



Title	特異モデルに対する双対平坦構造とその応用
Author(s)	中島, 直道
Citation	北海道大学. 博士(情報科学) 甲第15534号
Issue Date	2023-03-23
DOI	10.14943/doctoral.k15534
Doc URL	http://hdl.handle.net/2115/89375
Type	theses (doctoral)
File Information	Naomichi_Nakajima.pdf



[Instructions for use](#)

博士論文

特異モデルに対する双対平坦構造とその応用

中島 直道

2023年3月

北海道大学大学院情報科学院情報科学専攻
情報理工学コース 応用特異点論研究室

特異モデルに対する双対平坦構造とその応用

中島 直道

目次

第1章 導入	3
第2章 ヘッセ構造	7
2.1 ヘッセ多様体の理論	7
2.1.1 ヘッセ構造	7
2.1.2 双対ヘッセ構造	10
2.1.3 ルジャンドル変換と双対ポテンシャル	12
2.1.4 ヘッセ構造のダイバージェンス	14
2.2 統計多様体と双対平坦構造	16
2.2.1 双対平坦構造	17
2.2.2 コントラスト関数	18
第3章 接触構造と波面	23
3.1 接触構造	23
3.2 ルジャンドル写像と波面	27
3.3 ルジャンドルファイブレーション	28
3.4 ルジャンドル写像のリーマン幾何的特徴づけと連接接束	31
第4章 概ヘッセ構造	35
4.1 概ヘッセ多様体の理論	35
4.1.1 ダブルファイブレーション構造と双対平坦構造	35
4.1.2 e/m -波面と e/m -コースティクス	38
4.1.3 e/m -連接接束	42
4.1.4 概ヘッセ多様体	47
4.1.5 3次テンソルと α -幾何学	48
4.2 アファイン座標系における e/m -波面の標準形	52
4.2.1 標準形	53
4.2.2 判定法の内在的表現	59
第5章 ダイバージェンス	61
5.1 拡張ピタゴラスの定理と射影定理	61
5.1.1 e/m -曲線	61

5.1.2	正準ダイバージェンスと拡張ピタゴラスの定理	63
5.2	弱コントラスト関数	66
5.3	e/m -波面の特異点の判定法と正準ダイバージェンス	67
第 6 章	応用可能性の検討	71
6.1	深層学習と統計的推定	71
6.2	非凸最適化問題	73
第 7 章	結論	75
	謝辞	77
	参考文献	79

第 1 章

導入

情報幾何学とは、1980年代に甘利によって創設された学問分野であり、統計モデルのパラメータ空間 M をフィッシャー・ラオ計量で与えられるリーマン計量 h とある種の双対性を満たす二種類の TM 上のアファイン接続 ∇, ∇^* を備えた空間 (M, h, ∇, ∇^*) とみなすことで、統計学や機械学習等における種々の解析を幾何学的な視点から行うものである。このような双対的な構造を持つ空間は統計多様体と呼ばれる [31]。統計多様体 (M, h, ∇, ∇^*) であって、接続 ∇, ∇^* が共に平坦であるものを双対平坦多様体といい、その上の構造 (h, ∇, ∇^*) を双対平坦構造という。甘利・長岡 [2] によって導入された双対平坦構造は情報幾何学において最も重要な空間概念であり、この空間構造を通すことで、情報幾何学は統計学や機械学習、凸最適化理論、量子情報理論などの様々な分野に対して統一的な幾何学的解釈をもたらす [1, 3, 9, 12]。

双対平坦多様体の概念を定義する場合には、統計多様体の特殊な場合として定義する文脈と、アファイン微分幾何学におけるヘッセ多様体として定義する文脈の二つがある。ヘッセ多様体 (M, h, ∇) とは平坦接続 ∇ を備えた擬リーマン多様体 (M, h) であって、計量 h が局所的にアファイン座標系上のある関数によるヘッセ行列としてかけるものである。このとき h はヘッセ計量と呼ばれる。ヘッセ多様体の幾何学はコシュール (Koszul) に端を発し、ケーラー幾何学や等質空間論などとの密接な関わりの中で研究されている [37]。いずれにしても双対平坦多様体とヘッセ多様体は同一の概念であり、本論文内では文脈に応じてこれらの用語を使い分けるが同じものを扱っていることに注意する。

双対平坦多様体 (M, h, ∇, ∇^*) の持つ最も顕著な特徴は、擬リーマン計量 h が局所アファイン座標系におけるある関数のヘッセ行列として表現されることである。このような関数を (局所) ポテンシャル関数と呼ぶ。さらに、二種類の平坦アファイン接続 ∇, ∇^* それぞれにおける局所ポテンシャル関数 f, φ はルジャンドル双対性により結びついている。このとき φ は f のルジャンドル変換あるいは双対ポテンシャル関数と呼ばれる。情報幾何学では特に計量が正定値である場合に興味が向けられ (フィッシャー・ラオ計量は

半正定値である), このときポテンシャル関数は凸関数である.

情報幾何学の応用は, 双対平坦多様体 M 上で成り立つ拡張ピタゴラスの定理と射影定理の二つの定理によって支えられている [1, 3]. しかしながら, 深層学習を含む様々な実応用においてフィッシャー計量の退化する状況がしばしば現れる [18, 43]. すなわち, 局所ポテンシャルは非凸で変曲的あるいは多価になり得る. この場合には情報幾何学における微分幾何学的な手法は適用できず, 双対平坦構造を定義することはできない. このような空間を特異モデルと呼ぶことにする.

本論文の目的は, 特異モデルに対して情報幾何学的な応用を可能とするべく, 甘利・長岡による理論をその基礎の部分から刷新することである. 我々は双対平坦多様体の一般化として退化し得る対称 $(0, 2)$ -テンソル (§4.1.3 にて概ヘッセ計量と呼ぶ) を備えた多様体である概ヘッセ多様体 (quasi-Hessian manifold) を導入することで, 双対平坦多様体の理論を抜本から刷新し, 計量が退化する場合における一般化した理論を構築する. 我々はこの理論を概ヘッセ多様体の理論としてまとめ, 双対平坦多様体における描像が, 計量が退化していてもなお概ヘッセ多様体上の描像として得られることを示す. そこでは拡張ピタゴラスの定理や射影定理が適切に定式化され, 従って従来のこれらの定理に基づく応用的観点は空間に特異点があってもなお正当化される. これは我々の理論がポテンシャル関数の凸性に依拠しない新しい理論であることを意味する. さらに, 概ヘッセ計量の退化性を特徴づける特異点の解析を行うことで, 我々の理論の深化を図る.

我々の構成は接触幾何学およびルジャンドル特異点論に立脚するものである. 接触幾何学は元々リー (Lie), ダルブー (Darboux), グルサ (Goursat) らによって 19 世紀中頃から 1 階線形偏微分方程式論を研究するために開発された分野であり [22], 局所的には変曲的あるいはより一般に非凸であったり, または多価であったりするようなポテンシャル関数 (のグラフ) を波面と呼ばれる特異点を持った超曲面として捉える幾何学である. またルジャンドル特異点論は 1980 年代にアーノルド (Arnol'd) やヘルマンダー (Hörmander), ザカリューキン (Zakalyukin) らによって確立された理論であり, 接触幾何学の枠組みで扱われる波面の特異性を記述するものである [22, 21]. 接触幾何学とルジャンドル特異点論は力学や幾何光学, 熱力学, 制御理論, カタストロフ理論等の実に多くの様々な分野との関連の中で発達してきたものである [6, 22, 5, 7, 8, 23, 34]. 我々の構成において本質的な役割を果たすものは波面, すなわち非凸ポテンシャルあるいは多価ポテンシャルに対しても依然として働くルジャンドル双対性であり, これは上記の枠組みで記述されるものである.

一般にポテンシャル関数が非凸であればその双対ポテンシャル関数のグラフはもはや滑らかな多様体にはならず (しかしながらこれは波面である), このとき接束上の接続は意味を持たない. 我々の提案する概ヘッセ多様体 M は退化し得る対称 $(0, 2)$ -テンソル h が定義された多様体であり, 局所的にルジャンドル双対性を満たす二種類の波面 (§4.1.2 に

て e/m -波面と呼ぶ) を備えている; これらは凸ポテンシャルとその双対ポテンシャルの役割を担う. このときこれらの波面から自然に定義される互いに双対的な二種類の連接接束 $(E, \Phi : TM \rightarrow E, \nabla^E)$, $(E', \Phi' : TM \rightarrow E', \nabla^{E'})$ を考える. ここで, E, E' はそれぞれ接束 TM の代替物として機能する M 上のベクトル束であり, TM との比較写像であるベクトル束準同型写像 Φ, Φ' が備わっている. さらにこれらベクトル束の上には平坦接続 ∇^E および $\nabla^{E'}$ が自然に定義される. 連接接束の概念は波面のリーマン幾何学を研究する上で佐治・梅原・山田によって導入されたものであり, 波面の持つ内在的な構造が定式化されたものである [35]. 我々はこの概念のいわばアファイン微分幾何学版と呼ぶべきものを導入し, これを接束の代替物として用いる. 概ヘッセ計量 h が M 上の全ての点で非退化であるとき束写像 Φ, Φ' は同型写像になり ($E = E' = TM$), 概ヘッセ多様体 $(M, h, (E, \Phi, \nabla^E), (E', \Phi', \nabla^{E'}))$ は従来の意味の双対平坦多様体を復元することが確かめられる (命題 4.1.17).

さらに, 双対平坦多様体における甘利・チェンツォフ (Chentsov) の 3 次テンソルは, 計量 h が退化していてもなお我々の概ヘッセ多様体上にて定義される対称 $(0, 3)$ -テンソル C に一般化される (定義 4.1.28). ここで TM 上の接続は存在しないことに注意する.

概ヘッセ多様体 M 上で定義されたテンソル h および C は江口によるコントラスト関数の理論と実に良く整合する [16]. 一般に多様体上のコントラスト関数とは, ある条件を満たす“距離関数”であり, コントラスト関数は対角集合上での高階微分により多様体上に様々なテンソルを導入する. 江口はこのようなコントラスト関数が導く多様体上の幾何学を徹底的に調べており, 特にコントラスト関数の 4 階までの微分を用いて定まるテンソルは重要であり, 2 階および 3 階微分を用いて定義されるテンソルは多様体上に自然に統計多様体の構造を誘導する [16, 31]. 双対平坦多様体の各アファイン座標系上の凸ポテンシャルはブレグマンダイバージェンスを定め, これは座標近傍上のコントラスト関数である. 各ブレグマンダイバージェンスはアファイン座標系の張り合わせに関して両立的であり, 従って, 双対平坦多様体上のコントラスト関数が自然に定まる (定義 2.1.13). これは双対平坦多様体のダイバージェンスと呼ばれ, 双対平坦多様体の幾何構造と密接に関わる. このダイバージェンスによる描像は, ダイバージェンスが概ヘッセ多様体 M 上の正準ダイバージェンス $\mathcal{D}_M : M \times M \rightarrow \mathbb{R}$ (§5.2 にて弱コントラスト関数と呼ばれる) に一般化されることで, 概ヘッセ多様体上においても極めて自然な形で得られる (定理 5.2.1, 5.3.1). さらに甘利・長岡による拡張ピタゴラスの定理および射影定理が我々の特異的設定においても定式化される (定理 5.1.7, 5.1.8).

概ヘッセ多様体に局所的に備わる e/m -波面, すなわちポテンシャルおよび双対ポテンシャルのグラフは一般にはアファイン空間内の特異点を持った超曲面である. 概ヘッセ計量 h の退化性はこれらの波面の特異点によって特徴づけられるものであり, 概ヘッセ多様体の一つの特徴づけとして e/m -波面の特異点を調べることは重要である. その解析の

一つとして本論文の中では e/m -波面の持つ典型的な特異点の C^∞ 型を特徴づける概ヘッセ多様体上の幾何学的諸量を用いた座標系によらない判定法を与え、アファイン座標系におけるその標準形（すなわちポテンシャル関数のグラフ）を導出する．この標準形の導出はまた次の二つの研究と関連するものである．一つ目は佐治・梅原・山田による波面のリーマン幾何学である．彼らはカスプ辺型やツバメの尾型の特異点に対する使い勝手の良い判定法を与え、特異曲率の概念などを導入することでこれらの特異点のリーマン幾何学的特徴づけを行っている [27, 35]．彼らの研究と同様にそのアファイン微分幾何学版も豊富な理論であることが考えられる．我々の導出するアファイン座標系における標準形は波面の特異点周りでのアファイン微分幾何学的な解析（および情報幾何学的な解析）を可能とするものであり、その理論を開拓する上での基礎になるだろう．二つ目は非凸最適化理論と変分法の文脈にてエクラン (Ekeland) によって 1970 年代に行われた変曲点における波面の標準形の導出に関する研究である [15]．実際、我々の導出する標準形は二つあり、一つ目の標準形 (定理 4.2.2) はカスプ辺型の特異点を記述するものであり、これはエクランによる結果 [15] の一般化かつ精密化を行ったものである．二つ目の標準形 (定理 4.2.5) は概ヘッセ計量が半正定値であるときに現れる特異点を記述するものである（例えば、フィッシャー・ラオ計量は半正定値である）．

このように我々の理論は本来のモチベーションである情報幾何学的な実応用に関する分野のみならず、数学における様々な領域とも密接に関わりを持つ理論である．

本論文の構成は以下の通りである．第2章ではヘッセ多様体の理論をまとめる．この章で扱う内容が我々の理論と対になるものである．ここではヘッセ多様体のポテンシャル関数が導入する双対構造やダイバージェンスについて議論する．また §2.2 では、統計多様体の文脈から導入される双対平坦多様体とヘッセ多様体の同値性が与えられ、統計多様体の幾何学の観点からの特徴づけを確認する．第3章では、我々の構成において重要な役割を果たす接触幾何学についてのまとめを行う．この章で定義される接触多様体とルジャンドル部分多様体は概ヘッセ多様体を構成する上で極めて重要な概念である．第4章では、概ヘッセ多様体を導入し、その幾何学的な性質を調べる．この章の始めではヘッセ多様体における局所的な幾何構造をルジャンドル部分多様体におけるそれとして捉える．この視点が一般の退化した計量をもつ概ヘッセ多様体の構築に繋がる．第5章では、双対平坦多様体のダイバージェンスを概ヘッセ多様体の正準ダイバージェンスに一般化し、その性質や拡張ピタゴラスの定理および射影定理について議論する．第6章では我々の理論の応用可能性について議論し、第7章で結論を述べ、本論文の締めとする．

本論文を通して、太字は列ベクトルを表すとし (e.g., $\mathbf{x} = (x_1, \dots, x_n)^T$)、プライム記号 \mathbf{x}' は単に \mathbf{x} と区別するために用いる (微分や転置のような操作は意味しない)．また、記号の簡略化のため、 $\frac{\partial f}{\partial \mathbf{x}}$ で $(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})^T$ を表すとする．また、多様体と写像のクラスは C^∞ を仮定する．

第 2 章

ヘッセ構造

この章ではヘッセ多様体の理論としてヘッセ構造の持つ性質や双対構造についてまとめる。ヘッセ多様体は特に二種類の局所的なポテンシャル関数を備え、これらはルジャンドル双対性により結びついている。また、ヘッセ多様体上にはダイバージェンスが定まり、拡張ピタゴラスの定理と射影定理が成り立つことを述べる。これらの議論の大部分は [37] に沿ったものである。さらに統計多様体の文脈からは、接続が平坦である統計多様体としてヘッセ多様体を得られ、この場合それは双対平坦多様体と呼ばれる。統計多様体の幾何学は江口によるコントラスト関数の理論が密接に関わっており [16]、最後の節ではこの理論と双対平坦多様体を関連を述べる。

以降、多様体や接続等については既知とする ([22, 37] 等を参照)。

2.1 ヘッセ多様体の理論

2.1.1 ヘッセ構造

M を n 次元多様体とし、 ∇ を接束 TM の接続とする。 ∇ の捩率テンソル T と曲率テンソル R が共に消えているとき、すなわち $T = 0$ かつ $R = 0$ なるとき ∇ を平坦アファイン接続といい、 (M, ∇) を平坦多様体と呼ぶ。平坦多様体には次のような性質がある。

定理 2.1.1 (M, ∇) を平坦多様体とする。このとき各 $p \in M$ に対し、局所座標系 $(U, (x_1, \dots, x_n))$ が存在し、 $\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = 0$ ($1 \leq i, j \leq n$) が成り立つ。また、このような座標系間の座標変換はアファイン変換である。このとき、局所座標系 (x_1, \dots, x_n) を p 周りの ∇ に関するアファイン座標系という。

以降、接続 ∇ に関する概念を接頭辞 “ ∇ -” をつけて表現する (e.g., ∇ -アファイン座標系)。

h を M 上の擬リーマン計量とし, ∇ を M 上の平坦アファイン接続とする.

定義 2.1.2 各点 $p \in M$ に対し, アファイン座標系 $(U, \psi = (x_1, \dots, x_n))$ と関数 $f: \psi(U) \rightarrow \mathbb{R}$ が存在して, $\psi(U)$ 上で次が成り立つとき (h, ∇) は M のヘッセ構造であるという:

$$h_{ij} := h\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) = \frac{\partial^2 f}{\partial x_i \partial x_j}. \quad (2.1)$$

またこのとき, h を ∇ に関するヘッセ計量, f を h の ∇ に関するポテンシャル, (M, h, ∇) をヘッセ多様体と呼ぶ.

注意 2.1.3 (1) の条件は局所的に $h = \nabla df$ と書けることと同値である. 実際, (x_1, \dots, x_n) を ∇ -アファイン座標系とすると, ∇ の平坦性から

$$\begin{aligned} (\nabla df)\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) &= (\nabla_{\frac{\partial}{\partial x_i}} df)\left(\frac{\partial}{\partial x_j}\right) \\ &= \frac{\partial}{\partial x_i}\left(df \frac{\partial}{\partial x_j}\right) - df\left(\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j}\right) \\ &= \frac{\partial^2 f}{\partial x_i \partial x_j}. \end{aligned}$$

命題 2.1.4 擬リーマン多様体 (M, h) とその上の平坦接続 ∇ に対して, 次は互いに同値である:

- (1) h はヘッセ計量である.
- (2) 3次テンソル ∇h は全対称である. すなわち M 上の任意のベクトル場 X, Y, Z に対して $(\nabla_X h)(Y, Z) = (\nabla_Y h)(X, Z)$ が成り立つ.
- (3) ∇ -アファイン座標系 (x_1, \dots, x_n) に対し, $\frac{\partial h_{ij}}{\partial x_k} = \frac{\partial h_{kj}}{\partial x_i}$ ($1 \leq i, j, k \leq n$) が成り立つ.

証明: $(U; \psi = (x_1, \dots, x_n))$ を ∇ -アファイン座標系とする. このとき任意の i, j, k に対し,

$$\begin{aligned} (\nabla_{\frac{\partial}{\partial x_i}} h)\left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_j}\right) &= \frac{\partial}{\partial x_i} h\left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_j}\right) - h\left(\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_j}\right) - h\left(\frac{\partial}{\partial x_k}, \nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j}\right) \\ &= \frac{\partial h_{kj}}{\partial x_i}, \end{aligned}$$

$$(\nabla_{\frac{\partial}{\partial x_k}} h)\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) = \frac{\partial h_{ij}}{\partial x_k}$$

より (2) と (3) の同値性が分かる.

