon}w)$ を満たす．これに記号を合わせて，骨格 ODE (10.0.3) の解をここでは $\Phi(h) := \varphi^h$ と書くことにする．命題 10.1.2 で見たように，$\Phi\colon \mathcal{H} \to \widetilde{\mathcal{C}}$ は連続である．

本節の前半では次の有界性条件を仮定に追加して，Freidlin-Wentzell の大偏差原理を証明する．

(BND) $\|\sigma\|_\infty := \sup_{x\in\mathbb{R}^e}|\sigma(x)| < \infty$ かつ $\|b\|_\infty := \sup_{x\in\mathbb{R}^e}|b(x)| < \infty$.

命題 10.2.1. 条件 **(LIP)** に加えて条件 **(BND)** も仮定する．このとき定理 10.0.1 で述べた大偏差原理が成り立つ．

証明. 本節の残りでは，**(LIP)** と **(BND)** の下で以下の２条件が成立することを確認する．

$$\lim_{m\to\infty}\varlimsup_{\varepsilon\searrow 0}\varepsilon\log\mathbb{P}\left(\|X^{\varepsilon,m} - X^\varepsilon\|_\infty > \delta\right) = -\infty, \qquad \delta \in (0,\infty), \tag{10.2.2}$$

[9] 例えば [11, 第 6.1 章，命題 6.1.2] には，**(LIP)** と **(BND)** の下で Euler・丸山近似が適切な意味で収束することが証明されている．(ただし本書ではその命題は用いない.)

$$\lim_{m\to\infty} \sup_{h\in\mathcal{H};\|h\|_{\mathcal{H}}\le r} \|\Phi^m(h)-\Phi(h)\|_\infty = 0, \qquad r\in[0,\infty). \tag{10.2.3}$$

これらの２条件により命題 1.6.5 が使えることを丁寧に確認しよう.（それから直ちに命題 10.2.1 が従う.）まず指数的に良い近似を導入した定義 1.6.3 における確率空間 $(\Omega, \mathcal{B}_\varepsilon, \mathbb{Q}_{\varepsilon,m})$ として，(ε, m) に依存しない $(\mathcal{C}_0, \mathcal{F}, \mathbb{P})$ を選んだと思うと，(10.2.2) は $\{X^{\varepsilon,m}\}$ が $\{X^\varepsilon\}$ を指数的に良く近似することを意味する．そこで以下のようにおくと，(10.2.3) により指数的に良い近似に関する命題 1.6.5 の仮定が全て満たされていることがわかる．$\mathcal{X}=\mathcal{C}_0$, $\mathcal{Y}=\widetilde{\mathcal{C}}$, $\mu_\varepsilon=\mathbb{P}^\varepsilon$（すなわち $(\sqrt{\varepsilon}w_t)_{t\in[0,T]}$ の法則），$\widetilde{\mu}_\varepsilon \overset{\mathrm{Law}}{=} X^\varepsilon$, $\psi_m=\Phi^m$ とし，最後に $\psi\colon \mathcal{X}\to\mathcal{Y}$ は Φ の任意の Borel 可測拡張とする. □

不等式 (10.2.3) の証明. 上記の２つの不等式のうち (10.2.3) は比較的容易なので，これから示す．本証明中では，σ, b と T のみに依存する正定数を C と書く.（C の値は行ごとに変わる.）

$\|h\|_{\mathcal{H}}\le r$ として，$\widehat{\gamma}=\Phi^m(h)$ および $\varphi=\Phi(h)$ と書く．(10.2.1) 式で w を h に置き換えたものから直ちに，任意の t と m に対して

$$|\widehat{\gamma}_t-\widehat{\gamma}_{\lfloor mt\rfloor/m}| \le \|b\|_\infty m^{-1}+\|\sigma\|_\infty r m^{-1/2} \le C(1+r)m^{-1/2}$$

となることがわかる．ここで **(BND)** と Schwarz の不等式を用いた.

次に $\varphi-\widehat{\gamma}$ を評価する．**(LIP)** と Schwarz の不等式により，

$$\begin{aligned}
|\varphi_t-\widehat{\gamma}_t|^2 &\le 2\left|\int_0^t \{b(\varphi_s)-b(\widehat{\gamma}_{\lfloor ms\rfloor/m})\}ds\right|^2 + 2\left|\int_0^t \{\sigma(\varphi_s)-\sigma(\widehat{\gamma}_{\lfloor ms\rfloor/m})\}h'_s ds\right|^2 \\
&\le C(1+r)^2\int_0^t \left|\varphi_s-\widehat{\gamma}_{\lfloor ms\rfloor/m}\right|^2 ds \\
&\le C(1+r)^2\left\{\int_0^t |\varphi_s-\widehat{\gamma}_s|^2 ds + \int_0^t \left|\widehat{\gamma}_s-\widehat{\gamma}_{\lfloor ms\rfloor/m}\right|^2 ds\right\} \\
&\le C(1+r)^2\int_0^t |\varphi_s-\widehat{\gamma}_s|^2 ds + C(1+r)^4 m^{-1}
\end{aligned}$$

となる．Gronwall の不等式により，上の評価から

$$\|\Phi(h)-\Phi^m(h)\|_\infty = \|\varphi-\widehat{\gamma}\|_\infty \le C(1+r)^4 m^{-1}\exp[C(1+r)^2T]$$

を得る．この評価から直ちに (10.2.3) がわかる. □

１つ目の評価 (10.2.2) を証明する前に，まず伊藤過程に関する補題を１つ示す.

補題 10.2.2. $\varepsilon\in(0,1]$，$z\in\mathbb{R}^e$ および $\{\mathcal{F}_t\}_{t\in[0,T]}$ 発展的可測な $\mathbb{R}^e$ 値過程 $(\beta_t)_{t\in[0,T]}$ と $\{\mathcal{F}_t\}_{t\in[0,T]}$ 発展的可測な有界 $\mathbb{R}^{e\times d}$ 値過程 $(\alpha_t)_{t\in[0,T]}$ に対して，

$$Z_t^\varepsilon := z+\int_0^t \beta_s ds + \sqrt{\varepsilon}\int_0^t \alpha_s dw_s, \qquad t\in[0,T].$$

とおく．ここで右辺第 3 項は伊藤積分である．また τ_1 を $[0,T]$ 値 $\{\mathcal{F}_t\}_{t\in[0,T]}$ 停止時刻とする．さらに，正定数 M, N, ρ が存在して，任意の $t \in [0,\tau_1]$ に対して

$$|\alpha_t| \le M(\rho^2 + |Z_t^\varepsilon|^2)^{1/2}, \qquad |\beta_t| \le N(\rho^2 + |Z_t^\varepsilon|^2)^{1/2} \tag{10.2.4}$$

を満たすことも仮定する．このとき，任意の $\delta > 0$ と $\varepsilon \in (0,1]$ に対して

$$\varepsilon \log \mathbb{P}\left(\sup_{0\le t\le\tau_1} |Z_t^\varepsilon| \ge \delta\right) \le T\{2N + M^2(2+\sqrt{e})\} + \log\Big(\frac{\rho^2+|z|^2}{\rho^2+\delta^2}\Big)$$

が成立する．

証明. $x = (x^1,\dots,x^e) \in \mathbb{R}^e$ に対して $\varphi(x) = (\rho^2 + |x|^2)^{1/\varepsilon}$ とおくと，$\varphi\colon \mathbb{R}^e \to (0,\infty)$ は明らかに C^2 級である．よって，$U_t^\varepsilon := \varphi(Z_t^\varepsilon)$ を伊藤の公式を用いて計算できる．[10]

まず Z^ε の各成分の 2 次変分過程を計算すると，

$$\begin{aligned} d\langle Z^{\varepsilon,i}, Z^{\varepsilon,j}\rangle_t &= d\Big\langle \sqrt{\varepsilon}\sum_k \int_0^\cdot \alpha_s^{ik} dw_s^k,\ \sqrt{\varepsilon}\sum_l \int_0^\cdot \alpha_s^{jl} dw_s^l \Big\rangle_t \\ &= \varepsilon \sum_{k,l} \alpha_t^{ik}\alpha_t^{jl}\delta_{kl} dt = \varepsilon \sum_k \alpha_t^{ik}\alpha_t^{jk} dt = \varepsilon(\alpha_t\alpha_t^\top)^{ij} dt \end{aligned}$$

となる．ここで δ_{kl} は Kronecker のデルタ記号である．また $A^\top$ は行列 A の転置行列を表す．伊藤の公式を用いて計算すると

$$\begin{aligned} dU_t^\varepsilon &= \sum_i \partial_i\varphi(Z_t^\varepsilon) dZ_t^{\varepsilon,i} + \frac{1}{2}\sum_{i,j} \partial_{ij}^2\varphi(Z_t^\varepsilon) d\langle Z^{\varepsilon,i}, Z^{\varepsilon,j}\rangle_t \\ &= \nabla\varphi(Z_t^\varepsilon)^\top dZ_t^\varepsilon + \frac{\varepsilon}{2}\mathrm{Trace}[\alpha_t\alpha_t^\top \nabla^2\varphi(Z_t^\varepsilon)] dt \\ &= \Big\{\nabla\varphi(Z_t^\varepsilon)^\top\beta_t + \frac{\varepsilon}{2}\mathrm{Trace}[\alpha_t\alpha_t^\top\nabla^2\varphi(Z_t^\varepsilon)]\Big\} dt + \sqrt{\varepsilon}\nabla\varphi(Z_t^\varepsilon)^\top\alpha_t dw_t \end{aligned} \tag{10.2.5}$$

となる．ここで Trace は正方行列のトレース（対角成分の和）を表す．また $\nabla\varphi$ と $\nabla^2\varphi$ はそれぞれ φ の勾配と Hesse 行列を表す．(簡単のため $\partial_i = \partial/\partial x^i$, $\partial_{ij}^2 = \partial^2/\partial x^i\partial x^j$ などと略記した.)

直接の微分計算により

$$\big(\partial_1\varphi(x),\dots,\partial_e\varphi(x)\big)^\top = \nabla\varphi(x) = \frac{2\varphi(x)}{\varepsilon(\rho^2+|x|^2)} x$$

となるので，(10.2.4) と合わせると，$t \in [0,\tau_1]$ のときに

$$|\nabla\varphi(Z_t^\varepsilon)^\top\beta_t| \le \frac{2\varphi(Z_t^\varepsilon)}{\varepsilon(\rho^2+|Z_t^\varepsilon|^2)}|Z_t^\varepsilon|\cdot N(\rho^2+|Z_t^\varepsilon|^2)^{1/2} \le \frac{2N}{\varepsilon}U_t^\varepsilon \tag{10.2.6}$$

[10] 伊藤の公式は SDE に関する全ての教科書に載っている．例えば [16, 第 2.4 節，定理 2.4.2] や [13, 第 4.2 節，定理 4.13] や [23, 第 2.2 節，定理 2.1] を見よ．

と評価できることがわかる．

次は２階微分を計算すると，$i,j \in \{1,\dots,e\}$ に対して

$$\partial_{ij}^2\varphi(x) = \frac{1}{\varepsilon}\Big(\frac{1}{\varepsilon}-1\Big)4x_i x_j(\rho^2+|x|^2)^{(1/\varepsilon)-2} + \delta_{ij}\frac{2}{\varepsilon}(\rho^2+|x|^2)^{(1/\varepsilon)-1}$$

と求まる．よって

$$\begin{aligned}|\nabla^2\varphi(x)| &\le \frac{4}{\varepsilon}\Big(\frac{1}{\varepsilon}-1\Big)(\rho^2+|x|^2)^{(1/\varepsilon)-2}|xx^\top| + \frac{2}{\varepsilon}(\rho^2+|x|^2)^{(1/\varepsilon)-1}|\mathrm{Id}_e| \\ &\le \frac{1}{\varepsilon}\Big(\frac{4}{\varepsilon}+2\sqrt{e}\Big)(\rho^2+|x|^2)^{(1/\varepsilon)-1}\end{aligned}$$

という評価を得る．再び (10.2.4) を用いて，$t \in [0,\tau_1]$ のときに

$$\begin{aligned}\frac{\varepsilon}{2}\left|\mathrm{Trace}[\alpha_t\alpha_t^\top\nabla^2\varphi(Z_t^\varepsilon)]\right| &\le \frac{\varepsilon}{2}|\alpha_t\alpha_t^\top||\nabla^2\varphi(Z_t^\varepsilon)| \\ &\le \frac{\varepsilon}{2}M^2(\rho^2+|Z_t^\varepsilon|^2)\frac{1}{\varepsilon}\Big(\frac{4}{\varepsilon}+2\sqrt{e}\Big)(\rho^2+|Z_t^\varepsilon|^2)^{(1/\varepsilon)-1} \\ &\le \frac{M^2(2+\sqrt{e})}{\varepsilon}U_t^\varepsilon \qquad (10.2.7)\end{aligned}$$

と評価できる．

さて $\delta > 0$ に対して，別の停止時刻を $\tau_2 := \inf\{t \in [0,T] \mid |Z_t^\varepsilon| \ge \delta\} \wedge \tau_1$ と定める．すると $t \le \tau_2$ のときは Z_t^ε は有界なので $t \mapsto \int_0^{t\wedge\tau_2}\nabla\varphi(Z_s^\varepsilon)^\top\alpha_s dw_s$ は L^2 マルチンゲールで，特に任意の t に対して $\int_0^{t\wedge\tau_2}\nabla\varphi(Z_s^\varepsilon)^\top\alpha_s dw_s$ の期待値は 0 である．したがって，(10.2.5) の期待値を取ると

$$\begin{aligned}\mathbb{E}\left[U_{t\wedge\tau_2}^\varepsilon\right] &= U_0^\varepsilon + \mathbb{E}\left[\int_0^{t\wedge\tau_2}\Big\{\nabla\varphi(Z_s^\varepsilon)^\top\beta_s + \frac{\varepsilon}{2}\mathrm{Trace}[\alpha_s\alpha_s^\top\nabla^2\varphi(Z_s^\varepsilon)]\Big\}ds\right] \\ &\le \varphi(z) + \frac{2N+M^2(2+\sqrt{e})}{\varepsilon}\mathbb{E}\left[\int_0^{t\wedge\tau_2}U_s^\varepsilon ds\right] \\ &= \varphi(z) + \frac{2N+M^2(2+\sqrt{e})}{\varepsilon}\mathbb{E}\left[\int_0^{t\wedge\tau_2}U_{s\wedge\tau_2}^\varepsilon ds\right] \\ &\le \varphi(z) + \frac{2N+M^2(2+\sqrt{e})}{\varepsilon}\int_0^t\mathbb{E}\left[U_{s\wedge\tau_2}^\varepsilon\right]ds\end{aligned}$$

となる．１つ目の不等式で (10.2.6) と (10.2.7) を用い，最後の不等式で U^ε の非負性を用いた．Gronwall の不等式により

$$\mathbb{E}\left[U_{\tau_2}^\varepsilon\right] = \mathbb{E}\left[U_{T\wedge\tau_2}^\varepsilon\right] \le \varphi(z)\exp(T\{2N+M^2(2+\sqrt{e})\}/\varepsilon)$$

を得る．

ここで $\psi(\delta) := (\rho^2+\delta^2)^{1/\varepsilon}$ とおく．明らかに $\varphi(x) = \psi(|x|)$ である．$\psi(\delta)$ は $\delta \in (0,\infty)$ の関数だと見ると単調増加なので，

$$\mathbb{P}\left(|Z_{\tau_2}^\varepsilon| \ge \delta\right) = \mathbb{P}\left(\varphi(Z_{\tau_2}^\varepsilon) \ge \psi(\delta)\right) \le \frac{\mathbb{E}\left[\varphi(Z_{\tau_2}^\varepsilon)\right]}{\psi(\delta)} = \frac{\mathbb{E}\left[U_{\tau_2}^\varepsilon\right]}{\psi(\delta)}$$

を得る．上では Chebyshev の不等式も使った．これら 2 つの不等式を合わせて次の評価を得る．

$$\begin{aligned}\varepsilon \log \mathbb{P}\left(\sup_{0\le t\le \tau_1} |Z_t^\varepsilon| \ge \delta\right) &\le \varepsilon \log \mathbb{P}\left(|Z_{\tau_2}^\varepsilon| \ge \delta\right)\\ &\le T\{2N + M^2(2+\sqrt{e})\} + \varepsilon \log \varphi(z) - \varepsilon \log \psi(\delta)\\ &= T\{2N + M^2(2+\sqrt{e})\} + \log\Big(\frac{\rho^2+|z|^2}{\rho^2+\delta^2}\Big).\end{aligned}$$

これで補題 10.2.2 が証明できた．　□

不等式 (10.2.2) **の証明**．任意の $\delta > 0$ を選び，以下では固定する．$Z_t^{\varepsilon,m} := X_t^{\varepsilon,m} - X_t^\varepsilon$ とおき，$z = 0$, $\beta_t := b(X_{\lfloor mt\rfloor/m}^{\varepsilon,m}) - b(X_t^\varepsilon)$ および $\alpha_t := \sigma(X_{\lfloor mt\rfloor/m}^{\varepsilon,m}) - \sigma(X_t^\varepsilon)$ に対して補題 10.2.2 を使いたい．任意の $\rho > 0$ に対して，停止時刻 $\tau_1 = \tau_1(\varepsilon, m, \rho)$ を

$$\tau_1 := \inf\{t \in [0,T] \mid |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \ge \rho\} \wedge T$$

と定める．**(LIP)** と 3 角不等式により，$t \le \tau_1$ のとき $|\alpha_t| \vee |\beta_t| \le L(\rho + |Z_t^{\varepsilon,m}|)$ だと簡単にわかるので，α と β は条件 (10.2.4) を満たす．これで補題 10.2.2 が使えて，

$$\varepsilon \log \mathbb{P}\left(\sup_{0\le t\le \tau_1} |Z_t^{\varepsilon,m}| \ge \delta\right) \le K + \log\Big(\frac{\rho^2}{\rho^2+\delta^2}\Big)$$

とわかる．ここで $K > 0$ は $\varepsilon, \delta, \rho, m$ には依存しないある正定数である．この式において $\varepsilon \searrow 0$ としてから $\rho \searrow 0$ とすると，次を得る．

$$\lim_{\rho\searrow 0} \sup_{m\ge 1} \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\sup_{0\le t\le \tau_1} |X_t^{\varepsilon,m} - X_t^\varepsilon| \ge \delta\right) = -\infty, \qquad \delta > 0. \tag{10.2.8}$$

さて，$\tau_1 = T$ のときは $\|X^{\varepsilon,m} - X^\varepsilon\|_\infty = \sup_{0\le t\le \tau_1} |X_t^{\varepsilon,m} - X_t^\varepsilon|$ なので，

$$\begin{aligned}\{\|X^{\varepsilon,m} - X^\varepsilon\|_\infty > \delta\} &\subset \{\tau_1 < T\} \cup \{\sup_{0\le t\le \tau_1} |X_t^{\varepsilon,m} - X_t^\varepsilon| \ge \delta\}\\ &\subset \{\sup_{0\le t\le T} |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \ge \rho\} \cup \{\sup_{0\le t\le \tau_1} |X_t^{\varepsilon,m} - X_t^\varepsilon| \ge \delta\}\end{aligned}$$

となることに注意せよ．したがって，任意の $\rho > 0$ に対して

$$\lim_{m\to\infty} \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\sup_{0\le t\le T} |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \ge \rho\right) = -\infty \tag{10.2.9}$$

を示すと不等式 (10.2.2) の証明が終わる．実際，補題 1.2.2 を用いると，各 m, ρ ごとに

$$\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\|X^{\varepsilon,m} - X^\varepsilon\|_\infty > \delta\right) \le \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\sup_{0\le t\le T} |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \ge \rho\right)$$

$$\vee \sup_{m\geq 1} \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\sup_{0\leq t\leq \tau_1} |X_t^{\varepsilon,m} - X_t^{\varepsilon}| \geq \delta\right)$$

を得るが，この式で $m\to\infty$ として (10.2.9) を用いると

$$\varlimsup_{m\to\infty} \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\|X^{\varepsilon,m} - X^{\varepsilon}\|_\infty > \delta\right) \leq \sup_{m\geq 1} \varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\sup_{0\leq t\leq \tau_1} |X_t^{\varepsilon,m} - X_t^{\varepsilon}| \geq \delta\right)$$

を得る．左辺は ρ に依存しないが，右辺は τ_1 を通じて依存する．ここで $\rho\searrow 0$ として (10.2.8) を用いると，目標の不等式 (10.2.2) を得る．

それでは (10.2.9) を確認する．以下では $C = \|\sigma\|_\infty \vee \|b\|_\infty$ とおく．($C=0$ のときは全てが自明なので，$C>0$ と仮定してよい．また記述の便宜上，(w_t) の定義域を $[0,T]$ から $[0,\lceil T\rceil]$ に拡張しておく．）簡単な計算で

$$\sup_{t\in[0,T]} |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \leq C\big(m^{-1} + \sqrt{\varepsilon} \max_{0\leq k\leq m\lceil T\rceil-1} \sup_{s\in[0,1/m]} |w_{s+k/m} - w_{k/m}|\big)$$

と評価できる．$m > C/\rho$ ならば,

$$\begin{aligned}
&\sup_{t\in[0,T]} |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \geq \rho \\
&\Longrightarrow \max_{0\leq k\leq m\lceil T\rceil-1} \sup_{s\in[0,1/m]} |w_{s+k/m} - w_{k/m}| \geq \frac{\rho - C/m}{C\sqrt{\varepsilon}} \\
&\Longrightarrow \text{ある } k\in\{0,\dots,m\lceil T\rceil-1\} \text{ に対し } \sup_{s\in[0,1/m]} |w_{s+k/m} - w_{k/m}| \geq \frac{\rho - C/m}{C\sqrt{\varepsilon}}
\end{aligned}$$

となる．任意の k,m に対して $(w_{s+k/m} - w_{k/m})$ は再び標準 Brown 運動である．よって

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[0,T]} |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \geq \rho\right) &\leq \mathbb{P}\left(\bigcup_{k=0}^{m\lceil T\rceil-1} \sup_{s\in[0,1/m]} |w_{s+k/m} - w_{k/m}| \geq \frac{\rho - C/m}{C\sqrt{\varepsilon}}\right) \\
&\leq m\lceil T\rceil \mathbb{P}\left(\sup_{s\in[0,1/m]} |w_s| \geq \frac{\rho - C/m}{C\sqrt{\varepsilon}}\right) \\
&\leq 2dm\lceil T\rceil \exp\left(-\frac{m(\rho - C/m)^2}{2C^2 d\varepsilon}\right).
\end{aligned}$$

最後の不等式で Brown 運動の脱出確率に関する補題 9.4.2 を用いた．したがって

$$\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\sup_{t\in[0,T]} |X_t^{\varepsilon,m} - X_{\lfloor mt\rfloor/m}^{\varepsilon,m}| \geq \rho\right) \leq -\frac{m(\rho - C/m)^2}{2C^2 d}$$

となるが，この式で $m\to\infty$ とすると右辺は $-\infty$ に発散する．これで (10.2.9) が示された． □

ここから条件 **(BND)** を仮定から外して，係数 σ と b が条件 **(LIP)** のみを満たす場合を考察する．

各 $m \in \mathbb{N}$ に対して，$\chi_m\colon \mathbb{R}^e \to [0,1]$ を $|x| \le m$ のときは $\chi_m(x) = 1$，$|x| \ge m+1$ のときは $\chi_m(x) = 0$ を満たす Lipschitz 連続関数とする．これを用いて $\sigma_m := \chi_m \cdot \sigma$ および $b_m := \chi_m \cdot b$ とおくと，これらは明らかに **(LIP)** も **(BND)** も満たす．

SDE (10.0.1) において係数をこれらで置き換えたものの解を $\widetilde{X}^{\varepsilon,m} = (\widetilde{X}^{\varepsilon,m}_t)_{t\in[0,T]}$ と書く．すなわち，式で書けば

$$\widetilde{X}^{\varepsilon,m}_t = a + \int_0^t b_m(\widetilde{X}^{\varepsilon,m}_s)ds + \sqrt{\varepsilon}\int_0^t \sigma_m(\widetilde{X}^{\varepsilon,m}_s)dw_s, \qquad t \in [0,T]$$

である．$h \in \mathcal{H}$ のとき，この SDE に対応する骨格 ODE を

$$\varphi^{m,h}_t = a + \int_0^t b_m(\varphi^{m,h}_s)ds + \int_0^t \sigma_m(\varphi^{m,h}_s)h'_s ds, \qquad t \in [0,T]$$

と書く．対応する良い速度関数を J_m と表す．

$$J_m(\xi) := \inf\{I(h) \mid h \in \mathcal{H},\ \ \xi = \varphi^{m,h}\}, \qquad \xi \in \widetilde{\mathcal{C}}.$$

すでに証明したように，各 m に対して $\{\widetilde{X}^{\varepsilon,m}\}_{\varepsilon\in(0,1]}$ は $\varepsilon \searrow 0$ のときに良い速度関数 J_m に対して大偏差原理を満たす．以下では $m \to \infty$ のときに，この大偏差原理から極限 $\{X^\varepsilon\}_{\varepsilon\in(0,1]}$ の大偏差原理が導けることを示す．

補題 10.2.3. 条件 **(LIP)** を仮定する．このとき，$\{\widetilde{X}^{\varepsilon,m}\}_{m\in\mathbb{N},\varepsilon\in(0,1]}$ は $\{X^\varepsilon\}_{\varepsilon\in(0,1]}$ の指数的に良い近似である．すなわち，任意の $\delta > 0$ に対して次が成立する．

$$\lim_{m\to\infty}\varlimsup_{\varepsilon\searrow 0} \varepsilon \log \mathbb{P}\left(\|\widetilde{X}^{\varepsilon,m} - X^\varepsilon\|_\infty > \delta\right) = -\infty.$$

証明. 停止時刻を $\tau_{\varepsilon,m} = \inf\{t \in [0,T] \mid |X^\varepsilon| \ge m\} \wedge T$ とおく．要するに，これは半径 m の球からの脱出時刻である．よく知られているように，$\tau_{\varepsilon,m}$ までは X^ε と $\widetilde{X}^{\varepsilon,m}$ は一致する．念のためこれを確認しよう．確率積分の停止操作に関する基本関係式

$$\int_0^{t\wedge\tau_{\varepsilon,m}} \sigma(X^\varepsilon_s)dw_s = \int_0^{t\wedge\tau_{\varepsilon,m}} \sigma(X^\varepsilon_s)\mathbf{1}_{\{s\le\tau_{\varepsilon,m}\}}dw_s = \int_0^t \sigma(X^\varepsilon_s)\mathbf{1}_{\{s\le\tau_{\varepsilon,m}\}}dw_s$$