次に (1) と (3) の同値性を示す. (1) を仮定すると (3) が成り立つことは明らかである. (3) を仮定する. アファイン座標系 (x_1, \dots, x_n) において, $h_j := \sum_{i=1}^n h_{ij} dx_i (1 \leq j \leq n)$ とおく.

$$\begin{aligned} dh_j &= \sum_{i=1}^n dh_{ij} dx_i \\ &= \sum_{i,r=1}^n \frac{\partial h_{ij}}{\partial x_r} dx_r \wedge dx_i \\ &= \sum_{i < r} \left(\frac{\partial h_{ij}}{\partial x_r} - \frac{\partial h_{rj}}{\partial x_i} \right) dx_r \wedge dx_i \end{aligned}$$

であり (3) より $dh_j = 0$ となる. ポアンカレの補題より $\psi(U)$ 上の関数 $f_j \in C^\infty(\psi(U))$ が存在して $df_j = h_j$ となる. また, $\tilde{h} := \sum_{j=1}^n f_j dx_j$ とおくと,

$$\begin{aligned} d\tilde{h} &= \sum_{j=1}^n df_j \wedge dx_j \\ &= \sum_{j=1}^n h_j \wedge dx_j \\ &= \sum_{j,i=1}^n h_{ij} dx_i \wedge dx_j \\ &= \sum_{i < j} (h_{ij} - h_{ji}) dx_i \wedge dx_j \\ &= 0. \end{aligned}$$

ここで最後の等式には擬リーマン計量 h の対称性を用いた; $h_{ij} = h_{ji}$. $d\tilde{h} = 0$ より, 再びポアンカレの補題からある $f \in C^\infty(\psi(U))$ が存在して $df = \tilde{h}$ となる. h_j, \tilde{h} の定義より

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = h_{ij}$$

が従い, h はヘッセ計量となる. □

注意 2.1.5 ヘッセ構造は上記のような特徴づけを持つが, (3) による特徴づけはアファイン座標系上でのポテンシャル関数の存在を示す可積分条件を表しており特に重要である. ここでポテンシャル関数の存在を示す上で h の対称性のみが用いられ, その非退化性は必要ないことに注意する.

2.1.2 双対ヘッセ構造

∇ を \mathbb{R}^n の標準平坦接続とし, $\mathbf{x} = (x_1, \dots, x_n)$ を ∇ のアファイン座標系とする. さらに, $U \subset \mathbb{R}^n$ を開集合とし (h, ∇) を U のヘッセ構造, $f: U \rightarrow \mathbb{R}$ をポテンシャルとする.

$p_i := \frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$ とおくと, f がポテンシャルであることから $\mathbf{p} = (p_1, \dots, p_n)$ は U の各点で非退化であり, U のもう一つの局所座標系を与えることが分かる. この座標系 \mathbf{p} を \mathbf{x} の双対座標系と呼ぶ. 双対座標系 $\mathbf{p}: U \rightarrow \mathbf{p}(U) := V$ は U 全体で定まっているものとする.

注意 2.1.6 ∇ -アファイン座標系とその上のポテンシャルの取り替えによって生じる双対座標系の座標変換はアファイン変換である. 実際, \mathbf{x} と \mathbf{y} を U 上の二つの ∇ -アファイン座標系とし, $f_1(\mathbf{x})$ と $f_2(\mathbf{y})$ をそれぞれの座標系上のポテンシャルとすると, アファイン座標系の定義から座標変換

$$\mathbf{y} = \psi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

を得る. ここで $A \in GL_n(\mathbb{R})$, $\mathbf{b} \in \mathbb{R}^n$ である. H_1 と H_2 をそれぞれポテンシャル $f_1(\mathbf{x})$ と $f_2(\mathbf{y})$ のヘッセ行列とすると, 任意の $p \in U$ において $H_1(p) = A^T H_2(p) A$ である. すなわち関数 $f_1(\mathbf{x})$ と $f_2 \circ \psi(\mathbf{x})$ の2階偏導関数は全て等しい. 従ってこれらは一次の項を除いて一致する:

$$f_2 \circ \psi(\mathbf{x}) = f_1(\mathbf{x}) + \mathbf{c}^T \mathbf{x} + d.$$

これより $A' := (A^T)^{-1}$ とおくと \mathbf{x} の双対座標系 $\mathbf{p} = \frac{\partial f_1}{\partial \mathbf{x}}$ は $\mathbf{p} \mapsto A' \mathbf{p} + A' \mathbf{c}$ によって変換されることがわかる.

定理 2.1.7 座標系 \mathbf{p} をアファイン座標系とする U 上の平坦接続を ∇^* とするとき, 次が成り立つ:

- (1) h の \mathbf{x} に関する成分行列を $[h_{ij}]$ とし, その逆行列を $[h^{ij}]$ とするとき, $h(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial x_j}) = \delta_{ij}$, $h(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}) = h^{ij}$ が成り立つ.
- (2) U 上の任意のベクトル場 X, Y, Z に対し, $Xh(Y, Z) = h(\nabla_X Y, Z) + h(Y, \nabla_X^* Z)$.
- (3) (h, ∇^*) は U 上のヘッセ構造である.

証明: ベクトル場の変換

$$\frac{\partial}{\partial x_i} = \sum_{l=1}^n h_{li} \frac{\partial}{\partial p_l}, \quad \frac{\partial}{\partial p_i} = \sum_{l=1}^n h^{li} \frac{\partial}{\partial x_l}$$

に注意する. 任意の $1 \leq i, j \leq n$ に対し,

$$\begin{aligned}
h\left(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial x_j}\right) &= h\left(\sum_{k=1}^n h^{ki} \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_j}\right) \\
&= \sum_{k=1}^n h^{ki} h_{kj} = \delta_{ij}, \\
h\left(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}\right) &= h\left(\sum_{k=1}^n h^{ki} \frac{\partial}{\partial x_k}, \sum_{l=1}^n h^{lj} \frac{\partial}{\partial x_l}\right) \\
&= \sum_{k,l=1}^n h^{ki} h^{lj} h_{kl} \\
&= \sum_{k=1}^n h^{ki} \delta_{kj} = h^{ij}.
\end{aligned}$$

を得る. 従って, (1) が成り立つ.

また,

$$\begin{aligned}
h\left(\frac{\partial}{\partial x_j}, \nabla_{\frac{\partial}{\partial x_i}}^* \frac{\partial}{\partial x_k}\right) &= h\left(\frac{\partial}{\partial x_j}, \sum_{l,r=1}^n h_{li} \nabla_{\frac{\partial}{\partial p_l}}^* h_{rk} \frac{\partial}{\partial p_r}\right) \\
&= h\left(\frac{\partial}{\partial x_j}, \sum_{l,r=1}^n h_{li} \frac{\partial h_{rk}}{\partial p_l} \frac{\partial}{\partial p_r}\right) \\
&= \sum_{l,r=1}^n h_{li} \frac{\partial h_{rk}}{\partial p_l} h\left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial p_r}\right) \\
&= \sum_{l=1}^n h_{li} \frac{\partial h_{jk}}{\partial p_l} \\
&= \left(\sum_{l=1}^n h_{li} \frac{\partial}{\partial p_l}\right) h_{jk} \\
&= \frac{\partial}{\partial x_i} h\left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k}\right).
\end{aligned}$$

$\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = 0$ であるため, 上式は (2) の等式を意味する.

最後に (3) を示す. V 上の関数 $\varphi : V \rightarrow \mathbb{R}$ を $\varphi(\mathbf{p}) = \mathbf{x}^T \mathbf{p} - f(\mathbf{x})$ で定める. このとき, 任意の $1 \leq i, j \leq n$ に対し,

$$\frac{\partial \varphi}{\partial p_i}(\mathbf{p}) = \sum_{k=1}^n \frac{\partial x_k}{\partial p_i} p_k + x_i - \sum_{k=1}^n \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial p_i} = x_i$$

および

$$\frac{\partial^2 \varphi}{\partial p_i \partial p_j}(\mathbf{p}) = \frac{\partial x_i}{\partial p_j} = h^{ij} = h\left(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}\right).$$

従って, (h, ∇^*) は φ をポテンシャルとする U のヘッセ構造である. \square

以上の構成は容易に一般のヘッセ多様体へ拡張できる. 注意 2.1.6 より, ヘッセ多様体における局所双対座標系はアファイン変換で貼り合わされる. 従って, 上記定理で局所的に定義された ∇^* は ∇ -アファイン座標系 \mathbf{x} やその上のポテンシャル $f(\mathbf{x})$ の取り方にはよらずヘッセ構造 (h, ∇) のみから定まることがわかる. このため接続 ∇^* は M 上で大域的に定義される.

系 2.1.8 (M, h, ∇) をヘッセ多様体とする. 上記で定まる M 上の接続 ∇^* について次が成り立つ:

- (1) ∇^* は平坦接続である.
- (2) M 上の任意のベクトル場 X, Y, Z に対し, $Xh(Y, Z) = h(\nabla_X Y, Z) + h(Y, \nabla_X^* Z)$.
- (3) (h, ∇^*) は M のヘッセ構造である.
- (4) M 上の任意のベクトル場 X, Y, Z に対し, $\nabla h(X, Y, Z) = h(Y, \nabla_X^* Z) - h(Y, \nabla_X Z)$.

証明: ∇^* の定義から直ちに (1) が従う. (2) と (3) は局所的な議論であるため, 定理 2.1.7 から従う. また, X, Y, Z を M 上のベクトル場とすると, (2) より

$$\nabla h(X, Y, Z) = Xh(Y, Z) - h(\nabla_X Y, Z) - h(Y, \nabla_X Z) = h(Y, \nabla_X^* Z) - h(Y, \nabla_X Z)$$

となる \square

上記の系において特に, (4) は (2) のみから従うことに注意する. 3次テンソル ∇h については §2.2 にて再度議論する.

定義 2.1.9 ∇^* を ∇ の h に関する双対平坦接続と呼ぶ. また, (h, ∇^*) を (h, ∇) の双対ヘッセ構造と呼ぶ.

2.1.3 ルジャンドル変換と双対ポテンシャル

ヘッセ構造 (h, ∇) とその双対ヘッセ構造 (h, ∇^*) のポテンシャルの関係性について述べる. $\mathbf{x} = (x_1, \dots, x_n)$ を ∇ -アファイン座標系, $\mathbf{p} = (p_1, \dots, p_n)$ を ∇^* -アファイン座標系とする. $f(\mathbf{x})$ をヘッセ構造 (h, ∇) のポテンシャルとする. 定理 2.1.7(3) の証明における双対ヘッセ構造 (h, ∇^*) のポテンシャル $\varphi = \mathbf{x}^T \mathbf{p} - f$ は次の2つを満たしていた:

$$\frac{\partial \varphi}{\partial p_i} = x_i, \quad \frac{\partial^2 \varphi}{\partial p_i \partial p_j} = h\left(\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}\right) = h^{ij}.$$

実は上の二つを満たす関数は定数項の違いを除いて φ のみである。実際、関数 $\tilde{\varphi}(\mathbf{p})$ が上の二つを満たすとすると、

$$\begin{aligned} \frac{\partial^2 f}{\partial p_i \partial p_j} &= \sum_{k,l=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \frac{\partial x_k}{\partial p_i} \frac{\partial x_l}{\partial p_j} + \sum_{k=1}^n p_k \frac{\partial x_l}{\partial p_i \partial p_j} \\ &= \sum_{k,l=1}^n h_{ij} \frac{\partial}{\partial p_i} \left(\frac{\partial \tilde{\varphi}}{\partial p_k} \right) \frac{\partial}{\partial p_j} \left(\frac{\partial \tilde{\varphi}}{\partial p_l} \right) + \sum_{k=1}^n p_k \frac{\partial x_l}{\partial p_i \partial p_j} \\ &= \sum_{k,l=1}^n h_{ij} h^{ik} h^{jl} + \sum_{k=1}^n p_k \frac{\partial x_l}{\partial p_i \partial p_j} \\ &= \frac{\partial^2 \tilde{\varphi}}{\partial p_i \partial p_j} + \sum_{k=1}^n p_k \frac{\partial x_l}{\partial p_i \partial p_j} \end{aligned}$$

および

$$\begin{aligned} \frac{\partial^2}{\partial p_i \partial p_j} \left(\sum_{k=1}^n p_k x_k \right) &= \frac{\partial}{\partial p_i} \left(x_j + \sum_{k=1}^n p_k \frac{\partial x_k}{\partial p_j} \right), \\ &= \frac{\partial^2 \tilde{\varphi}}{\partial p_i \partial p_j} + \frac{\partial x_i}{\partial p_j} + \sum_{k=1}^n p_l \frac{\partial^2 x_k}{\partial p_i \partial p_j}, \\ &= 2 \frac{\partial^2 \tilde{\varphi}}{\partial p_i \partial p_j} + \sum_{k=1}^n p_l \frac{\partial^2 x_k}{\partial p_i \partial p_j} \end{aligned}$$

であることから

$$\frac{\partial^2}{\partial p_i \partial p_j} \left(f + \tilde{\varphi} - \sum_{k=1}^n x_k p_k \right) = 0$$

が成り立つ。従って、 $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, $b \in \mathbb{R}$ とすると、

$$\tilde{\varphi} = \mathbf{x}^T \mathbf{p} - f + \mathbf{a}^T \mathbf{p} + b = \varphi + \mathbf{a}^T \mathbf{p} + b.$$

とかける。仮定より $\frac{\partial \tilde{\varphi}}{\partial p_i} = x_i = \frac{\partial \varphi}{\partial p_i}$ であるため、 $\mathbf{a} = 0$ 。以上より $\tilde{\varphi}$ は次を満たす：

$$\tilde{\varphi} = \varphi + b.$$

双対ヘッセ構造 (h, ∇^*) のポテンシャル $\varphi = \mathbf{x}^T \mathbf{p} - f$ を f のルジャンドル変換あるいは f の双対ポテンシャルと呼ぶ。ポテンシャル f と φ はルジャンドル双対性

$$f + \varphi - \sum_{i=1}^n x_i p_i = 0$$

によって結び付けられていることがわかる。

注意 2.1.10 f のルジャンドル変換 φ は ∇ -アフィン座標系 \mathbf{x} の線形座標変換によって不変である. 実際, 注意 2.1.6 において, $\mathbf{b} = \mathbf{c} = 0$, $d = 0$ とすると, ∇ -アフィン座標系 \mathbf{x} とその双対座標系 \mathbf{p} はそれぞれ $\mathbf{x} \mapsto A\mathbf{x}$, $\mathbf{p} \mapsto A'\mathbf{p}$ によって変換され,

$$(A\mathbf{x})^T(A'\mathbf{p}) = \mathbf{x}^T A^T (A')^{-1} \mathbf{p} = \mathbf{x}^T \mathbf{p}.$$

2.1.4 ヘッセ構造のダイバージェンス

(M, h, ∇) を n 次元ヘッセ多様体とし, (U, \mathbf{x}) を \mathbf{x} が ∇ -アフィン座標系であるような座標近傍とする. $f(\mathbf{x})$ をポテンシャルとし, ∇^* を双対接続, $\varphi(\mathbf{p})$ を f の双対ポテンシャルとする. 関数 $\mathcal{D} : U \times U \rightarrow \mathbb{R}$ を

$$\mathcal{D}(p, q) = f(\mathbf{x}(p)) + \varphi(\mathbf{p}(q)) - \mathbf{x}(p)^T \mathbf{p}(q) \quad (2.2)$$

で定義する. 同じことであるが

$$\begin{aligned} \mathcal{D}(p, q) &= f(\mathbf{x}(p)) + \varphi(\mathbf{p}(q)) - \mathbf{x}(p)^T \mathbf{p}(q) \\ &= f(\mathbf{x}(p)) - f(\mathbf{x}(q)) + (\mathbf{x}(q) - \mathbf{x}(p))^T \mathbf{p}(q). \end{aligned}$$

この関数 \mathcal{D} をポテンシャル f から定まるブレグマンダイバージェンスと呼ぶ [1]. 点 $p \in U$ を固定し関数 $F = \mathcal{D}(p, \cdot) : U \rightarrow \mathbb{R}$ を考えると

$$\frac{\partial F}{\partial x_i} = -\frac{\partial f}{\partial x_i} + p_i + (\mathbf{x} - \mathbf{x}(p))^T \frac{\partial \mathbf{p}}{\partial x_i} = (\mathbf{x} - \mathbf{x}(p))^T \frac{\partial \mathbf{p}}{\partial x_i}, \quad (2.3)$$

$$\frac{\partial F}{\partial x_i \partial x_j} = \frac{\partial p_i}{\partial x_j} + (\mathbf{x} - \mathbf{x}(p))^T \frac{\partial^2 \mathbf{p}}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j} + (\mathbf{x} - \mathbf{x}(p))^T \frac{\partial^2 \mathbf{p}}{\partial x_i \partial x_j}. \quad (2.4)$$

である. これより以下が直ちに従う.

命題 2.1.11 ブレグマンダイバージェンス \mathcal{D} と点 $p \in U$ に対して関数 $F := \mathcal{D}(p, \cdot)$ を定める. このとき次が成り立つ:

- (1) $\mathcal{D}(p, p) = 0$.
- (2) U 上の任意のベクトル場 X に対し, $X(p)F = 0$.
- (3) U 上の任意のベクトル場 X, Y に対し, $X(p)Y(p)F = h(X(p), Y(p))$.

証明: (1) は f と φ のルジャンドル双対性より従う. (2) と (3) はそれぞれ式 (2.3), (2.4) より従う. \square

命題 2.1.12 U 上のブレグマンダイバージェンス \mathcal{D} の値は ∇ -アフィン座標系 \mathbf{x} やポテンシャル $f(\mathbf{x})$ の取り方によらない.

証明： \mathbf{y} をもう一つの ∇ -アファイン座標系とすると，座標変換 $\mathbf{y} = \psi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ を得る．ここで $A \in GL_n(\mathbb{R})$, $\mathbf{b} \in \mathbb{R}^n$ である． $\tilde{f}(\mathbf{y})$ をポテンシャルとすると注意 2.1.6 より，

$$\tilde{f} \circ \psi(\mathbf{x}) = f(\mathbf{x}) + \mathbf{c}^T \mathbf{x} + d$$

とかける．ここで $\mathbf{c} \in \mathbb{R}^n, d \in \mathbb{R}$ である． \mathbf{q} を \tilde{f} によって定まる双対座標系とすると $\mathbf{q} = A'\mathbf{p} + A'\mathbf{c}$ ($A' := (A^T)^{-1}$)．従って，任意の $p, q \in U$ に対し

$$\begin{aligned} & \tilde{f}(\mathbf{y}(p)) - \tilde{f}(\mathbf{y}(q)) + (\mathbf{y}(q) - \mathbf{y}(p))^T \mathbf{q}(q) \\ &= f(\mathbf{x}(p)) - f(\mathbf{x}(q)) + \mathbf{c}^T (\mathbf{x}(p) - \mathbf{x}(q)) + \{A(\mathbf{x}(q) - \mathbf{x}(p))\}^T (A'\mathbf{p}(q) + A'\mathbf{c}) \\ &= f(\mathbf{x}(p)) - f(\mathbf{x}(q)) + (\mathbf{x}(q) - \mathbf{x}(p))^T \mathbf{p}(q). \end{aligned}$$

が成り立つ。 □

上記命題より，ブレグマンダイバージェンスはヘッセ構造 (h, ∇) のみから定まることがわかる． $\{(U_\alpha, \mathbf{x}_\alpha)\}_{\alpha \in \Lambda}$ をヘッセ多様体 (M, h, ∇) のアトラスで各 \mathbf{x}_α が ∇ -アファイン座標系であるものとし， \mathcal{D}_α をそれぞれの座標近傍 $(U_\alpha, \mathbf{x}_\alpha)$ 上で定義されたブレグマンダイバージェンスとする．このとき，対角集合 $\Delta_M = \{(p, p) \in M \times M\}$ の開近傍 $U(\Delta_M) := \{(p, q) \in M \times M \mid p, q \in U_\alpha\}$ 上でブレグマンダイバージェンスは定義される．

定義 2.1.13 任意の $(p, q) \in U(\Delta_M)$ に対し， $(p, q) \in U_\alpha$ であるとき $\mathcal{D}_M(p, q) := \mathcal{D}_\alpha(p, q)$ と定める．命題 2.1.12 より，well-defined に

$$\mathcal{D}_M : U(\Delta_M) \rightarrow \mathbb{R}$$

が定義される．この \mathcal{D}_M をヘッセ多様体 (M, h, ∇) のダイバージェンスと呼ぶ．

注意 2.1.14 ブレグマンダイバージェンスは \mathbb{R}^n の領域とその上の凸関数から定まる概念である．[1, 3] 等に従い，用語としてブレグマンダイバージェンスとヘッセ多様体のダイバージェンスとを区別して用いることにする．

命題 2.1.15 \mathcal{D}'_M を双対ヘッセ構造 (h, ∇^*) から定まるダイバージェンスとする．このとき $\mathcal{D}'_M(p, q) = \mathcal{D}_M(q, p)$ が成り立つ．

証明：式 (2.2) より明らか。 □

注意 2.1.16 上記の命題より， $F = \mathcal{D}(\cdot, p)$ と定めた場合にも（第二変数を固定していることに注意），命題 2.1.11 の結果は成り立つことに注意する．

次の二つの定理は情報幾何学の応用において重要な役割を果たすものである．以下ではダイバージェンス \mathcal{D}_M は M 上で定義されているものとする．

定理 2.1.17 $c_e, c_m : [0, 1] \rightarrow M$ をそれぞれ ∇, ∇^* に関する測地線とし, $c_e(0) = c_m(0)$ とする. $p := c_e(1), q := c_e(0), r := c_m(1)$ とおく. q において c_e と c_m が直交する時, すなわち $h(\dot{c}_e(0), \dot{c}_m(0)) = 0$ なる時, $\mathcal{D}_M(p, r) = \mathcal{D}_M(p, q) + \mathcal{D}_M(q, r)$ が成り立つ. ここで, $\dot{c}_e(0), \dot{c}_m(0) \in T_q M$ はそれぞれ c_e, c_m の速度ベクトルである.