を思い出そう．さらに $s \le \tau_{\varepsilon,m}$ のときは $\sigma(X^\varepsilon_s) = \sigma_m(X^\varepsilon_s)$ などが成り立つので

$$\begin{aligned}
\widetilde{X}^{\varepsilon,m}_{t\wedge\tau_{\varepsilon,m}} - X^\varepsilon_{t\wedge\tau_{\varepsilon,m}} &= \int_0^{t\wedge\tau_{\varepsilon,m}} \left\{b_m(\widetilde{X}^{\varepsilon,m}_{s\wedge\tau_{\varepsilon,m}}) - b(X^\varepsilon_{s\wedge\tau_{\varepsilon,m}})\right\}ds \\
&\qquad + \sqrt{\varepsilon}\int_0^{t\wedge\tau_{\varepsilon,m}} \left\{\sigma_m(\widetilde{X}^{\varepsilon,m}_{s\wedge\tau_{\varepsilon,m}}) - \sigma(X^\varepsilon_{s\wedge\tau_{\varepsilon,m}})\right\}dw_s \\
&= \int_0^t \left\{b_m(\widetilde{X}^{\varepsilon,m}_{s\wedge\tau_{\varepsilon,m}}) - b_m(X^\varepsilon_{s\wedge\tau_{\varepsilon,m}})\right\}\mathbf{1}_{\{s\le\tau_{\varepsilon,m}\}}ds
\end{aligned}$$

$$+\sqrt{\varepsilon}\int_0^t \left\{\sigma_m(\widetilde{X}^{\varepsilon,m}_{s\wedge\tau_{\varepsilon,m}})-\sigma_m(X^{\varepsilon}_{s\wedge\tau_{\varepsilon,m}})\right\}\mathbf{1}_{\{s\le\tau_{\varepsilon,m}\}}dw_s$$

となる．この式を 2 乗して σ_m と b_m の Lipschitz 連続性を使うと，ある正定数 $c=c(m)$ に対して

$$\mathbb{E}\left[|\widetilde{X}^{\varepsilon,m}_{t\wedge\tau_{\varepsilon,m}}-X^{\varepsilon}_{t\wedge\tau_{\varepsilon,m}}|^2\right]\le c\int_0^t \mathbb{E}\left[|\widetilde{X}^{\varepsilon,m}_{s\wedge\tau_{\varepsilon,m}}-X^{\varepsilon}_{s\wedge\tau_{\varepsilon,m}}|^2\right]ds$$

となることが簡単に示せる．各 t に対して $\mathbb{E}\left[|\widetilde{X}^{\varepsilon,m}_{t\wedge\tau_{\varepsilon,m}}-X^{\varepsilon}_{t\wedge\tau_{\varepsilon,m}}|^2\right]=0$ であることが Gronwall の不等式から直ちに従う．t に関する連続性を考慮に入れると，

$$\mathbb{P}\left(\sup_{0\le t\le T}\left|\widetilde{X}^{\varepsilon,m}_{t\wedge\tau_{\varepsilon,m}}-X^{\varepsilon}_{t\wedge\tau_{\varepsilon,m}}\right|=0\right)=1$$

である．これで脱出時刻まで二つの過程が一致することが確認できた．従って，特に次がわかる．

$$\|\widetilde{X}^{\varepsilon,m}-X^{\varepsilon}\|_\infty>\delta \quad\Longrightarrow\quad \tau_{\varepsilon,m}<T \quad\Longrightarrow\quad \|\widetilde{X}^{\varepsilon,m}\|_\infty\ge m. \tag{10.2.10}$$

次は $r>0$ に対して，命題 10.1.1 で調べた量を

$$A_r:=\sup\{\|\varphi^h\|_\infty \mid h\in\overline{B}_{\mathcal{H}}(0,r)\}<\infty \tag{10.2.11}$$

と表す．$m>|a|$ を満たす $m\in\mathbb{N}$ に対して，

$$r_m:=\sup\{r>0\mid A_r\le m\}\ \in(0,\infty]$$

とおく．命題 10.1.2 により $\|h\|_{\mathcal{H}}\le r_m$ ならば $\|\varphi^h\|_\infty\le m$ であるので，$\varphi^h=\varphi^{m,h}$ となる．明らかに $\{r_m\}$ は m に関して広義単調増加であり，さらに $(r_\infty:=)\lim_{m\to\infty}r_m=\infty$ となる．実際，$r_\infty<\infty$ ならば $A_{r_\infty+1}=\infty$ となるため，矛盾が生ずる．

さて (10.2.10) と $\{\widetilde{X}^{\varepsilon,m}\}_{\varepsilon\in(0,1]}$ に対する大偏差原理（命題 10.2.1）の上からの評価を用いると，

$$\begin{aligned}
\varlimsup_{\varepsilon\searrow0}\varepsilon\log\mathbb{P}\left(\|\widetilde{X}^{\varepsilon,m}-X^{\varepsilon}\|_\infty>\delta\right) &\le \varlimsup_{\varepsilon\searrow0}\varepsilon\log\mathbb{P}\left(\|\widetilde{X}^{\varepsilon,m}\|_\infty\ge m\right)\\
&\le -\inf\{J_m(\xi)\mid \|\xi\|_\infty\ge m\}\\
&= -\inf\left\{\|h\|_{\mathcal{H}}^2/2 \mid h\in\mathcal{H},\ \|\varphi^{m,h}\|_\infty\ge m\right\}\\
&\le -r_{m-1}^2/2
\end{aligned}$$

と評価できるが，最右辺は $m\to\infty$ のときに $-\infty$ に発散するので，補題の証明が終わる．なお最後の不等式では，$\|h\|_{\mathcal{H}}\le r_{m-1}$ のとき $\varphi^h=\varphi^{m-1,h}=\varphi^{m,h}\in\overline{B}_{\widetilde{\mathcal{C}}}(0,m-1)$ となることを用いた． □

補題 10.2.4. 条件 **(LIP)** を仮定する．このとき，

$$J(\xi) = \sup_{\delta>0} \varliminf_{m\to\infty} \inf\{J_m(\eta) \mid \eta \in B_{\widetilde{\mathcal{C}}}(\xi,\delta)\}, \qquad \xi \in \widetilde{\mathcal{C}} \tag{10.2.12}$$

が成り立つ．さらに任意の閉集合 $F \subset \widetilde{\mathcal{C}}$ に対して

$$\inf\{J(\xi) \mid \xi \in F\} \le \varliminf_{m\to\infty} \inf\{J_m(\xi) \mid \xi \in F\} \tag{10.2.13}$$

も成り立つ．

証明．まず (10.2.12) を示す．$\xi \in \widetilde{\mathcal{C}}$ を任意に取る．δ に関する単調性により，右辺の $\sup_{\delta>0}$ は $\lim_{\delta\searrow 0}$ に置き換えられる．特に $\delta \le 1$ と仮定してよいので，$\eta \in B_{\widetilde{\mathcal{C}}}(\xi,\delta)$ ならば $\|\eta\|_\infty \le 1 + \|\xi\|_\infty$ である．すると $m > 1 + \|\xi\|_\infty$ のときは $J_m(\eta) = J(\eta)$ となる．したがって，

$$J(\xi) = \lim_{\delta\searrow 0} \inf\{J(\eta) \mid \eta \in B_{\widetilde{\mathcal{C}}}(\xi,\delta)\}$$

を示せば十分であるが，命題 10.1.3 で J が下半連続であることを見たので，実はこの式は (1.2.5) そのものである．

続いて (10.2.13) を示す．簡単のために $l := \inf\{J(\xi) \mid \xi \in F\}$, $l(m) := \inf\{J_m(\xi) \mid \xi \in F\}$ および $l(\infty) := \varliminf_{m\to\infty} l(m)$ と書く．背理法を用いる．$l > l(\infty)$ と仮定して矛盾を導く．$s \in (l(\infty), l)$ を１つ取る．$\{l(m_k)\}_{k\in\mathbb{N}}$ を $\{l(m)\}$ の部分列で $\lim_{k\to\infty} l(m_k) = l(\infty)$ となるものとする．このとき，J_m の定義により $\{h(m_k)\}_{k\in\mathbb{N}} \subset \mathcal{H}$ で，任意の十分大きな k に対して $\|h(m_k)\|_{\mathcal{H}}^2/2 \le s$ かつ $\varphi^{m_k,h(m_k)} \in F$ となるものが存在する．補題 9.1.2 により，$\{h(m_k)\}_{k\in\mathbb{N}}$ のさらなる部分列（再び同じ記号で表す）で弱収束するものが存在する．この収束先を $h(\infty)$ と書くと，補題 9.1.2 により $\|h(\infty)\|_{\mathcal{H}}^2/2 \le s$ である．係数行列の切り落とし方により $m_k > A_{\sqrt{2s}}$ であれば，$\varphi^{m_k,h(m_k)} = \varphi^{h(m_k)}$ である．（$A_{\sqrt{2s}}$ は (10.2.11) で定義された．）命題 10.1.2 により，$\widetilde{\mathcal{C}}$ 内で $\lim_{k\to\infty} \varphi^{h(m_k)} = \varphi^{h(\infty)} \in F$ であることもわかる．F が閉であることも用いた．よって $l \le J(\varphi^{h(\infty)}) \le \|h(\infty)\|_{\mathcal{H}}^2/2 \le s$ となり矛盾が生じた． □

指数的に良い近似に関する命題 1.6.4 を用いて，本章の主題である Freidlin-Wentzell の大偏差原理（定理 10.0.1）の証明を仕上げる．

定理 10.0.1 の証明． 補題 10.2.3 で $\{\widetilde{X}^{\varepsilon,m}\}_{m\in\mathbb{N},\varepsilon\in(0,1]}$ が $\{X^\varepsilon\}_{\varepsilon\in(0,1]}$ の指数的に良い近似であることを示した．各 m に対して，$\{\widetilde{X}^{\varepsilon,m}\}_{\varepsilon\in(0,1]}$ が大偏差原理を満たすことを命題 10.2.1 で示した．上の補題 10.2.4 で示した (10.2.12) と (10.2.13) がそれぞれ命題 1.6.4 の仮定 (1.6.1) と (1.6.2) である．J が良い速度関数であることは命題 10.1.3 で確かめた．以上により，命題 1.6.4 が適用できて，目標とする定理 10.0.1 の証明が終わる． □

10.3　大偏差評価を用いる定理 10.0.1 の証明

本節では大偏差評価を経由して Freidlin-Wentzell の大偏差原理（定理 10.0.1）を証明する方法を解説する.[11] この手法については第 1.7 節で一般的にまとめた．ここで言う大偏差評価とは (1.7.1) や (10.3.13) などの評価式を意味する．

以下では命題 1.7.1 を Freidlin-Wentzell の大偏差原理の証明に適用する．その場合，確率空間は Wiener 空間 $(\mathcal{C}_0, \mathcal{F}, \mathbb{P})$，距離空間は $\mathcal{X} = \mathcal{C}_0$ および $\mathcal{Y} = \widetilde{\mathcal{C}}$，確率変数はそれぞれ $\Phi_\varepsilon = \sqrt{\varepsilon} w$ と $\Lambda_\varepsilon = X^\varepsilon$ となる．$\mathcal{C}_0$ 上の速度関数 I は (9.0.3) で定義されるものであり，そのとき $\widetilde{\mathcal{C}}$ 上の速度関数 J は (10.0.4) で定義されるものと一致する．このとき $I^{-1}([0,\infty)) = \mathcal{H}$ であり，$\psi\colon I^{-1}([0,\infty)) \to \widetilde{\mathcal{C}}$ として骨格 ODE(10.0.3) の解写像 $h \mapsto \varphi^h$ を取る．

$r \in (0,\infty)$ のとき，$\psi\colon I^{-1}([0,r]) = \overline{B}_{\mathcal{H}}(0, \sqrt{2r}) \to \widetilde{\mathcal{C}}$ が $\mathcal{C}_0$ の位相に関して連続であることを確認しておく．$\{h\} \cup \{h(n)\}_{n\in\mathbb{N}} \subset \overline{B}_{\mathcal{H}}(0,\sqrt{2r})$ が $\lim_{n\to\infty} \|h(n) - h\|_\infty = 0$ を満たすとする．このとき補題 9.1.2 により，$\{h(n)\}_{n\in\mathbb{N}}$ は弱位相に関して相対コンパクトな点列である．弱収束列は各点収束するため,[12] その任意の集積点は h に一致する．よって，弱位相に関して $\lim_{n\to\infty} h(n) = h$ である．このとき命題 10.1.2 により $\lim_{n\to\infty} \|\varphi^{h(n)} - \varphi^h\|_\infty = 0$ であるので，所望の連続性が示せた．

したがって，この状況において命題 1.7.1 の仮定中にある大偏差評価 (1.7.1) を証明すれば，Freidlin-Wentzell の大偏差原理が得られる．引き続き 4 つ組 $(\mathcal{C}_0, \mathcal{F}, \mathbb{P}; \{\mathcal{F}_t\}_{t\in[0,T]})$ 上で議論する．

大偏差評価を証明する前に，まず伊藤過程の脱出確率に関する補題を 1 つ示す．これは Brown 運動の脱出確率を評価した補題 9.4.2 の拡張である．

補題 10.3.1. $\mathbb{R}^{e\times d}$ 値の確率過程 $(A_t)_{t\in[0,T]}$ は $\{\mathcal{F}_t\}_{t\in[0,T]}$ に関して発展的可測かつ有界であるとし，正定数 C を $|A|$ の上界とする．また $0 \le S < T$ とする．このとき次が成立する．

$$\mathbb{P}\left(\sup_{S\le t\le T}\left|\int_S^t A_s dw_s\right| > R\right) \le 2e\exp\left(-\frac{R^2}{2eC^2(T-S)}\right), \quad R > 0.$$

証明. $S = 0$ の場合を示す．一般の S の場合も同様に証明できる．$v \in \mathbb{R}^e$ に対して

$$N_t^v := v \cdot \int_0^t A_s dw_s = \int_0^t v \cdot A_s dw_s = \int_0^t (A_s^\top v) \cdot dw_s$$

と定めると，その 2 次変分は $\langle N^v\rangle_t = \int_0^t |A_s^\top v|^2 ds \le C^2|v|^2 t$ となる．したがって明らかに N^v は Novikov 条件 $\mathbb{E}\left[\exp(\langle N^v\rangle_T/2)\right] < \infty$ を満たすので,

$$M_t^v := \exp\left(N_t^v - \frac{1}{2}\langle N^v\rangle_t\right) = \exp\left(v\cdot\int_0^t A_s dw_s - \frac{1}{2}\int_0^t |A_s^\top v|^2 ds\right)$$

---

[11] この方法は Azencott [48] により創始された

[12] 各 $s \in [0,T]$ に対して，$t \mapsto (s\wedge t)e_i$ と弱収束列の内積を考えればすぐにわかる．なお $\{e_i\}_{i=1}^d$ は $\mathbb{R}^d$ の標準基底である．

と定めると，$M^v = (M_t^v)_{t\in[0,T]}$ は真のマルチンゲールになる．非負値マルチンゲールなので Doob の不等式[13]

$$\mathbb{P}\left(\sup_{0\le t\le T} M_t^v \ge r\right) \le \frac{1}{r}\mathbb{E}\left[M_T^v\right] = \frac{1}{r}, \qquad r>0$$

が成り立つ．

以下 $|v|=1$ を仮定する．もし $\sup_{0\le t\le T} N_t^v > R/\sqrt{e}$ だとすると，任意の $\lambda>0$ に対して

$$\sup_{0\le t\le T} M_t^{\lambda v} \ge \sup_{0\le t\le T} \exp\left(\lambda N_t^v\right)\exp\left(-\frac{\lambda^2 C^2 T}{2}\right) \ge \exp\left(\frac{\lambda R}{\sqrt{e}} - \frac{\lambda^2 C^2 T}{2}\right)$$

が成り立つ．したがって

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le T} N_t^v > \frac{R}{\sqrt{e}}\right) &\le \mathbb{P}\left(\sup_{0\le t\le T} M_t^{\lambda v} \ge \exp\left(\frac{\lambda R}{\sqrt{e}} - \frac{\lambda^2 C^2 T}{2}\right)\right)\\ &\le \exp\left(-\frac{\lambda R}{\sqrt{e}} + \frac{\lambda^2 C^2 T}{2}\right)\end{aligned}$$

を得るが，特に $\lambda = R/(C^2T\sqrt{e})$ を代入して次の評価を得る．

$$\mathbb{P}\left(\sup_{0\le t\le T} N_t^v > \frac{R}{\sqrt{e}}\right) \le \exp\left(-\frac{R^2}{2eC^2T}\right).$$

さて $\{e_i\}_{1\le i\le e}$ を $\mathbb{R}^e$ の標準基底とし，$v=e_i$ として上の評価を使う．e 次元ベクトルのノルムが R より大ならば，その成分の絶対値のうち少なくとも一つは $R/\sqrt{e}$ より大きい．したがって

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le T}\left|\int_0^t A_s dw_s\right| > R\right) &\le \sum_{i=1}^{e} \mathbb{P}\left(\sup_{0\le t\le T} |N_t^{e_i}| > \frac{R}{\sqrt{e}}\right)\\ &\le \sum_{i=1}^{e}\left\{\mathbb{P}\left(\sup_{0\le t\le T} N_t^{e_i} > \frac{R}{\sqrt{e}}\right) + \mathbb{P}\left(\sup_{0\le t\le T} N_t^{-e_i} > \frac{R}{\sqrt{e}}\right)\right\}\\ &\le 2e\exp\left(-\frac{R^2}{2eC^2T}\right)\end{aligned}$$

と評価できるので，本補題の証明が終わる．□

SDE (10.0.1) の解 $X^\varepsilon = (X_t^\varepsilon)_{t\in[0,T]}$ に対する大偏差評価 (1.7.1) をいきなり証明するのはかなり難しいため，次の命題では Lipschitz 条件 **(LIP)** に加えて有界性条件 **(BND)** も仮定する．

[13] 確率過程の径数が連続時間である場合の Doob の不等式については [11, 第 3.3 節，定理 3.3.1], [13, 第 2.2 節，定理 2.4], [16, 第 1.6 節，定理 1.6.1] などを参照せよ．

命題 10.3.2. SDE(10.0.1) の係数 σ と b に対して，条件 **(LIP)** と **(BND)** を仮定する．このとき，任意の $R, \delta, r \in (0,\infty)$ に対して，ある $\rho \in (0,\infty)$ と $\varepsilon_0 \in (0,1]$ が存在して

$$\mathbb{P}\left(\|X^\varepsilon - \varphi^h\|_\infty > \delta,\ \|\sqrt{\varepsilon}w - h\|_\infty \le \rho\right) \le \exp(-R/\varepsilon), \quad h \in \overline{B}_{\mathcal{H}}(0,r),\ \varepsilon \in (0,\varepsilon_0] \tag{10.3.1}$$

を満たす．

証明. 本証明中では $r \in (0,\infty)$ は任意とし，$h \in \overline{B}_{\mathcal{H}}(0,r)$ とする．また部分区間 $[0,t]$ 上の一様ノルムを $\|\cdot\|_{\infty,[0,t]}$ と表す．

【第 1 段】まず X^ε と φ^h の差を評価しよう．簡単な計算で

$$\begin{aligned} X_t^\varepsilon - \varphi_t^h &= \int_0^t \sigma(X_s^\varepsilon)(\sqrt{\varepsilon}dw_s - h_s' ds) \\ &\quad + \int_0^t \{\sigma(X_s^\varepsilon) - \sigma(\varphi_s^h)\}h_s' ds + \int_0^t \{b(X_s^\varepsilon) - b(\varphi_s^h)\}ds \end{aligned}$$

を得る．両辺の $\mathbb{R}^e$ ノルムを取ると，**(LIP)** により

$$|X_t^\varepsilon - \varphi_t^h| \le \left|\int_0^t \sigma(X_s^\varepsilon)(\sqrt{\varepsilon}dw_s - h_s' ds)\right| + L\int_0^t |X_s^\varepsilon - \varphi_s^h|(|h_s'|+1)ds$$

を得る．時刻に関する上限を取ると

$$\begin{aligned} \|X^\varepsilon - \varphi^h\|_{\infty,[0,t]} &\le \left\|\int_0^\cdot \sigma(X_s^\varepsilon)(\sqrt{\varepsilon}dw_s - h_s' ds)\right\|_{\infty,[0,t]} \\ &\quad + L\int_0^t \|X^\varepsilon - \varphi^h\|_{\infty,[0,s]}(|h_s'|+1)ds \end{aligned}$$

を得るが，これを 2 乗して Schwarz の不等式と条件 $\|h\|_{\mathcal{H}} \le r$ を使うと

$$\begin{aligned} \|X^\varepsilon - \varphi^h\|_{\infty,[0,t]}^2 &\le 2\left\|\int_0^\cdot \sigma(X_s^\varepsilon)(\sqrt{\varepsilon}dw_s - h_s' ds)\right\|_\infty^2 \\ &\quad + 4L^2(r^2+T)\int_0^t \|X^\varepsilon - \varphi^h\|_{\infty,[0,s]}^2 ds \end{aligned}$$

を得る．Gronwall の不等式により

$$\|X^\varepsilon - \varphi^h\|_\infty \le \sqrt{2}\exp[2L^2(r^2+T)T]\left\|\int_0^\cdot \sigma(X_s^\varepsilon)(\sqrt{\varepsilon}dw_s - h_s' ds)\right\|_\infty$$

という評価を得る． したがって (10.3.1) を示すには，次の評価を示せば十分である．

任意の $R,\delta,r\in(0,\infty)$ に対して，ある $\rho\in(0,\infty)$ と $\varepsilon_0\in(0,1]$ が存在して，任意の $h\in\overline{B}_{\mathcal{H}}(0,r)$ と $\varepsilon\in(0,\varepsilon_0]$ に対して

$$\mathbb{P}\underbrace{\Big(\Big\|\int_0^\cdot \sigma(X_s^\varepsilon)(\sqrt{\varepsilon}dw_s-h_s'ds)\Big\|_\infty>\delta,\ \|\sqrt{\varepsilon}w-h\|_\infty\le\rho\Big)}_{=:\widetilde{A}_\varepsilon}\le\exp(-R/\varepsilon). \qquad (10.3.2)$$

以下では (10.3.2) の左辺に現れる事象を $\widetilde{A}_\varepsilon$ と書く．

【第 2 段】本段では Cameron-Martin の定理を使って，Wiener 空間 $\mathcal{C}_0$ 上でずらし変換を行う．我々の目的のためには，ずらして得られる確率過程 U^ε のほうが元の X^ε よりも計算しやすい．本段では $\varepsilon\in(0,1]$ と $h\in\overline{B}_{\mathcal{H}}(0,r)$ とする．以下では ε や h に依存する量がいくつか現れるが，ε や h への依存性は必ずしも明示しない．

まず新しい確率測度 $\widetilde{\mathbb{P}}$ を

$$d\widetilde{\mathbb{P}}:=M_T\,d\mathbb{P},\qquad M_t:=\exp\left(\frac{1}{\sqrt{\varepsilon}}\int_0^t h_s'\cdot dw_s-\frac{1}{2\varepsilon}\int_0^t|h_s'|^2ds\right)$$

と定め，$\widetilde{\mathbb{P}}$ に関する期待値を $\widetilde{\mathbb{E}}$ と表す．Cameron-Martin の定理により $\widetilde{w}_t:=w_t-h_t/\sqrt{\varepsilon}$ とおくと，$\widetilde{w}$ は $\widetilde{\mathbb{P}}$ の下では d 次元の標準 Brown 運動になる．また Schwarz の不等式により

$$\mathbb{P}\left(\widetilde{A}_\varepsilon\right)=\widetilde{\mathbb{E}}\left[\mathbf{1}_{\widetilde{A}_\varepsilon}M_T^{-1}\right]\le\widetilde{\mathbb{P}}\left(\widetilde{A}_\varepsilon\right)^{\frac12}\widetilde{\mathbb{E}}\left[M_T^{-2}\right]^{\frac12}$$

となるが，

$$\begin{aligned}\widetilde{\mathbb{E}}\left[M_T^{-2}\right]&=\mathbb{E}\left[M_T^{-1}\right]\\&=\mathbb{E}\left[\exp\left(-\frac{1}{\sqrt{\varepsilon}}\int_0^T h_s'\cdot dw_s+\frac{\|h\|_{\mathcal{H}}^2}{2\varepsilon}\right)\right]\\&=\mathbb{E}\left[\exp\left(-\frac{1}{\sqrt{\varepsilon}}\int_0^T h_s'\cdot dw_s-\frac{\|h\|_{\mathcal{H}}^2}{2\varepsilon}\right)\right]\exp\left(\frac{\|h\|_{\mathcal{H}}^2}{\varepsilon}\right)=\exp\left(\frac{\|h\|_{\mathcal{H}}^2}{\varepsilon}\right)\end{aligned}$$

であるので，結局

$$\mathbb{P}\left(\widetilde{A}_\varepsilon\right)\le\exp[r^2/(2\varepsilon)]\,\widetilde{\mathbb{P}}\left(\widetilde{A}_\varepsilon\right)^{\frac12} \qquad (10.3.3)$$

と評価される．

ここで新しい SDE を導入しよう．

$$U_t^{\varepsilon,h}=a+\int_0^t\{b(U_s^{\varepsilon,h})+\sigma(U_s^{\varepsilon,h})h_s'\}ds+\sqrt{\varepsilon}\int_0^t\sigma(U_s^{\varepsilon,h})dw_s. \qquad (10.3.4)$$

(LIP) により，この SDE は強い意味の一意解を持ち，それは Picard 流の逐次近似の極限列として得られる．特に $U^{\varepsilon,h}=(U_t^{\varepsilon,h})_{t\in[0,T]}$ は Brown 運動 $w=(w_t)_{t\in[0,T]}$ の可測写像