証明: c_e と c_m はそれぞれ ∇ および ∇^* に関する測地線であるため, \mathbf{x} を ∇ -アフィン座標系, \mathbf{p} を ∇^* -アフィン座標系とすると, 任意の $t \in [0, 1]$ に対し,

$$\mathbf{x}(c_e(t)) = \mathbf{x}(q) + (\mathbf{x}(p) - \mathbf{x}(q))t, \quad \mathbf{p}(c_m(t)) = \mathbf{p}(q) + (\mathbf{p}(p) - \mathbf{p}(q))t$$

とかける. このとき,

$$\begin{aligned} \mathcal{D}_M(p, r) - \mathcal{D}_M(p, q) - \mathcal{D}_M(q, r) &= -(\mathbf{x}(p) - \mathbf{x}(q))^T (\mathbf{p}(r) - \mathbf{p}(q)) \\ &= -h(\dot{c}_e(0), \dot{c}_m(0)). \end{aligned}$$

仮定より c_e と c_m は q で直交するため, 示すべき等式が従う. \square

定理 2.1.18 $S \subset M$ を部分多様体とし, $c : [0, 1] \rightarrow M$ を $q := c(1) \in S$ なる ∇^* -測地線とする. $p := c(0)$ とおき, 関数 $F := \mathcal{D}_M(-, p) : S \rightarrow \mathbb{R}$ を定める. このとき, 点 q において関数 F が極値をとることと, q において S と c が h に関して直交することが必要十分である. また, c を ∇ -測地線に, F を $\mathcal{D}_M(p, \cdot)$ に置き換えることでも同様の結果が成り立つ.

証明: $\gamma : (-\varepsilon, \varepsilon) \rightarrow S$ を $\gamma(0) = q$ なる曲線とする.

$$F(\gamma(s)) = f(\mathbf{x}(\gamma(s))) - f(\mathbf{x}(p)) + (\mathbf{x}(p) - \mathbf{x}(\gamma(s)))^T \mathbf{p}(p)$$

より

$$\frac{d}{ds} F(\gamma(0)) = (\mathbf{p}(q) - \mathbf{p}(p))^T \frac{d\mathbf{x} \circ \gamma}{ds}(0)$$

を得る. 従って, q が F の極値を与えることと, c が S と q において h に関して直交することは同値である. 後半についても同様に示せる. \square

2.2 統計多様体と双対平坦構造

甘利・長岡により導入された双対平坦多様体は以下で定める統計多様体の特別な場合に相当し, これはヘッセ多様体と同一の概念である. 統計多様体の幾何学は江口によるコントラスト関数の理論と密接に関わる. 本節では統計多様体の文脈から双対平坦多様体あるいはヘッセ多様体について議論する.

2.2.1 双対平坦構造

(M, h) を擬リーマン多様体とする. ∇ を接束 TM の接続とし, ∇ は捩れを持たないものとする.

定義 2.2.1 ([31]) 3次テンソル $T := \nabla h$ が全対称である, すなわち M 上の任意のベクトル場 X, Y, Z に対して

$$(\nabla_X h)(Y, Z) = (\nabla_Y h)(X, Z)$$

が成り立つとき (M, h, ∇) を統計多様体と呼ぶ.

統計多様体 (M, h, ∇) の (h に関する) 双対接続 ∇^* は次で定義される:

$$Xh(Y, Z) = h(\nabla_X Y, Z) + h(Y, \nabla_X^* Z). \quad (2.5)$$

上記の統計多様体の定義では接続 ∇ が与えられたとき 3次テンソル T と双対接続 ∇^* が定められたが, 同値な定義として擬リーマン多様体 (M, h) 上に対称な $(0, 3)$ -テンソル T を与えるラウリッツェン (Lauritzen) によるものがある [28]. すなわち, $\bar{\nabla}$ を h のレビ・チビタ接続とし, M 上の接続 ∇, ∇^* を次で定義すると ∇ と ∇^* は双対的であり, (M, h, ∇) は統計多様体となる:

$$h(\nabla_X Y, Z) = h(\bar{\nabla}_X Y, Z) - \frac{1}{2}C(X, Y, Z), \quad h(\nabla_X^* Y, Z) = h(\bar{\nabla}_X Y, Z) + \frac{1}{2}C(X, Y, Z).$$

いずれの定義にしても, 統計多様体の幾何学を考える上で, (h, T, ∇, ∇^*) を考えることが重要である.

(M, h, ∇) を統計多様体とするとき, 双対接続 ∇^* は捩れを持たず, $\nabla^* h$ もまた全対称である. さらに, ∇ が平坦であるとき, ∇^* もまた平坦である.

定義 2.2.2 ([1, 3]) 統計多様体 (M, h, ∇) において, 接続 ∇ が平坦であるとき (M, h, ∇, ∇^*) を双対平坦多様体と呼び, (h, ∇, ∇^*) を双対平坦構造と呼ぶ.

命題 2.1.4 より確かにヘッセ多様体と双対平坦多様体は同一の概念であることが分かる.

例 2.2.3 指数型分布族 M とは, 次の形をした確率密度関数の族である:

$$p(\mathbf{u}|\theta) = \exp(\mathbf{u}^T \theta - \psi(\theta)).$$

ここで, $\mathbf{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$ は確率変数 (その測度は $d\mu$) であり, $\theta = (\theta_1, \dots, \theta_n)^T \in U \subset \mathbb{R}^n$ はパラメータ (U は開集合) である. 正規化関数 $\psi(\theta) = \log \int \exp(\mathbf{u}^T \theta) d\mu$ はこの族のポテンシャルと呼ばれる. U のアファイン構造を固定し, $\partial_i = \frac{\partial}{\partial \theta_i}$ とおく. こ

のとき期待値は対応する双対座標を与えることが分かる：

$$\eta_i := \mathbf{E}[u_i|\theta] = \partial_i \psi(\theta).$$

また、(共)分散行列は

$$h_{ij} := \mathbf{V}[\mathbf{u}|\theta]_{ij} = \partial_i \partial_j \psi(\theta) = \mathbf{E}[(\partial_i \log p)(\partial_j \log p)]$$

によって記述される．ここで最右辺はフィッシャー情報量を意味する． $h = [h_{ij}]$ が正定値で、 θ, η をそれぞれ ∇, ∇^* -アファイン座標系とみなすとき、 (M, h, ∇, ∇^*) は双対平坦多様体になる．正規分布や有限離散分布は典型的な例である．

さらに、双対平坦多様体の理論では、ヘッセ計量 h のみならず、対称 $(0, 3)$ テンソル $T = \nabla h$ も重要な役割を果たす [3]．

定義 2.2.4 ([3]) 双対平坦多様体 (M, h, ∇, ∇^*) に対し、対称 $(0, 3)$ テンソル $T = \nabla h$ を甘利・チェンツォフの3次テンソルと呼ぶ．

系 2.1.8 の (2) より3次テンソル T は任意のベクトル場 X, Y, Z に対し、

$$T(X, Y, Z) = h(Y, \nabla_X^* Z) - h(Y, \nabla_X Z) = h(\nabla_X^* Y, Z) - h(\nabla_X Y, Z)$$

を満たしていることが分かる．また、この3次テンソルは局所的にポテンシャル関数の3階微分によって表現される．すなわち、次が成り立つ．

命題 2.2.5 (M, h, ∇, ∇^*) を n 次元双対平坦多様体とし、 $\mathbf{x} = (x_1, \dots, x_n)$ を ∇ -アファイン座標系、 $f(\mathbf{x})$ をその上のポテンシャル関数とする． $\partial_i := \frac{\partial}{\partial x_i}$ ($1 \leq i \leq n$) とおく．このとき任意の i, j, k に対して次が成り立つ：

$$T(\partial_i, \partial_j, \partial_k) = \partial_i \partial_j \partial_k f.$$

証明： ∇ の平坦性と $h(\partial_j, \partial_k) = \partial_j \partial_k f$ であることに注意して、

$$\begin{aligned} T(\partial_i, \partial_j, \partial_k) &= h(\partial_j, \nabla_{\partial_i}^* \partial_k) - h(\partial_j, \nabla_{\partial_i} \partial_k) \\ &= \partial_i h(\partial_j, \partial_k) \\ &= \partial_i \partial_j \partial_k f. \end{aligned}$$

□

2.2.2 コントラスト関数

多様体 M の対角集合 Δ_M の開近傍 $U \subset M \times M$ 上の関数 $\rho : U \rightarrow \mathbb{R}$ が与えられたとき、この関数の微分を考えることで M 上にはさまざまなテンソルが導入される．特に

ρ がコントラスト関数であるときには M 上に統計多様体としての構造を自然に定める。コントラスト関数による統計多様体の幾何学の分析は江口により徹底的になされている [16]。ここではまず [16] におけるコントラスト関数の一般論をまとめ、その後双対平坦多様体について議論を行う。

M 上のベクトル場 $X_1, \dots, X_k, Y_1, \dots, Y_l$ に対して、関数 $\rho[X_1 \cdots X_k | Y_1 \cdots Y_l] : M \rightarrow \mathbb{R}$,

$$\rho[X_1 \cdots X_k | Y_1 \cdots Y_l](r) = (X_1)_p \cdots (X_k)_p (Y_1)_q \cdots (Y_l)_q (\rho(p, q))|_{p=q=r}. \quad (2.6)$$

を定める。

定義 2.2.6 関数 $\rho : U \rightarrow \mathbb{R}$ が次を満たすとき ρ は M 上のコントラスト関数と呼ばれる：任意の $r \in M$ およびベクトル場 X, Y に対し

- (1) $\rho[-|-](r) = \rho(r, r) = 0$,
- (2) $\rho[X|-](r) = \rho[-|X](r) = 0$,
- (3) $-\rho[X|Y]$ は M 上の擬リーマン計量である。

関数 ρ がコントラスト関数であるとき、定義 2.2.6 における条件 (3) から

$$h(X, Y) := -\rho[X|Y], \quad T(X, Y, Z) := -\rho[Z|XY] + \rho[XY|Z] \quad (2.7)$$

とおくことにより M 上に擬リーマン計量 h と対称 3 次テンソル T が定義される。ここで X, Y, Z は M 上の任意のベクトル場である。これより、 (M, h, T) は統計多様体となる。逆に、統計多様体 (M, h, T) が与えられたときに、上式 (2.7) により h と C を復元するコントラスト関数の存在が松本 [29] により示されている（存在は一意的ではなく統計多様体とコントラスト関数の間には 1 対多の関係がある）。また、 M 上のアファイン接続 ∇, ∇^* を

$$h(\nabla_X Y, Z) = -\rho[XY|Z], \quad h(\nabla_X^* Y, Z) = -\rho[Z|XY]$$

によって定義すると、これらは捩れの無い接続であり互いに双対的である：

$$Xh(Y, Z) = h(\nabla_X Y, Z) + h(Y, \nabla_X^* Z).$$

二つの微分演算子 $D, D^* : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$ を次で定義する：

$$h(D_{X,Y} Z, W) = -\rho[XYZ|W], \quad h(D_{X,Y}^* Z, W) = -\rho[W|XYZ].$$

ここで $\mathfrak{X}(M)$ は M 上のベクトル場全体の集合である。 D, D^* を用いて M 上の (1,3) テンソル B, B^* を定義する [16]：

$$B(X, Y) = D_{X,Y} - \nabla_X \nabla_Y, \quad B^*(X, Y) = D_{X,Y}^* - \nabla_X^* \nabla_Y^*. \quad (2.8)$$

ここで X, Y は M 上のベクトル場である. テンソル B と B^* はそれぞれ [30] においてバートレットテンソル, 双対バートレットテンソルと呼ばれており, コントラスト関数 ρ から誘導される統計多様体 (M, h, ∇, ∇^*) における ∇, ∇^* および h の曲率と関係する [30].

定理 2.2.7 [16, 定理 1] (M, ρ) に対し, 次は同値である:

- (1) $B = 0$.
- (2) $B^* = 0$.
- (3) M の各点においてある局所座標系 (x_1, \dots, x_n) が存在して

$$\nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} = 0, \quad D_{\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}} \frac{\partial}{\partial x_k} = 0 \quad (1 \leq i, j, k \leq n).$$

- (4) M の各点においてある局所座標系 (p_1, \dots, p_n) が存在して

$$\nabla_{\frac{\partial}{\partial p_i}}^* \frac{\partial}{\partial p_j} = 0, \quad D_{\frac{\partial}{\partial p_i}, \frac{\partial}{\partial p_j}}^* \frac{\partial}{\partial p_k} = 0 \quad (1 \leq i, j, k \leq n).$$

証明: [16] をみよ. □

等式 $h(\nabla_Y Z, W) = -\rho[YZ|W]$ と $h(W, \nabla_Y^* Z) = -\rho[W|YZ]$ のそれぞれに対し両辺に X を施すことで

$$\begin{aligned} h(\nabla_X \nabla_Y Z, W) + h(\nabla_Y Z, \nabla_X^* W) &= h(D_{X,Y} Z, W) - \rho[YZ|XW], \\ h(\nabla_X W, \nabla_Y^* Z) + h(W, \nabla_X^* \nabla_Y^* Z) &= -\rho[XW|YZ] + h(D_{X,Y}^* Z, W) \end{aligned}$$

を得る. これより

$$\begin{aligned} h(B(X, Y)Z, W) &= h(\nabla_Y Z, \nabla_X^* W) + \rho[YZ|XW], \\ h(B^*(X, Y)Z, W) &= h(\nabla_X W, \nabla_Y^* Z) + \rho[XW|YZ] \end{aligned}$$

であるため, 定理 2.2.7 の (3) および (4) における二つ目の条件はそれぞれ, 任意の $1 \leq i, j, k, l \leq n$ に対し

$$\rho\left[\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \middle| \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_l}\right] = 0, \quad \rho\left[\frac{\partial}{\partial p_i} \frac{\partial}{\partial p_j} \middle| \frac{\partial}{\partial p_k} \frac{\partial}{\partial p_l}\right] = 0$$

であることに置き換えても良いことがわかる.

双対平坦多様体とその上のダイバージェンスについてテンソル B, B^* を考える. (M, h, ∇, ∇^*) を n 次元双対平坦多様体, $U(\Delta_M)$ を対角集合 Δ_M の開近傍, $\mathcal{D}: U(\Delta_M) \rightarrow \mathbb{R}$ をダイバージェンスとする. このとき双対平坦多様体の幾何構造はダイバージェンスによって次のように特徴づけられる.

定理 2.2.8 双対平坦多様体 (M, h, ∇, ∇^*) 上のダイバージェンス \mathcal{D} は M 上のコントラスト関数であり, 計量 h と甘利・チェンツォフテンソル $T = \nabla h$ は \mathcal{D} により復元される. すなわち M 上の任意のベクトル場 X, Y, Z に対して

$$h(X, Y) = -\mathcal{D}[X|Y], \quad T(X, Y, Z) = -\mathcal{D}[Z|XY] + \mathcal{D}[XY|Z].$$

証明: 命題 2.1.11 より直ちに従う. □

テンソル B, B^* をそれぞれ \mathcal{D} から生成されるバートレットテンソルおよび双対バートレットテンソルとする.

定理 2.2.9 双対平坦多様体 (M, h, ∇, ∇^*) とその上のダイバージェンス \mathcal{D} について $B = B^* = 0$ である. すなわち M 上の任意のベクトル場 X, Y, Z, W に対して

$$h(\nabla_X \nabla_Y Z, W) = -\mathcal{D}[XYZ|W], \quad (2.9)$$

$$h(\nabla_X^* \nabla_Y^* Z, W) = -\mathcal{D}[W|XYZ]. \quad (2.10)$$

証明: 直接計算により $B = 0$ を示す. $(U, \mathbf{x} = (x_1, \dots, x_n))$ を \mathbf{x} が ∇ -アフィン座標系である座標近傍とし, $f(\mathbf{x})$ をポテンシャル, $\mathbf{p} = \frac{\partial f}{\partial \mathbf{x}}$ を双対座標系とする. このとき, 任意の $1 \leq i, j, k, l \leq n$ に対し,

$$\mathcal{D}\left[\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \middle| \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_l}\right] = 0$$

を示す. $p, q \in U$ に対し

$$\mathcal{D}(p, q) = f(\mathbf{x}(p)) - f(\mathbf{x}(q)) + (\mathbf{x}(q) - \mathbf{x}(p))^T \mathbf{p}(q)$$

であるから,

$$\left(\frac{\partial}{\partial x_i}\right)_p \left(\frac{\partial}{\partial x_j}\right)_p \mathcal{D}(p, q) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}(p)).$$

従って, $\mathcal{D}\left[\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \middle| \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_l}\right] = 0$ である. これより, 定理 2.2.7 から $B = 0$ および $B^* = 0$ が従う. □

第 3 章

接触構造と波面

この章では、第 4 章以降で用いられる接触幾何学の諸概念についてのまとめを与える。まず始めに接触構造の定義と簡単な性質についてのまとめを行う。その後、ある種の特異点を持つ超曲面のクラスである波面と、その波面を扱う上で重要な概念であるルジャンドルファイブレーションについての議論を行う。最後に、我々がこの後導入する概ヘッセ多様体の構成において重要な役割を果たす佐治・梅原・山田による連接接束を取り上げる。我々の構成にはリーマン計量は登場しないが（リーマン計量が退化する場合を考えている）、アイデアの元となった彼らの概念との差を述べるために、波面と連接接束についてリーマン計量を用いた特徴付けを行う。なお、シンプレクティック多様体についての諸概念は既知とし、詳しくは [35] の付録 D を参考にして欲しい。

3.1 接触構造

M を $m = 2n + 1$ 次元の多様体とする。各点 $p \in M$ に対し $T_p M$ における超平面 $K_p \subset T_p M$ （すなわち余次元 1 の部分空間）を対応づける $K : M \ni p \mapsto K_p \subset TM$ を超平面場という。 $U \subset M$ を開集合とする。 U 上のベクトル場 X に対して各点 $p \in U$ で $X_p \in K_p$ なるとき、 X は K に属するという。また、 K が $p \in M$ において滑らかであるとは、 p の近傍 $U \subset M$ とその上の各点で一次独立なベクトル場 $X_1, \dots, X_{2n} \in \mathfrak{X}(U)$ が存在して X_1, \dots, X_{2n} がそれぞれ K に属するときをいう。このとき $\{X_1, \dots, X_{2n}\}$ を K の（局所）基底場という。任意の点 $p \in M$ で K が滑らかなとき、単に K は滑らかであるという。以下超平面場を考えるときには滑らかなもののみ考える。

定義 3.1.1 $U \subset M$ を開集合とし、 θ を U 上の一次微分形式とする。 θ が K に（ U 上で）同伴するとは、任意の $p \in U$ に対し、 $K_p = \ker \theta_p$ ときをいう。

命題 3.1.2 $U \subset M$ を開集合とし、 $\{X_1, \dots, X_{2n}\}$ を U 上の K の基底場とする。 U 上

のベクトル場 Y を $\{X_1, \dots, X_{2n}, Y\}$ が U の各点で基底となるようにとる. このとき, 双対基底 $\{X_1^*, \dots, X_{2n}^*, Y^*\}$ に対して, $\theta := Y^*$ は K に同伴する. また, この θ は 0 を取らない関数倍を除き, $\{X_1, \dots, X_{2n}, Y\}$ の選び方によらない.

証明: 各点 $p \in U$ に対し, $T_p M = \mathbb{R}(X_1)_p \oplus \mathbb{R}(X_2)_p \oplus \dots \oplus \mathbb{R}(X_{2n})_p \oplus \mathbb{R}Y_p$ であり, θ の定義から $\ker \theta_p = \mathbb{R}(X_1)_p \oplus \mathbb{R}(X_2)_p \oplus \dots \oplus \mathbb{R}(X_{2n})_p = K_p$ となり, θ は K に同伴することが分かる. 次に一意性を示す. ω を K に同伴する U 上の一次微分形式とする: $\ker \omega_p = K_p = \mathbb{R}(X_1)_p \oplus \mathbb{R}(X_2)_p \oplus \dots \oplus \mathbb{R}(X_{2n})_p$. $f_i \in C^\infty(U)$ ($1 \leq i \leq 2n+1$) とすると,

$$\begin{aligned} \omega \left(\sum_{i=1}^{2n} f_i X_i + f_{2n+1} Y \right) &= \omega(f_{2n+1} Y) = f_{2n+1} \omega(Y), \\ \theta \left(\sum_{i=1}^{2n} f_i X_i + f_{2n+1} Y \right) &= f_{2n+1}. \end{aligned}$$

従って, $\omega = \omega(Y)\theta$ が成り立ち, Y の作り方から $\omega(Y) \in C^\infty(U)$ は U の各点で 0 にはならない. すなわち, K に同伴する任意の一次微分形式 ω は θ の (各点で 0 にならない) 関数倍となっている. \square

上記の命題より, 滑らかな超平面場 K には同伴する一次微分形式が局所的に必ず存在し, かつその存在は (ゼロを取らない関数倍を除いて) 一意的であることがわかる. 従って次の定義は超平面場に同伴する一次微分形式の取り方によらない.

定義 3.1.3 K を M 上の滑らかな超平面場とし, θ を K に同伴する開集合 $U \subset M$ 上の一次微分形式とする. (M, K) が接触多様体であるとは, 各点 $p \in U$ で

$$\theta_p \wedge (d\theta_p)^n \neq 0 \quad (3.1)$$

なるときをいう. このとき, K を M の接触構造といい, θ を (局所) 接触形式と呼ぶ.

命題 3.1.4 超平面場 K に同伴する U 上の一次微分形式 θ に対して次は互いに同値である:

- (1) θ は接触形式である.
- (2) 各 $p \in U$ において, $d\theta_p|_{K_p}$ は K_p 上のシンプレクティック形式である. すなわち, $d\theta_p|_{K_p}$ は K_p 上で非退化な交代形式である.

証明: $p \in M$ とする. $T_p M \setminus K_p$ のベクトルを考えることにより, $T_p M$ において $\theta_p \wedge (d\theta_p)^n \neq 0$ なることと, K_p において $(d\theta_p)^n \neq 0$ なることは同値であることが分かる. すなわち, (1) は $(d\theta_p|_{K_p})^n \neq 0$ を意味する. また, $d\theta_p|_{K_p}$ は定義から明らかに交代

閉形式であるから、(2) は $d\theta_p|_{K_p}$ の非退化性を意味する. このようにして、(1) と (2) の同値性が従う. \square

接触多様体の定義では、接触形式の大域的な存在は要請していない. しかし、多様体上に非可積分条件 (3.1) を満たす 1 形式が存在するとき、それは大域的な接触形式を与える.

命題 3.1.5 M 上の一次微分形式 θ が M の各点 $p \in M$ で $\theta_p \wedge (d\theta_p)^n \neq 0$ を満たすなら、 $K_p := \ker \theta_p$ とすることで (M, K) は接触多様体になる.

証明: K の滑らかさを示す. M 上のリーマン計量 g を一つ固定する. $\Phi_g : TM \rightarrow X_p \mapsto g_p(X_p, \cdot) \in T^*M$ は微分同相かつ束同型を与える. $X := \Phi_g^{-1} \circ \theta \in \mathfrak{X}(M)$ とおくと、任意の $Y \in \mathfrak{X}(M)$ に対し、 $\theta(Y) = \Phi_g(X)(Y) = g(X, Y)$ となる. $p \in M$ とし、 $(U, (y_1, \dots, y_{2n+1}))$ を p 周りの局所座標系とする. 必要であれば p の近傍 U を小さくすることで、 $\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_{2n+1}}$ の中から $2n$ 個のベクトル $\frac{\partial}{\partial y_{i_1}}, \dots, \frac{\partial}{\partial y_{i_{2n}}}$ を選んで、 $\{X, \frac{\partial}{\partial y_{i_1}}, \dots, \frac{\partial}{\partial y_{i_{2n}}}\}$ が $\mathfrak{X}(U)$ の基底になるようにできる. この基底変換によって与えられる局所座標系 $(x_1, x_2, \dots, x_{2n+1})$ では $\frac{\partial}{\partial x_1} = X$ となっている. 基底 $\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_{2n+1}}\}$ にグラム・シュミットの直交化をこの順番で行い、得られた正規直交基底を $\{e_1 = \frac{X}{\|X\|}, e_2, \dots, e_{2n+1}\}$ とする. ただし、 $\|X\|$ は g による X のノルムである. 以上より $\theta(Y) = 0$ であることと $g(X, Y) = 0$ は同値であり、更にこれは $Y \in \mathbb{R}e_2 \oplus \dots \oplus \mathbb{R}e_{2n+1}$ と同値であることがわかる. 従って、 $\{e_2, \dots, e_{2n+1}\}$ は $K = \ker \theta$ の滑らかな基底場となる. \square

例 3.1.6 (奇数次元ユークリッド空間) $(x_1, \dots, x_n, p_1, \dots, p_n, z)$ を \mathbb{R}^{2n+1} の標準的な座標系とする.

$$\theta := dz - \sum_{i=1}^n p_i dx_i$$

により \mathbb{R}^{2n+1} 上の一次微分形式 θ を定義する. $d\theta = -\sum_{i=1}^n dp_i \wedge dx_i$ であるから、

$$\begin{aligned} (d\theta)^n &= \left(-\sum_{i_1=1}^n dp_{i_1} \wedge dx_{i_1} \right) \wedge \dots \wedge \left(-\sum_{i_n=1}^n dp_{i_n} \wedge dx_{i_n} \right) \\ &= (-1)^n n! \cdot dp_1 \wedge dx_1 \wedge dp_2 \wedge dx_2 \wedge \dots \wedge dp_n \wedge dx_n. \end{aligned}$$

従って、 $\theta \wedge (d\theta)^n = (-1)^n n! \cdot dz \wedge dp_1 \wedge dx_1 \wedge dp_2 \wedge dx_2 \wedge \dots \wedge dp_n \wedge dx_n$ は \mathbb{R}^{2n+1} の各点でゼロにはならない. 従って、命題 3.1.5 より $(\mathbb{R}^{2n+1}, K := \ker \theta)$ は接触多様体となる. この接触多様体を標準接触多様体と呼び、 θ を標準接触形式と呼ぶ.

標準接触多様体は接触多様体の基本的な例である一方で、実は任意の接触多様体の局所的な構造は標準接触形式により決定される; この事実はダルブーの定理として知られている.

定理 3.1.7 (接触構造に関するダルブーの定理) (M, K) を $2n + 1$ 次元接触多様体とし, θ を開集合 $U \subset M$ 上の接触形式とする. このとき U 上に, ある局所座標系 $(V, (x_1, \dots, x_n, p_1, \dots, p_n, z))$ が存在して, $\theta|_V = dz - \sum_{i=1}^n p_i dx_i$ とかける.

接触多様体の重要な例を二つ挙げる.

例 3.1.8 (接触要素の作る多様体) M を n 次元多様体とする. 余接束 T^*M の射影化を $\mathbb{P}(T^*M)$ とする. すなわち, 各点 $p \in M$ において, $\eta, \xi \in T_p^*M \setminus \{0\}$ はある $\lambda \in \mathbb{R} \setminus \{0\}$ が存在して $\eta = \lambda\xi$ なるとき ($\ker \eta = \ker \xi$) に同値であると定め, この同値関係 \sim_p による商空間 $\mathbb{P}(T_p^*M) := (T_p^*M \setminus \{0\}) / \sim_p$ を T_p^*M の射影化と呼び, $\mathbb{P}(T^*M) := \bigcup_{p \in M} \mathbb{P}(T_p^*M)$ を T^*M の射影化と呼ぶ. これには自然に $2n - 1$ 次元多様体の構造が入る.

以下で接触構造を与える. $\pi : \mathbb{P}(T^*M) \rightarrow M$ を自然な射影とする. $[\eta_p] \in \mathbb{P}(T_p^*M) (\subset \mathbb{P}(T^*M))$ において,

$$\begin{aligned} K_{[\eta_p]} &:= (d\pi_{[\eta_p]})^{-1}(\ker \eta_p), \\ &= \ker (\eta_p \circ d\pi_{[\eta_p]} : T_{[\eta_p]}(\mathbb{P}(T^*M)) \rightarrow \mathbb{R}) \end{aligned}$$

により超平面場を定める ($\dim K_{[\eta_p]} = 2n - 2$). \sim_p の定義から集合 $\ker \eta_p$ は代表元の取り方によらないため $K_{[\eta_p]}$ は well-defined である. K が $\mathbb{P}(T^*M)$ の接触構造であることを示す.

具体的に接触形式を構成する. $(U, \varphi = (x_1, \dots, x_n))$ を p の周りの M の局所座標系とし, T^*U 上の正準座標系を $(x_1, \dots, x_n, r_1, \dots, r_n)$ とする. $\eta_p \in T_p^*M \setminus \{0\}$ を固定し, $r_n(\eta_p) \neq 0$ であるとする. このとき, $V := \{r_n \neq 0\}$, $p_i := -\frac{r_i}{r_n}$ ($i \neq n$) とおくことで, $(V, (x_1, \dots, x_n, p_1, \dots, p_{n-1}))$ は $[\eta_p]$ 周りの $\mathbb{P}(T^*M)$ の局所座標系を与える.

$\theta := dx_n - \sum_{i=1}^{n-1} p_i dx_i$ とおくと, 各 $[\eta] \in V$ において $K_{[\eta]} = \ker \theta_{[\eta]}$ である. 実際, 座標系 $(x_1, \dots, x_n, r_1, \dots, r_n)$ において $\eta = r_n dx_n + \sum_{i=1}^{n-1} r_i dx_i = r_n(dx_n - \sum_{i=1}^{n-1} p_i dx_i)$ と書けるため, $\ker(\eta \circ d\pi_{[\eta]}) = \ker \theta_{[\eta]}$ である. θ が接触形式を与えることは例 3.1.6 の通りである. この接触多様体 $(\mathbb{P}(T^*M), K)$ を接触要素の作る多様体と呼ぶ.

例 3.1.9 (リーマン多様体の単位余接束) (M, g) を n 次元リーマン多様体とする. $S(TM) := \{v \in TM \mid g(v, v) = 1\}$ とおく. 微分同相写像 $\Phi_g : TM \ni X \mapsto g(X, \cdot) \in T^*M$ により, $2n - 1$ 次元多様体 $S(T^*M) := \Phi_g(S(TM))$ を定める. $S(T^*M)$ を M の単位余接束という.

$S(T^*M)$ に接触構造を定める. $\pi^{S*} : S(T^*M) \rightarrow M$ を自然な射影とする. $\eta \in S(T^*M)$ において, 次で超平面場 K^{S*} を定める:

$$\begin{aligned} K_\eta^{S*} &:= (d\pi_\eta^{S*})^{-1}(\ker \eta), \\ &= \ker (\eta \circ d\pi_\eta^{S*} : T_\eta(S(T^*M)) \rightarrow \mathbb{R}). \end{aligned}$$

K^{S^*} が接触構造であることを示す. $\Psi : S(T^*M) \rightarrow \mathbb{P}(T^*M)$ を射影とすると, これは $2 : 1$ のはめ込みかつ束写像である. $(\mathbb{P}(T^*M), K)$ を接触要素の作る多様体とすると, $d\Psi(K^{S^*}) = K$ より K^{S^*} の滑らかさが従い, K に同伴する接触形式 θ の Ψ による引き戻し $\Psi^*\theta$ により $S(T^*M)$ 上に接触形式が定まる. 以上により $(S(T^*M), K^{S^*})$ は接触多様体となる.

$$\begin{array}{ccc} S(T^*M) & \xrightarrow{\Psi} & \mathbb{P}(T^*M) \\ & \searrow \pi^{S^*} & \downarrow \pi \\ & & M \end{array}$$

また, $S(T^*M)$ 上には大域的な接触形式が定まる. $\tilde{\theta}$ を T^*M 上のリュウビル形式とする:

$$\tilde{\theta}_{\eta_p} := (\pi_1)^*\eta_p = \eta_p \circ (d\pi_1)_{\eta_p} \quad (\eta_p \in T_p^*M).$$

ここで, $\pi_1 : T^*M \rightarrow M$ は射影である. このときリュウビル形式の制限 $\theta := \tilde{\theta}|_{S(T^*M)}$ は $(S(T^*M), K^{S^*})$ の接触形式を与える. 実際, $\pi^{S^*} = \pi_1|_{S(T^*M)}$ であるため, 各 $\eta_p \in S(T^*M)$ において, $K_{\eta_p}^{S^*} = \ker(\eta_p \circ d\pi_{\eta_p}^{S^*}) = \ker \theta_{\eta_p}$.

定義 3.1.10 (M_i, K_i) ($i = 1, 2$) を接触多様体とし, $\dim M_1 = \dim M_2$ とする. $f : M_1 \rightarrow M_2$ が局所接触微分同相写像であるとは, f ははめ込みであり, $df(K_1) = K_2$ なるときをいう. f が微分同相写像であるとき, f を接触微分同相写像, または接触同型写像と呼ぶ.

3.2 ルジャンドル写像と波面

(W, K) を $2n + 1$ 次元接触多様体とし, M を多様体とする. $f : M \rightarrow W$ が等方的 (isotropic) であるとは, 各点 $p \in M$ に対して $df_p(T_pM) \subset K_{f(p)}$ が成り立つときをいう. θ を接触形式とすると, f が等方的であることと, $f^*\theta = 0$ であることは互いに同値である ($K_{f(p)} = \ker \theta_{f(p)}$).

命題 3.2.1 はめ込み $f : M \rightarrow W$ が等方的ならば, $\dim M \leq n$.

証明: $p \in M$ とし, θ を $f(p)$ 周りの接触形式とする. 命題 3.1.4 より $d\theta_{f(p)}$ は $K_{f(p)}$ のシンプレクティック形式を与える. 一方, 等方性から $f^*\theta = 0$ であるため $f^*d\theta = d(f^*\theta) = 0$. ゆえに $d\theta_{f(p)}$ は $df_p(T_pM)$ 上ゼロである. 従って, $\dim T_pM = \dim df_p(T_pM) \leq \frac{1}{2} \dim K_{f(p)} = n$. \square

定義 3.2.2 $\dim M = n$ のとき等方的はめ込み $f : M \rightarrow W$ をルジャンドルはめ込みと

定義 3.2.3 M を n 次元多様体, N を $n+1$ 次元多様体とする. $f: M \rightarrow N$ がルジャンドル写像であるとは, ルジャンドルはめ込み $L: M \rightarrow \mathbb{P}(T^*N)$ が存在し, $\pi \circ L = f$ が成り立つときをいう. ここで, $\pi: \mathbb{P}(T^*N) \rightarrow N$ は自然な射影である. このとき L を f のルジャンドル持ち上げと呼ぶ. また, 上記の定義において, L がルジャンドルはめ込みとは限らず単に等方的写像であるとき f をフロントルと呼び, このとき L を f の等方的持ち上げと呼ぶ.

$$\begin{array}{ccc} & & \mathbb{P}(T^*N) \\ & \nearrow \exists L & \downarrow \pi \\ M & \xrightarrow{f} & N \end{array}$$

注意 3.2.4 与えられたルジャンドル写像に対して, そのルジャンドル持ち上げは一般的に一意的ではない. 例えば, $f: \mathbb{R} \rightarrow \mathbb{R}^2$ を原点への定値写像とする. $h(t)$ を 1 変数関数とすると,

$$L: \mathbb{R} \ni t \mapsto [\cos h(t)(dx)_o + \sin h(t)(dy)_o] \in \mathbb{P}(T^*\mathbb{R}^2)$$

は $h(t)$ が零点を持たない限り f のルジャンドルはめ込みになっている.

命題 3.2.5 n 次元多様体 M の $n+1$ 次元多様体 N へのはめ込み $f: M \rightarrow N$ はルジャンドル写像である.

証明: [36] を見よ. □

ルジャンドル写像 $f: M \rightarrow N$ は一般に特異点を持つ. すなわち, ある点 $p \in M$ において f の微分写像の階数が真に落ちる. ルジャンドル写像の特異点をルジャンドル特異点と呼ぶ [6, 22]. ルジャンドル写像 $f: M \rightarrow N$ の像 $f(M) \subset N$ は一般に波面 (wavefront) と呼ばれる. f がはめ込みであるときその波面は N における滑らかな超曲面である一方で, f が特異点を持つときその波面は N における“特異点を持った超曲面”となる.

3.3 ルジャンドルファイブレーション

(W, K) を $2n+1$ 次元接触多様体とする. 部分多様体 $L \subset W$ がルジャンドル部分多様体であるとは, 各点 $p \in L$ において $T_p L \subset K_p$ であり, $\dim L = n$ なるときをいう. 前半の条件は包含写像 $\iota: L \rightarrow W$ がルジャンドルはめ込みであることと同値である.

例 3.3.1 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ とする. 標準接触多様体 \mathbb{R}^{2n+1} における (df, f) のグラフ

$$L := \left\{ (\mathbf{x}, \mathbf{p}, z) \in \mathbb{R}^{2n+1} \mid p_i = \frac{\partial f}{\partial x_i}, z = f(\mathbf{x}) \right\}$$

はルジャンドル部分多様体となる. 実際, $df - \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i = 0$.

定義 3.3.2 (W, K) を $2n+1$ 次元接触多様体とし, N を $n+1$ 次元多様体とする. ファイバー束 $\pi : W \rightarrow N$ がルジャンドルファイブレーションであるとは, 各点 $p \in N$ に対し, ファイバー $\pi^{-1}(p) \subset W$ がルジャンドル部分多様体であるときをいう.

例 3.3.3 $T^*\mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{2n+1}$ において $(\mathbf{x}, \mathbf{p}, z)$ を標準的な座標系とする. このとき標準接触形式 $\theta = dz - \mathbf{p}^T d\mathbf{x}$ により $(\mathbb{R}^{2n+1}, \theta)$ は接触多様体となった (例 3.1.6). このときファイバー束 $\pi : \mathbb{R}^{2n+1} \ni (\mathbf{x}, \mathbf{p}, z) \mapsto (\mathbf{x}, z) \in \mathbb{R}^n \times \mathbb{R}$ はルジャンドルファイブレーションとなる. 実際, 各 $p = (\mathbf{x}, z) \in \mathbb{R}^n \times \mathbb{R}$ においてファイバーは $\pi^{-1}(p) = \{\mathbf{x}\} \times \mathbb{R}^n \times \{z\}$ となり, この上で $dx_i = dz = 0$ ($1 \leq i \leq n$) である.