の像として得られる.[14] したがって特に SDE(10.3.4) において，Brown 運動 w を別の（同法則の）Brown 運動に取り替えても，$(U^{\varepsilon,h}, w) = (U_t^{\varepsilon,h}, w_t)_{t\in[0,T]}$ の法則は変わらない.[15] $\widetilde{\mathbb{P}}$ の下では X^ε は (10.3.4) において w を $\widetilde{w}$ に取り替えた SDE を満たしているため，

$$\widetilde{\mathbb{P}}\text{ の下での }(X^\varepsilon, \widetilde{w}) \overset{\text{Law}}{=} \mathbb{P}\text{ の下での }(U^{\varepsilon,h}, w)$$

という法則の一致を見る．これから直ちに

$$\widetilde{\mathbb{P}}\left(\widetilde{A}_\varepsilon\right) = \mathbb{P}\left(\left\|\int_0^\cdot \sigma(U_s^{\varepsilon,h})\sqrt{\varepsilon}dw_s\right\|_\infty > \delta,\ \|\sqrt{\varepsilon}w\|_\infty \le \rho\right) \tag{10.3.5}$$

を得る．(10.3.3) と (10.3.5) を合わせて考えると，(10.3.2) を（したがって本命題を）証明するためには次の評価を示せば十分であることがわかる．

任意の $R, \delta, r \in (0,\infty)$ に対して，ある $\rho \in (0,\infty)$ と $\varepsilon_0 \in (0,1]$ が存在して，任意の $h \in \overline{B}_{\mathcal{H}}(0,r)$ と $\varepsilon \in (0,\varepsilon_0]$ に対して

$$\mathbb{P}\underbrace{\left(\left\|\int_0^\cdot \sigma(U_s^{\varepsilon,h})\sqrt{\varepsilon}dw_s\right\|_\infty > \delta,\ \|\sqrt{\varepsilon}w\|_\infty \le \rho\right)}_{=:A_\varepsilon} \le \exp(-R/\varepsilon). \tag{10.3.6}$$

以下では (10.3.6) の左辺に現れる事象を A_ε と書く．

【第 3 段】本段では区間 $[0,T]$ を幅 $1/n$ の小区間に分割することにより，前段で導入した SDE の解 $U^{\varepsilon,h}$ を近似する単過程 $U^{\varepsilon,h,n}$ を導入する．$n \in \mathbb{N}$ とし，

$$U_t^{\varepsilon,h,n} := U_{k/n}^{\varepsilon,h}, \qquad t \in [k/n,\ (k+1)/n) \cap [0,T],\ 0 \le k \le \lfloor nT \rfloor$$

とおく．新しく導入するパラメーター $\delta_1 \in (0,\infty)$ に対して

$$\begin{aligned}
B_\varepsilon &= \left\{\left\|\int_0^\cdot \{\sigma(U_s^{\varepsilon,h}) - \sigma(U_s^{\varepsilon,h,n})\}\sqrt{\varepsilon}dw_s\right\|_\infty > \frac{\delta}{2}, \quad \|U^{\varepsilon,h} - U^{\varepsilon,h,n}\|_\infty \le \delta_1\right\}, \\
C_\varepsilon &= \left\{\left\|\int_0^\cdot \sigma(U_s^{\varepsilon,h,n})\sqrt{\varepsilon}dw_s\right\|_\infty > \frac{\delta}{2}, \quad \|\sqrt{\varepsilon}w\|_\infty \le \rho\right\}, \\
D_\varepsilon &= \left\{\|U^{\varepsilon,h} - U^{\varepsilon,h,n}\|_\infty > \delta_1\right\}
\end{aligned}$$

と 3 つの事象を定めると，

$$A_\varepsilon \subset B_\varepsilon \cup C_\varepsilon \cup D_\varepsilon \tag{10.3.7}$$

という包含関係が成り立つ．次の段落でこれを示す．

---

[14] この可測写像は伊藤写像と呼ばれる．仮にこれを $\mathcal{I}^{\varepsilon,h}\colon \mathcal{C}_0 \to \widetilde{\mathcal{C}}$ と書くと，$U^{\varepsilon,h}$ は $U^{\varepsilon,h} = \mathcal{I}^{\varepsilon,h}(w)$ という形をしている．

[15] $(U^{\varepsilon,h}, w)$ の法則は Wiener 測度 $\mathbb{P}$ の下での可測写像 $w \mapsto (\mathcal{I}^{\varepsilon,h}(w), w)$ の像測度として一意に決まる．

記号を簡単にするために

$$W := \int_0^{\cdot} \sigma(U_s^{\varepsilon,h})\sqrt{\varepsilon}dw_s,\ \widehat{W} := \int_0^{\cdot} \sigma(U_s^{\varepsilon,h,n})\sqrt{\varepsilon}dw_s,\ Y := \sqrt{\varepsilon}w,\ Z := U^{\varepsilon,h} - U^{\varepsilon,h,n}$$

と表す．すると

$$\begin{aligned}
A_\varepsilon &= \{\|W\|_\infty > \delta,\ \|Y\|_\infty \le \rho\} \\
&\subset \{\|W\|_\infty > \delta,\ \|Y\|_\infty \le \rho,\ \|Z\|_\infty \le \delta_1\} \cup \{\|Z\|_\infty > \delta_1\} \\
&\subset \{(\|W - \widehat{W}\|_\infty > \delta/2 \text{ または } \|\widehat{W}\|_\infty > \delta/2),\ \|Y\|_\infty \le \rho,\ \|Z\|_\infty \le \delta_1\} \\
&\qquad\qquad \cup \{\|Z\|_\infty > \delta_1\} \\
&\subset \{\|W - \widehat{W}\|_\infty > \delta/2,\ \|Y\|_\infty \le \rho,\ \|Z\|_\infty \le \delta_1\} \\
&\qquad\qquad \cup \{\|\widehat{W}\|_\infty > \delta/2,\ \|Y\|_\infty \le \rho,\ \|Z\|_\infty \le \delta_1\} \cup \{\|Z\|_\infty > \delta_1\} \\
&\subset \{\|W - \widehat{W}\|_\infty > \delta/2,\ \|Z\|_\infty \le \delta_1\} \cup \{\|\widehat{W}\|_\infty > \delta/2,\ \|Y\|_\infty \le \rho\} \cup \{\|Z\|_\infty > \delta_1\} \\
&= B_\varepsilon \cup C_\varepsilon \cup D_\varepsilon
\end{aligned}$$

となる.（ここで，条件文の間にあるカンマは「かつ」を意味する.）これで (10.3.7) が確認できた.

【第 4 段】本段では事象 B_ε と C_ε の確率を評価する．まず B_ε を調べる．**(LIP)** により

$$|U_s^{\varepsilon,h} - U_s^{\varepsilon,h,n}| \le \delta_1 \quad \Longrightarrow \quad |\sigma(U_s^{\varepsilon,h}) - \sigma(U_s^{\varepsilon,h,n})|\sqrt{\varepsilon} \le L\delta_1\sqrt{\varepsilon}$$

となることは明らかである．停止時刻を $\tau := \inf\{s \in [0,T] \mid |U_s^{\varepsilon,h} - U_s^{\varepsilon,h,n}| > \delta_1\} \wedge T$ とおくと，

$$\{\tau = T\} = \{\|U^{\varepsilon,h} - U^{\varepsilon,h,n}\|_\infty \le \delta_1\}$$

であるから

$$\begin{aligned}
B_\varepsilon &= \left\{\left\|\int_0^{\cdot} \{\sigma(U_s^{\varepsilon,h}) - \sigma(U_s^{\varepsilon,h,n})\}\sqrt{\varepsilon}dw_s\right\|_\infty > \frac{\delta}{2},\quad \tau = T\right\} \\
&= \left\{\left\|\int_0^{\cdot} \mathbf{1}_{\{s\le\tau\}}\{\sigma(U_s^{\varepsilon,h}) - \sigma(U_s^{\varepsilon,h,n})\}\sqrt{\varepsilon}dw_s\right\|_\infty > \frac{\delta}{2},\quad \tau = T\right\} \\
&\subset \left\{\left\|\int_0^{\cdot} \mathbf{1}_{\{s\le\tau\}}\{\sigma(U_s^{\varepsilon,h}) - \sigma(U_s^{\varepsilon,h,n})\}\sqrt{\varepsilon}dw_s\right\|_\infty > \frac{\delta}{2}\right\}
\end{aligned}$$

とわかる．最右辺にある確率積分の被積分関数のノルムは $L\delta_1\sqrt{\varepsilon}$ で評価できることに注意せよ．補題 10.3.1 を用いると

$$\mathbb{P}\left(B_\varepsilon\right) \le 2e\exp\left(-\frac{(\delta/2)^2}{2e(L\delta_1\sqrt{\varepsilon})^2T}\right) \le 2e\exp\left(-\frac{\delta^2}{8eL^2T\delta_1^2\varepsilon}\right)$$

を得る．よって

$$0 < \delta_1 < \frac{\delta}{L\sqrt{8eTR}} \implies \mathbb{P}\,(B_\varepsilon) \le 2e\exp(-R/\varepsilon). \tag{10.3.8}$$

次は C_ε を調べる．条件 $\|\sqrt{\varepsilon}w\|_\infty \le \rho$ の下では

$$\begin{aligned}\Bigl|\int_0^t \sigma(U_s^{\varepsilon,h,n})\sqrt{\varepsilon}dw_s\Bigr| &= \Bigl|\sum_{k=0}^{\lfloor nT\rfloor} \sigma(U_{k/n}^{\varepsilon,h})\sqrt{\varepsilon}(w_{\{(k+1)/n\}\wedge t} - w_{(k/n)\wedge t})\Bigr| \\ &\le 2\|\sqrt{\varepsilon}w\|_\infty \sum_{k=0}^{\lfloor nT\rfloor} |\sigma(U_{k/n}^{\varepsilon,h})| \\ &\le 2\|\sigma\|_\infty \rho(T+1)n\end{aligned}$$

が成り立つ．**(BND)** により $\|\sigma\|_\infty < \infty$ であることに注意せよ．よって，$\|\sigma\|_\infty > 0$ であれば

$$\rho < \frac{\delta}{4\|\sigma\|_\infty(T+1)n} \implies C_\varepsilon = \emptyset \implies \mathbb{P}\,(C_\varepsilon) = 0 \tag{10.3.9}$$

を得る.（なお $\|\sigma\|_\infty = 0$ のときは，明らかに $\mathbb{P}\,(C_\varepsilon) = 0$ である.）

【第 5 段】本段では **(BND)** を用いて，事象 D_ε の確率を評価する．**(BND)** を仮定したため，$\|\sigma\|_\infty \vee \|b\|_\infty < \infty$ であり，補題 10.3.1 を用いることができる．本段では記号を簡単にするため，$n \in \mathbb{N}$ と $0 \le k \le \lfloor nT\rfloor$ に対して $A(n,k) = [k/n,\,(k+1)/n) \cap [0,T]$ とおく.

以下しばらく $A(n,k)$ 上で計算する．$U_t^{\varepsilon,h,n}$ の定義から直ちに

$$|U_t^{\varepsilon,h} - U_t^{\varepsilon,h,n}| \le \int_{k/n}^t \{|b(U_s^{\varepsilon,h})| + |\sigma(U_s^{\varepsilon,h})|\,|h_s'|\}ds + \left|\sqrt{\varepsilon}\int_{k/n}^t \sigma(U_s^{\varepsilon,h})dw_s\right|$$

と評価できるので，事象の確率の評価としては

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in A(n,k)} |U_t^{\varepsilon,h} - U_t^{\varepsilon,h,n}| > \delta_1\right) &\le \mathbb{P}\left(\sup_{t\in A(n,k)} \int_{k/n}^t \{|b(U_s^{\varepsilon,h})| + |\sigma(U_s^{\varepsilon,h})|\,|h_s'|\}ds > \frac{\delta_1}{2}\right) \\ &\quad + \mathbb{P}\left(\sup_{t\in A(n,k)} \left|\sqrt{\varepsilon}\int_{k/n}^t \sigma(U_s^{\varepsilon,h})dw_s\right| > \frac{\delta_1}{2}\right)\end{aligned} \tag{10.3.10}$$

を得る．Schwarz の不等式を用いた簡単な計算で

$$\sup_{t\in A(n,k)} \int_{k/n}^t \{|b(U_s^{\varepsilon,h})| + |\sigma(U_s^{\varepsilon,h})|\,|h_s'|\}ds \le \frac{\|b\|_\infty}{n} + \frac{\|\sigma\|_\infty\|h\|_{\mathcal{H}}}{\sqrt{n}} \le \frac{\|b\|_\infty + \|\sigma\|_\infty r}{\sqrt{n}}$$

とわかるので，$\sqrt{n} > 2(\|b\|_\infty + \|\sigma\|_\infty r)/\delta_1$ のときは (10.3.10) の右辺第１項は消える．補題 10.3.1 を用いると，右辺第２項は

$$2e\exp\left(-\frac{(\delta_1/2)^2}{2e(\sqrt{\varepsilon}\|\sigma\|_\infty)^2 n^{-1}}\right) \le 2e\exp\left(-\frac{n\delta_1^2}{8e\|\sigma\|_\infty^2\varepsilon}\right)$$

と評価できる．(この評価は $\|\sigma\|_\infty = 0$ の場合も成り立つ．)

したがって，$\sqrt{n} > 2(\|b\|_\infty + \|\sigma\|_\infty r)/\delta_1$ のとき

$$\begin{aligned}\mathbb{P}(D_\varepsilon) &\le \mathbb{P}\left(\max_{0\le k\le\lfloor nT\rfloor}\sup_{t\in A(n,k)}|U_t^{\varepsilon,h} - U_t^{\varepsilon,h,n}| > \delta_1\right)\\ &\le 2e(T+1)n\exp\left(-\frac{n\delta_1^2}{8e\|\sigma\|_\infty^2\varepsilon}\right) \qquad (10.3.11)\end{aligned}$$

と評価できる．

【第 6 段】最後にパラメーターをどう決めるのかを明らかにして議論をまとめる．まず $r, \delta \in (0,\infty)$ は任意だとする．さらに十分大きな $R \in (0,\infty)$ も任意に与えられたとする．

まず δ_1 を (10.3.8) にあるように選ぶと，$\varepsilon \in (0,1]$ のとき $\mathbb{P}(B_\varepsilon) \le 2e\exp(-R/\varepsilon)$ が成り立つ．δ_1 は δ と R から決まる．次は n を決める．(10.3.11) の右辺が $2e(T+1)\exp(-R/\varepsilon)$ 以下であることは

$$n\exp\left[\varepsilon^{-1}\left(-\frac{n\delta_1^2}{8e\|\sigma\|_\infty^2} + R\right)\right] \le 1, \qquad (10.3.12)$$

と同値である．$n > 8e\|\sigma\|_\infty^2 R/\delta_1^2$ のときは，(10.3.12) において指数関数の中で ε^{-1} の係数は負なので $\varepsilon = 1$ の場合が一番大きくなる．さらに

$$\lim_{n\to\infty} n\exp\left(-\frac{n\delta_1^2}{8e\|\sigma\|_\infty^2} + R\right) = 0$$

は明らかであるから，n を十分大きく選べば (10.3.12) は任意の $\varepsilon \in (0,1]$ に対して成立する．以上をまとめると，既に決まった R, r, δ, δ_1 に対して十分大きな n を１つ選ぶと，$\mathbb{P}(D_\varepsilon) \le 2e(T+1)\exp(-R/\varepsilon)$ が任意の $\varepsilon \in (0,1]$ に対して成り立つ．最後に (10.3.9) にあるように ρ を選ぶと $\mathbb{P}(C_\varepsilon) = 0$ である．ρ は δ と n から決まることに注意せよ．

したがって，任意に与えられた r, δ, R に対して上の要領で ρ を選ぶと

$$\mathbb{P}(A_\varepsilon) \le \mathbb{P}(B_\varepsilon) + \mathbb{P}(C_\varepsilon) + \mathbb{P}(D_\varepsilon) \le 2e(T+2)\exp(-R/\varepsilon), \quad \varepsilon \in (0,1]$$

となる．最後に $\varepsilon_0 \in (0,1]$ を $2e(T+2)\exp(-1/\varepsilon_0) \le 1$ を満たすように取ると，

$$\mathbb{P}(A_\varepsilon) \le \exp[-(R-1)/\varepsilon], \quad \varepsilon \in (0,\varepsilon_0]$$

となる．R は任意の十分大きな正数なので，$R-1$ もそうである．これで (10.3.6) が証明できた．以上で命題 10.3.2 の証明が終わった． □

係数行列に対して標準的な切り落としの議論を適用することにより，命題 10.3.2 で得られた大偏差評価から有界性条件 **(BND)** を外す．下記の系 10.3.3 が本節の主結果である．

系 10.3.3. SDE(10.0.1) の係数 σ と b に対して条件 **(LIP)** を仮定する．このとき，任意の $R, \delta, r \in (0,\infty)$ に対して，ある $\rho \in (0,\infty)$ と $\varepsilon_0 \in (0,1]$ が存在して

$$\mathbb{P}\left(\|X^\varepsilon - \varphi^h\|_\infty > \delta,\ \|\sqrt{\varepsilon}w - h\|_\infty \le \rho\right) \le \exp(-R/\varepsilon), \quad h \in \overline{B}_{\mathcal{H}}(0,r),\ \varepsilon \in (0,\varepsilon_0] \tag{10.3.13}$$

を満たす.

証明. $r \in (0,\infty)$ を任意とする．命題 10.1.1 により，

$$\sup\{\|\varphi^h\|_\infty \mid h \in \mathcal{H},\ h \in \overline{B}_{\mathcal{H}}(0,r)\} \le m-1$$

となる $m = m_r \in \mathbb{N}$ が存在する．この m に対して，$\chi_m\colon \mathbb{R}^e \to [0,1]$ を $|x| \le m$ のときは $\chi_m(x) = 1$，$|x| \ge m+1$ のときは $\chi_m(x) = 0$ を満たす Lipschitz 連続関数とする．これを用いて係数を $\sigma_m := \chi_m \cdot \sigma$ および $b_m := \chi_m \cdot b$ と切り落とすと，これらは明らかに **(LIP)** も **(BND)** も満たす．新しい係数 σ_m, b_m に対応する SDE (10.0.1) と骨格 ODE (10.0.3) の解をそれぞれ $\widehat{X}^\varepsilon$ と $\widehat{\varphi}^h$ 表す．作り方から明らかに $h \in \overline{B}_{\mathcal{H}}(0,r)$ であれば $\widehat{\varphi}^h = \varphi^h$ である．

よく知られているように，半径 m の球から脱出するまでは X^ε と $\widehat{X}^\varepsilon$ は一致する．正確に書けば，$\tau := \{t \in [0,T] \mid |X_t^\varepsilon| \ge m\} \wedge T$ および $\widehat{\tau} := \{t \in [0,T] \mid |\widehat{X}_t^\varepsilon| \ge m\} \wedge T$ と停止時刻を定めると，$\mathbb{P}$-a.s. に $\tau = \widehat{\tau}$ かつ $(X^\varepsilon_{t\wedge\tau})_{t\in[0,T]} = (\widehat{X}^\varepsilon_{t\wedge\tau})_{t\in[0,T]}$ である.（これとほぼ同じ事実に対する詳しい説明が補題 10.2.3 の証明中にあるので，ここでは説明を省略する.）

$\delta \in (0,1)$ および $h \in \overline{B}_{\mathcal{H}}(0,r)$ を任意とする．このとき $\|X^\varepsilon - \varphi^h\|_\infty \le \delta$ であれば X^ε は半径 m の球から出ないので $\tau = T$ である.（X^ε を $\widehat{X}^\varepsilon$ に取り替えた場合も同様の事実が成り立つ.）したがって，この場合 X^ε と $\widehat{X}^\varepsilon$ は一致するので，

$$\{\|X^\varepsilon - \varphi^h\|_\infty \le \delta\} = \{\|\widehat{X}^\varepsilon - \widehat{\varphi}^h\|_\infty \le \delta\}, \quad \mathbb{P}\text{-a.s.}$$

である.[16] この補集合を取り，さらに $\{\|\sqrt{\varepsilon}w - h\|_\infty \le \rho\}$ との交わりを作ると

$$\{\|X^\varepsilon - \varphi^h\|_\infty > \delta, \|\sqrt{\varepsilon}w - h\|_\infty \le \rho\} = \{\|\widehat{X}^\varepsilon - \widehat{\varphi}^h\|_\infty > \delta, \|\sqrt{\varepsilon}w - h\|_\infty \le \rho\},\ \mathbb{P}\text{-a.s.}$$

右辺の事象の確率に対しては命題 10.3.2 が適用できる．以上で系の証明が終わった．□

注意 10.3.4. 上の系 10.3.3 を命題 1.7.1 と組み合わせると，条件 **(LIP)** の下での Freidlin-Wentzell の大偏差原理（定理 10.0.1）が直ちに得られる．

[16] 2 つの事象 A, B が $\mathbb{P}$-a.s. に等しいとは $\mathbf{1}_A = \mathbf{1}_B$, $\mathbb{P}$-a.s. という意味であるが，これは $\mathbb{P}(A \triangle B) = 0$ と同値である．通常どおり $A \triangle B := (A \setminus B) \cup (B \setminus A)$ は 2 つの事象の対称差である．

10.4　弱収束法による定理 10.0.1 の証明

本節では弱収束法を用いて Freidlin-Wentzell の大偏差原理（定理 10.0.1）を証明する.[17] この手法は比較的新しいが，すでに広く普及しており，現在ではこの手法を用いた論文が数多く生産されている.

弱収束法では大偏差原理を Laplace 原理の形で証明する．既に命題 10.1.3 で示したように速度関数 J は良いので，Varadhan の補題（定理 1.4.2）とその逆（定理 1.4.4）により，大偏差原理（定理 10.0.1）と Laplace 原理（命題 10.4.1）は同値である.

命題 10.4.1. 条件 **(LIP)** を仮定する．$\varepsilon \searrow 0$ とするとき，$\{X^\varepsilon\}_{\varepsilon\in(0,1]}$ は $\widetilde{\mathcal{C}}$ 上で良い速度関数 J に対して速度 $1/\varepsilon$ で Laplace 原理を満たす．すなわち，任意の有界連続関数 $F\colon \widetilde{\mathcal{C}} \to \mathbb{R}$ に対して，次が成立する.

$$\lim_{\varepsilon\searrow 0} \varepsilon \log \mathbb{E}\left[\exp\left(-\frac{1}{\varepsilon}F(X^\varepsilon)\right)\right] = -\inf_{\xi\in\widetilde{\mathcal{C}}}\{F(\xi)+J(\xi)\}. \tag{10.4.1}$$

本節の残りでこの命題を証明する．その中で鍵になるのは指数型 Wiener 汎関数の変分表示公式である．それを述べるために，まず記号をいくつか導入する.（なお本節の残りでは常に **(LIP)** を仮定し，Cameron-Martin 空間の閉球 $\overline{B}_{\mathcal{H}}(0,R)$ には弱位相を入れる.）

以下では本章の冒頭で導入した 4 つ組 $(\mathcal{C}_0, \mathcal{F}, \mathbb{P}; \{\mathcal{F}_t\}_{t\in[0,T]})$ 上で作業する．ここで $\mathcal{F}$ は Borel 加法族 $\mathcal{B}(\mathcal{C}_0)$ の $\mathbb{P}$ 完備化である．当然ながら，$\mathbb{P}$ の下では座標過程 $(w_t)_{0\le t\le T}$ は d 次元標準 Brown 運動である．この 4 つ組上で定義された発展的可測な $\mathbb{R}^d$ 値確率過程 $v=(v_t)_{t\in[0,T]}$ で，条件

$$\mathbb{E}\left[\int_0^T |v_t|^2 dt\right] < \infty$$

を満たすのものの全体を $\mathscr{A}$ と表す．また各 $R\in(0,\infty)$ に対して，

$$\mathscr{A}_R := \left\{ v\in\mathscr{A} \;\middle|\; \int_0^T |v_t|^2 dt \le R^2,\ \mathbb{P}\text{-a.s.}\right\}, \qquad \mathscr{A}_{\mathrm{b}} := \bigcup_{R>0} \mathscr{A}_R$$

と定める．$\mathscr{A}, \mathscr{A}_R, \mathscr{A}_{\mathrm{b}}$ に属する過程の不定積分として得られる過程全体がなす集合をそれぞれ $\mathscr{A}', \mathscr{A}'_R, \mathscr{A}'_{\mathrm{b}}$ と名付ける．正確には

$$\mathscr{A}' := \left\{ \int_0^{\cdot} v_s ds = \left(\int_0^t v_s ds\right)_{t\in[0,T]} \;\middle|\; v\in\mathscr{A} \right\}$$

と定め，$\mathscr{A}'_R$ と $\mathscr{A}'_{\mathrm{b}}$ も同じ要領で定める．このとき $\mathscr{A}'_{\mathrm{b}} = \bigcup_{R>0}\mathscr{A}'_R$ である．また $\mathscr{A}'_R$ の元は $\overline{B}_{\mathcal{H}}(0,R)$ 値の確率変数である．補題 9.1.2 により $\overline{B}_{\mathcal{H}}(0,R)$ はコンパクト距離空間

[17] 弱収束法をこの種の大偏差原理に用いるのは [49] が初出である．この手法に関しては [29, 第 3.2 節] に手際よくまとめられている.

なので，$\mathscr{A}'_R$ を $\overline{B}_{\mathcal{H}}(0,R)$ 値確率変数の族と見なした場合，$\mathscr{A}'_R$（およびその任意の部分集合）は自動的に緊密である.[18]

それでは指数型 Wiener 汎関数の変分表示公式を正確に述べる.（なお注意 C.0.3 で詳しく述べるが，$\eta\in\mathscr{A}'$ のとき，$G=\widetilde{G}$, $\mathbb{P}$-a.s であれば $G(\cdot+\eta)=\widetilde{G}(\cdot+\eta)$, $\mathbb{P}$-a.s である.）

命題 10.4.2. $G\colon \mathcal{C}_0\to\mathbb{R}$ を有界な $\mathcal{F}$ 可測関数とする．このとき

$$-\log\mathbb{E}\left[\exp(-G)\right]=\inf_{\eta\in\mathscr{A}'}\mathbb{E}\left[G(\,\cdot\,+\eta)+\frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right]$$

が成立する.