例 3.1.8, 3.1.9 における接触多様体も自然な射影に関してルジャンドルファイブレーションの構造を持つ.

命題 3.3.4 $\pi : \mathbb{P}(T^*N) \rightarrow N$ はルジャンドルファイブレーションの構造を与える.

証明: $\dim N = n+1$ とすると, $\pi : \mathbb{P}(T^*N) \rightarrow N$ は n 次元射影空間をファイバーとするファイバー束である. 例 3.1.8 における $\mathbb{P}(T^*N)$ の接触超平面場の定義より, ファイバー方向の接ベクトルは全て接触超平面に含まれるため, 全てのファイバーはルジャンドル部分多様体である. 従って, $\pi : \mathbb{P}(T^*N) \rightarrow N$ はルジャンドルファイブレーションであることが分かる. \square

同様にして次も分かる.

命題 3.3.5 $\pi^{S*} : S(T^*N) \rightarrow N$ はルジャンドルファイブレーションの構造を与える.

$\Pi : (W, K) \rightarrow N$ をルジャンドルファイブレーションとする. $c \in W$ とし, $p := \Pi(c) \in N$ とする. 各ファイバーはルジャンドル部分多様体であることから $T_c\Pi^{-1}(p) \subset K_c$ である. 従って, 超平面場 K の射影 $d\Pi_c(K_c) \subset T_pN$ は底空間 N における超平面を与える. この超平面 $d\Pi_c(K_c)$ を核として与える $[\eta] \in \mathbb{P}(T_p^*N)$ が K_c から一意的に決まる. この $[\eta]$ を $[d\Pi_c(K_c)]$ とかくことにする. 次の定理により, 任意のルジャンドルファイブレーションは局所的に, 命題 3.3.4 におけるルジャンドルファイブレーションの構造と一致することが分かる.

定理 3.3.6 ルジャンドルファイブレーション $\Pi : (W, K) \rightarrow N$ に対し, $\mathcal{L} : W \ni c \mapsto [d\Pi_c(K_c)] \in \mathbb{P}(T^*N)$ は局所接触微分同相写像であり, $\Pi = \pi \circ \mathcal{L}$ を満たす. 但し, $\pi : \mathbb{P}(T^*N) \rightarrow N$ は自然な射影である.

$$\begin{array}{ccc}
 W & \xrightarrow{\mathcal{L}} & \mathbb{P}(T^*N) \\
 & \searrow \Pi & \swarrow \pi \\
 & & N
 \end{array}$$

証明： $\dim W = 2n + 1$ とする。 \mathcal{L} が局所微分同相写像であることを示せば他の性質は明らかである。 $c \in W$, $p := \Pi(c) \in N$ とする。ファイバー束の定義より、 c の周りのある局所座標系 $(U, (x_0, x_1, \dots, x_n, y_1, \dots, y_n))$ が存在し、 (x_0, \dots, x_n) が p 周りの N の局所座標系を与え、これらの座標系に関して $\Pi(x_0, x_1, \dots, x_n, y_1, \dots, y_n) = (x_0, \dots, x_n)$ となる。 $\Pi : (W, K) \rightarrow N$ はルジャンドルファイブレーションであるから、その接触形式 θ は一般に、座標系 $(x_0, \dots, x_n, y_1, \dots, y_n)$ において、次のように書ける：

$$\theta = \sum_{i=0}^n p_i dx_i.$$

ここで、各 p_i は U 上の関数である。 θ は零点を持たないため、座標系 (x_0, \dots, x_n) を入れ替えることで、 $p_0(c) \neq 0$ として良い。また、必要であれば U を小さくすることで U 上で $p_0 \neq 0$ とできる。このとき、 $\theta' := \frac{1}{p_0} \theta$ もまた (W, K) の接触形式を与えていることに注意する。各 $1 \leq i \leq n$ に対して、 $-p_i/p_0$ を改めて p_i と書くことにすると、 $\theta' = dx_0 - \sum_{i=1}^n p_i dx_i$ 。 θ' は接触形式であることから、

$$\theta' \wedge (d\theta')^n = (-1)^n n! dx_0 \wedge (dp_1 \wedge dx_1) \wedge \dots \wedge (dp_n \wedge dx_n) \neq 0.$$

これより、 $dx_0, dx_1, \dots, dx_n, dp_1, \dots, dp_n$ は一次独立であることが従い、必要であれば U をさらに小さくすることにより、 $(x_0, \dots, x_n, p_1, \dots, p_n)$ は c 周りの W の局所座標系を与えることが分かる。 (x_0, \dots, x_n) は N 上の局所座標系を与えていることに注意すると、この c 周りの W の局所座標系を用いることで、 \mathcal{L} は局所的に、次のように表されることが分かる：

$$\mathcal{L}(c') = \left[(dx_0)_{\Pi(c')} - \sum_{i=1}^n p_i(c')(dx_i)_{\Pi(c')} \right] \in \mathbb{P}(T^*N) \quad (c' \in U).$$

この表示から \mathcal{L} は滑らかであること、さらに局所微分同相写像であることが分かる。 \square

ルジャンドルファイブレーション $\Pi : (W, K) \rightarrow N$ のルジャンドル部分多様体 $L \subset W$ への射影の制限 $\Pi|_L : L \rightarrow N$ は次のようにしてルジャンドル写像となる。

系 3.3.7 写像 $\Pi|_L : L \rightarrow N$ はルジャンドル写像になる。

証明： $\mathcal{L} : W \rightarrow \mathbb{P}(T^*N)$ を定理 3.3.6 の局所接触微分同相写像とする。 $\iota : L \rightarrow W$ を包含写像とすると、 $\mathcal{L} \circ \iota : L \rightarrow \mathbb{P}(T^*N)$ は $\Pi|_L$ に対する $\mathbb{P}(T^*N)$ へのルジャンドルは

め込みである.

$$\begin{array}{ccc}
 W & \xrightarrow{\mathcal{L}} & \mathbb{P}(T^*N) \\
 \uparrow \iota & \nearrow \mathcal{L} \circ \iota & \downarrow \pi \\
 L & \xrightarrow{\Pi|_L} & N
 \end{array}$$

□

従って, 任意のルジャンドルファイブレーションのルジャンドル部分多様体への制限写像はルジャンドル写像になることが分かる. また逆に, ルジャンドル写像が与えられたとき, 射影化された余接束へのルジャンドルはめ込みにより, 局所的にルジャンドル部分多様体として埋め込み, 同一視することで, ルジャンドル写像は局所的に, 射影化された余接束におけるルジャンドル部分多様体への射影の制限と一致することが分かる.

3.4 ルジャンドル写像のリーマン幾何的特徴づけと接続接束

波面のアンビエントな空間がリーマン多様体であるとき, ルジャンドル写像にはより具体的な特徴づけが与えられる. この節で扱う内容は我々の構成には直接用いないが, 構成の違いを確かめるために佐治・梅原・山田によるリーマン計量を用いたルジャンドル写像の特徴づけと接続接束の構成をみる. この節で扱う命題や定理の証明は [36] を参照されたい.

定義 3.4.1 M を n 次元多様体, N を $n + 1$ 次元多様体とする. $f : M \rightarrow N$ が余向き付け可能なルジャンドル写像であるとは, N の単位余接束 $S(T^*N)$ へのルジャンドルはめ込み $\tilde{L} : M \rightarrow S(T^*N)$ が存在し, $\pi^{S^*} \circ \tilde{L} = f$ が成り立つときをいう. ここで, $\pi^{S^*} : S(T^*N) \rightarrow N$ は自然な射影である. このとき \tilde{L} を余向き付け可能なルジャンドル写像 f の $S(T^*N)$ へのルジャンドル持ち上げという.

g を N のリーマン計量とし, $\Phi_g : TN \rightarrow T^*N$ を g が与える同型とする. K^{S^*} を単位余接束 $S(T^*N)$ の接触構造とし, $K^S := d\Phi_g^{-1}(K^{S^*})$ により $(S(TN), K^S)$ を接触多様体とみなす.

命題 3.4.2 M を n 次元多様体とし, (N, g) を $n + 1$ 次元リーマン多様体とする. $f : M \rightarrow N$ とする. このとき, f がルジャンドル写像 (resp. フロントル) であることと, 任意の $p \in M$ に対して, その近傍 $U_p \subset M$ とはめ込み (resp. C^∞ 写像) $\nu_p : U_p \rightarrow S(TN)$ が存在して, 族 $\{(U_p, \nu_p)\}_{p \in M}$ が以下の条件を満たすことは同値である:

- (1) 任意の $p \in M$ に対して $\pi^S \circ \nu_p = f|_{U_p}$;

- (2) 任意の $p \in M$, $q \in U_p$, $v \in df_q(T_q M)$ に対して $g_q(\nu_p(q), v) = 0$;
 (3) 任意の $p, p' \in M$ に対し, $U_p \cap U_{p'} \neq \emptyset$ ならば $\nu_p = \nu_{p'}$ もしくは $\nu_p = -\nu_{p'}$ が成り立つ.

ここで $\pi^S : S(TN) \rightarrow N$ は自然な射影である. この ν_p を f の (局所) 単位法線ベクトル場という. 更に, ν_p が滑らかに M 全体に拡張できることと, f が余向き付け可能であることは同値である.

定義 3.4.3 (接続接束 [35]) M を n 次元多様体とし, $\pi : E \rightarrow M$ を n 階ベクトル束とする. h を E 上の計量とし, ∇ を計量 h と両立的な E の接続とする. また, $\Phi : TM \rightarrow E$ を束写像とする. (E, M, h, ∇, Φ) が接続接束であるとは, M 上の任意のベクトル場 X, Y に対して, 次が成り立つときをいう :

$$\nabla_X \Phi(Y) - \nabla_Y \Phi(X) - \Phi([X, Y]) = 0.$$

多様体間の写像 $f : M \rightarrow N$ による, N 上の階数 r のベクトル束 $\pi : E \rightarrow N$ の引き戻し (f^*E, π^*, M) が以下で定義される :

$$\begin{aligned} f^*E &:= \{(x, \xi) \in M \times E \mid f(x) = \pi(\xi)\}, \\ \pi^* : f^*E \ni (x, \xi) &\mapsto x \in M. \end{aligned}$$

このとき (f^*E, π^*, M) は M 上の階数 r のベクトル束となり, $\iota_f : f^*E \rightarrow E, \iota_f(x, \xi) = \xi$ はベクトル束の準同型写像を与え, 各ファイバーでは線形同型写像 $\iota_f|_{f^*E_p} : f^*E_p \rightarrow E_{f(p)}$ を与える. さらに, $f^*E_p = \pi^{*-1}(p) = \{p\} \times \pi^{-1}(f(p))$ であるから, $\Gamma(E) \ni \xi \mapsto \tilde{\xi} = (\cdot, \xi \circ f) \in \Gamma(f^*E)$ が定義される. ここで $\Gamma(E)$ はベクトル束 $\pi : E \rightarrow N$ の切断全体の集合である. $\tilde{\xi}$ と $\xi \circ f$ を同一視する.

定理 3.4.4 g を N 上のリーマン計量, ∇ を g に関するレビ・チビタ接続, $f : M \rightarrow N$ とする. このとき, 以下の3つを満たす f^*TN 上の接続 $\tilde{\nabla}$ が一意的に存在する :

- (1) 任意の $\xi \in \Gamma(TN)$ と $v \in TM$ に対して, $\tilde{\nabla}_v \tilde{\xi} = \nabla_{df(v)}(\xi \circ f)$ が成り立つ ;
 (2) $\tilde{\nabla}$ は ι_f による g の誘導計量 g^f と両立する ;
 (3) $\Phi_f : TM \ni v_p \mapsto (p, df_p(v_p)) \in f^*TN$ に対して次が成り立つ :

$$\tilde{\nabla}_X \Phi_f(Y) - \tilde{\nabla}_Y \Phi_f(X) - \Phi_f([X, Y]) = 0 \quad (X, Y \in \Gamma(TM)).$$

M を n 次元多様体, (N, g) を $n+1$ 次元リーマン多様体とする. また, $f : M \rightarrow N$ をルジャンドル写像とし, $\{(U_p, \nu_p)\}_{p \in M}$ を f に付随する単位法線ベクトル場とする.

$$\begin{aligned} \mathcal{E}_f &:= \{(q, v) \in f^*TN \mid g_{f(q)}(v, \nu_p(q)) = 0, q \in U_p\}, \\ \pi_f : \mathcal{E}_f \ni (q, v) &\mapsto q \in M \end{aligned}$$

を定めると, $(\mathcal{E}_f, \pi_f, M)$ には自然に M 上のベクトル束としての構造が入り, この定義は単位法線ベクトル場の取り方に依存しない. $(\mathcal{E}_f, \pi_f, M)$ に接続接束としての構造を与える. $(q, v), (q, w) \in \mathcal{E}_f$ に対し,

$$\begin{aligned} h((q, v), (q, w)) &:= g_{f(q)}(v, w) \\ &= g_{f(q)}(\iota_f(q, v), \iota_f(q, w)) \end{aligned}$$

で \mathcal{E}_f 上に計量 h を与える. また, $\Phi_f: TM \rightarrow f^*TN$ を $\Phi_f(v_q) = (q, df_q(v_q))$ で定めると, $g(df_q(v_q), \nu_p(q)) = 0$ よりベクトル束準同型 $\Phi_f: TM \rightarrow \mathcal{E}_f$ が誘導される. さらに, 各 U_p 上で

$$\nabla_X^f \xi := \tilde{\nabla}_X \xi - g(\tilde{\nabla}_X \xi, \nu_p) \nu_p \quad (X \in \mathfrak{X}(M), \xi \in \Gamma(\mathcal{E}_f))$$

を考えると, $\nabla_X^f \xi \in \mathcal{E}_f$ である. ここで, 上の定義は ν_p の土の不定性にはよらないことに注意する. これより, \mathcal{E}_f 上の大域的な接続 ∇^f が定まる. ここで, $\tilde{\nabla}$ は g に関するレビ・チビタ接続 ∇ から誘導される f^*TN 上の接続であり, 上の定義式では $\iota_f: f^*TN \ni (p, v) \rightarrow v \in TN$ による同一視を用いている.

定理 3.4.5 $(\mathcal{E}_f, M, h, \nabla^f, \Phi_f)$ は接続接束である.

このように, 佐治・梅原・山田はリーマン計量による直交射影を用いて波面に付随する接続接束を定めた. ここでリーマン計量を用いている点は接続 ∇^f の直交射影による構成のみであることに注意する. 次章以降で議論される概ヘッセ多様体ではこの接続接束が本質的にその構成に関わる. 但し, 我々はリーマン計量が退化する状況を考えているため, リーマン計量の代わりにアファイン構造を利用して同様の概念を定める. §4.1.3 にてその詳細な定義を与える.

第 4 章

概ヘッセ構造

この章では，退化した計量を許容する双対平坦構造の一般化として概ヘッセ多様体を導入し，その理論について述べる．一般化にあたり，まずルジャンドル双対性および従来の意味の双対平坦構造を接触幾何学の枠組みで捉え直す．その後，アファイン構造を用いて連接接束を定め，概ヘッセ多様体を導入しその性質を探る．

4.1 概ヘッセ多様体の理論

4.1.1 ダブルファイブレーション構造と双対平坦構造

まずヘッセ多様体におけるルジャンドル双対性を接触幾何学の枠組みで捉え直す． $(\mathbb{R}^{2n+1} = T^*\mathbb{R}^n \times \mathbb{R} = J^1(\mathbb{R}^n, \mathbb{R}), \xi = \ker \theta)$ を例 3.1.6 における標準接触多様体とし，その座標系を $(\mathbf{x}, \mathbf{p}, z)$ で表すことにする ($J^1(\mathbb{R}^n, \mathbb{R})$ は n 変数関数の 1-ジェット空間である)．底空間を \mathbb{R}_x^n ，ファイバーを \mathbb{R}_p^n と書き区別することにする．例 3.3.3 では， $\pi : \mathbb{R}^{2n+1} \ni (\mathbf{x}, \mathbf{p}, z) \mapsto (\mathbf{x}, z) \in \mathbb{R}_x^n \times \mathbb{R}$ がルジャンドルファイブレーションを与えることを確かめた．次で与えられる変換 $\mathcal{L} : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}^{2n+1}$ を考える：

$$(\mathbf{x}', \mathbf{p}', z') = \mathcal{L}(\mathbf{x}, \mathbf{p}, z) := (\mathbf{p}, \mathbf{x}, \mathbf{p}^T \mathbf{x} - z).$$

この \mathcal{L} は接触微分同相写像である．すなわち， \mathcal{L} は超平面場 ξ を保つ；実際， $\mathcal{L}^* \theta = \mathcal{L}^*(dz' - \mathbf{p}'^T d\mathbf{x}') = -\theta$ ．

$$\pi' := \pi \circ \mathcal{L} : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}_p^n \times \mathbb{R}, (\mathbf{x}, \mathbf{p}, z) \mapsto (\mathbf{p}, \mathbf{p}^T \mathbf{x} - z)$$

とおくと，これもまたルジャンドルファイブレーションである．このとき，標準接触多様体のダブルファイブレーション構造とは次で定義される図式である：

$$\mathbb{R}_x^n \times \mathbb{R} \xleftarrow{\pi} \mathbb{R}^{2n+1} \xrightarrow{\pi'} \mathbb{R}_p^n \times \mathbb{R} \quad (dL)$$

$\Pi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ を第一成分への射影とし,

$$\pi_1 = \Pi \circ \pi : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}_x^n, \quad \pi'_1 = \Pi \circ \pi' : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}_p^n$$

とおく. L を \mathbb{R}^{2n+1} のルジャンドル部分多様体とする. L が正則モデルであるとは, 射影 π_1, π'_1 によって L がそれぞれある開集合 $U \subset \mathbb{R}_x^n$ と $V \subset \mathbb{R}_p^n$ に微分同相であるときをいう [33]. このとき, U 上のある関数 $z = f(\mathbf{x})$ と V 上のある関数 $z' = \varphi(\mathbf{p})$ が存在して

- $L \subset \mathbb{R}^{2n+1} = J^1(\mathbb{R}_x^n, \mathbb{R})$ は (df, f) のグラフであり;
- $\mathcal{L}(L) \subset \mathbb{R}^{2n+1} = J^1(\mathbb{R}_p^n, \mathbb{R})$ は $(d\varphi, \varphi)$ のグラフである.

これより, 次の関係式を得る:

$$\mathbf{p} = \frac{\partial f}{\partial \mathbf{x}}, \quad \mathbf{x} = \frac{\partial \varphi}{\partial \mathbf{p}}, \quad f(\mathbf{x}) + \varphi(\mathbf{p}) - \mathbf{p}^T \mathbf{x} = 0.$$

座標変換は勾配写像 $U \rightarrow V$, $\mathbf{x} \mapsto \mathbf{p} = \frac{\partial f}{\partial \mathbf{x}}$ である. これは微分同相写像であり, このため任意の $x \in U$ において $f(\mathbf{x})$ のヘッセ行列は非退化であることが分かる. また, 逆写像 $V \rightarrow U$ は $\mathbf{p} \mapsto \mathbf{x} = \frac{\partial \varphi}{\partial \mathbf{p}}$ で与えられ, そのヘッセ行列は $f(\mathbf{x})$ のヘッセ行列の逆行列である. $z' = \varphi(\mathbf{p})$ を $z = f(\mathbf{x})$ のルジャンドル変換と呼び, 逆もまたそのように呼ぶ. また f をポテンシャル関数, φ をその双対ポテンシャル関数と呼ぶ. この関係がルジャンドル双対性であり, ヘッセ多様体の局所的構造はまさにこの正則モデルに対応する.

注意 4.1.1 (射影双対) ルジャンドル双対性は射影双対性のアファイン的表現である (cf. [22]). $\mathbb{P}^{n+1} := \mathbb{R}P^{n+1}$ を実射影空間とし, $\mathbb{P}^{n+1*} := \mathbb{R}P^{n+1*}$ をその双対射影空間とする. N を $\mathbb{P}^{n+1} \times \mathbb{P}^{n+1*}$ における $p \in \ker \lambda$ を満たす組み (p, λ) からなる部分多様体とする. すなわち N は余次元 1 の部分多様体 ($\dim N = 2n + 1$) で方程式

$$p_0 x_0 + p_1 x_1 + \cdots + p_{n+1} x_{n+1} = 0$$

によって定義される. ここで $p = [x_0 : x_1 : \cdots : x_{n+1}] \in \mathbb{P}^{n+1}$, $\lambda = [p_0 : p_1 : \cdots : p_{n+1}] \in \mathbb{P}^{n+1*}$ である. N は射影化された余接束 $\mathbb{P}(T^*\mathbb{P}^{n+1})(= \mathbb{P}(T^*\mathbb{P}^{n+1*}))$ (cf. 例 3.1.8) と自然に同一視され, このとき N は $2n + 1$ 次元の接触多様体になる. $x_{n+1} \neq 0$ と $p_0 \neq 0$ によって定義される N の開集合 O_N を考える. $x_{n+1} = p_0 = -1$, $z = x_0$, $z' = p_{n+1}$ とおくことで, 定義方程式は次のように書ける:

$$z + z' - \mathbf{p}^T \mathbf{x} = 0.$$

明らかに O_N は二つの座標系 $(\mathbf{x}, \mathbf{p}, z)$, $(\mathbf{p}, \mathbf{x}, z')$ を持ち, これらの間の座標変換はまさに, N の接触構造を保つ $\mathcal{L} : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}^{2n+1}$ である. 射影幾何学においては, ダブルルジャンドルファイブレーション

$$\mathbb{P}^{n+1} \xleftarrow{\pi} N \xrightarrow{\pi'} \mathbb{P}^{n+1*}$$

は点と超平面における双対原理を表している．ここで， π, π' は自然な射影である．この関式を O_N に制限し，座標 $(\mathbf{x}, \mathbf{p}, z)$ を用いて O_N を \mathbb{R}^{2n+1} と同一視することで，関式 (dL) を得る．例えば， $n = 1$ のとき，パラメータ付けられた平面曲線

$$C : (x, z) := (x, f(x)) \in \mathbb{R}^2 \subset \mathbb{P}^2$$

を考える．このとき，接線の軌跡はその射影双対をなす：

$$C^* : (p, z') = \left(\frac{df}{dx}(x), x \frac{df}{dx}(x) - f(x) \right) \in \mathbb{R}^{2*} \subset \mathbb{P}^{2*}.$$

ここで， C が凸であるとき， C^* もまたそうである．また， C が変曲点を持つとき，例えば， $f(x) = \frac{1}{3}x^3 + \dots$ であるとき， C^* は対応する点， $(p, z') = (x^2 + \dots, \frac{2}{3}x^3 + \dots)$ ，においてカスプを持つ．従って，この場合は， C^* は局所的に 2 値関数 $z' = \pm \frac{2}{3}p^{3/2} + \dots$ ($p \geq 0$) のグラフで表される．

アファインルジャンドル同値とは， $\mathbb{R}_x^n \times \mathbb{R}$ と $\mathbb{R}_p^n \times \mathbb{R}$ 上のアファイン変換

$$\begin{aligned} F(\mathbf{x}, z) &= (A\mathbf{x} + \mathbf{b}, z + \mathbf{c}^T \mathbf{x} + d), \\ F^*(\mathbf{p}, z') &= (A'\mathbf{p} + \mathbf{b}', z' + \mathbf{c}'^T \mathbf{p} + d') \end{aligned}$$

によって得られるアファイン変換 $\mathcal{L}_F : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}^{2n+1}$ ，

$$\mathcal{L}_F(\mathbf{x}, \mathbf{p}, z) = (A\mathbf{x} + \mathbf{b}, A'\mathbf{p} + \mathbf{b}', z + \mathbf{c}^T \mathbf{x} + d)$$

のことである．ここで A は正則行列であり，

$$A' = (A^T)^{-1}, \quad \mathbf{b}' = A'\mathbf{c}, \quad \mathbf{b} = A\mathbf{c}', \quad d' = \mathbf{b}'^T \mathbf{b} - d.$$

\mathcal{L}_F は F (あるいは F^*) によって決まることに注意する．簡単に分かるように， \mathcal{L}_F は接触形式とダブルファイブレーション (dL) を保つ： $\mathcal{L}_F^* \theta = \theta$ ．また，次の関式を可換にする．

$$\begin{array}{ccccc} \mathbb{R}_x^n \times \mathbb{R}_z & \xleftarrow{\pi} & \mathbb{R}^{2n+1} & \xrightarrow{\pi'} & \mathbb{R}_p^n \times \mathbb{R}_{z'} \\ \downarrow F \simeq & & \downarrow \mathcal{L}_F & & \downarrow F^* \\ \mathbb{R}_x^n \times \mathbb{R}_z & \xleftarrow{\pi} & \mathbb{R}^{2n+1} & \xrightarrow{\pi'} & \mathbb{R}_p^n \times \mathbb{R}_{z'} \end{array}$$

定義 4.1.2 \mathbb{R}^{2n+1} の二つのルジャンドル部分多様体 L_1, L_2 がアファインルジャンドル同値であるとは，ある \mathcal{L}_F が存在して $\mathcal{L}_F(L_1) = L_2$ なるときをいう．

ヘッセ多様体における張り合わせの条件は，アファインルジャンドル同値そのものであり (cf. 注意 2.1.6)，ポテンシャル f のヘッセ行列から定まる擬リーマン計量 h と射影 π_1, π'_1 を通して，固定された $\mathbb{R}_x^n, \mathbb{R}_p^n$ のアファイン構造から誘導される平坦アファイン接続 ∇, ∇^* からなる L 上の双対平坦構造 (h, ∇, ∇^*) はアファインルジャンドル同値で不変である．従って，次のような言い換えが成り立つ：

命題 4.1.3 任意の双対平坦多様体あるいはヘッセ多様体は \mathbb{R}^{2n+1} のいくつかの正則モデル L_α のアファインルジャンドル同値による張り合わせで構成される. その計量 h とアファイン接続 ∇, ∇^* の組みもまた, 各 L_α の双対平坦構造 $(h_\alpha, \nabla_\alpha, \nabla_\alpha^*)$ から大域的に定義され, この定義は well-defined である.

この張り合わせによる構成は §4.1.4 における定義 4.1.22 で, より一般的な形式で再び述べる.

注意 4.1.4 各張り合わせ写像は, \mathbb{R}^{2n+1} における正則モデルの近傍上に作用するため, 上記の張り合わせによる構成は, ある接触多様体のルジャンドル部分多様体として実現される; 同様にあるシンプレクティック多様体のラグランジュ部分多様体 (部分多様体であって, 包含写像によるシンプレクティック形式の引き戻しが消えている最大次元のもの) を構成する. (M, h, ∇, ∇^*) を双対平坦多様体とし, $h = \nabla df$ なる大域的なポテンシャル $f: M \rightarrow \mathbb{R}$ が存在する場合を考える. ∇ -アファイン座標系の局所座標系 U_α をとる. このとき, 局所ポテンシャル $f|_{U_\alpha}$ は 1-ジェット空間 $J^1(U_\alpha, \mathbb{R}) = T^*U_\alpha \times \mathbb{R}$ における正則モデル L_α を定め, これらは $c = 0, d = 0$ なるアファインルジャンドル同値 \mathcal{L}_F によって貼り合わされる. 逆に, このアファインルジャンドル同値によって貼り合わされた局所モデルは大域的なポテンシャルを持つ双対平坦多様体を構成する. 条件を弱め, 閉 1 次微分形式 σ が存在し $h = \nabla\sigma$ を満たすとする (ポアンカレの補題より局所的に $\sigma = df$ なるポテンシャル f が存在する). このとき, M はコシュール型であるという [37]. この場合は, $c = 0$, 場合によっては $d \neq 0$ なるアファインルジャンドル同値 \mathcal{L}_F によって正則モデルを張り合わせることに対応する.

4.1.2 e/m -波面と e/m -コースティクス

一般に $\pi: N \rightarrow B$ をルジャンドルファイブレーション, $L \subset N$ をルジャンドル部分多様体としたとき, 波面 $\pi(L) \subset B$ を $W(L)$ と書くことにする.

ルジャンドル部分多様体 $L \subset \mathbb{R}^{2n+1}$ が正則モデルであるとき, ダブルファイブレーション構造 (dL) による二つの波面 $\pi(L), \pi'(L)$ はそれぞれ双対平坦多様体におけるポテンシャル関数と双対ポテンシャル関数のグラフに対応していた. 一般にルジャンドル写像は特異点を持ち得る ($T_p L$ がファイバーに接する) が, このとき波面はもはや滑らかな部分多様体にはならない.

\mathbb{R}^{2n+1} 上のダブルファイブレーション π, π' の図式 (dL) と任意のルジャンドル部分多様体 $L \subset \mathbb{R}^{2n+1}$ を考える. 2つのルジャンドル写像

$$\pi^e := \pi \circ \iota: L \rightarrow \mathbb{R}_x^n \times \mathbb{R}_z, \quad \pi^m := \pi' \circ \iota: L \rightarrow \mathbb{R}_p^n \times \mathbb{R}_{z'}$$

を考え、これらを情報幾何学の慣習に従い、 e -、 m -ルジャンドル写像と呼ぶ（接頭辞“ e -”と“ m -”は統計学、すなわち指数型と混合型に由来する [3]）。

定義 4.1.5 (e/m -波面 [33])

$$W_e(L) := \pi^e(L) \subset \mathbb{R}_x^n \times \mathbb{R}_z, \quad W_m(L) := \pi^m(L) \subset \mathbb{R}_p^n \times \mathbb{R}_z,$$

とおき、これらをそれぞれ L に付随する e/m -波面と呼ぶ。

e/m -波面は互いに、点と超平面の射影双対におけるルジャンドル双対の関係にある（注意 4.1.1）。

通常、 $T^*\mathbb{R}^n$ のラグランジュ部分多様体の底空間への射影はラグランジュ写像と呼ばれる [6, 22]。用語の濫用により次の 2 つをそれぞれ e -、 m -ラグランジュ写像と呼ぶ：

$$\pi_1^e = \Pi \circ \pi^e : L \rightarrow \mathbb{R}_x^n, \quad \pi_1^m = \Pi \circ \pi^m : L \rightarrow \mathbb{R}_p^n.$$

簡単に分かるように、 $p \in L$ における次の 2 つは互いに同値である：

- p はルジャンドル写像 $\pi^e : L \rightarrow \mathbb{R}_x^n \times \mathbb{R}_z$ の特異点である；
- p はラグランジュ写像 $\pi_1^e : L \rightarrow \mathbb{R}_x^n$ の特異点である。

実際、任意の $v \in T_p L$ は $dz_p(v) - \mathbf{p}(p)^T d\mathbf{x}_p(v) = 0$ を満たし、 $d\mathbf{x}_p(v) = 0$ であるとき、 $dz_p(v) = 0$ となる。すなわち、ルジャンドル部分多様体 $L \subset T^*\mathbb{R}^n \times \mathbb{R}$ は z -軸と接しない。

定義 4.1.6 (e/m -コースティクス [33]) e -臨界点集合 $C(\pi_1^e) \subset L$ とは、 e -ラグランジュ写像 $\pi_1^e : L \rightarrow \mathbb{R}_x^n$ の全ての特異点からなる集合であり、その像 $\pi_1^e C(\pi_1^e) \subset \mathbb{R}_x^n$ を L に付随する e -コースティクスと呼ぶ。 m -に相当する概念も全て同様に定義する。

定義 4.1.7 ([33]) L が $p \in L$ の周りで局所正則モデルであるとは、 p の L における開近傍が存在して、それが \mathbb{R}^{2n+1} の正則モデルとなるときをいう。すなわち、 p は e -臨界点でも m -臨界点でもどちらでもない。

$p \in L$ が e -臨界点ではなく、 m -臨界点である場合を考える（以下の例 4.1.8, 4.1.9 を見よ）。この場合では、 p の周りの L の局所座標系として \mathbf{x} をとることができ、 $W_e(L)$ は $\pi_1^e(p)$ の近くで定義されるある局所ポテンシャル関数 $z = f(\mathbf{x})$ のグラフとなっている。このとき、座標系 \mathbf{x} に関して e -ラグランジュ写像は恒等写像として記述され、 e -コースティクスは空集合である。一方で、 m -ラグランジュ写像 $\pi^m : L \rightarrow \mathbb{R}_p^n$ は勾配写像 $\mathbf{p} = \nabla f(\mathbf{x})$ として記述される。 p は π^m の臨界点であるから、 $W_m(L)$ は特異的である。この場合に、 L を退化ポテンシャルを持つ特異モデルと呼ぶことにする。特にこの場合は、 f が変曲点を

持つことを許容し、従って $W_m(L)$ は \mathbb{R}_p^n の m -コースティクスに沿って分岐する多価関数 $z' = \varphi(\mathbf{p})$ のグラフとなっている。

例 4.1.8 (A_2 -特異点)

$$f(\mathbf{x}) = \frac{x_1^3}{3} + \frac{x_2^2}{2}$$

とおく。このとき、 $\mathbf{p} = (p_1, p_2) = (x_1^2, x_2)$ であり、ヘッセ行列 $h = \nabla^2 f$ の退化集合 Σ は $x_1 = 0$ によって定義される。さらに、図 4.1 の左側に示されるように、

- e -波面 $W_e(L)$ は滑らかであり、 Σ に沿って変曲点を持つ。 e -コースティクスは存在しない；
- m -波面 $W_m(L)$ はカusp辺を持つ特異曲面である。これは $p_1 \geq 0$ において定義される 2 値双対ポテンシャル

$$z' = \mathbf{p}^T \mathbf{x} - z = \frac{2}{3}x_1^3 + \frac{1}{2}x_2^2 = \pm \frac{2}{3}p_1^{3/2} + \frac{1}{2}p_2^2$$

のグラフであり、 m -コースティクス $p_1 = 0$ に沿って分岐する。

この特異点は、ヘッセ行列が非負であるときには現れない。例えば、全ての統計モデルに対してフィッシャー・ラオ行列は非負である。

例 4.1.9 (A_3 -特異点)

$$f(\mathbf{x}) = \frac{x_1^4}{4} + \frac{x_2^2}{2}$$

とおく。このとき、 $\mathbf{p} = (p_1, p_2) = (x_1^3, x_2)$ であり、ヘッセ行列 $h = \nabla^2 f$ の退化集合 Σ は $x_1 = 0$ によって定義される。さらに、図 4.1 の右側に示されるように、

- e -波面 $W_e(L)$ は滑らかであり、凸である。 e -コースティクスは存在しない；
- m -波面 $W_m(L)$ は特異曲面である。これは大域的に定義される双対ポテンシャル

$$z' = \mathbf{p}^T \mathbf{x} - z = \frac{3}{4}x_1^4 + \frac{1}{2}x_2^2 = \frac{3}{4}p_1^{4/3} + \frac{1}{2}p_2^2$$

のグラフであるが、 m -コースティクス $p_1 = 0$ に沿って特異的である。

これは関数の極小値の分類に現れる典型的な特異点型である [42]。

さらに、 $p \in L$ が同時に e -臨界点と m -臨界点であることも起き得る。このとき、 $W_e(L)$ と $W_m(L)$ はそれぞれ、 $\pi^e(p)$ と $\pi^m(p)$ において特異的になる。一般に、陰関数定理により、分割 $I \sqcup J = \{1, \dots, n\}$ ($I \cap J = \emptyset$) が存在し、座標 $\mathbf{x}_I = (x_i)^T$ と $\mathbf{p}_J = (p_j)^T$ ($i \in I, j \in J$) によって L は p の周りで局所的にパラメータ付けされる。実際、ある関数

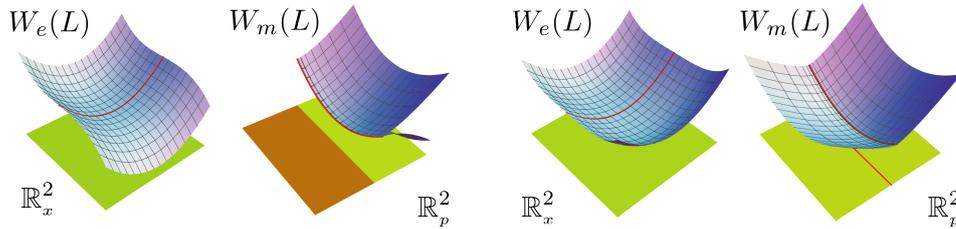


図 4.1 e/m -波面と e/m -コースティクス (例 4.1.8, 4.1.9).

$g(\mathbf{x}_I, \mathbf{p}_J)$ が存在し, p の近くで L を次のように表現することができる:

$$\mathbf{p}_I = \frac{\partial g}{\partial \mathbf{x}_I}, \quad \mathbf{x}_J = -\frac{\partial g}{\partial \mathbf{p}_J}, \quad z = \mathbf{p}_J^T \mathbf{x}_J + g(\mathbf{x}_I, \mathbf{p}_J).$$

ここで, $\frac{\partial g}{\partial \mathbf{x}_I} = (\frac{\partial g}{\partial x_{i_1}}, \dots)^T$ ($I = (i_1, i_2, \dots)$). これは §3.3 の例 3.3.1 における形式と, 接触構造を保つ標準的な変換

$$\mathbb{R}^{2n+1} \rightarrow \mathbb{R}^{2n+1} \quad (\mathbf{x}, \mathbf{p}, z) \mapsto (\mathbf{x}_I, \mathbf{p}_J, \mathbf{p}_I, -\mathbf{x}_J, -\mathbf{p}_J^T \mathbf{x}_J + z)$$

から従う. 通常, $g(\mathbf{x}_I, \mathbf{p}_J)$ は L の p 周りにおける母関数と呼ばれる [6, §20]. 特に, $J = \emptyset$ (resp. $I = \emptyset$) の場合, 母関数はポテンシャル $z = f(\mathbf{x})$ (resp. 双対ポテンシャル $z' = \varphi(\mathbf{p})$ のマイナス倍) である. e/m -ルジャンドル写像は局所的に次のように表現できる:

$$\begin{aligned} \pi^e : (L, p) &\rightarrow \mathbb{R}_x^n \times \mathbb{R}_z, & (\mathbf{x}_I, \mathbf{p}_J) &\mapsto \left(\mathbf{x}_I, -\frac{\partial g}{\partial \mathbf{p}_J}, -\mathbf{p}_J^T \frac{\partial g}{\partial \mathbf{p}_J} + g(\mathbf{x}_I, \mathbf{p}_J) \right), \\ \pi^m : (L, p) &\rightarrow \mathbb{R}_p^n \times \mathbb{R}_{z'}, & (\mathbf{x}_I, \mathbf{p}_J) &\mapsto \left(\frac{\partial g}{\partial \mathbf{x}_I}, \mathbf{p}_J, \mathbf{x}_I^T \frac{\partial g}{\partial \mathbf{x}_I} - g(\mathbf{x}_I, \mathbf{p}_J) \right). \end{aligned}$$

e/m -ラグランジュ写像 π_1^e, π_1^m もまた, 最後の z, z' 座標を無視することでそれぞれ得られる.