証明. 見かけ上少しだけ異なるが，この命題は付録において証明する命題 C.0.2 （および注意 C.0.3）と事実上同じである.（$v\mapsto\int_0^{\cdot}v_s ds$ が $L^2([0,T],\mathbb{R}^d)$ から $\mathcal{H}$ へのユニタリ同型であることに注意すれば直ちに同値性がわかる.） □

補題 10.4.3. 任意の $\delta\in(0,1)$ と $K\in(0,\infty)$ に対してある $R\in(0,\infty)$ が存在して，以下の条件を満たす．任意の $\varepsilon\in(0,1]$ と $\|G\|_\infty\le K$ となる有界な $\mathcal{F}$ 可測関数 $G\colon\mathcal{C}_0\to\mathbb{R}$ に対して

$$-\varepsilon\log\mathbb{E}\left[\exp(-G/\varepsilon)\right]\ge\inf_{\eta\in\mathscr{A}'_R}\mathbb{E}\left[G(\,\cdot\,+\eta/\sqrt{\varepsilon})+\frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right]-\delta \tag{10.4.2}$$

が成り立つ．特に命題 10.4.2 中の等式において，inf を取る範囲を $\mathscr{A}'$ から $\mathscr{A}'_{\mathrm{b}}$ に取り替えても右辺の値は変わらない．すなわち次が成立する.

$$-\log\mathbb{E}\left[\exp(-G)\right]=\inf_{\eta\in\mathscr{A}'_{\mathrm{b}}}\mathbb{E}\left[G(\,\cdot\,+\eta)+\frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right]. \tag{10.4.3}$$

証明. (10.4.3) は命題 10.4.2 と (10.4.2) から直ちに従う．以下では (10.4.2) を示す.

まず命題 10.4.2 により，

$$-\varepsilon\log\mathbb{E}\left[\exp(-G/\varepsilon)\right]=\inf_{\eta\in\mathscr{A}'}\mathbb{E}\left[G(\,\cdot\,+\eta/\sqrt{\varepsilon})+\frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right]$$

となる．任意の $\kappa\in(0,1]$ に対して，ある $\widetilde{\eta}^{\varepsilon,\kappa}\in\mathscr{A}'$ が存在して，

$$\|G\|_\infty\ge\inf_{\eta\in\mathscr{A}'}\mathbb{E}\left[G(\,\cdot\,+\eta/\sqrt{\varepsilon})+\frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right]\ge\mathbb{E}\left[G(\,\cdot\,+\widetilde{\eta}^{\varepsilon,\kappa}/\sqrt{\varepsilon})+\frac{1}{2}\|\widetilde{\eta}^{\varepsilon,\kappa}\|_{\mathcal{H}}^2\right]-\kappa$$

を満たす．これから直ちに次の評価を得る.

$$\infty>2(2\|G\|_\infty+1)\ge\sup_{\varepsilon\in(0,1]}\mathbb{E}\left[\|\widetilde{\eta}^{\varepsilon,\kappa}\|_{\mathcal{H}}^2\right]. \tag{10.4.4}$$

[18] より正確には「$\mathscr{A}'_R$ の元から誘導される法則の全体が緊密」と言うべきであろう.

次は任意の $R \in (0,\infty)$ に対して停止時刻を

$$\tau_R^{\varepsilon,\kappa} := \inf\left\{t \in [0,T] \,\middle|\, \int_0^t |(\widetilde{\eta}^{\varepsilon,\kappa})'_s|^2 ds \geq R^2\right\} \wedge T$$

と定め，$\eta^{\varepsilon,\kappa} = \eta^{\varepsilon,\kappa,R}$ を $\eta_t^{\varepsilon,\kappa} := \widetilde{\eta}^{\varepsilon,\kappa}_{t\wedge\tau_R^{\varepsilon,\kappa}}$ と定める．Lebesgue 測度に関してほとんど全ての t に対して $(\eta^{\varepsilon,\kappa})'_t = (\widetilde{\eta}^{\varepsilon,\kappa})'_t \mathbf{1}_{[0,\tau_R^{\varepsilon,\kappa}]}(t)$ であるから，$\eta^{\varepsilon,\kappa} \in \mathscr{A}'_R$ かつ確率 1 で $\|\eta^{\varepsilon,\kappa}\|_{\mathcal{H}} \leq \|\widetilde{\eta}^{\varepsilon,\kappa}\|_{\mathcal{H}}$ となる．このとき明らかに

$$\begin{aligned}\mathbb{E}\left[G(\,\cdot\,+\widetilde{\eta}^{\varepsilon,\kappa}/\sqrt{\varepsilon}) + \frac{1}{2}\|\widetilde{\eta}^{\varepsilon,\kappa}\|_{\mathcal{H}}^2\right] &\geq \mathbb{E}\left[G(\,\cdot\,+\eta^{\varepsilon,\kappa}/\sqrt{\varepsilon}) + \frac{1}{2}\|\eta^{\varepsilon,\kappa}\|_{\mathcal{H}}^2\right] \\ &\quad + \mathbb{E}\left[G(\,\cdot\,+\widetilde{\eta}^{\varepsilon,\kappa}/\sqrt{\varepsilon}) - G(\,\cdot\,+\eta^{\varepsilon,\kappa}/\sqrt{\varepsilon})\right]\end{aligned}$$

である．$\{\tau_R^{\varepsilon,\kappa} < T\} = \{\|\widetilde{\eta}^{\varepsilon,\kappa}\|_{\mathcal{H}}^2 > R^2\}$ に注意しつつ，(10.4.4) と Chebyshev の不等式を用いると，

$$\begin{aligned}\mathbb{E}\left[|G(\,\cdot\,+\widetilde{\eta}^{\varepsilon,\kappa}/\sqrt{\varepsilon}) - G(\,\cdot\,+\eta^{\varepsilon,\kappa}/\sqrt{\varepsilon})|\right] &\leq 2\|G\|_\infty \mathbb{P}\left(\tau_R^{\varepsilon,\kappa} < T\right) \\ &\leq \frac{4\|G\|_\infty(2\|G\|_\infty + 1)}{R^2}\end{aligned}$$

が従う．$R^2 = 4K(2K+1)/\delta$ と取れば，この右辺は δ 以下である．以上をまとめて

$$\begin{aligned}\inf_{\eta\in\mathscr{A}'}\mathbb{E}\left[G(\,\cdot\,+\eta/\sqrt{\varepsilon}) + \frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right] &\geq \mathbb{E}\left[G(\,\cdot\,+\eta^{\varepsilon,\kappa}/\sqrt{\varepsilon}) + \frac{1}{2}\|\eta^{\varepsilon,\kappa}\|_{\mathcal{H}}^2\right] - \kappa - \delta \\ &\geq \inf_{\eta\in\mathscr{A}'_R}\mathbb{E}\left[G(\,\cdot\,+\eta/\sqrt{\varepsilon}) + \frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right] - \kappa - \delta\end{aligned}$$

となる．κ は任意なので，これで (10.4.2) が証明できた． □

ここで $\eta \in \mathscr{A}'_{\mathrm{b}}$ に対して，SDE (10.0.1) において，w を $w + \eta(w)/\sqrt{\varepsilon}$ でおき換えて得られる SDE を導入する．

$$X_t^{\varepsilon,\eta} = a + \int_0^t b(X_s^{\varepsilon,\eta})ds + \sqrt{\varepsilon}\int_0^t \sigma(X_s^{\varepsilon,\eta})dw_s + \int_0^t \sigma(X_s^{\varepsilon,\eta})\eta'_s ds. \tag{10.4.5}$$

なお η はランダムな $\mathcal{H}$ の元なので，右辺の最後の項は各見本 w ごとに 1 次元の Lebesgue 積分 $\int_0^t \sigma(X^{\varepsilon,\eta}(w)_s)\eta(w)'_s ds$ を行なったものである．(ちなみに **(LIP)** の下でこの SDE の解の存在と一意性が成り立つことは，$\eta \equiv 0$ の場合と同様の議論で証明できるので省略する．) 当然ながら，$X^\varepsilon(w+\eta(w)/\sqrt{\varepsilon}) = X^{\varepsilon,\eta}(w)$ が $\mathbb{P}$ に関してほとんど全ての w に対して成り立つ．したがって補題 10.4.3 中の (10.4.3) により，任意の有界連続関数 $F\colon \widetilde{\mathcal{C}} \to \mathbb{R}$ に対して，

$$-\varepsilon\log\mathbb{E}\left[\exp\left(-\frac{F(X^\varepsilon)}{\varepsilon}\right)\right] = \varepsilon\inf_{\eta\in\mathscr{A}'_{\mathrm{b}}}\mathbb{E}\left[F(X^{\varepsilon,\eta})/\varepsilon + \frac{1}{2}\|\eta/\sqrt{\varepsilon}\|_{\mathcal{H}}^2\right]$$

$$= \inf_{\eta \in \mathscr{A}'_{\mathrm{b}}} \mathbb{E}\left[F(X^{\varepsilon,\eta}) + \frac{1}{2}\|\eta\|^2_{\mathcal{H}}\right] \tag{10.4.6}$$

となる．この式を見ると，問題は SDE (10.4.5) の解 $X^{\varepsilon,\eta}$ の研究に帰着されそうである．そこでまず $X^{\varepsilon,\eta}$ の性質に関する補題を 2 つ証明しよう．

補題 10.4.4. 条件 **(LIP)** を仮定する．$R \in (0,\infty)$ とし，$\{\eta(\varepsilon)\}_{\varepsilon\in(0,1]} \subset \mathscr{A}'_R$ とする．このとき $\widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R)$ 値の確率変数の族 $\{(X^{\varepsilon,\eta(\varepsilon)}, \eta(\varepsilon))\}_{\varepsilon\in(0,1]}$ は緊密である．

証明. 本証明中では $c > 0$ は $s, t, \varepsilon, R, \{\eta(\varepsilon)\}, a$ に依存しない定数とし，行ごとにその値は変わってよいとする．まず $\{X^{\varepsilon,\eta(\varepsilon)}\}_{\varepsilon\in(0,1]}$ が $\widetilde{\mathcal{C}}$ において緊密であることを示す．

まず Schwarz の不等式と σ の線形増大性により

$$\begin{aligned}
\mathbb{E}\left[\left|\int_0^t \sigma(X_s^{\varepsilon,\eta(\varepsilon)})\eta(\varepsilon)'_s ds\right|^4\right] &\le c\mathbb{E}\left[\left\{\int_0^t (1+|X_s^{\varepsilon,\eta(\varepsilon)}|)^2 ds \int_0^t |\eta(\varepsilon)'_s|^2 ds\right\}^2\right] \\
&\le cR^4 \mathbb{E}\left[\int_0^t (1+|X_s^{\varepsilon,\eta(\varepsilon)}|)^4 ds\right] \\
&\le cR^4 \left(1 + \int_0^t \mathbb{E}\left[|X_s^{\varepsilon,\eta(\varepsilon)}|^4\right] ds\right)
\end{aligned}$$

となる．b の線形増大性を使うとより簡単な議論で

$$\mathbb{E}\left[\left|\int_0^t b(X_s^{\varepsilon,\eta(\varepsilon)}) ds\right|^4\right] \le c\left(1 + \int_0^t \mathbb{E}\left[|X_s^{\varepsilon,\eta(\varepsilon)}|^4\right] ds\right)$$

となることもわかる．Burkholder の不等式と σ の線形増大性により，(10.4.5) の確率積分項も以下のように同様の評価を満たす．

$$\mathbb{E}\left[\left|\int_0^t \sigma(X_s^{\varepsilon,\eta(\varepsilon)}) dw_s\right|^4\right] \le c\mathbb{E}\left[\left\{\int_0^t |\sigma(X_s^{\varepsilon,\eta(\varepsilon)})|^2 ds\right\}^2\right] \le c\left(1 + \int_0^t \mathbb{E}\left[|X_s^{\varepsilon,\eta(\varepsilon)}|^4\right] ds\right).$$

SDE(10.4.5) の両辺を 4 乗して，上の三つの不等式を用いると

$$\mathbb{E}\left[|X_t^{\varepsilon,\eta(\varepsilon)}|^4\right] \le c(|a|^4 + R^4 + 1) + c(R^4+1)\int_0^t \mathbb{E}\left[|X_s^{\varepsilon,\eta(\varepsilon)}|^4\right] ds, \qquad t \in [0,T]$$

という評価を得る．Gronwall の不等式により，

$$\sup_{0\le t\le T} \mathbb{E}\left[|X_t^{\varepsilon,\eta(\varepsilon)}|^4\right] \le c(|a|^4 + R^4 + 1)\exp(c(R^4+1)T) \tag{10.4.7}$$

を得る．(以下ではこの右辺の値を $C = C_{R,a}$ と表す.)

ほぼ同様の議論により，次は $X^{\varepsilon,\eta(\varepsilon)}$ の差分を評価する．$0 \le s \le t \le T$ のとき，

$$\begin{aligned}\mathbb{E}\left[|X_t^{\varepsilon,\eta(\varepsilon)} - X_s^{\varepsilon,\eta(\varepsilon)}|^4\right] &\le c\mathbb{E}\left[\left|\int_s^t b(X_u^{\varepsilon,\eta(\varepsilon)})du\right|^4\right] \\ &\quad + c\mathbb{E}\left[\left|\int_s^t \sigma(X_u^{\varepsilon,\eta(\varepsilon)})dw_u\right|^4\right] + c\mathbb{E}\left[\left|\int_s^t \sigma(X_u^{\varepsilon,\eta(\varepsilon)})\eta(\varepsilon)'_u du\right|^4\right] \\ &\le c(1+R^4)(t-s)\int_s^t \mathbb{E}\left[1+|X_u^{\varepsilon,\eta(\varepsilon)}|^4\right]du \\ &\le c(1+R^4)(1+C)(t-s)^2\end{aligned}$$

と評価できる．最後に (10.4.7) を使った．右辺の $(t-s)^2$ の前の正定数は s,t,ε に依存しないので，$\{X^{\varepsilon,\eta(\varepsilon)}\}_{\varepsilon\in(0,1]}$ は Kolmogorov の緊密性条件を満たしており，$\widetilde{\mathcal{C}}$ において緊密であることがわかる.[19]

緊密性の定義により，任意の $\delta > 0$ に対してあるコンパクト部分集合 $K_\delta \subset \widetilde{\mathcal{C}}$ が存在して $\mathbb{P}\left(X^{\varepsilon,\eta(\varepsilon)} \notin K_\delta\right) < \delta$ となる．このとき $K_\delta \times \overline{B}_{\mathcal{H}}(0,R)$ は $\widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R)$ のコンパクト部分集合であり，$\mathbb{P}\left((X^{\varepsilon,\eta(\varepsilon)},\eta(\varepsilon)) \notin K_\delta \times \overline{B}_{\mathcal{H}}(0,R)\right) = \mathbb{P}\left(X^{\varepsilon,\eta(\varepsilon)} \notin K_\delta\right) < \delta$ を満たすので，$\{(X^{\varepsilon,\eta(\varepsilon)},\eta(\varepsilon))\}_{\varepsilon\in(0,1]}$ が緊密であることが示せた．□

補題 10.4.5. 条件 **(LIP)** を仮定する．$R \in (0,\infty)$ および $\{\eta(\varepsilon)\}_{\varepsilon\in(0,1]} \subset \mathscr{A}'_R$ とし，$(X^{\varepsilon,\eta(\varepsilon)},\eta(\varepsilon))$ は $\varepsilon \searrow 0$ のときにある $\widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R)$ 値の確率変数 $(\widetilde{X},\eta)$ に法則収束すると仮定する．このとき，$\mathbb{P}$ に関して確率 1 で

$$\widetilde{X}_t = a + \int_0^t b(\widetilde{X}_s)ds + \int_0^t \sigma(\widetilde{X}_s)\eta'_s ds, \qquad t \in [0,T]$$

が成り立つ（すなわち $\widetilde{X} = \varphi^\eta$ となる）．

証明. 各 t を固定するごとに

$$X_t^{\varepsilon,\eta(\varepsilon)} - a - \int_0^t b(X_s^{\varepsilon,\eta(\varepsilon)})ds - \int_0^t \sigma(X_s^{\varepsilon,\eta(\varepsilon)})\eta(\varepsilon)'_s ds = \sqrt{\varepsilon}\int_0^t \sigma(X_s^{\varepsilon,\eta(\varepsilon)})dw_s \qquad (10.4.8)$$

の両辺が，$\mathbb{R}^e$ 値の確率変数として $\varepsilon \searrow 0$ のときに法則収束することを見る．

まず右辺は定値確率変数 0 に法則収束（すなわち右辺の法則が Dirac 測度 δ_0 に弱収束）する．これを示すには右辺が 0 に L^2 収束していることを確認すればよい．実際，適当な正定数 c に対して

$$\mathbb{E}\left[\left|\sqrt{\varepsilon}\int_0^t \sigma(X_s^{\varepsilon,\eta(\varepsilon)})dw_s\right|^2\right] \le c\varepsilon\int_0^t (1+\mathbb{E}\left[|X_s^{\varepsilon,\eta(\varepsilon)}|^2\right])ds \le c(1+\sqrt{C})T\varepsilon$$

[19] Kolmogorov の緊密性条件については，例えば [23, 第 1 章，定理 1.2] または [24, 第 9.5 節] を参照せよ．

を満たす．ここで (10.4.7) も使った．

左辺を処理するために，写像 $Q_t\colon \widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R) \to \mathbb{R}^e$ を

$$Q_t(\xi, h) := \xi_t - a - \int_0^t b(\xi_s)ds - \int_0^t \sigma(\xi_s)h'_s ds$$

と定める．(10.4.8) の左辺は $Q_t(X^{\varepsilon,\eta(\varepsilon)}, \eta(\varepsilon))$ に等しい．したがって，Q_t の連続性を認めると左辺は $\varepsilon \searrow 0$ のときに

$$Q_t(\widetilde{X}, \eta) = \widetilde{X}_t - a - \int_0^t b(\widetilde{X}_s)ds - \int_0^t \sigma(\widetilde{X}_s)\eta'_s ds$$

に法則収束する．しかしこの法則が δ_0 であるため，この確率変数は確率 1 で 0 である．t に関する連続性と合わせて考えると，これで補題の主張にある ODE が成立することがわかる．

最後に Q_t の連続性を示そう．補題 9.1.2 により $\widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R)$ は距離空間なので，定義域の収束点列を Q_t で送ったものが値域で収束することを示せばよい．$\{(\xi(n), h(n))\}_{n\in\mathbb{N}} \subset \widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R)$ かつ $\lim_{n\to\infty}(\xi(n), h(n)) = (\xi(\infty), h(\infty))$ とする．$n \to \infty$ のとき

$$\xi(n)_t - a - \int_0^t b(\xi(n)_s)ds \ \to \ \xi(\infty)_t - a - \int_0^t b(\xi(\infty)_s)ds$$

となることはほぼ自明である．一方で

$$\int_0^t \sigma(\xi(n)_s)h(n)'_s ds \ \to \ \int_0^t \sigma(\xi(\infty)_s)h(\infty)'_s ds$$

となることは自明ではないが，事実上同じ計算が命題 10.1.2 の証明中の (10.1.3) 附近にある．以上により，$Q_t(\xi(n), h(n)) \to Q_t(\xi(\infty), h(\infty))$ とわかり，Q_t の連続性が得られた． □

命題 10.4.1 の証明. F が $\widetilde{\mathcal{C}}$ 上の有界連続関数であるとき，(10.4.1) を以下の 2 つの不等式に分解して示す．

$$\varlimsup_{\varepsilon\searrow 0} -\varepsilon \log \mathbb{E}\left[\exp\left(-\frac{1}{\varepsilon}F(X^\varepsilon)\right)\right] \le \inf_{\xi\in\widetilde{\mathcal{C}}}\{F(\xi) + J(\xi)\}, \tag{10.4.9}$$

$$\varliminf_{\varepsilon\searrow 0} -\varepsilon \log \mathbb{E}\left[\exp\left(-\frac{1}{\varepsilon}F(X^\varepsilon)\right)\right] \ge \inf_{\xi\in\widetilde{\mathcal{C}}}\{F(\xi) + J(\xi)\}. \tag{10.4.10}$$

まず (10.4.9) を示す．(10.4.9) の右辺は有限だと仮定してよい．任意の $\delta \in (0,1)$ に対して，ある $\xi^* \in \widetilde{\mathcal{C}}$ で次の条件を満たすものが存在する．

$$F(\xi^*) + J(\xi^*) \le \inf_{\xi\in\widetilde{\mathcal{C}}}\{F(\xi) + J(\xi)\} + \delta.$$

J の定義により，ある $h \in \mathcal{H}$ が存在して，$\xi^* = \varphi^h$ かつ $\|h\|_{\mathcal{H}}^2/2 \le J(\xi^*) + \delta$ を満たす．よって，$R = \sqrt{2\{J(\xi^*)+1\}}$ とおけば $h \in \overline{B}_{\mathcal{H}}(0,R) \subset \mathscr{A}'_R$ である．(なお右の包含では，$\overline{B}_{\mathcal{H}}(0,R)$ の元を定値確率変数と見なしている．)

さて (10.4.5) において $\eta = h$ とした場合の解を $X^{\varepsilon,h}$ と書くと，補題 10.4.4 と補題 10.4.5 により，$\{(X^{\varepsilon,h},h)\}_{\varepsilon\in(0,1]}$ は緊密であり，かつ法則収束の位相での集積点は (φ^h,h) のみなので，実は $\lim_{\varepsilon\searrow 0}(X^{\varepsilon,h},h) = (\varphi^h,h)$ と法則収束している．特に $\lim_{\varepsilon\searrow 0} X^{\varepsilon,h} = \varphi^h = \xi^*$ と $\widetilde{\mathcal{C}}$ において法則収束する．

ここで (10.4.6) を用いると，

$$\begin{aligned}
\varlimsup_{\varepsilon\searrow 0} -\varepsilon \log \mathbb{E}\left[\exp\left(-\frac{1}{\varepsilon}F(X^\varepsilon)\right)\right] &= \varlimsup_{\varepsilon\searrow 0} \inf_{\eta\in\mathscr{A}'_{\mathrm{b}}} \mathbb{E}\left[F(X^{\varepsilon,\eta}) + \frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right] \\
&\le \varlimsup_{\varepsilon\searrow 0} \mathbb{E}\left[F(X^{\varepsilon,h}) + \frac{1}{2}\|h\|_{\mathcal{H}}^2\right] \\
&= F(\xi^*) + \frac{1}{2}\|h\|_{\mathcal{H}}^2 \\
&\le F(\xi^*) + J(\xi^*) + \delta \le \inf_{\xi\in\widetilde{\mathcal{C}}}\{F(\xi)+J(\xi)\} + 2\delta
\end{aligned}$$

という不等式を示せる．最後に $\delta \searrow 0$ として (10.4.9) を得る．

次は (10.4.10) を示す．$\delta \in (0,1)$ を任意に取る．補題 10.4.3 を $G = F(X^\varepsilon)$ に適用すると，ε に依存しないある定数 $R \in (0,\infty)$ と $\{\eta(\varepsilon)\}_{\varepsilon\in(0,1]} \subset \mathscr{A}'_R$ が存在して，

$$-\varepsilon \log \mathbb{E}\left[\exp\left(-\frac{1}{\varepsilon}F(X^\varepsilon)\right)\right] \ge \mathbb{E}\left[F(X^{\varepsilon,\eta(\varepsilon)}) + \frac{1}{2}\|\eta(\varepsilon)\|_{\mathcal{H}}^2\right] - 2\delta \tag{10.4.11}$$

を満たす．補題 10.4.4 により，0 に収束する狭義単調減少列 $\{\varepsilon_k\}_{k\in\mathbb{N}}$ で，次の 2 条件を満たすものが存在する．

- $\varliminf_{\varepsilon\searrow 0} -\varepsilon \log \mathbb{E}\left[\exp(-F(X^\varepsilon)/\varepsilon)\right] = \lim_{k\to\infty} -\varepsilon_k \log \mathbb{E}\left[\exp(-F(X^{\varepsilon_k})/\varepsilon_k)\right]$,
- $k \to \infty$ のとき，$(X^{\varepsilon_k,\eta(\varepsilon_k)}, \eta(\varepsilon_k))$ は $\widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R)$ において法則収束する．

補題 10.4.5 により，法則収束の意味での極限を $\lim_{k\to\infty}(X^{\varepsilon_k,\eta(\varepsilon_k)},\eta(\varepsilon_k)) = (\varphi^\eta,\eta)$ と書いてよい（ここで $\eta = \lim_{k\to\infty}\eta(\varepsilon_k)$ である）．

次は

$$\varliminf_{k\to\infty} \mathbb{E}\left[F(X^{\varepsilon_k,\eta(\varepsilon_k)}) + \frac{1}{2}\|\eta(\varepsilon_k)\|_{\mathcal{H}}^2\right] \ge \mathbb{E}\left[F(\varphi^\eta) + \frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right] \tag{10.4.12}$$

を示す．これは Fatou の補題の一種であるが，今は法則収束しかしていないため一工夫が必要になる．Skorokhod の表現定理により，[20] ある確率空間 $(\widehat{\Omega},\widehat{\mathcal{F}},\widehat{\mathbb{P}})$ 上に以下の条件を満たす $\widetilde{\mathcal{C}} \times \overline{B}_{\mathcal{H}}(0,R)$ 値の確率変数列 $\{(\widehat{X}(k),\widehat{\eta}(k))\}_{k\in\mathbb{N}\cup\{\infty\}}$ が存在する．

$$(\varphi^\eta,\eta) \overset{\mathrm{Law}}{=} (\widehat{X}(\infty),\widehat{\eta}(\infty)),$$

[20] Skorokhod の表現定理については，[23, 付録 I の定理 1.7] や [27, 定理 6.7] などを参照せよ．

$$
\begin{aligned}
&(X^{\varepsilon_k,\eta(\varepsilon_k)}, \eta(\varepsilon_k)) \overset{\text{Law}}{=} (\widehat{X}(k), \widehat{\eta}(k)), && k \in \mathbb{N}, \\
&\lim_{k\to\infty} (\widehat{X}(k), \widehat{\eta}(k)) = (\widehat{X}(\infty), \widehat{\eta}(\infty)), && \widehat{\mathbb{P}}\text{-a.s.}
\end{aligned}
$$

この概収束列を用いると通常の Fatou の補題が使えるため,

$$
\begin{aligned}
\varliminf_{k\to\infty} \mathbb{E}\left[F(X^{\varepsilon_k,\eta(\varepsilon_k)}) + \frac{1}{2}\|\eta(\varepsilon_k)\|_{\mathcal{H}}^2\right] &= \varliminf_{k\to\infty} \widehat{\mathbb{E}}\left[F(\widehat{X}(k)) + \frac{1}{2}\|\widehat{\eta}(k)\|_{\mathcal{H}}^2\right] \\
&\geq \widehat{\mathbb{E}}\left[\varliminf_{k\to\infty} \left\{F(\widehat{X}(k)) + \frac{1}{2}\|\widehat{\eta}(k)\|_{\mathcal{H}}^2\right\}\right] \\
&\geq \widehat{\mathbb{E}}\left[F(\widehat{X}(\infty)) + \frac{1}{2}\|\widehat{\eta}(\infty)\|_{\mathcal{H}}^2\right] \\
&= \mathbb{E}\left[F(\varphi^\eta) + \frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right]
\end{aligned}
$$

と評価できる．なお $\widehat{\mathbb{E}}$ は $\widehat{\mathbb{P}}$ に関する期待値である.[21] これで (10.4.12) が得られた．なお２つ目の不等式では，F の連続性と $\|\cdot\|_{\mathcal{H}}^2$ が $\mathcal{H}$ の弱位相に関して下半連続であることも使った.（$\|h\|_{\mathcal{H}}^2 = \sup_{\|e\|_{\mathcal{H}}=1}\langle h, e\rangle_{\mathcal{H}}^2$ と書けるため，補題 1.2.5 により $\|\cdot\|_{\mathcal{H}}^2$ の下半連続性がわかる.）

部分列 $\{\varepsilon_k\}$ の取り方と (10.4.11)–(10.4.12) をまとめると

$$
\begin{aligned}
\varliminf_{\varepsilon\searrow 0} -\varepsilon \log \mathbb{E}\left[\exp\left(-\frac{1}{\varepsilon}F(X^\varepsilon)\right)\right] &\geq \mathbb{E}\left[F(\varphi^\eta) + \frac{1}{2}\|\eta\|_{\mathcal{H}}^2\right] - 2\delta \\
&\geq \mathbb{E}\left[F(\varphi^\eta) + J(\varphi^\eta)\right] - 2\delta \\
&\geq \inf_{\xi\in\widetilde{\mathcal{C}}}\{F(\xi) + J(\xi)\} - 2\delta
\end{aligned}
$$

を得る．$\delta > 0$ は任意なので (10.4.10) が示せた．これで命題 10.4.1 の証明が終わった．□

[21] この部分の議論は補題 5.0.4 と事実上同じだが，読者の便宜のためにもう一度書いた.

付 録A　凸関数に関する基本事項

本章ではEuclid空間内の凸集合上で定義された凸関数に関する基本的な事実を紹介する．本章の目的は凸関数に関する知識を系統的に解説することではなく，本文中で用いた事実に簡潔な説明をすることである．

Euclid空間 $\mathbb{R}^d$ $(d \in \mathbb{N})$ の部分集合 C が凸集合であるとは，以下の条件を満たすことをいう．

$$x, y \in C,\ t \in (0,1) \quad \Longrightarrow \quad tx + (1-t)y \in C.$$

$\{C_\lambda\}_{\lambda\in\Lambda}$ が凸集合の族ならば，共通部分 $\cap_{\lambda\in\Lambda} C_\lambda$ も凸集合である．したがって，任意の $A \subset \mathbb{R}^d$ に対して，A を含む最小の凸集合が存在する．これを A の凸包と呼ぶ．なお $d = 1$ のときは，空集合または1点集合でない限り，C が凸であることと区間であることとは同値である．

$C \subset \mathbb{R}^d$ を空でない凸集合とする．関数 $f\colon C \to (-\infty, \infty]$ が以下の条件を満たすとき，凸関数と呼ぶ．

$$x, y \in C,\ t \in (0,1) \quad \Longrightarrow \quad f(tx + (1-t)y) \le tf(x) + (1-t)f(y). \tag{A.0.1}$$

なお任意の相異なる x と y に対して，上の不等式 (A.0.1) において "$\le$"を "$<$"に変えたものが成立する場合は，f を狭義凸関数と呼ぶ．空でない添字集合 Λ で添字付けられた C 上の凸関数の族 $\{f_\lambda\}_{\lambda\in\Lambda}$ に対して，明らかに $\sup_{\lambda\in\Lambda} f_\lambda$ も C 上の凸関数である．f の実効定義域を $\mathrm{dom}(f) := \{x \in C \mid f(x) \in \mathbb{R}\}$ と表す．

A.1　1変数凸関数の基本事項

本節では $d = 1$ の場合を扱う．$I \subset \mathbb{R}$ を区間とし，凸関数 $f\colon I \to \mathbb{R}$ の基本性質を列挙する．[1]

命題 A.1.1. $a < b < c$ となる $a, b, c \in I$ に対して

$$\frac{f(b)-f(a)}{b-a} \le \frac{f(c)-f(a)}{c-a} \le \frac{f(c)-f(b)}{c-b} \tag{A.1.1}$$

が成り立つ．

[1] 本節では空集合と1点集合を区間の定義から外す．f が $(-\infty, \infty]$ 値の場合は，f を $\mathrm{dom}(f)$ に制限したものに対して本節の結果を適用できる．

証明. $\lambda := (b-a)/(c-a) \in (0,1)$ とおくと，$b = \lambda a + (1-\lambda)c$ である．f の凸性により，

$$f(b) = f(\lambda a + (1-\lambda)c) \leq \lambda f(a) + (1-\lambda)f(c) = \frac{c-b}{c-a}f(a) + \frac{b-a}{c-a}f(c)$$

を得る．この不等式を移項して (A.1.1) を得る． □

命題 A.1.2. 任意の $s \in I^\circ$ に対して，s における右微分係数 $f'_+(s)$ と左微分係数 $f'_-(s)$ が実数として存在し，$f'_-(s) \leq f'_+(s)$ を満たす．ここで通常どおり

$$f'_+(s) = \lim_{x \searrow s} \frac{f(x)-f(s)}{x-s}, \qquad f'_-(s) = \lim_{x \nearrow s} \frac{f(x)-f(s)}{x-s}$$

と定めた．また f'_+ と f'_- はともに I° 上で広義単調増加である．

証明. 命題 A.1.1 を繰り返し用いる．$s \in I^\circ$ に対して，x, x_0, x_2 を $x_0 < s < x < x_2$ となるように取る．

$$\psi(x) := \frac{f(x)-f(s)}{x-s}, \qquad x \in I^\circ \cap (s, \infty)$$

とおくと，(A.1.1) により ψ は広義単調増加であり，さらに

$$-\infty < \frac{f(s)-f(x_0)}{s-x_0} \leq \psi(x) \leq \frac{f(x_2)-f(s)}{x_2-s} < \infty$$

を満たす．よって $f'_+(s) = \lim_{x \searrow s} \psi(x) \in \mathbb{R}$ は存在する．同様に $f'_-(s) \in \mathbb{R}$ も存在する．再び (A.1.1) により，$f'_+(s)$ と $f'_-(s)$ の定義式において片側極限を取る前の差分商に大小関係があるため，$f'_-(s) \leq f'_+(s)$ である．f'_+ と f'_- が広義単調増加であることも (A.1.1) から直ちに従う． □

命題 A.1.3. f は I° 上で連続である．すなわち，f の不連続点は仮に存在するにしても I の端点のみである．

証明. 命題 A.1.2 により，f は各 $s \in I^\circ$ において右（左）微分可能なので，右（左）連続である．よって，f は各 s において連続である． □

命題 A.1.1 の系として，凸関数の増減に関する簡単な事実を紹介する．この系は不等式 (A.1.1) から直ちに従う．

系 A.1.4. 区間 I 上で定義された凸関数 $f\colon I \to \mathbb{R}$ が $a \in I$ において最小値を実現する（すなわち $f(a) = \inf_{x \in I} f(x)$ となる）と仮定する．このとき $\{x \in I \cap (a, \infty) \mid f(x) > f(a)\}$ が空または１点集合でなければ，この集合は区間であり，f はこの区間上で狭義単調増加である．同様に，仮に $\{x \in I \cap (-\infty, a) \mid f(x) > f(a)\}$ が空または１点集合でなければ，この集合は区間であり，f はこの区間上で狭義単調減少である．

A.2　多変数凸関数の基本事項

本節では $\mathbb{R}^d$ 内の凸集合とその上の凸関数の基本性質を列挙する．まず相対内部の定義を思い出す．空でない凸集合 $C \subset \mathbb{R}^d$ に対して，C の相対内部 $\mathrm{ri}\,(C)$ は次で定義される．

$$\mathrm{ri}\,(C) := \{y \in C \mid \text{任意の}\, x \in C \,\text{に対してある}\, \varepsilon > 0 \,\text{が存在して}\, y - \varepsilon(x-y) \in C\}.$$

次の命題の証明では $\mathbb{R}^d$ 内の単体が基本的な役割を果たすので，命題に進む前に単体の定義と凸集合の "本質的な" 次元について述べる．$k = 0, \ldots, d$ に対して，$\mathbb{R}^d$ の空でない部分集合 $\Delta \subset \mathbb{R}^d$ が k 単体であるとは，ある $x_0, \ldots, x_k \in \mathbb{R}^d$ が存在して次の 2 条件を満たすことをいう．

(i) $x_1 - x_0, \ldots, x_k - x_0$ は線形独立である．ただし，$k = 0$ のときはこの条件は考えない．

(ii) $\Delta = \left\{ \sum_{i=0}^{k} c_i x_i \;\middle|\; c_i \in [0,1]\ (i = 0, \ldots, k), \sum_{i=0}^{k} c_i = 1 \right\}$.

このとき相対内部の定義により

$$\mathrm{ri}\,(\Delta) = \left\{ \sum_{i=0}^{k} c_i x_i \;\middle|\; c_i \in (0,1)\ (i = 0, \ldots, k), \sum_{i=0}^{k} c_i = 1 \right\}$$

が簡単にわかる．

$C \subset \mathbb{R}^d$ を空でない凸集合とする．0 単体は $\mathbb{R}^d$ 内の 1 点集合であることに注意すると，C は空でないので

$$k_C := \max\{k \in \{0, \ldots, d\} \mid \Delta \subset C \,\text{となる}\, k \,\text{単体}\, \Delta \,\text{が存在する}\,\} \tag{A.2.1}$$

は問題なく定まる．Δ が k 単体であれば $k_\Delta = k$ であることに注意せよ．

命題 A.2.1. 空でない凸集合 $C \subset \mathbb{R}^d$ に対して次が成立する．

(1) $\mathrm{ri}\,(C)$ は空でない．

(2) 任意の $x \in C$ と $y \in \mathrm{ri}\,(C)$ および $0 < \alpha \leq 1$ に対して，$(1-\alpha)x + \alpha y \in \mathrm{ri}\,(C)$ が成立する．

証明. (1) k_C を (A.2.1) により与えられるものとすると，k_C の定義により $\Delta \subset C$ となる k_C 単体 Δ と，(i) と (ii) を満たす $x_0, \ldots, x_{k_C} \in C$ が存在する．このとき

$$C \subset \left\{ \sum_{i=0}^{k_C} c_i x_i \;\middle|\; c_i \in \mathbb{R}\ (i = 0, \ldots, k_C), \sum_{i=0}^{k_C} c_i = 1 \right\} \tag{A.2.2}$$

が成立することを示す.（よって C は k_C 次元アファイン部分空間に含まれる.）(A.2.2) を示すために任意に $x \in C$ を取る．C は凸なので

$$\left\{ cx + \sum_{i=0}^{k_C} c_i x_i \;\middle|\; c \in [0,1], c_i \in [0,1]\ (i = 0, \dots, k_C), c + \sum_{i=0}^{k_C} c_i = 1 \right\} \subset C$$

が成立する．このとき $x - x_0, x_1 - x_0, \dots, x_{k_C} - x_0$ が線型独立であると仮定すると，$x_0, x, x_1, \dots, x_{k_C}$ により生成される $k_C + 1$ 単体は C に含まれる．しかしこれは k_C の定義に矛盾するので，$x - x_0, x_1 - x_0, \dots, x_{k_C} - x_0$ は線形従属である．よって少なくとも１つは 0 でないある $c, c_1, \dots, c_{k_C} \in \mathbb{R}$ が存在して

$$c(x - x_0) + \sum_{i=1}^{k_C} c_i (x_i - x_0) = 0$$

を満たす．$c = 0$ とすると (i) により $c_0 = \cdots = c_{k_C} = 0$ となるので，$c \neq 0$ である．よって

$$x = \left(1 + \sum_{i=0}^{k_C} \frac{c_i}{c} \right) x_0 - \sum_{i=1}^{k_C} \frac{c_i}{c} x_i$$

と書けるので，(A.2.2) が成立する．

$\mathrm{ri}\,(\Delta) \subset \mathrm{ri}\,(C)$ を示す．$\widetilde{x} = \sum_{i=0}^{k_C} \widetilde{c}_i x_i \in \mathrm{ri}\,(\Delta)$ と $x \in C$ を任意に取る．(A.2.2) によりある $c_i \in \mathbb{R}$ $(i = 0, \dots, k_C)$ が存在して $x = \sum_{i=0}^{k_C} c_i x_i$ と書ける．このとき $\varepsilon > 0$ に対して

$$\widetilde{x} - \varepsilon(x - \widetilde{x}) = \sum_{i=0}^{k_C} \left(\widetilde{c}_i - \varepsilon \left(c_i - \widetilde{c}_i \right) \right) x_i =: \sum_{i=0}^{k_C} d_i x_i$$

となる．各 $i = 0, \dots, k_C$ に対して $\widetilde{c}_i > 0$ なので，$\varepsilon > 0$ が十分小さければ $d_i > 0$ $(i = 0, \dots, k_C)$ となる．さらに $\sum_{i=0}^{k_C} d_i = 1$ が成立するので $\widetilde{x} - \varepsilon(x - \widetilde{x}) \in C$ を得る．よって $\mathrm{ri}\,(C)$ の定義により $\widetilde{x} \in \mathrm{ri}\,(C)$ となり，$\mathrm{ri}\,(\Delta) \subset \mathrm{ri}\,(C)$ が示された．明らかに $\mathrm{ri}\,(\Delta) \neq \emptyset$ なので，$\mathrm{ri}\,(C) \neq \emptyset$ である．

(2) 任意の $y \in \mathrm{ri}\,(C)$ に対して，ある k_C 単体 $\Delta_y \subset C$ が存在して $y \in \mathrm{ri}\,(\Delta_y)$ となることを示す．$k_C = 0$ のときは明らかなので，以下 $k_C \geq 1$ とする．$y \in \mathrm{ri}\,(C)$ なので，各 $i = 0, \dots, k_C$ に対してある $\varepsilon_i > 0$ が存在して $y - \varepsilon_i (x_i - y) \in C$ が成立する．$\varepsilon =: \min_{i=0,\dots,k_C} \varepsilon_i > 0$ とおくと，C は凸なので各 $i = 0, \dots, k_C$ に対して $y - \varepsilon(x_i - y) \in C$ が成立する．ここで $\delta := \varepsilon\{4(1+\varepsilon)\}^{-1}$ とおき，$i \neq j$ を満たす $i, j = 0, \dots, k_C$ に対して

$$c_i = \widetilde{c}_j = \frac{1}{2(1+\varepsilon)}, \quad d_i = \varepsilon c_i + \delta, \quad \widetilde{d}_j = \varepsilon c_j - \delta$$

とおく．$c_i + \widetilde{c}_j + d_i + \widetilde{d}_j = 1$ であり C は凸なので，単純計算により

$$c_i(y - \varepsilon(x_i - y)) + d_i x_i + \widetilde{c}_j (y - \varepsilon(x_j - y)) + \widetilde{d}_j x_j = y + \delta(x_i - x_j) \in C \qquad \text{(A.2.3)}$$

がわかる．

$x_* \in \mathrm{ri}\,(\Delta)$ を $x_0, \ldots, x_{k_C}$ の重心とする，すなわち

$$x_* := \frac{x_0 + \cdots + x_{k_C}}{k_C + 1}$$

である．(A.2.3) により各 $i = 0, \ldots, k_C$ に対して $z_i := y + \delta(x_i - x_*) \in C$ が成立する．$z_i - z_0 = \delta(x_i - x_0)$ なので，(i) により $\{z_i - z_0\}_{i=1}^{k_C}$ は線形独立である．ここで $\Delta_y \subset C$ を $z_0, \ldots, z_{k_C}$ により定義される k_C 単体とする．すなわち

$$\begin{aligned}\Delta_y :&= \left\{ \sum_{i=0}^{k_C} c_i z_i \;\middle|\; c_i \in [0,1]\ (i = 0, \ldots, k), \sum_{i=0}^{k_C} c_i = 1 \right\} \\ &= \left\{ y + \sum_{i=0}^{k_C} c_i \delta(x_i - x_*) \;\middle|\; c_i \in [0,1]\ (i = 0, \ldots, k), \sum_{i=0}^{k_C} c_i = 1 \right\}\end{aligned}$$

により定義する．このとき $c_0 = \cdots = c_{k_C} = (k_C + 1)^{-1}$ とすることにより，$y \in \mathrm{ri}\,(\Delta_y)$ がわかる．

任意の $x \in C$ と $y \in \mathrm{ri}\,(C)$ および $0 < \alpha \le 1$ に対して，$(1-\alpha)x + \alpha y \in \mathrm{ri}\,(C)$ が成立することを示す．Δ_y を $y \in \mathrm{ri}\,(\Delta_y) \subset C$ となる k_C 単体とする．このとき

$$(1-\alpha)x + \alpha y \in \mathrm{ri}\,((1-\alpha)x + \alpha\Delta_y)$$

が成立するので，(1) の最後の段落で行った議論により $\mathrm{ri}\,((1-\alpha)x + \alpha\Delta_y) \subset \mathrm{ri}\,(C)$ を得るので，$(1-\alpha)x + \alpha y \in \mathrm{ri}\,(C)$ が成立する．　□

注意 A.2.2. $C \subset \mathbb{R}^d$ を空でない凸集合で，１点集合でもないとする（つまり $k_C \ge 1$ とする）．各 $y \in \mathrm{ri}\,(C)$ に対して $\Delta_y \subset C$ を命題 A.2.1(2) の証明中で与えられる k_c 単体とすれば,

$$\mathrm{ri}\,(C) = \bigcup_{y \in \mathrm{ri}\,(C)} \mathrm{ri}\,(\Delta_y)$$

が成立する．$\mathrm{ri}\,(\Delta_y)$ は $\mathbb{R}^{k_C}$ における開球と同相かつ C は k_C 次元アファイン空間に含まれるので，$\mathrm{ri}\,(C)$ は $\mathbb{R}^{k_C}$ のある開集合と同相である．

次の命題は命題 A.1.3 の多変数版である．

命題 A.2.3. $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ を凸関数とする．Γ を $\mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ 上に制限すると，Γ は相対位相に関して $\mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ 上連続である．

証明. $\mathrm{dom}\,\Gamma$ は空でないと仮定して示す．さらに $\mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ が１点集合のとき命題は明らかなので，$\mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ は１点集合でもないとする．$k = k_{\mathrm{dom}\,\Gamma}$ とおくと注意 A.2.2 によ

り，$\mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ は $\mathbb{R}^k$ のある開集合と同相である．Γ をその開集合上の関数と同一視することにより，$\mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ が $\mathbb{R}^k$ の開集合である場合に命題を示せば十分である．

$x \in \mathrm{ri}\,(\mathrm{dom}\,\Gamma) \subset \mathbb{R}^k$ を任意に取る．$x \in \mathrm{ri}\,(\Delta)$ となる C に含まれる k 単体

$$\Delta = \left\{ \sum_{i=0}^{k} c_i x_i \;\middle|\; c_i \in [0,1] \ (i = 0, \ldots, k), \sum_{i=0}^{k} c_i = 1 \right\}$$

と，$\overline{B}(x,\delta) \subset \Delta$ となる $\delta > 0$ を任意に取り固定する．まず Γ は $\overline{B}(x,\delta)$ 上有界であることを示す．任意の $y \in \overline{B}(x,\delta)$ に対して，Γ の凸性により

$$\Gamma(y) \leq \max_{i=0,\ldots,k} \Gamma(x_i) < \infty$$

が成立する．また

$$z_1 := 2x - y \quad \Longleftrightarrow \quad x = \frac{1}{2}z_1 + \frac{1}{2}y$$

とおくと，$z_1 \in \overline{B}(x,\delta)$ なので Γ の凸性により $\Gamma(x) \leq \Gamma(z_1)/2 + \Gamma(y)/2$ を得る．よって

$$\Gamma(y) \geq 2\Gamma(x) - \Gamma(z_1) \geq 2\Gamma(x) - \max_{i=0,\ldots,k} \Gamma(x_i)$$

が成立する．以上により

$$\sup_{y \in B(x,\delta)} |\Gamma(y)| \leq 2|\Gamma(x)| + \max_{i=0,\ldots,k} |\Gamma(x_i)| =: M < \infty$$

を得る．

次に $y \in \overline{B}(x,\delta), y \neq x$ に対して

$$|\Gamma(x) - \Gamma(y)| \leq \frac{2M}{\delta}|x-y|$$

が成立することを示す．まず

$$z := x + \frac{\delta}{|x-y|}(x-y) \quad \Longleftrightarrow \quad x = \frac{|x-y|}{\delta + |x-y|}z + \frac{\delta}{\delta + |x-y|}y$$

とおく．このとき $z \in \overline{B}(x,\delta)$ なので Γ の凸性により

$$\Gamma(x) - \Gamma(y) \leq \frac{|x-y|}{\delta + |x-y|}\left(\Gamma(z) - \Gamma(y)\right) \leq \frac{2M}{\delta}|x-y|$$

が成立する．次に

$$w := x - \frac{\delta}{|x-y|}(x-y) \quad \Longleftrightarrow \quad y = \frac{|x-y|}{\delta}w + \frac{\delta - |x-y|}{\delta}x$$

とおけば, $w \in \overline{B}(x,\delta)$ なので Γ の凸性により

$$\Gamma(y) - \Gamma(x) \leq \frac{|x-y|}{\delta}\left(\Gamma(w) - \Gamma(x)\right) \leq \frac{2M}{\delta}|x-y|$$

を得る．よって Γ は $x \in \mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ で連続である．$\square$

凸関数 $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ に対して，相対内部の定義により $(\mathrm{dom}\,\Gamma)^\circ \subset \mathrm{ri}\,(\mathrm{dom}\,\Gamma)$ が成立するので，命題 A.2.3 により直ちに次の系を得る．

系 A.2.4. 凸関数 $\Gamma : \mathbb{R}^d \to (-\infty, \infty]$ は $(\mathrm{dom}\,\Gamma)^\circ$ 上連続である．

付 録B　正則条件付き確率に関する基本事項

距離空間 S に値を取る確率変数 ξ が与えられたとき，条件付き確率 “$\mathbb{P}(\cdot \mid \xi = x)$” は確率論においてしばしば重要になる．$S$ が非可算集合のときはこの条件付き確率を定めるのはそれほど容易ではないが，この件に関しては次の事実がよく知られている．本命題中の $\{p(x, A) \mid x \in S,\ A \in \mathcal{B}(\Omega)\}$ を「確率変数 ξ を与えたときの（あるいは条件 “$\xi = x$” の下での）$\mathbb{P}$ の正則条件付き確率」と呼ぶ．

命題 B.0.1. $(\Omega, \mathcal{B}(\Omega))$ を可分完備距離空間とその上の Borel 加法族とし，$\mathbb{P}$ をこの可測空間上の確率測度とする．また $(S, \mathcal{B}(S))$ も可分完備距離空間とその上の Borel 加法族とし，$\xi\colon \Omega \to S$ を $\mathcal{B}(\Omega)/\mathcal{B}(S)$ 可測な写像とする．

このとき，系 $\{p(x, A) \mid x \in S,\ A \in \mathcal{B}(\Omega)\}$ で以下の 3 条件を満たすものが存在する．

(1) 各 $x \in S$ に対して，$A \in \mathcal{B}(\Omega) \mapsto p(x, A)$ は $(\Omega, \mathcal{B}(\Omega))$ 上の確率測度である．

(2) 各 $A \in \mathcal{B}(\Omega)$ に対して，$x \in S \mapsto p(x, A)$ は $\mathcal{B}(S)/\mathcal{B}(\mathbb{R})$ 可測である．

(3) 各 $A \in \mathcal{B}(\Omega)$ と各 $E \in \mathcal{B}(S)$ に対して，次が成立する．

$$\mathbb{P}(A \cap \{\xi \in E\}) = \int_E p(x, A)\, \mathbb{P} \circ \xi^{-1}(dx).$$

この系は次の意味で一意的である．仮に系 $\{\widehat{p}(x, A) \mid x \in S,\ A \in \mathcal{B}(\Omega)\}$ も上記の 3 条件を満たすとすると，$\mathbb{P} \circ \xi^{-1}(N) = 0$ となる $N \in \mathcal{B}(S)$ が存在して，$x \notin N$ であれば

$$p(x, A) = \widehat{p}(x, A), \qquad A \in \mathcal{B}(\Omega)$$

が成立する．

さらに $\mathbb{P} \circ \xi^{-1}(N') = 0$ となる $N' \in \mathcal{B}(S)$ が存在して，$x \notin N'$ であれば

$$p(x, \{\xi \in E\}) = \mathbf{1}_E(x), \qquad E \in \mathcal{B}(S)$$

となる．特に $x \notin N'$ であれば $p(x, \{\xi = x\}) = 1$ である．

付 録C　指数型Brown汎関数の変分表示公式

第 10.4 節において，弱収束法を用いて Freidlin-Wentzell の大偏差原理を証明したが，その中で鍵になったのが正値 Brown 汎関数に対する変分表示公式であった．ここではその変分表示公式に証明を与える．

$T>0$ と $d \in \mathbb{N}$ を任意とし，第 9 章と同じく時間区間 $[0,T]$ 上の $\mathbb{R}^d$ 値標準 Brown 運動に関する Wiener 空間と Cameron-Martin 空間を $(\mathcal{C}_0, \mathcal{B}(\mathcal{C}_0), \mathbb{P})$ および $\mathcal{H}$ と書く．すなわち $\mathcal{C}_0$ と $\mathcal{H}$ はそれぞれ (9.0.1) と (9.0.2) で定義される経路空間で，$\mathbb{P}$ は d 次元の Wiener 測度，$\mathcal{B}(\mathcal{C}_0)$ は $\mathcal{C}_0$ の Borel 加法族である.[1]