例 4.1.10

$$g(x_1, p_2) = \frac{x_1^3}{3} + \frac{p_2^4}{4}$$

を母関数とするルジャンドル部分多様体 $L \subset \mathbb{R}^5$ を考える. e/m -ルジャンドル写像 π^e と π^m はそれぞれ (x_1, p_2) を

$$(x_1, x_2, z) = \left(x_1, -p_2^3, \frac{x_1^3}{3} - \frac{3p_2^4}{4} \right), \quad (p_1, p_2, z') = \left(x_1^2, p_2, \frac{2x_1^3}{3} - \frac{p_2^4}{4} \right)$$

へ写す. これらの像 $W_e(L)$ と $W_m(L)$ は特異曲面であり, e/m -コースティクスは \mathbb{R}_x^2 上では $x_2 = 0$ によって定義され, \mathbb{R}_p^2 上では $p_1 = 0$ によって定義される (図 4.2 を見よ).

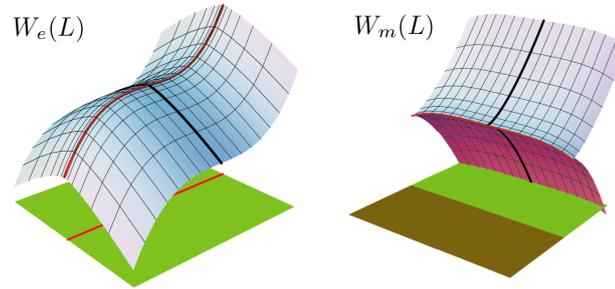


図 4.2 e/m -波面が共に特異的である場合 (例 4.1.10).

注意 4.1.11 (階層的構造) ポテンシャル $z = f(\mathbf{x})$ を持つ双対平坦多様体 L (すなわち正則モデル) に対し, 2つの座標系 \mathbf{x} と $\mathbf{p} (= \frac{\partial f}{\partial \mathbf{x}})$ はそれぞれ ∇ -平坦, ∇^* -平坦であり, これらは階層的構造を生み出す – 新しい座標系 $(\mathbf{x}_I, \mathbf{p}_J)$ をとり, これを甘利における混合座標と呼び [1], これは, L 上に2つの葉層構造を作り出し, それらは $\{\mathbf{x}_I = \text{定数}\}$ と $\{\mathbf{p}_J = \text{定数}\}$ によって定義される; これらの葉はそれぞれ ∇^* -平坦と ∇ -平坦であり, f に付随するヘッセ計量に関してそれらは互いに直交する. この構造は応用上有用であり, 詳しくは [1] を参照されたい. 任意のルジャンドル部分多様体 L に対して, 大域的なポテンシャルは一般には存在しないが, 上で見たように, 任意の $p \in L$ に対してその近傍 U 上の母関数 $g(\mathbf{x}_I, \mathbf{p}_J)$ はいつでも見つけることができる (p に依存して分割 $I \sqcup J$ の選び方には自由度がある). このことから, 混合座標 $(\mathbf{x}_I, \mathbf{p}_J)$ と U 上の2つの葉層構造が局所的に定義される (以下の注意 4.1.18 を見よ). 通常, e/m -コースティクスが存在 (すなわち h が退化すること) により, これらの座標系を L 上で大域的に拡張することはできないが, そうであっても, この葉層構造は大域的な整合性を持つ. 次の小節にて適切な定式化を与える.

4.1.3 e/m -連接接束

L を \mathbb{R}^{2n+1} のルジャンドル部分多様体とする. 前小節で見たように, e -波面 $W_e(L)$ および m -波面 $W_m(L)$ は一般には滑らかな多様体にはならないが, その“接束”の代わりに連接接束をそこに付随させる. §3.4 ではリーマン計量を用いて連接接束を定義したが, 計量を用いることができないため, 代わりに波面の周りの固定されたアファイン構造を利用し, 言わば連接接束のアファイン幾何学版を導入する. 任意の点 $p \in L$ は $T_{\pi^e(p)}(\mathbb{R}_x^n \times \mathbb{R}_z) = \mathbb{R}_x^n \times \mathbb{R}_z$ における超平面 E_p を定め, そのような超平面の族が階数 n の L 上のベクトル束をつくる:

$$E(= E_L) := \{ (p, w) \in L \times (\mathbb{R}_x^n \times \mathbb{R}_z) \mid dz_p(w) - \mathbf{p}(p)^T d\mathbf{x}_p(w) = 0 \}.$$

L はルジャンドル部分多様体であるため、

$$T_p L \subset \xi_p = \ker \theta_p = (d\pi_p)^{-1}(E_p)$$

であり、このことから $d\pi^e(T_p L) \subset E_p$. ここで E_p は p における E のファイバーである. このとき、束写像

$$\Phi : TL \rightarrow E, \quad v_p \mapsto d\pi_p^e(v_p)$$

が定まる. Φ が同型であることと $\pi^e : L \rightarrow \mathbb{R}_x^n \times \mathbb{R}_z$ がはめ込みとなることが同値であることに注意する.

注意 4.1.12 E は e -波面 $W_e(L)$ の接束の“極限”であることに注意する. $d\pi_p : T_p \mathbb{R}^{2n+1} \rightarrow T_{\pi(p)}(\mathbb{R}_x^n \times \mathbb{R})$ の核は $\{\frac{\partial}{\partial p_1}, \dots, \frac{\partial}{\partial p_n}\}$ によって張られている. $p \in L$ が e -ルジャンドル写像 $\pi^e : L \rightarrow \mathbb{R}_x^n \times \mathbb{R}$ の正則点であるとき, $T_p L \cap \ker d\pi_p = \{0\}$ であり,

$$E_p = \text{Im } d\pi_p^e(T_p L) = T_{\pi^e(p)} W_e(L).$$

実際、この場合には、 π^e は p の周りではめ込みであり、 $W_e(L)$ は $\pi^e(p)$ 周りの滑らかな部分多様体である. π^e の正則点 $p_i \in L$ の列が臨界点 p に収束するとき、ベクトル束 E の連続性より、 $T_{p_i} L$ の像は E_p へ収束する ($\mathbb{R}^{n+1} = \mathbb{R}_x^n \times \mathbb{R}$ の n -平面のグラスマン多様体において). この場合には、 $W_e(L)$ は $\pi^e(p)$ の周りで特異的であるためこの点において接空間は定義されず、しかしながらその代替物として接空間の極限 E_p を持つ. E_p のもう一つの特徴付けは

$$E_p = \ker [d\pi'_p : T_p \mathbb{R}^{2n+1} \rightarrow T_{\pi^m(p)}(\mathbb{R}_p^n \times \mathbb{R}_{z'})].$$

ここで、 $E_p \subset \mathbb{R}_x^n \oplus \{0\} \oplus \mathbb{R}_z \subset \mathbb{R}^{2n+1} = T_p \mathbb{R}^{2n+1}$ とみなしている. 実際、 p における π' のヤコビ行列は

$$\begin{bmatrix} O & E & 0 \\ \mathbf{p}(p)^T & \mathbf{x}(p)^T & -1 \end{bmatrix}$$

であるため、その核は $dz_p - \mathbf{p}(p)^T d\mathbf{x}_p = 0$ と $d\mathbf{p}_p = 0$ によって得られる. ここで接触超平面場は次の直和分解を持つことに注意する:

$$\xi_p = \ker d\pi'_p \oplus \ker d\pi_p.$$

$\tilde{\nabla}$ をアファイン空間 $\mathbb{R}_x^n \times \mathbb{R}$ の平坦接続とし、任意の $p \in L$ に対して、 $\psi_p : \mathbb{R}_x^n \times \mathbb{R} \rightarrow E_p$ を z -軸に沿った線形射影とする. このとき、 L 上のベクトル束 E の接続 ∇^E は自然に

$$\nabla_X^E \xi(p) := \psi_p \circ \tilde{\nabla}_X \xi(p)$$

によって定義される. ここで、 ξ は E の任意の切断であり、 X は L の p 周りのベクトル場である.

補題 4.1.13 ([33]) 上記で定義される接続 ∇^E は平坦であり, 相対的な振率が消えている. すなわち, 任意の L 上のベクトル場 X, Y に対して次が成り立つ:

$$\nabla_X^E(\Phi(Y)) - \nabla_Y^E(\Phi(X)) = \Phi([X, Y]).$$

証明: $s_i(p) := \frac{\partial}{\partial x_i} + p_i \frac{\partial}{\partial z} \in E_p$ ($1 \leq i \leq n$) とおくと, これらは E の大域的な平坦切断を与える:

$$\nabla_X^E s_i = \psi_p(\tilde{\nabla}_X s_i) = \psi_p(X(p_i) \frac{\partial}{\partial z}) = 0.$$

従って ∇^E は平坦である. 次に, Φ が $\pi_1^e = (f_1, \dots, f_n) : L \rightarrow \mathbb{R}_x^n$ のヤコビ行列によって表現される点に注意する. すなわち, L の局所座標系 (t_1, \dots, t_n) において $\Phi(\partial_j) = \sum_{i=1}^n (\partial_j f_i) s_i$ である. ここで, $\partial_j = \frac{\partial}{\partial t_j}$ である. $X = \sum_{k=1}^n a_k \partial_k$, $Y = \sum_{j=1}^n b_j \partial_j$ とおくと,

$$\nabla_X^E \Phi(Y) = \sum_{i,j,k} ((\partial_k \partial_j f_i) a_k b_j + (\partial_j f_i) a_k (\partial_k b_j)) s_i.$$

残りは直接計算することによって分かる. \square

定義 4.1.14 ([33]) (E, Φ, ∇^E) を e -波面 $W_e(L)$ に付随する接続接束と呼ぶ.

全て同様の方法で, m -波面 $W_m(L)$ に対して, 接続接束 E' および束写像 $\Phi' := d\pi^m : TL \rightarrow E'$, 平坦接続 $\nabla^{E'}$ が共に定義される:

$$E' (= E'_L) := \{ (p, w) \in L \times (\mathbb{R}_p^n \times \mathbb{R}_{z'}^n) \mid dz'(w) - \mathbf{x}(p)^T dp(w) = 0 \}.$$

実際, \mathbb{R}^{2n+1} の異なる座標系 $(\mathbf{p}, \mathbf{x}, z')$ を用いることで, ダブルルジャンドルフアイブレーション (dL) は写像の組 $(\pi \circ \mathcal{L}^{-1}, \pi)$ として捉え直すことができ, このとき, この双対的表現にこれまで議論してきた構成を当てはめると, $(E', \Phi', \nabla^{E'})$ が定まる. 特に, E'_p は $\ker d\pi_p^e$ と同一視される (注意 4.1.12 を見よ).

我々は E と E' を L 上のベクトル束として定義したが, これらは実際のところ周囲の空間である \mathbb{R}^{2n+1} 上で定義される. 接触超平面場 ξ は次の直和分解を持つ:

$$\xi_p = \ker d\pi'_p \oplus \ker d\pi_p \simeq E_p \oplus E'_p \simeq \mathbb{R}_x^n \oplus \mathbb{R}_p^n.$$

ここで, E, E' 両方に対して平坦接続の標準枠

$$s_i(p) = \frac{\partial}{\partial x_i} + p_i \frac{\partial}{\partial z} \in E_p, \quad s'_i(p) = \frac{\partial}{\partial p_i} + x_i \frac{\partial}{\partial z'} \in E'_p \quad (4.1)$$

がとれ, これにより E' を E の双対と同一視しでき, 逆もまたそうであることがわかる. また, z, z' -軸に沿った射影により自然な対応 $s_i \leftrightarrow \frac{\partial}{\partial x_i}$, $s'_i \leftrightarrow \frac{\partial}{\partial p_i}$ が存在する. 超平面場 ξ は, シンプレクティック形式 $\omega = \sum_{i=1}^n dx_i \wedge dp_i$ のみならず,

$$\tau := \sum_{i=1}^n dx_i dp_i = \frac{1}{2} \sum_{i=1}^n (dx_i \otimes dp_i + dp_i \otimes dx_i)$$

から誘導される (n, n) 型の擬リーマン計量を持つ。枠 s_i と s_j^* を用いることで、 E_p と E'_p のベクトルを列ベクトル \mathbf{u} と \mathbf{u}' でそれぞれ書き、このとき

$$\tau(\mathbf{u} \oplus \mathbf{u}', \mathbf{v} \oplus \mathbf{v}') = \frac{1}{2} [\mathbf{u}^T \ \mathbf{u}'^T] \begin{bmatrix} O & E \\ E & O \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{v}' \end{bmatrix} = \frac{1}{2} (\mathbf{u}^T \mathbf{v}' + \mathbf{u}'^T \mathbf{v})$$

であり、また $\omega(\mathbf{u} \oplus \mathbf{u}', \mathbf{v} \oplus \mathbf{v}') = \frac{1}{2} (\mathbf{u}^T \mathbf{v}' - \mathbf{u}'^T \mathbf{v})$.

任意のアファインリジャンドル同値 \mathcal{L}_F は ξ 上の ω と τ を保つ。実際これは \mathcal{L}_F が $\mathbf{u} \oplus \mathbf{u}'$ を $A\mathbf{u} \oplus A'\mathbf{u}'$ に写すことによる。ここで、 $A' = (A^T)^{-1}$.

定義 4.1.15 ([33]) τ の引き戻しにより L の概ヘッセ計量を定義する:

$$h(Y, Z) := \tau(\iota_* Y, \iota_* Z) \quad (Y, Z \in TL).$$

ここで、 $\iota_* = \Phi \oplus \Phi' : TL \hookrightarrow \xi = E \oplus E'$ は包含写像である。

h は退化し得る対称 $(0, 2)$ -テンソルであるが、用語の濫用により“計量”という言葉を用いていることに注意する。 Φ が同型であるとき、 h はポテンシャル $z = f(\mathbf{x})$ のヘッセ行列が定めるヘッセ計量に一致する； $T_p L$ の任意のベクトルは、 $H = [h_{ij}]$ をヘッセ行列とすると、 $\mathbf{u} \oplus H\mathbf{u} \in \xi_p$ と書くことができ、このとき

$$h(\mathbf{u}, \mathbf{v}) = \tau(\mathbf{u} \oplus H\mathbf{u}, \mathbf{v} \oplus H\mathbf{v}) = \mathbf{u}^T H\mathbf{v}.$$

一般に、 h の局所表示は次で与えられる。

補題 4.1.16 ([33]) $g(\mathbf{x}_I, \mathbf{p}_J)$ を母関数とする。このとき、

$$h = \sum_{i,k \in I} \frac{\partial^2 g}{\partial x_i \partial x_k} dx_i dx_k - \sum_{j,l \in J} \frac{\partial^2 g}{\partial p_j \partial p_l} dp_j dp_l.$$

証明：直接計算により分かる：

$$\begin{aligned} \tau &= d\mathbf{x}_I^T d\mathbf{p}_I + d\mathbf{x}_J^T d\mathbf{p}_J \\ &= d\mathbf{x}_I^T d(\partial_I g) - d(\partial_J g)^T d\mathbf{p}_J \\ &= d\mathbf{x}_I^T g_{II} d\mathbf{x}_I + d\mathbf{x}_I^T g_{IJ} d\mathbf{p}_J - (g_{JI} d\mathbf{x}_I)^T d\mathbf{p}_J - (g_{JJ} d\mathbf{p}_J)^T d\mathbf{p}_J \\ &= d\mathbf{x}_I^T g_{II} d\mathbf{x}_I - d\mathbf{p}_J^T g_{JJ} d\mathbf{p}_J. \end{aligned}$$

ここで、例えば $\partial_I g = \frac{\partial g}{\partial \mathbf{x}_I}$ 、 $g_{II} = [\frac{\partial^2 g}{\partial x_i \partial x_j}]_{i,j \in I}$ であり、また、1形式の対称積の記法を用いており、 $(g_{JI})^T = g_{IJ}$ である。□

補題 4.1.17 ([33]) $p \in L$ とする。次の性質は全て同値である:

(1) h は p で非退化である；

- (2) p は e -臨界点でも m -臨界点でもない;
- (3) L は局所的に p の周りで正則モデルである;
- (4) h は p の近くの局所ポテンシャル $z = f(\mathbf{x})$ に付随するヘッセ計量である;
- (5) Φ と Φ' は共に p において同型である.

証明：補題 4.1.16 により, (1) は g_{II} と g_{JJ} の両方が非退化であることを意味する. このとき, §4.1.2 の最後で記述された e/m -ラグランジュ写像 π_1^e と π_1^m の標準形を用いることで, これらの写像が局所微分同相写像であることが陰関数定理から分かり, これはまさに (2) と (3) である. またこれは, 母関数として局所ポテンシャル $z = f(\mathbf{x})$ が取れることを意味し, これはすなわち (4) と同値である. さらに Φ と Φ' は e/m -ラグランジュ写像のヤコビ行列によって表現することができるため, (2) と (5) は同じである. \square

注意 4.1.18 注意 4.1.11 で指摘したように, 局所的にはいつでも混合座標系 $(\mathbf{x}_I, \mathbf{p}_J)$ を見つけることができる. 補題 4.1.16 により, h が退化しているときでさえ, 葉 $\{\mathbf{p}_J = \text{定数}\}$ と $\{\mathbf{x}_I = \text{定数}\}$ は直交している: $h(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial p_j}) = 0$ ($i \in I, j \in J$).

命題 4.1.17 により概ヘッセ計量 h の退化性は e/m -ルジャンドル写像の特異点によって特徴づけられることがわかる.

定義 4.1.19 ([33]) $\Sigma (= \Sigma_{L,h})$ を h の退化する点 $p \in L$ 全てからなる集合とする (同じことであるが, Φ もしくは Φ' が退化する点からなる集合とする):

$$\Sigma = C(\pi^e) \cup C(\pi^m).$$

Σ を概ヘッセ計量 h の退化集合と呼ぶ.

L はルジャンドル部分多様体であるため, $T_p L$ はシンプレクティックベクトル空間 $\xi_p = E_p \oplus E'_p$ のラグランジュ部分空間である. Φ (resp. Φ') は $T_p L$ の E_p (resp. E'_p) 成分への線形射影であることに注意し, 特に, $\dim T_p L \cap E_p \geq 1$ (resp. $\dim T_p L \cap E'_p \geq 1$) であることと, p が m -臨界点 (resp. e -臨界点) であることは同値である. 特に, h の零空間は次のように分解される:

$$\text{null } h_p := \{v \in T_p L \mid h(v, \cdot) = 0\} = \ker \Phi'_p \oplus \ker \Phi_p = (T_p L \cap E_p) \oplus (T_p L \cap E'_p).$$

定義 4.1.20 ([33]) 任意のルジャンドル部分多様体 $L \subset \mathbb{R}^{2n+1}$ に対して, $(h, (E, \nabla^E, \Phi), (E', \nabla^{E'}, \Phi'))$ を L の双対平坦構造と呼ぶ.

$$\begin{array}{ccc} & TL & \\ \Phi \swarrow & & \searrow \Phi' \\ E_L & & E'_L \end{array}$$

注意 4.1.21 正則モデル L が与えられたとき, Φ と Φ' が同型であるような L の双対平坦構造 (h, E, E') を得る. これは, 従来の (§2.2 の意味の) 双対平坦構造を復元する; 実際, TL 上の ∇ と ∇^* は一意的に

$$\Phi(\nabla_X Y) = \nabla_X^E \Phi(Y), \quad \Phi'(\nabla_X^* Y) = \nabla_X^{E'} \Phi'(Y)$$

によって決定される. ここで X, Y は L 上の任意のベクトル場である. 一方で, L が退化ポテンシャル $z = f(\mathbf{x})$ を持つ特異モデルである場合は, Φ は同型であるが Φ' は同型でない場合に相当する. このとき, 上記と同じように Φ により ∇^E から TL 上の接続 ∇ は得られる一方で, ∇^* は存在しない. Φ と Φ' の両方が同型でないとき, TL 上には接続は存在しない.