$\mathcal{N}$ を $\mathbb{P}$ 零集合の全体とし，$\mathcal{F} = \mathcal{B}(\mathcal{C}_0) \vee \mathcal{N}$ を Borel 加法族の $\mathbb{P}$ 完備化とする．また $0 \le t \le T$ のとき，$\mathcal{G}_t = \sigma\{w_s \mid 0 \le s \le t\}$ および $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{N}$ とおく．ここで $\mathcal{G}_T = \mathcal{B}(\mathcal{C}_0)$ であることに注意せよ.[2]情報系 $\{\mathcal{F}_t\}_{t\in[0,T]}$ は右連続なので，４つ組 $(\mathcal{C}_0, \mathcal{F}, \mathbb{P}; \{\mathcal{F}_t\}_{t\in[0,T]})$ はいわゆる「通常の条件」を満たす．本章では基本的にこの４つ組の上で作業する．

$\mathbb{P}$ の下では座標過程 $(w_t)_{0\le t\le T}$ は d 次元標準 Brown 運動である．本章においてはこの座標過程をあえて

$$W_t(w) = w_t, \qquad t \in [0,T],\, w \in \mathcal{C}_0$$

と表すことにする．

注意 C.0.1. 同じ対象に２つの記号 w と W を用意する理由は，単に記述を簡単にするためである．確率論の習慣では，期待値 $\mathbb{E}[\cdot]$ の中には確率空間の元（この場合は w）を書かないため，W という別の記号を用意しないと本章に登場する式が書きづらくなる．（例えば変分表示公式の右辺が典型例である．）しかし W は結局 $\mathcal{C}_0$ の恒等写像のことなので，$\mathcal{C}_0$ 上の実数値またはベクトル値関数 F に対して $\mathbb{E}[F]$ と $\mathbb{E}[F(W)]$ は同じ意味であり，どちらも定義により $\int_{\mathcal{C}_0} F(w)\mathbb{P}(dw)$ に等しい．w と W の使い分けは気分的なもので，確率測度と関係のない決定論的な議論をしているときは w を使い，確率論的な議論をしているときは W を使う．

---

[1]本章では Euclid 空間の標準内積を $\langle v, v'\rangle_{\mathbb{R}^d}$ ではなく $v \cdot v'$ と書く（$v, v' \in \mathbb{R}^d$）．ノルムは $|v| = \sqrt{v\cdot v}$ と書く．一方，２本線のノルム $\|\cdot\|$ は無限次元空間のノルムに対してのみ使う．

[2]これに関する説明が補題 C.2.1 の証明中にある．

4 つ組 $(\mathcal{C}_0, \mathcal{F}, \mathbb{P}; \{\mathcal{F}_t\}_{t\in[0,T]})$ 上で定義された発展的可測な $\mathbb{R}^d$ 値確率過程 $v = (v_t)_{t\in[0,T]}$ で，条件

$$\|v\|_{\mathscr{A}} := \mathbb{E}\left[\int_0^T |v_t|^2 dt\right]^{\frac{1}{2}} < \infty$$

を満たすのものの全体を $\mathscr{A}$ と表す．この $\mathscr{A}$ は Brown 運動に対する確率積分を定義するときに出てくる関数空間であり，ごく自然に Hilbert 空間の構造が入る．$\mathscr{A}$ には自然に内積が定まるが，それを $\langle\cdot,\cdot\rangle_{\mathscr{A}}$ と表す．また各 $R \in (0,\infty)$ に対して，

$$\mathscr{A}_R := \left\{ v \in \mathscr{A} \,\middle|\, \int_0^T |v_t|^2 dt \le R^2,\ \mathbb{P}\text{-a.s.} \right\}, \qquad \mathscr{A}_{\mathrm{b}} := \bigcup_{R>0}^{\infty} \mathscr{A}_R$$

と定める．

それでは本章の主題である指数型 Wiener 汎関数の変分表示公式を正確に述べる．[3] 本章の残りではこの命題を証明する．上からの評価は補題 C.3.4 で示し，下からの評価は補題 C.4.5 で示す．

命題 C.0.2. $F\colon \mathcal{C}_0 \to \mathbb{R}$ を有界な Borel 可測関数とする．このとき

$$-\log \mathbb{E}\left[e^{-F}\right] = \inf_{v\in\mathscr{A}} \mathbb{E}\left[F\left(W + \int_0^{\cdot} v_s ds\right) + \frac{1}{2}\int_0^T |v_s|^2 ds\right]$$

が成立する．

上の命題に関して簡単な注意をする．本書ではこの命題を SDE の解に対して用いるつもりなので，この種のことを確認するべきである．SDE 理論は基本的に L^2 理論であるため，解は各 w で定義されているわけではなく，Wiener 測度に関する同値類としてしか与えられていないことを思い出そう．

注意 C.0.3. 命題 C.0.2 において，関数 F を $\mathbb{P}$-a.s. に等しい別の関数 $\widetilde{F}$ に取り替えても左辺の値が変わらないことは自明であるが，実は右辺の値も変わらない．理由を述べる．以下の補題 C.1.2 で示すように，$v \in \mathscr{A}$ のときは $W + \int_0^{\cdot} v_s ds$ の法則は Brown 運動 W の法則 $\mathbb{P}$ に対して絶対連続である．したがって，$\mathbb{P}$-a.s. に $F = \widetilde{F}$，すなわち $\mathbb{P}\left(F(W) \ne \widetilde{F}(W)\right) = 0$ ならば $\mathbb{P}\left(F(W + \int_0^{\cdot} v_s ds) \ne \widetilde{F}(W + \int_0^{\cdot} v_s ds)\right) = 0$ である．以上により，命題 C.0.2 における F は各点で定義された「関数」でなく，「$\mathbb{P}$ に関する関数の同値類」だと考えてもよいことがわかる．また各点で定義された F を考える場合でも，F の可測性に関する条件を「Borel 可測」から「$\mathcal{F}$ 可測」に緩めてよいこともわかる．

[3] この公式の初出は [49] である．ただし本書で紹介する証明法はむしろ [54] の方法に近い．（補題 C.2.5 で導入する $\widetilde{v}$ と $\overline{v}$ や下からの評価に Clark-Ocone 公式を使う点などがこの方法の特徴である．）

C.1 Wiener 空間上のずらしと指数マルチンゲール

$v \in \mathscr{A}$ に対して

$$\mathcal{E}_t^v := \exp\left(\int_0^t v_s \cdot dW_s - \frac{1}{2}\int_0^t |v_s|^2 ds\right), \qquad t \in [0,T]$$

と定める．座標を使って書くと $\int_0^t v_s \cdot dW_s = \int_0^t \sum_{j=1}^d v_s^j dW_s^j$ である．伊藤の公式により，$d\mathcal{E}_t^v = \mathcal{E}_t^v v_t \cdot dW_t$ となるので，$(\mathcal{E}_t^v)_{t\in[0,T]}$ は $\{\mathcal{F}_t\}$ 局所マルチンゲールである．

Girsanov の定理を思い出そう.[4] v による Wiener 空間 $\mathcal{C}_0$ 上のずらしを

$$\mathcal{T}_t^v(w) := w_t + \int_0^t v_s(w)ds, \qquad t \in [0,T]$$

とおく.[5]もし $\mathbb{P}$ の下で $(\mathcal{E}_t^v)_{t\in[0,T]}$ がマルチンゲールならば，新しい確率測度 $d\mathbb{P}^v := \mathcal{E}_T^v d\mathbb{P}$ の下で $(\mathcal{T}_t^{-v}(W))_{t\in[0,T]}$ は d 次元 $\{\mathcal{F}_t\}$-Brown 運動になる．言い換えると，任意の有界な Borel 可測関数 $F\colon \mathcal{C}_0 \to \mathbb{R}$ に対して

$$\mathbb{E}^v\left[F(\mathcal{T}^{-v}(W))\right] = \mathbb{E}\left[F(W)\right] \tag{C.1.1}$$

が成立する．ここで $\mathbb{E}^v$ は $\mathbb{P}^v$ に関する期待値である．

さて $v \in \mathscr{A}_{\mathrm{b}}$ の場合，すなわちある $R \in (0,\infty)$ に対して $v \in \mathscr{A}_R$ となる場合を考える．このとき

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\left\langle\int_0^\cdot v_s \cdot dW_s\right\rangle_T\right)\right] = \mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T |v_s|^2 ds\right)\right] \le e^{R^2/2} < \infty$$

であるため Novikov 条件は満たされている.[6] よって $(\mathcal{E}_t^v)_{t\in[0,T]}$ はマルチンゲールである．したがって $v \in \mathscr{A}_{\mathrm{b}}$ のときは Girsanov の定理 (C.1.1) は成立する．（一方，一般の $v \in \mathscr{A}$ に対しては成立する保証がない.）さらに $\mathcal{E}_t^v$ の L^p ノルムに関して次の評価がある．

補題 C.1.1. $v \in \mathscr{A}$ とする．このとき任意の $p \in [1,\infty)$ と $t \in [0,T]$ に対して次が成り立つ．

$$(\mathcal{E}_t^v)^p = \mathcal{E}_t^{pv} \exp\left(\frac{p(p-1)}{2}\int_0^t |v_s|^2 ds\right).$$

さらに $v \in \mathscr{A}_{\mathrm{b}}$ のときには $\mathbb{E}\left[(\mathcal{E}_t^v)^p\right] \le e^{R^2 p(p-1)/2}$ と評価できる．ここで R は $v \in \mathscr{A}_R$ を満たす正定数である．

---

[4]Girsanov の定理については [8, 第 3.5.A 節]，[16, 定理 2.6.3]，[23, 第 4.4 節] などを参照せよ．

[5]正確には $w \mapsto \int_0^t v_s(w)ds$ の Borel 可測版を取って，$\mathcal{T}^v\colon \mathcal{C}_0 \to \mathcal{C}_0$ が Borel 可測写像だと思っている．

[6]Novikov 条件は Girsanov の定理に対する最も有名な十分条件である．例えば [8, 第 3.5.D 節] や [16, 命題 2.7.7] や [23, 第 4.4 節の命題 4.3] を見よ．

証明. $(\mathcal{E}_t^v)^p$ の表示式は $\mathcal{E}_t^v$ からほぼ自明である．後半を示す．$v \in \mathscr{A}_{\mathrm{b}}$ のときには $pv \in \mathscr{A}_{\mathrm{b}}$ なので，$(\mathcal{E}_t^{pv})$ はマルチンゲールで，特に $\mathbb{E}\,[\mathcal{E}_t^{pv}] = 1$ である．したがって $\int_0^T |v_s|^2 ds \le R^2$ を仮定すれば，求める評価は簡単に得られる．　□

さらに別の確率測度を導入する．$v \in \mathscr{A}$ に対して，確率過程 $\mathcal{T}^v(W) = W + \int_0^\cdot v_s(W)ds$ が $\mathbb{P}$ の下で誘導する法則を $\mathbb{Q}^v$ と定義する．すなわち $\mathbb{Q}^v := \mathbb{P} \circ (\mathcal{T}^v)^{-1}$ である．

補題 C.1.2. $v \in \mathscr{A}$ に対して

$$H(\mathbb{Q}^v \mid \mathbb{P}) \le \frac{1}{2}\|v\|_{\mathscr{A}}^2 < \infty$$

が成り立つ．ここで H は $\mathcal{C}_0$ 上の 2 つの確率測度から決まる相対エントロピーである.[7] 特に $\mathbb{Q}^v$ は Wiener 測度 $\mathbb{P}$ に対して絶対連続である．

証明. まず $v \in \mathscr{A}_{\mathrm{b}}$ の場合を示す．この場合は Girsanov の定理 (C.1.1) が使えることに注意せよ．過程 $(\mathcal{T}_t^v(W))_{t\in[0,T]}$ が誘導する法則は $\mathbb{P}$ の下では定義により $\mathbb{Q}^v$ だが，$\mathbb{P}^{-v}$ の下では Girsanov の定理により $\mathbb{P}$ である．よって

$$H(\mathbb{Q}^v \mid \mathbb{P}) = H(\mathbb{P} \circ (\mathcal{T}^v)^{-1} \mid \mathbb{P}^{-v} \circ (\mathcal{T}^v)^{-1}) \le H(\mathbb{P} \mid \mathbb{P}^{-v})$$

となる．ここで定理 4.1.2（Donsker-Varadhan の変分表示公式の有界可測関数版）を用いた．この右辺は

$$\begin{aligned}\int_{\mathcal{C}_0} \log\left(\frac{d\mathbb{P}}{d\mathbb{P}^{-v}}\right) d\mathbb{P} &= \mathbb{E}\left[\log\{(\mathcal{E}_T^{-v})^{-1}\}\right] \\ &= \mathbb{E}\left[\int_0^T v_s \cdot dW_s + \frac{1}{2}\int_0^T |v_s|^2 ds\right] = \frac{1}{2}\mathbb{E}\left[\int_0^T |v_s|^2 ds\right]\end{aligned}$$

となるので，$v \in \mathscr{A}_{\mathrm{b}}$ の場合は証明できた．

一般の $v \in \mathscr{A}$ の場合を示す．$\{v^n\}_{n\in\mathbb{N}} \subset \mathscr{A}_{\mathrm{b}}$ で $\lim_{n\to\infty} \|v^n - v\|_{\mathscr{A}} = 0$ となる列を取る．このとき

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\mathcal{T}_t^{v^n}(W) - \mathcal{T}_t^v(W)|\right] \le \mathbb{E}\left[\int_0^T |v_s^n - v_s| ds\right] \le \sqrt{T}\|v^n - v\|_{\mathscr{A}}$$

となるので，$n \to \infty$ のとき $\mathcal{T}^{v^n}(W)$ は $\mathcal{C}_0$ 値確率変数列として $\mathcal{T}^v(W)$ に L^1 収束する．特に $\mathbb{Q}^{v^n}$ は $\mathbb{Q}^v$ に弱収束する．定理 4.1.3（Donsker-Varadhan の変分表示公式の有界連続関数版）から直ちに得られる相対エントロピーの下半連続性により，

$$H(\mathbb{Q}^v \mid \mathbb{P}) \le \varliminf_{n\to\infty} H(\mathbb{Q}^{v^n} \mid \mathbb{P}) \le \varliminf_{n\to\infty} \frac{1}{2}\|v^n\|_{\mathscr{A}}^2 = \frac{1}{2}\|v\|_{\mathscr{A}}^2$$

となる．これで補題 C.1.2 の証明が終わった．　□

---

[7] H の正確な定義は第 4 章の冒頭にある．

C.2　柱状関数と単過程

まず柱状関数の定義を思い出そう．関数 $F\colon \mathcal{C}_0 \to \mathbb{R}$ が $[0,T]$ のある分割 $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_m \leq T\}$ とある Borel 可測関数 $f\colon (\mathbb{R}^d)^m \to \mathbb{R}$ を用いて

$$F(w) = f(w_{t_1}, \ldots, w_{t_m}), \qquad \mathbb{P}\text{-a.s.} \tag{C.2.1}$$

と書けるとき，F を柱状関数という．ある分割 $\mathcal{P}$ および $\|f\|_\infty \vee \|\nabla f\|_\infty < \infty$ を満たすある C^1 級関数 $f\colon (\mathbb{R}^d)^m \to \mathbb{R}$ に対して F が (C.2.1) の表示を持つならば，F を C^1_{b} 級柱状関数とよび，$F \in \mathcal{F}C^1_{\mathrm{b}}$ と表す.（区間が $[0,T]$ であることを明記したいときは $\mathcal{F}_T C^1_{\mathrm{b}}$ と表すことにする.）

補題 C.2.1. $1 \leq p < \infty$ とする．このとき $\mathcal{F}C^1_{\mathrm{b}}$ は $L^p(\mathbb{P})$ において稠密である．

証明. まず念のために $\mathcal{G}_T$ と Borel 集合族 $\mathcal{B}(\mathcal{C}_0)$ が一致することを確認する．自明な射影 $\mathcal{C}_0 \ni w \mapsto w_t \in \mathbb{R}^d$ は連続なので，$\mathcal{G}_T \subset \mathcal{B}(\mathcal{C}_0)$ は明らか．逆の包含関係を示す．

任意の $r > 0$ と $z \in \mathcal{C}_0$ に対して，閉球 $\overline{B}_{\mathcal{C}_0}(z,r)$ が $\mathcal{G}_T$ 可測であることは，

$$\overline{B}_{\mathcal{C}_0}(z,r) = \{w \in \mathcal{C}_0 \mid \sup_{0 \leq t \leq T} |w_t - z_t| \leq r\} = \bigcap_{t \in [0,T] \cap \mathbb{Q}} \{w \in \mathcal{C}_0 \mid |w_t - z_t| \leq r\} \ \in \mathcal{G}_T$$

と変形すればすぐわかる．これから直ちに開球 $B_{\mathcal{C}_0}(z,r) = \cup_{n \in \mathbb{N}} \overline{B}_{\mathcal{C}_0}(z, r - 1/n)$ の $\mathcal{G}_T$ 可測性がわかる．$\mathcal{C}_0$ は可分な距離空間なので，その任意の開部分集合は開球の高々可算和の形に書けるので，$\mathcal{G}_T$ 可測である．これは $\mathcal{B}(\mathcal{C}_0) \subset \mathcal{G}_T$ を意味する．

稠密性を示す．$G \in L^p(\mathbb{P})$ を任意に取る.（G は $\mathcal{G}_T$ 可測版を持つ.）$G_n := (G \wedge n) \vee (-n)$ とおくと $\{G_n\}_{n \in \mathbb{N}}$ は G に L^p 収束するので，有界な関数の全体は $L^p(\mathbb{P})$ において稠密である．さらに単調族定理（の関数形）を用いた標準的な議論により，任意の $\mathcal{G}_T$ 可測な有界関数は有界柱状関数の列で L^p 近似できることが示せる．

したがって，問題は任意の柱状関数 F を C^1_{b} 級柱状関数の列で L^p 近似できるかどうかに帰着される．以下では F は (C.2.1) の形であるとし，分割 $\mathcal{P}$ は固定して議論する．まず f を $f_n := f \cdot \mathbf{1}_{B(0,n)}$ で置き換えると，F は $\{F_n\}_{n \in \mathbb{N}}$ で L^p 近似できることはすぐわかる.（ここで $B(0,n)$ は $(\mathbb{R}^d)^m$ の開球である.）よって f の台が有界の場合だけを示せば十分である．軟化子を用いた標準的な議論により，$(\mathbb{R}^d)^m$ 上の Lebesgue 測度 $d\lambda$ に関して f に L^p 収束する有界な台を持つ滑らかな関数の列 $\{f_n\}_{n \in \mathbb{N}}$ が作れる．$(w_{t_1}, \ldots, w_{t_m})$ が $(\mathbb{R}^d)^m$ に誘導する法則は非退化な正規分布なため $d\lambda$ に対して有界な密度関数 ρ を持つ．したがって，$n \to \infty$ のときに

$$\mathbb{E}\left[|F_n - F|^p\right] = \int_{(\mathbb{R}^d)^m} |f_n - f|^p \rho \, d\lambda \leq \|\rho\|_\infty \int_{(\mathbb{R}^d)^m} |f_n - f|^p d\lambda \to 0$$

となる．これで補題 C.2.1 の証明が終わった．　□

注意 C.2.2. 補題 C.2.1 の証明とまったく同様に，$0 < u \leq T$ かつ $1 \leq p < \infty$ のとき，$\mathcal{F}_u C_{\mathrm{b}}^1$ が $\{G \in L^p(\mathbb{P}) \mid G$ は $\mathcal{F}_u$ 可測 $\}$ において稠密であることが証明できる．

補題 C.2.3. $1 \leq p < \infty$ とし，$G\colon \mathcal{C}_0 \to \mathbb{R}$ は有界な Borel 可測関数とする．このとき，一様有界な $\{G_n\}_{n\in\mathbb{N}} \subset \mathcal{F}C_{\mathrm{b}}^1$ で G に $L^p(\mathbb{P})$ 収束かつ $\mathbb{P}$-a.s. 収束するものが存在する．

証明. K を $|G|$ の本質的上限とする．$K = 0$ ならば自明なので，$K > 0$ と仮定して示す．$\chi_K\colon \mathbb{R} \to \mathbb{R}$ を広義単調増加する C^1 級奇関数で，$[0, K]$ 上で $\chi_K(u) = u$ かつ $[2K, \infty)$ 上で $\chi_K(u) = 2K$ を満たすものとする．このとき χ_K と χ_K' は有界である．補題 C.2.1 により，G に L^p 収束する $\{\widetilde{G}_n\}_{n\in\mathbb{N}} \subset \mathcal{F}C_{\mathrm{b}}^1$ が存在する．$G_n = \chi_K \circ \widetilde{G}_n$ とおくと再び $\mathcal{F}C_{\mathrm{b}}^1$ に属し，$n \to \infty$ のとき

$$\mathbb{E}\left[|G_n - G|^p\right] = \mathbb{E}\left[|\chi_K \circ \widetilde{G}_n - \chi_K \circ G|^p\right] \leq \|\chi_K'\|_\infty^p \mathbb{E}\left[|\widetilde{G}_n - G|^p\right] \to 0$$

となる．作り方から $\sup_{n\in\mathbb{N}} \|G_n\|_\infty \leq 2K$ なので，$\mathbb{P}$-a.s. 収束する $\{G_n\}$ の部分列が求める近似列である． □

ここで $\mathbb{R}^d$ 値単過程の定義を導入する．$\mathcal{C}_0$ 上で定義された発展的可測な確率過程 $v = (v_t)_{t\in[0,T]}$ のうち，$[0, T]$ のある分割 $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_m = T\}$，ある $\xi_0 \in \mathbb{R}^d$ およびある有界 $\mathcal{G}_{t_k}$ 可測関数 $\xi_k\colon \mathcal{C}_0 \to \mathbb{R}^d$ $(1 \leq k \leq m-1)$ を用いて

$$v_t(w) = \xi_0 \mathbf{1}_{[t_0,t_1]}(t) + \sum_{k=1}^{m-1} \xi_k(w)\mathbf{1}_{(t_k,t_{k+1}]}(t), \qquad t \in [0, T],\, w \in \mathcal{C}_0 \tag{C.2.2}$$

という形に書けるものを（$\mathbb{R}^d$ 値）単過程と呼ぶ．本章ではこのような単過程の全体を $\mathcal{S}$ と書く．Wiener 測度 $\mathbb{P}$ の下で考えるので自然に $\mathcal{S} \subset \mathscr{A}_{\mathrm{b}} \subset \mathscr{A}$ と思えるが，確率積分の入門で学ぶように $\mathcal{S}$ は $\mathscr{A}$ において稠密である．[8]

注意 C.2.4. 通常の確率積分の入門書では，ξ_k が $\mathcal{F}_{t_k}$ 可測であることを要請しているが，任意の $\mathcal{F}_{t_k}$ 可測関数は $\mathbb{P}$ 零集合上で値を修正することにより $\mathcal{G}_{t_k}$ 可測になるので，上のように定義しても結局は同値である．$\mathcal{G}_{t_k}$ 可測であるため，ξ_k は実は（$w = (w_t)_{t\in[0,T]}$ そのものの関数というよりも）$w\restriction_{[0,t_k]} = (w_t)_{t\in[0,t_k]}$ の可測関数であること[9] に注意せよ．

補題 C.2.5. $v \in \mathcal{S}$ とする．このとき，ある $\widetilde{v} \in \mathcal{S}$ と $\overline{v} \in \mathcal{S}$ が存在して，それぞれ以下の条件を任意の $w \in \mathcal{C}_0$ に対して満たす．

$$\widetilde{v}(w) = v(\mathcal{T}^{-\widetilde{v}}(w)), \qquad \mathcal{T}^v \circ \mathcal{T}^{-\widetilde{v}}(w) = w, \tag{C.2.3}$$

$$v(w) = \overline{v}(\mathcal{T}^{-v}(w)), \qquad \mathcal{T}^{\overline{v}} \circ \mathcal{T}^{-v}(w) = w. \tag{C.2.4}$$

[8] この稠密性については，例えば [16, 補題 2.1.3]，[23, 補題 1.1]，[11, 命題 4.1.2] を見よ．

[9] この事実は直観的には明らかだと思う．測度論における標準的な議論を用いてこれを厳密に証明することもそれほど難しくない．

証明. 所与の v は (C.2.2) の形の表示を持つとする．まず条件 (C.2.3) を満たす $\widetilde{v}$ を以下のように番号の小さな小区間から順に具体的に構成する．

$$
\begin{aligned}
\widetilde{\xi}_0 &:= \xi_0, & \widetilde{v}_t(w) &= \widetilde{\xi}_0, & t &\in [t_0, t_1],\\
\widetilde{\xi}_1(w) &:= \xi_1\left(\left(w_t - \int_0^t \widetilde{v}_s(w)ds\right)_{t\in[0,t_1]}\right), & \widetilde{v}_t(w) &= \widetilde{\xi}_1(w), & t &\in (t_1, t_2],\\
&\cdots\cdots\cdots & &\cdots\cdots\cdots & &\cdots\cdots\\
\widetilde{\xi}_{m-1}(w) &:= \xi_{m-1}\left(\left(w_t - \int_0^t \widetilde{v}_s(w)ds\right)_{t\in[0,t_{m-1}]}\right), & \widetilde{v}_t(w) &= \widetilde{\xi}_{m-1}(w), & t &\in (t_{m-1}, t_m].
\end{aligned}
$$

なお $(w_t - \int_0^t \widetilde{v}_s(w)ds)_{t\in[0,t_k]}$ の法則は $(w_t)_{t\in[0,t_k]}$ の法則に対して絶対連続なので，$\xi_k = \zeta_k$ ($\mathbb{P}$-a.s) であれば $\widetilde{\xi}_k = \widetilde{\zeta}_k$ ($\mathbb{P}$-a.s) であることを注意しておく．明らかに $\widetilde{\xi}_k$ は有界かつ $\mathcal{G}_{t_k}$ 可測であり，よって $\widetilde{v} \in \mathcal{S}$ である．作り方から $\widetilde{v}(w) = v(\mathcal{T}^{-\widetilde{v}}(w))$ はほぼ自明であり，これを用いると

$$
\mathcal{T}^v\big(\mathcal{T}^{-\widetilde{v}}(w)\big) = \left(w - \int_0^\cdot \widetilde{v}_s(w)ds\right) + \int_0^\cdot v_s(\mathcal{T}^{-\widetilde{v}}(w))ds = w
$$

となる．これで (C.2.3) が確認できた．

同様の議論で次は条件 (C.2.4) を満たす $\overline{v}$ を以下のように具体的に構成する．

$$
\begin{aligned}
\overline{\xi}_0 &:= \xi_0, & \overline{v}_t(w) &= \overline{\xi}_0, & t &\in [t_0, t_1],\\
\overline{\xi}_1(w) &:= \xi_1\left(\left(w_t + \int_0^t \overline{v}_s(w)ds\right)_{t\in[0,t_1]}\right), & \overline{v}_t(w) &= \overline{\xi}_1(w), & t &\in (t_1, t_2],\\
&\cdots\cdots\cdots & &\cdots\cdots\cdots & &\cdots\cdots\\
\overline{\xi}_{m-1}(w) &:= \xi_{m-1}\left(\left(w_t + \int_0^t \overline{v}_s(w)ds\right)_{t\in[0,t_{m-1}]}\right), & \overline{v}_t(w) &= \overline{\xi}_{m-1}(w), & t &\in (t_{m-1}, t_m].
\end{aligned}
$$

上と同じ理由により，$\xi_k = \zeta_k$ ($\mathbb{P}$-a.s) であれば $\overline{\xi}_k = \overline{\zeta}_k$ ($\mathbb{P}$-a.s) である．明らかに $\overline{\xi}_k$ は有界かつ $\mathcal{G}_{t_k}$ 可測であり，よって $\overline{v} \in \mathcal{S}$ である．

まず $v(w) = \overline{v}(\mathcal{T}^{-v}(w))$ を確認する．$[t_0, t_1]$ 上では $v_t = \overline{v}_t = \xi_0$ であるため，$v_t(w) = \overline{v}_t(\mathcal{T}^{-v}(w)) = \xi_0$ は明らかである．k $(1 \le k \le m-1)$ に対して，$[t_0, t_k]$ 上で $v_t(w) = \overline{v}_t(\mathcal{T}^{-v}(w))$ が成り立つとする．このとき，

$$
\overline{\xi}_k(\mathcal{T}^{-v}(w)) = \xi_k\Big(\Big\{w_t - \int_0^t v_s(w)ds\Big\} + \int_0^t \overline{v}_s(\mathcal{T}^{-v}(w))ds,\, t \le t_k\Big) = \xi_k(w)
$$

となるので，$[t_0, t_{k+1}]$ 上で $v_t(w) = \overline{v}_t(\mathcal{T}^{-v}(w))$ が成り立つことがわかる．帰納法により $v(w) = \overline{v}(\mathcal{T}^{-v}(w))$ がわかる．これを用いると，

$$
\mathcal{T}^{\overline{v}}\big(\mathcal{T}^{-v}(w)\big) = \left(w - \int_0^\cdot v_s(w)ds\right) + \int_0^\cdot \overline{v}_s(\mathcal{T}^{-v}(w))ds = w
$$

となる．これで (C.2.4) が確認できた．以上で補題 C.2.5 が証明できた．　□

補題 C.2.6. $R \in (0,\infty)$ とする．このとき，$\mathscr{A}$ の位相に関して $\mathscr{A}_R \cap \mathcal{S}$ は $\mathscr{A}_R$ において稠密である．

証明. $u \in \mathscr{A}_R$ とする．任意の $\delta > 0$ に対して，$\lim_{n\to\infty} \|v^{n,\delta} - u\|_{\mathscr{A}} = 0$ となる列 $\{v^{n,\delta}\}_{n\in\mathbb{N}} \subset \mathscr{A}_{R+\delta} \cap \mathcal{S}$ が存在することを示せば十分である．実際このとき，任意の $m \in \mathbb{N}$ に対してある $n(m)$ が存在して $\|v^{n(m),1/m} - u\|_{\mathscr{A}} \leq 1/m$ となるが，作り方から

$$\{R(R+1/m)^{-1} v^{n(m),1/m}\}_{m\in\mathbb{N}} \subset \mathscr{A}_R \cap \mathcal{S}$$

であり，かつ $m \to 0$ のときに

$$\begin{aligned}\|R(R+1/m)^{-1} v^{n(m),1/m} - u\|_{\mathscr{A}} &\leq \left|\frac{R}{R+1/m} - 1\right| \|v^{n(m),1/m}\|_{\mathscr{A}} + \|v^{n(m),1/m} - u\|_{\mathscr{A}} \\ &\leq \left|\frac{R}{R+1/m} - 1\right| (R+1/m) + 1/m \\ &= 2/m \ \to 0\end{aligned}$$

となるからである．

以下では固定した R, δ に対して u を近似する列 $\{v^n\}_{n\in\mathbb{N}} \subset \mathscr{A}_{R+\delta} \cap \mathcal{S}$ を見つける．（ここからは v の上付き添字から δ を省略する．）まず $\mathcal{S}$ は $\mathscr{A}$ で稠密なので，$\{u^n\}_{n\in\mathbb{N}} \subset \mathcal{S}$ で $\lim_{n\to\infty} \|u^n - u\|_{\mathscr{A}} = 0$ となる列が存在する．各 u^n は (C.2.2) の形の表示を持つ．（以下では n ごとにこの種の表示を 1 つ選んで議論する．）

$$u_t^n(w) = \xi_0^n \mathbf{1}_{[t_0^n, t_1^n]}(t) + \sum_{k=1}^{m_n-1} \xi_k^n(w) \mathbf{1}_{(t_k^n, t_{k+1}^n]}(t).$$

ここで $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{m_n}^n = T\}$ は $[0,T]$ の分割であり，$\xi_0^n \in \mathbb{R}^d$ であり，さらに $\xi_k^n \colon \mathcal{C}_0 \to \mathbb{R}^d$ $(1 \leq k \leq m_n - 1)$ は有界 $\mathcal{G}_{t_k^n}$ 可測関数である．

さて

$$\eta_k^n(w) := \xi_k^n(w) \mathbf{1}_{Z_k^n}(w), \quad Z_k^n := \left\{ \int_0^{t_{k+1}^n} |u_s^n|^2 ds = \sum_{j=0}^{k} |\xi_j^n|^2 (t_{j+1}^n - t_j^n) \leq (R+\delta)^2 \right\}$$

とおくと，Z_k^n も η_k^n も $\mathcal{G}_{t_k^n}$ 可測である．したがって

$$v_t^n(w) = \eta_0^n \mathbf{1}_{[t_0^n, t_1^n]}(t) + \sum_{k=1}^{m_n-1} \eta_k^n(w) \mathbf{1}_{(t_k^n, t_{k+1}^n]}(t)$$

とおくと，作り方により $\mathbb{P}$-a.s. に $\|v^n\|_{L^2([0,T])}^2 := \int_0^T |v_s^n|^2 ds \leq (R+\delta)^2$ となるので，$v^n \in \mathscr{A}_{R+\delta} \cap \mathcal{S}$ である．

この $\{v^n\}$ が u に収束することを確認しよう．まず明らかに

$$\mathbb{E}\left[\int_0^T |u_s - v_s^n|^2 ds\right] = \mathbb{E}\left[\int_0^T |u_s - v_s^n|^2 ds\,;\, \int_0^T |u_s^n|^2 ds \le (R+\delta)^2\right] + \mathbb{E}\left[\int_0^T |u_s - v_s^n|^2 ds\,;\, \int_0^T |u_s^n|^2 ds > (R+\delta)^2\right] =: J_1^n + J_2^n.$$

J_1^n 内の条件の下では $v^n = u^n$ であることに注意すると，$J_1^n \le \|u - u^n\|_{\mathscr{A}}^2$ と評価できるので，J_1^n は $n \to \infty$ のとき 0 に収束する．J_2^n 内の被積分関数は $\mathbb{P}$-a.s. に $4(R+\delta)^2$ 以下であるので，$J_2^n \le 4(R+\delta)^2 \mathbb{P}\left(\|u^n\|_{L^2([0,T])} > R+\delta\right)$ と評価できる．３角不等式により $\|u^n\|_{L^2([0,T])} > R+\delta$ のとき，

$$\|u^n - u\|_{L^2([0,T])} \ge \|u^n\|_{L^2([0,T])} - \|u\|_{L^2([0,T])} > (R+\delta) - R = \delta$$

が成り立つので，

$$\mathbb{P}\left(\|u^n\|_{L^2([0,T])} > R+\delta\right) \le \mathbb{P}\left(\|u^n - u\|_{L^2([0,T])} > \delta\right) \le \delta^{-2}\|u^n - u\|_{\mathscr{A}}^2$$

と評価される．最右辺は 0 に収束するので，J_2^n も $n \to \infty$ のとき 0 に収束する．以上で補題 C.2.6 の証明が終わった． □

C.3　上からの評価

本節ではいくつかの補題を示した後，変分表示公式（命題 C.0.2）の上からの評価を証明する．証明の鍵は Girsanov の定理である．

補題 C.3.1. $F\colon \mathcal{C}_0 \to \mathbb{R}$ を有界な Borel 可測関数とし，$v \in \mathscr{A}_{\mathrm{b}}$ とする．このとき次が成立する．

$$-\log \mathbb{E}\left[e^{-F(W)}\right] \le \mathbb{E}^v\left[F(W) + \frac{1}{2}\int_0^T |v_s|^2 ds\right].$$

証明. $v \in \mathscr{A}_{\mathrm{b}}$ なので，Girsanov の定理が使える．$u > 0$ に対して $\log u \ge 1 - u^{-1}$ が成り立つことに注意すると

$$\begin{aligned}\log \mathbb{E}\left[e^{-F(W)}\right] + \mathbb{E}^v\left[F(W) + \log \mathcal{E}_T^v\right] &= \mathbb{E}^v\left[\log\left(\mathbb{E}\left[e^{-F(W)}\right] e^{F(W)} \mathcal{E}_T^v\right)\right] \\ &\ge \mathbb{E}^v\left[1 - \frac{e^{-F(W)}}{\mathbb{E}\left[e^{-F(W)}\right]\mathcal{E}_T^v}\right] \\ &= \mathbb{E}\left[\left(1 - \frac{e^{-F(W)}}{\mathbb{E}\left[e^{-F(W)}\right]\mathcal{E}_T^v}\right)\mathcal{E}_T^v\right] = 1 - 1 = 0\end{aligned}$$

とわかる．簡単な計算で

$$\mathbb{E}^v\left[\log \mathcal{E}_T^v\right] = \mathbb{E}^v\left[\int_0^T v_s \cdot dW_s - \int_0^T |v_s|^2 ds\right] + \mathbb{E}^v\left[\frac{1}{2}\int_0^T |v_s|^2 ds\right]$$

$$= \mathbb{E}\left[M_T\right] + \mathbb{E}^v\left[\frac{1}{2}\int_0^T |v_s|^2 ds\right]$$

となる．ここで以下のように定めた．

$$M_t := \left(\int_0^t v_s \cdot dW_s - \int_0^t |v_s|^2 ds\right)\mathcal{E}_t^v.$$

後は $\mathbb{E}\left[M_T\right] = 0$ を証明すれば十分である．$M_0 = 0$ であり，補題 C.1.1 により $M_t \in L^2(\mathbb{P})$ なので，伊藤の公式を用いて (M_t) が局所マルチンゲールであることを確かめればよい．$d\mathcal{E}_t^v = \mathcal{E}_t^v v_t \cdot dW_t$ に注意して計算すると，

$$\begin{aligned} dM_t &= \mathcal{E}_t^v\left(v_t \cdot dW_t - |v_t|^2 dt\right) \\ &\quad + \left(\int_0^t v_s \cdot dW_s - \int_0^t |v_s|^2 ds\right)\mathcal{E}_t^v v_t \cdot dW_t + d\left\langle \int_0^\cdot v_s \cdot dW_s,\ \int_0^\cdot \mathcal{E}_s^v v_s \cdot dW_s \right\rangle_t \\ &= \left(1 + \int_0^t v_s \cdot dW_s - \int_0^t |v_s|^2 ds\right)\mathcal{E}_t^v v_t \cdot dW_t \end{aligned}$$

を得る．有界変動項が消えているので，(M_t) は局所マルチンゲールだとわかる． □

補題 C.3.2. $F\colon \mathcal{C}_0 \to \mathbb{R}$ を有界な Borel 可測関数とし，$v \in \mathcal{S}$ とする．このとき次が成立する．

$$-\log \mathbb{E}\left[e^{-F(W)}\right] \le \mathbb{E}\left[F(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right].$$

証明. v に対して，$\widetilde{v}$ を補題 C.2.5 の条件 (C.2.3) を満たす単過程とする．(C.1.1) を用いて右辺を変形すると，

$$\begin{aligned} \mathbb{E}\left[F(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right] &= \mathbb{E}^{\widetilde{v}}\left[F(\mathcal{T}^v \circ \mathcal{T}^{-\widetilde{v}}(W)) + \frac{1}{2}\int_0^T |v_s(\mathcal{T}^{-\widetilde{v}}(W))|^2 ds\right] \\ &= \mathbb{E}^{\widetilde{v}}\left[F(W) + \frac{1}{2}\int_0^T |\widetilde{v}_s|^2 ds\right] \\ &\ge -\log \mathbb{E}\left[e^{-F(W)}\right] \end{aligned}$$

となる．ここで (C.2.3) と補題 C.3.1 も使った． □

補題 C.3.3. $F\colon \mathcal{C}_0 \to \mathbb{R}$ を有界な連続関数とし，$v \in \mathscr{A}$ とする．このとき次が成立する．

$$-\log \mathbb{E}\left[e^{-F(W)}\right] \le \mathbb{E}\left[F(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right].$$

証明. $\mathcal{S}$ は $\mathscr{A}$ において稠密なので，$\lim_{n\to\infty} \|v^n - v\|_{\mathscr{A}} = 0$ となる列 $\{v^n\}_{n\in\mathbb{N}} \subset \mathcal{S}$ が取れる．すると $n \to \infty$ とするとき，

$$\mathbb{E}\left[\left\|\int_0^\cdot v_s^n ds - \int_0^\cdot v_s ds\right\|_\infty\right] \le \mathbb{E}\left[\int_0^T |v_s^n - v_s| ds\right] \le \sqrt{T}\|v^n - v\|_{\mathscr{A}} \to 0$$

となるので，必要なら部分列（再び同じ記号で表す）を取れば

$$\lim_{n\to\infty}\left\|\int_0^\cdot v_s^n ds - \int_0^\cdot v_s ds\right\|_\infty = 0, \qquad \mathbb{P}\text{-a.s.}$$

が成り立つとしてよい．特に $\lim_{n\to\infty}\|\mathcal{T}^{v^n}(W)-\mathcal{T}^v(W)\|_\infty = 0$ が $\mathbb{P}$-a.s. に成り立つ．補題 C.3.2 から

$$-\log\mathbb{E}\left[e^{-F(W)}\right] \le \mathbb{E}\left[F(\mathcal{T}^{v^n}(W))\right] + \frac{1}{2}\|v^n\|_{\mathscr{A}}^2$$

であるが，右辺第 2 項は明らかに $\|v\|_{\mathscr{A}}^2/2$ に収束する．第 1 項は上で指摘した概収束と F の連続性と Lebesgue の収束定理により $\mathbb{E}\left[F(\mathcal{T}^v(W))\right]$ に収束することがわかる．これで補題 C.3.3 の証明が終わった．□

次の補題が本節の主目的である．この中で変分表示公式の上からの評価を証明する．

補題 C.3.4. $F\colon \mathcal{C}_0 \to \mathbb{R}$ を有界な Borel 可測関数とし，$v \in \mathscr{A}$ とする．このとき次が成立する．

$$-\log\mathbb{E}\left[e^{-F(W)}\right] \le \mathbb{E}\left[F(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right].$$

証明．補題 C.2.3 で示したように，一様有界な連続関数の列 $\{F_n\}_{n\in\mathbb{N}}$ で F に $\mathbb{P}$-a.s. 収束するものが存在する．各 n に対して補題 C.3.3 を使うと

$$-\log\mathbb{E}\left[e^{-F_n(W)}\right] \le \mathbb{E}\left[F_n(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right]$$

を得る．$\{F_n\}$ の一様有界性により，Lebesgue の収束定理を左辺に適用できるため，左辺が $-\log\mathbb{E}\left[e^{-F(W)}\right]$ に収束することがわかる．また補題 C.1.2 により，$\mathcal{T}^v(W)$ の法則は W の法則（すなわち Wiener 測度 $\mathbb{P}$）に対して絶対連続なので，$F_n(\mathcal{T}^v(W))$ は $F(\mathcal{T}^v(W))$ に $\mathbb{P}$-a.s. 収束する．一様有界性と Lebesgue の収束定理により $\mathbb{E}\left[F_n(\mathcal{T}^v(W))\right]$ は $\mathbb{E}\left[F(\mathcal{T}^v(W))\right]$ に収束する．□

C.4　下からの評価

本節では有名な Clark-Ocone の公式を用いて，変分表示公式（命題 C.0.2）の下からの評価を証明する．[10]

まず Clark-Ocone の公式を思い出そう．以下では $F(w) = f(w_{t_1},\dots,w_{t_m})$ は表示式 (C.2.1) の形をした C_{b}^1 級柱状関数だとする．このとき $N_t = \mathbb{E}\left[F|\mathcal{F}_t\right]$ とおくと，$(N_t)_{t\in[0,T]}$ は有界な連続 $\{\mathcal{F}_t\}_{t\in[0,T]}$ マルチンゲールになる．[11]

[10] この公式については，より一般的な主張と証明が [14, P. 201] にある．

[11] 伊藤の表現定理により，任意の 2 乗可積分な $\{\mathcal{F}_t\}_{t\in[0,T]}$ マルチンゲールは連続である．（これに関しては後で述べる．）

命題 C.4.1. C^1_{b} 級柱状関数 F と有界な $\{\mathcal{F}_t\}_{t\in[0,T]}$ マルチンゲール $(N_t)_{t\in[0,T]}$ を上の通りとする．このとき

$$z_t = \sum_{k=1}^{m} \mathbf{1}_{[0,t_k]}(t)\mathbb{E}\left[\nabla_{x^k} f(W_{t_1},\dots,W_{t_m})|\mathcal{F}_t\right], \qquad t\in[0,T]$$

とおくと，明らかに $(z_t)_{t\in[0,T]} \in \mathscr{A}_{\mathrm{b}}$ であり，さらに $\mathbb{P}$-a.s. に

$$N_t - N_0 = \int_0^t z_s \cdot dW_s, \qquad t\in[0,T]$$

が成立する．なお ∇_{x^k} は第 k 変数に関する勾配作用素を表す．よって $\nabla_{x^k} f$ は $\mathbb{R}^d$ 値の有界連続関数である．

証明. 記号を簡単にするために $d=1$ の場合を示す．一般の d の場合も事実上同じ議論で証明できる．本証明中では，Wiener 空間上の実数値および可分 Hilbert 空間 $\mathcal{K}$ 値の 2 乗可積分関数の全体をそれぞれ $L^2(\mathbb{P})$ および $L^2(\mathbb{P};\mathcal{K})$ で表す．また $[0,T]$ 上の Lebesgue 測度に関する実数値の 2 乗可積分関数の全体を $L^2([0,T])$ と表す．$g\in L^2([0,T])$ に対して $Ug := \int_0^\cdot g_s ds$ と書くと，U は $L^2([0,T])$ から $\mathcal{H}$ へのユニタリ同型写像であることを思い出そう．$h\in\mathcal{H}$ および $s\in[0,T]$ のとき，$\langle h, U\mathbf{1}_{[0,s]}\rangle_{\mathcal{H}} = h_s$ となることに注意せよ．

Malliavin 解析の意味での勾配作用素（$\mathcal{H}$ 微分ともいう）D を導入する．実数値柱状関数 $F\in\mathcal{F}C^1_{\mathrm{b}}$ と $h\in\mathcal{H}$ に対して

$$D_hF(w) = \lim_{u\to 0}\frac{F(w+uh)-F(w)}{u} \tag{C.4.1}$$

として方向微分 D_h を定め，Riesz の表現定理により

$$\langle DF(w), h\rangle_{\mathcal{H}} = D_hF(w), \qquad h\in\mathcal{H}$$

として勾配作用素 D を定める．DF は $\mathcal{H}$ 値であることに注意せよ．F に対する表示式 (C.2.1) を用いて書き下すと

$$D_hF(w) = \sum_{k=1}^{m}(\partial_k f)(w_{t_1},\dots,w_{t_m})h_{t_k}, \quad DF(w) = \sum_{k=1}^{m}(\partial_k f)(w_{t_1},\dots,w_{t_m})\,U\mathbf{1}_{[0,t_k]} \tag{C.4.2}$$

と簡単に計算できる.[12] 本来 (C.4.1) によって定義されているので，これらは F の表示に使われている C^1_{b} 級関数 f の選び方には依存しない．D は $L^2(\mathbb{P})$ から $L^2(\mathbb{P};\mathcal{H})$ への線形作用素で，その定義域は $\mathcal{F}C^1_{\mathrm{b}}$ とする．なお (C.4.1) は各 w ごとに微分している式だが，もちろん $L^2(\mathbb{P})$ における微分だとも解釈できる．実際，平均値の定理により

$$\frac{F(\cdot+uh)-F}{u} - D_hF$$

[12] 1 次元なので，$\nabla_{x^k} = \partial_k$ である．

$$= \int_0^1 \sum_{k=1}^m \{(\partial_k f)(W_{t_1} + \theta u h_{t_1}, \ldots, W_{t_m} + \theta u h_{t_m}) - (\partial_k f)(W_{t_1}, \ldots, W_{t_m})\} h_{t_k} d\theta$$

となるので，有界収束定理により右辺は $u \to 0$ のとき $L^2(\mathbb{P})$ において 0 に収束する．よって (C.4.1) の右辺も $L^2(\mathbb{P})$ において収束する．

次は $(D, \mathcal{F}C_{\mathrm{b}}^1)$ の双対作用素 D^* を導入する．$\Xi \in L^2(\mathbb{P}; \mathcal{H})$ と $G \in L^2(\mathbb{P})$ に依存する条件

$$\mathbb{E}[\langle DF, \Xi \rangle_{\mathcal{H}}] = \mathbb{E}[FG], \qquad F \in \mathcal{F}C_{\mathrm{b}}^1 \tag{C.4.3}$$

を考える．与えられた $\Xi \in L^2(\mathbb{P}; \mathcal{H})$ に対して，条件 (C.4.3) を満たす G が存在する場合に Ξ は D^* の定義域に属すとし，その場合そのような G は一意的に決まるので $D^*\Xi = G$ と定める.[13] そのような G が存在しない場合は Ξ は D^* の定義域に属さないとする．稠密な定義域を持つ作用素の双対なので，D^* は閉作用素である．

Ξ が $\mathcal{H}$ 値の定数関数の場合は簡単に D^* を計算できる．$h \in \mathcal{H}$ のとき，Cameron-Martin 公式により

$$\begin{aligned}
\mathbb{E}[\langle DF, h \rangle_{\mathcal{H}}] &= \mathbb{E}[D_h F] = \lim_{u \to 0} \mathbb{E}\left[\frac{1}{u}\{F(\cdot + uh) - F\}\right] \\
&= \lim_{u \to 0} \mathbb{E}\left[F \frac{1}{u}\left\{\exp\left(u \int_0^T h_s' dW_s - \frac{u^2}{2}\|h\|_{\mathcal{H}}^2\right) - 1\right\}\right] \\
&= \mathbb{E}\left[F \int_0^T h_s' dW_s\right], \qquad F \in \mathcal{F}C_{\mathrm{b}}^1
\end{aligned}$$

となる．この式は $D^*h = \int_0^T h_s' dW_s$ を意味する.[14] なお最後の等号は Lebesgue の収束定理によるが，F の有界性，指数関数に対する平均値の定理，$\int_0^T h_s' dW_s$ が平均 0，分散 $\|h\|_{\mathcal{H}}^2$ の正規分布に従うという事実も併せて使っている．D が Leibniz 則を満たすので，

$$D^*(Fh) = FD^*h - \langle DF, h \rangle_{\mathcal{H}}, \qquad h \in \mathcal{H},\ F \in \mathcal{F}C_{\mathrm{b}}^1 \tag{C.4.4}$$

となることも簡単に示せる．

C_{b}^1 級柱状関数 F が $F(w) = f(w_{t_1}, \ldots, w_{t_m})$ と表示され，$t_m \leq s \leq t \leq T$ だとすると $D^*(F \cdot U\mathbf{1}_{[s,t]}) = F \cdot (w_t - w_s)$ である．実際，DF は $U\mathbf{1}_{[0,t_k]}$（$1 \leq k \leq m$）の 1 次結合で書けるが，$\langle U\mathbf{1}_{[0,t_k]}, U\mathbf{1}_{[s,t]} \rangle_{\mathcal{H}} = \langle \mathbf{1}_{[0,t_k]}, \mathbf{1}_{[s,t]} \rangle_{L^2([0,T])} = 0$ であるため，(C.4.4) を用いる際に右辺第 2 項が消えるからである．

単過程 $v = (v_t)_{t \in [0,T]} \in \mathcal{S}$ は (C.2.2) の形だとし，さらに各 k に対して $\xi_k \in \mathcal{F}_{t_k}C_{\mathrm{b}}^1$ とする．すると

$$D^*(Uv) = \sum_{k=0}^{m-1} \xi_k(w)(w_{t_{k+1}} - w_{t_k}) = \int_0^T v_s dW_s$$

[13] この G の一意性は，補題 C.2.1 で見たように $\mathcal{F}C_{\mathrm{b}}^1$ が $L^2(\mathbb{P})$ において稠密であることから従う．
[14] この場合は h を $L^2(\mathbb{P}; \mathcal{H})$ の元だと見なしている．

となる．上の条件を満たす単過程の全体は $\mathscr{A}$ で稠密なので（$\mathcal{S}$ の $\mathscr{A}$ における稠密性と注意 C.2.2 による），各 $v \in \mathscr{A}$ に対し上の条件を満たす単過程の列 $\{v^n\}$ で $\lim_{n\to\infty} \|v^n - v\|_{\mathscr{A}} = 0$ となるものが存在する．このとき定義から明らかに $\lim_{n\to\infty} \|Uv^n - Uv\|_{L^2(\mathbb{P};\mathcal{H})} = 0$ であり，伊藤の等長性により

$$\lim_{n\to\infty} \left\| \int_0^T v_s^n dW_s - \int_0^T v_s dW_s \right\|_{L^2(\mathbb{P})} = \lim_{n\to\infty} \|v^n - v\|_{\mathscr{A}} = 0$$

である．したがって，D^* の閉性により

$$D^*(Uv) = \int_0^T v_s dW_s, \qquad v \in \mathscr{A}$$

と求まる．

これで準備が整ったので，ここから本命題の証明を始める．(N_t) は有界な $\{\mathcal{F}_t\}$ マルチンゲールなので，伊藤の表現定理により,[15] ある $g \in \mathscr{A}$ が存在して，$\mathbb{P}$-a.s. に

$$N_t - N_0 = \mathbb{E}\left[F|\mathcal{F}_t\right] - \mathbb{E}\left[F\right] = \int_0^t g_s dW_s, \qquad t \in [0,T]$$

を満たす．$\mathscr{A}$ の元として $g = z$ であることを示すことが目標である．そのためには任意の $v \in \mathscr{A}$ に対して $\langle g, v\rangle_{\mathscr{A}} = \langle z, v\rangle_{\mathscr{A}}$ となることを見ればよい．

伊藤の等長性および D^* に関する上記の基本的な事実を用いて

$$\begin{aligned}\langle g, v\rangle_{\mathscr{A}} &= \mathbb{E}\left[\int_0^T g_s v_s ds\right] = \mathbb{E}\left[\int_0^T g_s dW_s \int_0^T v_s dW_s\right] = \mathbb{E}\left[(F - \mathbb{E}\left[F\right])D^*(Uv)\right] \\ &= \mathbb{E}\left[\langle D(F - \mathbb{E}\left[F\right]),\, Uv\rangle_{\mathcal{H}}\right] = \mathbb{E}\left[\langle DF,\, Uv\rangle_{\mathcal{H}}\right] = \mathbb{E}\left[\langle U^{-1}(DF),\, v\rangle_{L^2([0,T])}\right]\end{aligned}$$

を得る．(C.4.2) により $U^{-1}(DF) = \sum_{k=1}^m (\partial_k f)(w_{t_1}, \ldots, w_{t_m})\,\mathbf{1}_{[0,t_k]}$ となることに注意せよ．以下ではこの確率過程を $\widetilde{z}$ とおく．$\widetilde{z}$ は $\{\mathcal{F}_t\}$ 適合または発展的可測ではない．明らかに $\mathbb{E}\left[\widetilde{z}_s|\mathcal{F}_s\right] = z_s$ である．v_s の $\mathcal{F}_s$ 可測性により $\mathbb{E}\left[\widetilde{z}_s v_s\right] = \mathbb{E}\left[\mathbb{E}\left[\widetilde{z}_s v_s|\mathcal{F}_s\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\widetilde{z}_s|\mathcal{F}_s\right]v_s\right]$ が成り立つことに注意すると，

$$\mathbb{E}\left[\langle \widetilde{z},\, v\rangle_{L^2([0,T])}\right] = \int_0^T \mathbb{E}\left[\widetilde{z}_s v_s\right] ds = \mathbb{E}\left[\int_0^T \mathbb{E}\left[\widetilde{z}_s|\mathcal{F}_s\right] v_s ds\right] = \langle z, v\rangle_{\mathscr{A}}$$

を得る．以上をまとめて $\langle g, v\rangle_{\mathscr{A}} = \langle z, v\rangle_{\mathscr{A}}$ を得る．以上で命題 C.4.1 の証明が完成した．□

[15] 伊藤の表現定理については [8, 第 3.4.D 節]，[14, 第 2.6 節]，[13, 第 4.4 節] などを参照せよ．

補題 C.4.2. C_{b}^1 級柱状関数 F を (C.2.1) のとおりとし，$v=(v_t)_{t\in[0,T]}\in\mathscr{A}_{\mathrm{b}}$ を以下のように定める．

$$v_t=-\sum_{k=1}^m \mathbf{1}_{[0,t_k]}(t)\frac{\mathbb{E}\left[e^{-F(W)}\nabla_{x^k}f(W_{t_1},\dots,W_{t_m})|\mathcal{F}_t\right]}{\mathbb{E}\left[e^{-F(W)}|\mathcal{F}_t\right]}. \tag{C.4.5}$$

このとき，任意の $t\in[0,T]$ に対して

$$\mathbb{E}\left[e^{-F(W)}|\mathcal{F}_t\right]=\mathbb{E}\left[e^{-F(W)}\right]\mathcal{E}_t^v,\qquad \mathbb{P}\text{-a.s.} \tag{C.4.6}$$

が成立し，さらに

$$-\log\mathbb{E}\left[e^{-F(W)}\right]=\mathbb{E}^v\left[F(W)+\frac{1}{2}\int_0^T|v_s|^2ds\right] \tag{C.4.7}$$

も成立する．

証明. $M_t=\mathbb{E}\left[e^{-F(W)}|\mathcal{F}_t\right]$ とおく．(M_t) に対して命題 C.4.1 を適用すると，

$$M_t-M_0=\int_0^t u_s\cdot dW_s,\qquad \mathbb{P}\text{-a.s.}$$

が成り立つ．ただし以下のように定めた．

$$u_t=-\sum_{k=1}^m \mathbf{1}_{[0,t_k]}(t)\mathbb{E}\left[e^{-F(W)}\nabla_{x^k}f(W_{t_1},\dots,W_{t_m})|\mathcal{F}_t\right].$$

正定数 $0<c\le C<\infty$ が存在して $0<c\le\sup_{t\in[0,T]}M_t\le C$ が $\mathbb{P}$-a.s. に成り立つことに注意せよ．したがって $\log M_t$ を伊藤の公式を用いて

$$d\log M_t=\frac{dM_t}{M_t}-\frac{1}{2}\frac{d\langle M,M\rangle_t}{M_t^2}=\frac{u_t}{M_t}\cdot dW_t-\frac{1}{2}\frac{|u_t|^2}{M_t^2}dt$$

と計算できる．したがって，各 t に対して

$$\log(M_t/M_0)=\log M_t-\log M_0=\int_0^t\frac{u_s}{M_s}\cdot dW_s-\frac{1}{2}\int_0^t\frac{|u_s|^2}{M_s^2}ds\qquad \mathbb{P}\text{-a.s.}$$

となる．$M_0=\mathbb{E}\left[e^{-F(W)}\right]$ と $u_s/M_s=v_s$ に注意すると (C.4.6) が得られる．

次は (C.4.6) において $t=T$ を代入した後に対数を取る．$e^{-F(W)}$ は $\mathcal{F}_T$ 可測なので，

$$-F(W)=\log\mathbb{E}\left[e^{-F(W)}\right]+\log\mathcal{E}_T^v$$

となる．これを移項して $\mathbb{P}^v$ に関する期待値を取ると

$$-\log\mathbb{E}\left[e^{-F(W)}\right]=\mathbb{E}^v\left[F(W)\right]+\mathbb{E}^v\left[\log\mathcal{E}_T^v\right]=\mathbb{E}^v\left[F(W)\right]+\frac{1}{2}\mathbb{E}^v\left[\int_0^T|v_s|^2ds\right]$$

となる．最後の式変形は補題 C.3.1 の証明中の計算と全く同じである．これで (C.4.7) も得られた． □

以下では $F\colon \mathcal{C}_0 \to \mathbb{R}$ を任意の有界 Borel 可測関数とする．この F に対して $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{F}C_{\mathrm{b}}^1$ を F に $L^1(\mathbb{P})$ 収束する一様有界な関数列とする．(このような柱状関数列の存在は補題 C.2.3 で示した．) 各 F_n に対して $v^n \in \mathscr{A}_{\mathrm{b}}$ を (C.4.5) のように定める．すると補題 C.4.2 により，(C.4.7) において (F, v) を (F_n, v^n) に置き換えた式が全ての n で成立する．その結果，

$$-\log \mathbb{E}\left[e^{-F(W)}\right] = I_n^1 + I_n^2 + I_n^3 \tag{C.4.8}$$

と分解できる．ただし，次のように定めた．

$$\begin{aligned}
I_n^1 &:= -\log \mathbb{E}\left[e^{-F(W)}\right] + \log \mathbb{E}\left[e^{-F_n(W)}\right],\\
I_n^2 &:= \mathbb{E}^{v^n}\left[F_n(W) - F(W)\right],\\
I_n^3 &:= \mathbb{E}^{v^n}\left[F(W) + \frac{1}{2}\int_0^T |v_s^n|^2 ds\right].
\end{aligned}$$

補題 C.4.3. 上の状況において，$\lim_{n\to\infty} I_n^1 = 0 = \lim_{n\to\infty} I_n^2$ となる．

証明. 一様有界性を仮定したので，$M := \sup_{n\in\mathbb{N}} \|F_n\|_\infty \vee \|F\|_\infty < \infty$ である．簡単な計算で

$$|e^{z_1} - e^{z_2}| \le e^M |z_1 - z_2|, \qquad z_1, z_2 \in [-M, M]$$

と評価できるので，$n \to \infty$ のとき

$$|\mathbb{E}\left[e^{-F(W)}\right] - \mathbb{E}\left[e^{-F_n(W)}\right]| \le e^M \mathbb{E}\left[|F(W) - F_n(W)|\right] = e^M \|F - F_n\|_{L^1(\mathbb{P})} \to 0$$

となる．log の連続性により $\lim_{n\to\infty} I_n^1 = 0$ がわかる．

次は I_n^2 の極限を計算する．十分大きな n に対して $\mathbb{E}\left[F_n(W)\right] \le \mathbb{E}\left[F(W)\right] + 1$ となる．指数関数の凸性と Jensen の不等式により

$$\mathbb{E}\left[e^{-F_n(W)}\right] \ge e^{\mathbb{E}[-F_n(W)]} \ge e^{-\mathbb{E}[F(W)]-1}$$

を得る．補題 C.4.2 の (C.4.6) を用いると

$$\mathcal{E}_T^{v^n} = \frac{e^{-F_n(W)}}{\mathbb{E}\left[e^{-F_n(W)}\right]} \le e^{M+\mathbb{E}[F(W)]+1}$$

となる．したがって，十分大きな n に対して

$$|I_n^2| \le \mathbb{E}\left[|F_n(W) - F(W)|\mathcal{E}_T^{v^n}\right] \le \|F - F_n\|_{L^1(\mathbb{P})} e^{M+\mathbb{E}[F(W)]+1}$$

と評価できるが，$\{F_n\}$ は F を $L^1(\mathbb{P})$ 位相で近似する列なので，右辺は $n \to \infty$ のとき 0 に収束する．これで $\lim_{n\to\infty} I_n^2 = 0$ も示せた． □

以上により変分表示公式の下からの評価を示すには，次の不等式

$$\inf_{n\in\mathbb{N}} I_n^3 \geq \inf_{v\in\mathscr{A}} \mathbb{E}\left[F(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right]$$

を示せば十分であることがわかる．少し主張を強めて，次の補題を証明しよう．

補題 C.4.4. $F\colon \mathcal{C}_0 \to \mathbb{R}$ を有界 Borel 可測関数とし，$v \in \mathscr{A}_{\mathrm{b}}$ とする．このとき次の評価が成立する．

$$\mathbb{E}^v\left[F(W) + \frac{1}{2}\int_0^T |v_s|^2 ds\right] \geq \inf_{u\in\mathcal{S}} \mathbb{E}\left[F(\mathcal{T}^u(W)) + \frac{1}{2}\int_0^T |u_s|^2 ds\right].$$

この補題を認めると本節の主目的である変分表示公式（命題 C.0.2）の下からの評価が直ちに得られることを確認しよう．

補題 C.4.5. $F\colon \mathcal{C}_0 \to \mathbb{R}$ を有界な Borel 可測関数とし，$v \in \mathscr{A}$ とする．このとき次が成立する．

$$-\log \mathbb{E}\left[e^{-F(W)}\right] \geq \inf_{v\in\mathscr{A}} \mathbb{E}\left[F(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right].$$

証明. 分解の式 (C.4.8) と補題 C.4.4 と包含 $\mathcal{S} \subset \mathscr{A}$ により

$$-\log \mathbb{E}\left[e^{-F(W)}\right] \geq I_n^1 + I_n^2 + \inf_{v\in\mathscr{A}} \mathbb{E}\left[F(\mathcal{T}^v(W)) + \frac{1}{2}\int_0^T |v_s|^2 ds\right]$$

となる．$n \to \infty$ として補題 C.4.3 を用いると証明が終わる． □

本節の残りでは補題 C.4.4 を証明する．（よって，以下では補題 C.4.4 の仮定の下で議論する．）そのための準備として，まず補題をいくつか示す．

$v \in \mathscr{A}_{\mathrm{b}}$ とする．これはある $R > 0$ に対して $v \in \mathscr{A}_R$ と同値である．このとき補題 C.2.6 により，ある $\{v^n\}_{n\in\mathbb{N}} \subset \mathscr{A}_R \cap \mathcal{S}$ が存在して $\lim_{n\to\infty} \|v^n - v\|_{\mathscr{A}} = 0$ となる．このとき補題 C.4.4 の左辺は以下のように分解できる．

$$\mathbb{E}^v\left[F(W) + \frac{1}{2}\int_0^T |v_s|^2 ds\right] = J_n^1 + \frac{1}{2}J_n^2 + J_n^3. \tag{C.4.9}$$

ただし，次のように定めた．

$$\begin{aligned}
J_n^1 &:= \mathbb{E}\left[F(W)(\mathcal{E}_T^v - \mathcal{E}_T^{v^n})\right], \\
J_n^2 &:= \mathbb{E}^v\left[\int_0^T |v_s|^2 ds\right] - \mathbb{E}^{v^n}\left[\int_0^T |v_s^n|^2 ds\right], \\
J_n^3 &:= \mathbb{E}^{v^n}\left[F(W) + \frac{1}{2}\int_0^T |v_s^n|^2 ds\right].
\end{aligned}$$

補題 C.4.6. 上記の設定の下で $\lim_{n\to\infty}\mathbb{E}\left[|\mathcal{E}_T^{v^n}-\mathcal{E}_T^{v}|\right]=0$ となる．特に $\lim_{n\to\infty}J_n^1=0$ である．

証明. $\{n'\}\subset\mathbb{N}$ を $\mathbb{N}$ の任意の部分列とするとき，さらなる部分列 $\{n''\}\subset\{n'\}$ が存在して $\lim_{n''\to\infty}\mathbb{E}\left[|\mathcal{E}_T^{v^{n''}}-\mathcal{E}_T^{v}|\right]=0$ となることを見れば十分である．なぜならば，これは $\{\mathcal{E}_T^{v^n}\}$ が $L^1(\mathbb{P})$ において相対コンパクト列であることと，集積点が $\mathcal{E}_T^v$ のみであることを意味するからである．

さて $\{v\}\cup\{v^n\}_{n\in\mathbb{N}}\subset\mathscr{A}_R$ および補題 C.1.1 により，$\{|\mathcal{E}_T^{v^n}-\mathcal{E}_T^{v}|\}_{n\in\mathbb{N}}$ は $L^2(\mathbb{P})$ において有界であるので，一様可積分である．したがって $\mathbb{P}$-a.s. に $\lim_{n''\to\infty}|\mathcal{E}_T^{v^{n''}}-\mathcal{E}_T^{v}|=0$ となることを示せば，本補題の１つ目の主張が従う.[16]

まず明らかに $n'\to\infty$ のとき

$$\mathbb{E}\left[\left|\int_0^T v_s^{n'}\cdot dW_s-\int_0^T v_s\cdot dW_s\right|^2\right]=\|v^{n'}-v\|_{\mathscr{A}}^2\to 0$$

となる．さらに

$$\begin{aligned}\left|\|v^{n'}\|_{L^2([0,T])}^2-\|v\|_{L^2([0,T])}^2\right|&\le\left(\|v^{n'}\|_{L^2([0,T])}+\|v\|_{L^2([0,T])}\right)\left|\|v^{n'}\|_{L^2([0,T])}-\|v\|_{L^2([0,T])}\right|\\&\le 2R\|v^{n'}-v\|_{L^2([0,T])},\qquad\mathbb{P}\text{-a.s.}\end{aligned}\tag{C.4.10}$$

であるため，$n'\to\infty$ のとき

$$\mathbb{E}\left[\left|\int_0^T|v_s^{n'}|^2ds-\int_0^T|v_s|^2ds\right|\right]\le 2R\mathbb{E}\left[\|v^{n'}-v\|_{L^2([0,T])}\right]\le 2R\|v^{n'}-v\|_{\mathscr{A}}\to 0$$

となる．２つ目の不等式は Schwarz の不等式と $\|\cdot\|_{\mathscr{A}}$ の定義による．したがって，ある部分列 $\{n''\}\subset\{n'\}$ を抜き出せば，$n''\to\infty$ のときに

$$\begin{aligned}\mathcal{E}_T^{v^{n''}}&=\exp\left(\int_0^T v_s^{n''}\cdot dW_s-\frac{1}{2}\int_0^T|v_s^{n''}|^2ds\right)\\&\to\mathcal{E}_T^v=\exp\left(\int_0^T v_s\cdot dW_s-\frac{1}{2}\int_0^T|v_s|^2ds\right),\qquad\mathbb{P}\text{-a.s.}\end{aligned}$$

となることがわかる．F は有界なので，$\lim_{n\to\infty}J_n^1=0$ は直ちに従う．　□

補題 C.4.7. 上記の設定の下で，$\lim_{n\to\infty}J_n^2=0$ となる．

証明. ３角不等式を用いた簡単な計算で

$$|J_n^2|\le\mathbb{E}\left[\left(\int_0^T|v_s^n|^2ds\right)|\mathcal{E}_T^{v^n}-\mathcal{E}_T^v|\right]+\mathbb{E}\left[\left|\int_0^T|v_s^n|^2ds-\int_0^T|v_s|^2ds\right|\mathcal{E}_T^v\right]$$

[16]一様可積分性に関する基本事項については [13, 第 1.2 節] や [11, 第 2.6 節] などを参照せよ．

を得る．右辺第 1 項は上から $R^2\mathbb{E}\left[|\mathcal{E}_T^{v^n}-\mathcal{E}_T^v|\right]$ で評価されるので，補題 C.4.6 により $n\to\infty$ のときに 0 に収束する．右辺第 2 項は補題 C.4.6 の証明中にある (C.4.10) とほとんど同じ計算をすることにより

$$2R\mathbb{E}\left[\|v^n-v\|_{L^2([0,T])}\mathcal{E}_T^v\right]\le 2R\mathbb{E}\left[\|v^n-v\|_{L^2([0,T])}^2\right]^{\frac{1}{2}}\mathbb{E}\left[(\mathcal{E}_T^v)^2\right]^{\frac{1}{2}}\le 2Re^{R^2/2}\|v^n-v\|_{\mathscr{A}}$$

と上から評価できるので，やはり $n\to\infty$ のときに 0 に収束する．ここで $\mathcal{E}_T^v$ の積率評価（補題 C.1.1）も使った． □

補題 C.4.4 の証明. 上記の $v^n\in\mathcal{S}$ に対して，補題 C.2.5 の (C.2.4) を満たす $\overline{v}^n\in\mathcal{S}$ を取る．すると

$$\begin{aligned}
J_n^3 &= \mathbb{E}^{v^n}\left[F(W)+\frac{1}{2}\int_0^T|v_s^n|^2ds\right]\\
&= \mathbb{E}^{v^n}\left[F(\mathcal{T}^{\overline{v}^n}\circ\mathcal{T}^{-v^n}(W))+\frac{1}{2}\int_0^T|\overline{v}_s^n(\mathcal{T}^{-v^n}(W))|^2ds\right]\\
&= \mathbb{E}\left[F(\mathcal{T}^{\overline{v}^n}(W))+\frac{1}{2}\int_0^T|\overline{v}_s^n|^2ds\right]\\
&\ge \inf_{u\in\mathcal{S}}\mathbb{E}\left[F(\mathcal{T}^u(W))+\frac{1}{2}\int_0^T|u_s|^2ds\right]
\end{aligned}$$

が成立する． 3 つ目の等号で (C.1.1) を用いた．この不等式と (C.4.9) から

$$\mathbb{E}^v\left[F(W)+\frac{1}{2}\int_0^T|v_s|^2ds\right]\ge J_n^1+\frac{1}{2}J_n^2+\inf_{u\in\mathcal{S}}\mathbb{E}\left[F(\mathcal{T}^u(W))+\frac{1}{2}\int_0^T|u_s|^2ds\right]$$

が従う．補題 C.4.6 と補題 C.4.7 により $n\to\infty$ のとき $J_n^1+(1/2)J_n^2\to 0$ となるので，これで補題 C.4.4 の証明が終わった． □

関連図書

[1] 池田信行，小倉幸雄，高橋陽一郎，眞鍋昭治郎,「確率論入門 II」，培風館，2015.

[2] 伊藤清,「確率論」，岩波書店，1991.

[3] 伊藤清三,「ルベーグ積分入門」，裳華房，1963.

[4] 稲垣宣生,「数理統計学」，裳華房，1990.

[5] 稲濱譲,「ラフパス理論と確率解析」，岩波書店，2022.

[6] 内田伏一,「集合と位相」，裳華房，1986.

[7] 梅原壽春,「情報理論の基礎－関数解析的展開－」，サイエンス社，1993.

[8] I. カラザス，S. E. シュレーブ,「ブラウン運動と確率積分」，シュプリンガー・フェアラーク東京，2001.

[9] 小谷眞一,「測度と確率１」，岩波書店，1997.

[10] 熊谷隆,「確率論」，共立出版，2003.

[11] 楠岡成雄,「確率解析」，知泉書館，2018.

[12] 重川一郎,「確率解析」，岩波書店，1998.

[13] 谷口説男,「確率微分方程式」，共立出版，2016.

[14] 谷口説男，松本裕行,「確率解析」，培風館，2013.

[15] 中島誠,「例と演習で学ぶ確率論」，講談社，2024.

[16] 長井英生,「確率微分方程式」，共立出版，1999.

[17] 福島正俊，竹田雅好,「マルコフ過程」，培風館，2008.

[18] 舟木直久,「確率微分方程式」，岩波書店，1997.

[19] 松坂和夫,「集合・位相入門」，岩波書店，1968.

[20] Rick Durrett 著,「デュレット確率論」, 森北出版, 2024.

[21] 森田茂之,「集合と位相空間」, 朝倉書店, 2002.

[22] 盛田健彦,「実解析と測度論の基礎」, 2004.

[23] 渡辺信三,「確率微分方程式」, 産業図書, 1975.

[24] 渡辺信三（編）, 重川一郎（編）,「確率論ハンドブック」, 2012.

[25] 日本数学会（編）,「岩波　数学辞典　第４版」, 岩波書店, 2007.

[26] Anderson, G. W., Guionnet, A. and Zeitouni, O.; An introduction to random matrices. Cambridge University Press, Cambridge, 2010. xiv+492 pp.

[27] Billingsley, P.; Convergence of probability measures. Second edition. John Wiley & Sons, Inc., New York, 1999. x+277 pp.

[28] Bogachev, V.; Gaussian measures. American Mathematical Society, Providence, RI, 1998. xii+433 pp.

[29] Budhiraja, A. and Dupuis, P.; Analysis and approximation of rare events. Springer, New York, 2019. xix+574 pp.

[30] Decreusefond, L.; Selected topics in Malliavin calculus—chaos, divergence and so much more. Bocconi University Press, Springer, Cham, 2022. xi+172 pp.

[31] den Hollander, F.; Large deviations. American Mathematical Society, Providence, RI, 2000. x+143 pp.

[32] Deuschel, J.D. and Stroock, D.W.; Large deviations. Academic Press, Inc., Boston, MA, 1989. xiv+307 pp.

[33] Dembo, A. and Zeitouni, O.; Large deviations techniques and applications. Second edition. Springer-Verlag, New York, 1998. xvi+396 pp.

[34] Dupuis, P. and Ellis, R.; A weak convergence approach to the theory of large deviations. John Wiley & Sons, Inc., New York, 1997. xviii+479 pp.

[35] Ellis, R.; Entropy, large deviations, and statistical mechanics. Springer-Verlag, New York, 1985. xiv+364 pp.

[36] Feng, J. and Kurtz, T.; Large deviations for stochastic processes. American Mathematical Society, Providence, RI, 2006. xii+410 pp.

[37] Friz, P. and Victoir, N.; Multidimensional stochastic processes as rough paths. Cambridge University Press, Cambridge, 2010. xiv+656 pp.

[38] Hiai, F. and Petz, D.; The semicircle law, free random variables and entropy. American Mathematical Society, Providence, RI, 2000. x+376 pp.

[39] Ikeda, N. and Watanabe, S.; Stochastic differential equations and diffusion processes. Second edition. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. xvi+555 pp.

[40] Klenke, A.; Probability theory—a comprehensive course. Third edition. Springer, Cham, 2020, 2020. xiv+716 pp.

[41] Olivieri, E. and Vares, M.E.; Large deviations and metastability. Cambridge University Press, Cambridge, 2005. xvi+512 pp.

[42] Rassoul-Agha, F. and Seppalainen, T.; A course on large deviations with an introduction to Gibbs measures. American Mathematical Society, Providence, RI, 2015. xiv+318 pp.

[43] Shwartz, A. and Weiss A.; Large deviations for performance analysis. Chapman & Hall, London, 1995. x+556 pp.

[44] Stroock, D.W.; Probability theory, an analytic view. Second edition. Cambridge University Press, Cambridge, 2011. xxii+527 pp.

[45] Stroock, D.W.; Gaussian measures in finite and infinite dimensions. Springer, Cham, 2023. xii+144 pp.

[46] Varadhan, S.R.S.; Large deviations and applications. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984, v+75 pp.

[47] Varadhan, S.R.S.; Large deviations. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2016. vii+104 pp.

[48] Azencott, R.; Grandes déviations et applications. Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978), pp. 1–176 Lecture Notes in Math., 774, Springer, Berlin, 1980.

[49] Boué, M. and Dupuis, P.; A variational representation for certain functionals of Brownian motion. Ann. Probab. **26** (1998), no. 4, 1641–1659.

[50] Csiszár, I.; A simple proof of Sanov's theorem. Bull. Braz. Math. Soc. (N.S.) **37** (2006), no. 4, 453–459.

[51] Hiai, F.; A concise exposition of large deviations. Real and stochastic analysis, 183–267. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014

[52] Itô, K. and Nisio, M.; On the convergence of sums of independent Banach space valued random variables. Osaka Math. J. **5** (1968), 35–48.

[53] Wu, L.; Large deviations, moderate deviations and LIL for empirical processes. Ann. Probab. **22** (1994), no. 1, 17–27.

[54] Zhang, X.; A variational representation for random functionals on abstract Wiener spaces. J. Math. Kyoto Univ. **49** (2009), no. 3, 475–490.