4.1.4 概ヘッセ多様体

定義 4.1.20 で与えた双対平坦構造の一般化はアファインルジャンドル同値と両立的である. すなわち, アファインルジャンドル同値 \mathcal{L}_F がルジャンドル部分多様体 L_1 と L_2 を同一視するとき, \mathcal{L}_F はその概ヘッセ計量を保つ: $\mathcal{L}_F^* h_2 = h_1$. さらにこのとき, 自然に接続接束間の束同型 $E_{L_1} \simeq E_{L_2}$, $E'_{L_1} \simeq E'_{L_2}$ が誘導され, この同型はそれぞれの上に導入されたアファイン平坦接続を同一視し, 次の図式を可換にする:

$$\begin{array}{ccccc} E_{L_1} & \xleftarrow{\Phi_1} & TL_1 & \xrightarrow{\Phi'_1} & E'_{L_1} \\ \simeq \downarrow & & \mathcal{L}_F \downarrow \simeq & & \downarrow \simeq \\ E_{L_2} & \xleftarrow{\Phi_2} & TL_2 & \xrightarrow{\Phi'_2} & E'_{L_2} \end{array}$$

このようにして, 張り合わせによる構成は意味をなすことがわかる. 厳密には, 次の性質を満たす族 $\{L_\alpha\}_{\alpha \in \Lambda}$ (Λ は可算集合) が与えられているとする:

- (1) 任意の $\alpha \in \Lambda$ に対し, L_α はルジャンドル部分多様体として \mathbb{R}^{2n+1} へ埋め込まれている開多様体 (コンパクトな連結成分を持たない境界なし多様体) である. 各 L_α を局所モデルと呼ぶ;
- (2) 任意の $\alpha, \beta \in \Lambda$ に対して, 開集合 $L_{\alpha\beta} \subset L_\alpha$ (同様に $L_{\beta\alpha} \subset L_\beta$) と微分同相写像 $\mathcal{L}_\beta^\alpha: L_{\alpha\beta} \rightarrow L_{\beta\alpha}$ が存在し, $L_{\alpha\beta}$ の各連結成分上で \mathcal{L}_β^α は周囲の空間 \mathbb{R}^{2n+1} のアファインルジャンドル同値によって得られる;
- (3) 任意の $\alpha, \beta, \gamma \in \Lambda$ に対し, $L_{\alpha\beta} \cap L_{\beta\alpha} \neq \emptyset$ であるとき, \mathbb{R}^{2n+1} 上の写像として $\mathcal{L}_\alpha^\gamma = \mathcal{L}_\beta^\gamma \circ \mathcal{L}_\alpha^\beta$ が成り立つ.

M を, 張り合わせの情報である $\mathcal{U} = \{L_\alpha, \mathcal{L}_\alpha^\beta\}$ から得られる商空間とする. M はハウ

スドルフ空間であることを仮定し、このとき M 自身は通常の意味で n 次元可微分多様体である。このとき M 上には自然に、退化し得る計量 h と互いに双対的な連接接束 $(E, \Phi : TM \rightarrow E, \nabla^E)$ および $(E', \Phi' : TM \rightarrow E', \nabla^{E'})$ が大域的に定義される。

定義 4.1.22 ([33]) $(h, (E, \nabla^E), (E', \nabla^{E'}))$ を備えた (M, \mathcal{U}) を概ヘッセ多様体と呼ぶ。退化集合 Σ を h が退化する M 上の全ての点からなる集合として定義する。

張り合わせ写像 \mathcal{L}_α^β は周囲の空間 \mathbb{R}^{2n+1} における $L_{\alpha\beta}$ の近傍上に作用し、接触構造を保つため、 M はある接触多様体のルジャンドル部分多様体として実現される。

注意 4.1.23 張り合わせの条件 (3) を $L_{\alpha\beta} \cap L_{\beta\alpha}$ 上の写像としての条件に弱めても、張り合わせに関して局所モデル上の構造は両立的であることに注意する。すなわち、 M 上において概ヘッセ計量 h や連接接束 (E, Φ, ∇^E) , $(E', \Phi', \nabla^{E'})$ は意味を持ち、 $E \oplus E'$ 上にはシンプレクティック形式 ω および (n, n) 型の擬リーマン計量 τ が定まる。

命題 4.1.3 と補題 4.1.17 より、次は明らかである：

命題 4.1.24 ([33]) M を概ヘッセ多様体とする。このとき h が全ての点で非退化であることは、 M がヘッセ多様体であることと同値である。

注意 4.1.25 より一般に、以上の構成は局所モデル L_α が多様体ではなく特異ルジャンドル多様体である場合にも意味を持つ。特異ルジャンドル多様体とは、接触構造に対する積分多様体で分割される閉部分集合のあるクラス (stratification) であり、例えば、フロンタルの等方的持ち上げの像は特異ルジャンドル多様体を定める (詳しくは石川 [20] を見よ)。標準接触多様体 \mathbb{R}^{2n+1} 内の部分集合が与えられると、それに沿って接触超平面場を射影することで原理的には“ベクトル束” E, E' が定義できる。フロンタルによって定義される特異ルジャンドル多様体は、多様体からのパラメータづけが与えられているため、 E, E' 上の接続も意味を持つ。このようにして、特異点を持つ概ヘッセ多様体が構成される。

概ヘッセ多様体の内在的定義もまた可能であろう。これはまたコダッチ構造の退化版を示唆しており (cf. [37]), 退化した計量を許容する統計多様体の概念に繋がっていくと考えられる。

4.1.5 3次テンソルと α -幾何学

§2.2 では双対平坦構造 (h, ∇, ∇^*) に対し甘利・チェンツォフの3次テンソル T を考えた。これは、接続 ∇ が存在しさえすれば、 h の非退化性にはよらずに常に空間全体で定義

されたことに注意する. 任意の概ヘッセ多様体

$$(M, h, (E, \Phi, \nabla^E), (E', \Phi', \nabla^{E'}))$$

に対して, 甘利・チェンツォフの3次テンソルを一般化する. TM の接続が存在しないため, 従来と同様に定義することはできない. しかし, これから定義する3次テンソルは極めて自然な一般化になっている (以下の命題 4.1.30).

補題 4.1.26 ([33]) 任意の M 上のベクトル場 X , E の切断 η , E' の切断 ζ' に対し, 次が成り立つ:

$$X\tau(\eta, \zeta') = \tau(\nabla_X^E \eta, \zeta') + \tau(\eta, \nabla_X^{E'} \zeta')$$

ここで簡単のために, $\tau(\eta, \zeta') := \tau(\eta \oplus 0, 0 \oplus \zeta')$ とおいた.

証明: 開集合 $U \subset M$ 上において, E, E' の平坦切断の局所フレーム s_i, s_j^* で $\tau(s_i, s_j^*) = \frac{1}{2}\delta_{ij}$ ($1 \leq i, j \leq n$) なるものをそれぞれとる. $\eta = \sum a_i s_i$, $\zeta' = \sum b_j s_j^*$ とおく. ここで, a_i, b_j は U 上の関数である. このとき

$$\nabla_X^E \eta = \sum X(a_i) s_i, \quad \nabla_X^{E'} \zeta' = \sum X(b_j) s_j^*, \quad \tau(\eta, \zeta') = \frac{1}{2} \sum a_i b_i.$$

これより等式が従う. □

任意の $Y, Z \in TM$ に対して

$$\eta = \Phi(Y), \quad \zeta = \Phi(Z) \in E, \quad \eta' = \Phi'(Y), \quad \zeta' = \Phi'(Z) \in E'$$

とおく. このとき,

$$h(Y, Z) = \tau(\eta \oplus \eta', \zeta \oplus \zeta') = \tau(\eta, \zeta) + \tau(\zeta, \eta').$$

補題 4.1.26 を用いることで, M 上の任意のベクトル場 X, Y, Z に対して

$$\begin{aligned} Xh(Y, Z) &= X(\tau(\eta, \zeta)) + X(\tau(\zeta, \eta')) \\ &= \tau(\nabla_X^E \eta, \zeta) + \tau(\eta, \nabla_X^{E'} \zeta') + \tau(\nabla_X^E \zeta, \eta') + \tau(\zeta, \nabla_X^{E'} \eta'). \end{aligned}$$

便宜的に, 第一項と第三項を ∇^E -パート, 残りを $\nabla^{E'}$ -パートと呼ぶことにする.

補題 4.1.27 M 上の任意のベクトル場 Y, Z に対して

$$\tau(\eta, \zeta') = \tau(\zeta, \eta').$$

証明: 直接計算により示される. 議論は局所的であるため M は局所モデル $L \subset \mathbb{R}^{2n+1}$ とする. $g(x_I, p_J)$ を L の母関数とし, 基底 $\{\frac{\partial}{\partial x_I}, \frac{\partial}{\partial p_J}\}$ に関して Y と Z をそれぞれ列ベ

クトル $[\mathbf{a}_I, \mathbf{a}_J]^T$, $[\mathbf{b}_I, \mathbf{b}_J]^T$ で表す. (4.1) の平坦フレーム s_i, s_j^* に関する成分表示として η, ζ' はそれぞれ

$$\eta = \begin{bmatrix} \mathbf{a}_I \\ -g_{JI}\mathbf{a}_I - g_{JJ}\mathbf{a}_J \end{bmatrix}, \quad \zeta' = \begin{bmatrix} g_{II}\mathbf{b}_I + g_{IJ}\mathbf{b}_J \\ \mathbf{b}_J \end{bmatrix}$$

と表現される. 従って

$$\tau(\eta, \zeta') = \frac{1}{2}(\mathbf{a}_I^T g_{II}\mathbf{b}_I - \mathbf{a}_J^T g_{JJ}\mathbf{b}_J) = \tau(\zeta, \eta')$$

より等式が従う. □

定義 4.1.28 ([33]) 概ヘッセ多様体 M に対して, 標準3次テンソル C を次の M 上の $(0, 3)$ -テンソルによって定義する:

$$C(X, Y, Z) := \tau(\eta, \nabla_X^E \zeta') + \tau(\zeta, \nabla_X^{E'} \eta') - \tau(\nabla_X^E \eta, \zeta') - \tau(\nabla_X^E \zeta, \eta').$$

注意 4.1.29 C は確かにテンソルである. 実際, C はベクトル場の和について分解し, また, $f \in C^\infty(M)$ に対し, 例えば

$$C(X, fY, Z) = fC(X, Y, Z) + (Xf)\{\tau(\zeta, \eta') - \tau(\eta, \zeta')\}$$

であるが第二項は補題 4.1.27 よりゼロとなる.

特に h が非退化なときは, $\Phi(\nabla_X Y) = \nabla_X^E \Phi(Y)$ であり (注意 4.1.21),

$$\tau(\nabla_X^E \eta, \zeta') = \tau(\Phi(\nabla_X Y), \Phi'(Z)) = \frac{1}{2}h(\nabla_X Y, Z)$$

などを得る. このようにして ∇^E -パートと $\nabla^{E'}$ -パートはそれぞれ,

$$\frac{1}{2}(h(\nabla_X Y, Z) + h(Y, \nabla_X Z)), \quad \frac{1}{2}(h(\nabla_X^* Y, Z) + h(Y, \nabla_X^* Z)).$$

と等しいことが分かる. 従って, C は甘利・チェンツォフの3次テンソル T に一致することが分かる:

$$\begin{aligned} C(X, Y, Z) &= \frac{1}{2}(h(\nabla_X^* Y, Z) - h(\nabla_X Y, Z)) + \frac{1}{2}(h(Y, \nabla_X^* Z) - h(Y, \nabla_X Z)) \\ &= T(X, Y, Z). \end{aligned}$$

局所座標系を用いることで, 次のようにテンソル C を書き下す. 局所モデル $L \subset \mathbb{R}^{2n+1}$ と $p \in L$ をとる. 前述のとおり, p の周りにおいて, L は, 母関数を $g(\mathbf{x}_I, \mathbf{p}_J)$ として, 局所座標系 $(\mathbf{x}_I, \mathbf{p}_J)$ によってパラメータ付けされる. 簡単のために, 各 $1 \leq k \leq n$ に対して,

$$\partial_k := \frac{\partial}{\partial x_k} \quad (k \in I) \quad \text{もしくは} \quad \partial_k := \frac{\partial}{\partial p_k} \quad (k \in J)$$

とおく. 次は命題 2.2.5 に対応するものである.

命題 4.1.30 ([33]) 標準 3 次テンソル C は局所的に、母関数 $g(\mathbf{x}_I, \mathbf{p}_J)$ の 3 次導関数である：任意の k, l, m に対し、

$$C(\partial_k, \partial_l, \partial_m) = \partial_k \partial_l \partial_m g.$$

特に C は対称である。

証明：直接計算していく。母関数は次で得られるラグランジュ埋め込み $L \rightarrow T^*\mathbb{R}^n$ を与える：

$$\iota : (\mathbf{x}_I, \mathbf{p}_J) \mapsto (\mathbf{x}_I, \mathbf{x}_J, \mathbf{p}_I, \mathbf{p}_J) := (\mathbf{x}_I, -\partial_J g, \partial_I g, \mathbf{p}_J).$$

従ってその微分 $\iota_* : T_p L \rightarrow T_p(T^*\mathbb{R}^n) = \mathbb{R}_x^n \oplus \mathbb{R}_p^n$ は

$$\iota_*(\partial_k) = \partial_k - \sum_{j \in J} (\partial_k \partial_j g) \frac{\partial}{\partial x_j} + \sum_{i \in I} (\partial_k \partial_i g) \frac{\partial}{\partial p_i}$$

と書ける。 s_i, s_i^* ($1 \leq i \leq n$) を E と E' の平坦切断とする； $\tau(s_i, s_j^*) = \frac{1}{2} \delta_{ij}$ 。このとき $k \in I$ に対して、

$$\Phi(\iota_* \partial_k) = s_k - \sum_{j \in J} (\partial_k \partial_j g) s_j, \quad \Phi'(\iota_* \partial_k) = \sum_{i \in I} (\partial_k \partial_i g) s_i^*$$

であり、 $k \in J$ に対して、

$$\Phi(\iota_* \partial_k) = - \sum_{j \in J} (\partial_k \partial_j g) s_j, \quad \Phi'(\iota_* \partial_k) = s_k^* + \sum_{i \in I} (\partial_k \partial_i g) s_i^*.$$

$\eta = \Phi(\iota_* \partial_l)$, $\eta' = \Phi'(\iota_* \partial_l)$, $\zeta = \Phi(\iota_* \partial_m)$, $\zeta' = \Phi'(\iota_* \partial_m)$, $X = \partial_k$ とおく。 $l \in I$, $m \in J$ と任意の k に対して、次を得る：

$$\begin{aligned} \tau(\eta, \nabla_X^{E'} \zeta') &= \tau(s_l - \sum_{j \in J} (\partial_l \partial_j g) s_j, \sum_{i \in I} (\partial_k \partial_m \partial_i g) s_i^*) = \frac{1}{2} \partial_k \partial_l \partial_m g, \\ \tau(\zeta, \nabla_X^{E'} \eta') &= \tau(- \sum_{j \in J} (\partial_m \partial_j g) s_j, \sum_{i \in I} (\partial_k \partial_l \partial_i g) s_i^*) = 0, \\ \tau(\nabla_X^E \eta, \zeta') &= \tau(- \sum_{j \in J} (\partial_k \partial_l \partial_j g) s_j, s_m^* + \sum_{i \in I} (\partial_m \partial_i g) s_i^*) = -\frac{1}{2} \partial_k \partial_l \partial_m g, \\ \tau(\nabla_X^E \zeta, \eta') &= \tau(- \sum_{j \in J} (\partial_k \partial_m \partial_j g) s_j, \sum_{i \in I} (\partial_l \partial_i g) s_i^*) = 0. \end{aligned}$$

このようにして、 $\nabla^{E'}$ -パートから ∇^E -パートを引くことで、 $C(\partial_k, \partial_l, \partial_m) = \partial_k \partial_l \partial_m g$ を得る。 $l, m \in I$ と任意の k に対して、次を得る：

$$\begin{aligned} \tau(\eta, \nabla_X^{E'} \zeta') &= \tau(s_l - \sum_{j \in J} (\partial_l \partial_j g) s_j, \sum_{i \in I} (\partial_k \partial_m \partial_i g) s_i^*) = \frac{1}{2} \partial_k \partial_l \partial_m g, \\ \tau(\zeta, \nabla_X^{E'} \eta') &= \tau(s_m - \sum_{j \in J} (\partial_m \partial_j g) s_j, \sum_{i \in I} (\partial_k \partial_l \partial_i g) s_i^*) = \frac{1}{2} \partial_k \partial_l \partial_m g, \\ \tau(\nabla_X^E \eta, \zeta') &= \tau(- \sum_{j \in J} (\partial_k \partial_l \partial_j g) s_j, \sum_{i \in I} (\partial_m \partial_i g) s_i^*) = 0, \\ \tau(\nabla_X^E \zeta, \eta') &= \tau(- \sum_{j \in J} (\partial_k \partial_m \partial_j g) s_j, \sum_{i \in I} (\partial_l \partial_i g) s_i^*) = 0. \end{aligned}$$

先ほどと同様に, $C(\partial_k, \partial_l, \partial_m) = \partial_k \partial_l \partial_m g$. $l, m \in J$ の場合も同様に成り立つことが分かる. \square

よく知られたように, 双対平坦多様体 M に対して, α -接続の族が次で定義される:

$$\nabla^{(\alpha)} = \frac{1+\alpha}{2} \nabla + \frac{1-\alpha}{2} \nabla^* \quad (\alpha \in \mathbb{R}).$$

あるいは同じことであるが, T を用いてレビ・チビタ接続 $\bar{\nabla}$ を線形に変形する:

$$h(\nabla_X^{(\alpha)} Y, Z) = h(\bar{\nabla}_X Y, Z) - \frac{\alpha}{2} T(X, Y, Z). \quad (4.2)$$

これは $\alpha = \pm 1$ のとき, ∇, ∇^* となっており, また $\nabla^{(\alpha)}$ と $\nabla^{(-\alpha)}$ は互いに双対的である:

$$Xh(Y, Z) = h(\nabla_X^\alpha Y, Z) + h(Y, \nabla_X^{-\alpha} Z).$$

接続 $\{\nabla^{(\alpha)}\}$ はいわゆる α -幾何学を形成する [3, 31]. 概ヘッセ多様体 M に対して, E と E' 上の接続が定義された. しかしこれらは TM 上のもではなく, α -幾何学の直接的なアナロジーは存在しない. ここで, 次の $N^{(\alpha)} : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \rightarrow C^\infty(M)$ を定義してみる (線形でないためテンソルではない):

$$\begin{aligned} N^{(\alpha)}(X, Y, Z) &:= (1+\alpha)\tau(\nabla_X^E \eta, \zeta') + (1-\alpha)\tau(\zeta, \nabla_X^{E'} \eta') \\ &= \tau(\nabla_X^E \eta, \zeta') + \tau(\zeta, \nabla_X^{E'} \eta') - \frac{\alpha}{2} C(X, Y, Z). \end{aligned}$$

このとき, N^α は双対的である:

$$Xh(Y, Z) = N^{(\alpha)}(X, Y, Z) + N^{(-\alpha)}(X, Z, Y).$$

さらに, 概ヘッセ計量 h が非退化であり M 上のヘッセ構造 (h, ∇, ∇^*) が復元されているとき, $2\tau(\nabla_X^E \eta, \zeta') = h(\nabla_X Y, Z)$ であることに注意すると, $N^{(\alpha)}(X, Y, Z)$ は式 (4.2) と一致することがわかる. すなわち M が双対平坦多様体であるとき, α -幾何学が完全に復元される.

4.2 アファイン座標系における e/m -波面の標準形

本節では, 概ヘッセ多様体上の幾何学的諸量を用いて e/m -波面に現れる特異点を特徴づける. すなわち, 典型的な特異点の C^∞ 型を捉える座標によらない判定法を与え, この際に現れる特異点型の標準形をアファイン座標系において導出する. さらに, 判定法の内在的表現について議論する.

4.2.1 標準形

標準形の導出にあたり中心的な役割を果たす道具が以下のマルグランジュの割り算定理である。

定理 4.2.1 (マルグランジュの割り算定理, e.g., [6, 17]) $f(t, \mathbf{x})$ を \mathbb{R}^{n+1} の原点の近傍で定義された実関数とする. ここで $t \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ である. 次数 d の多項式を $P(t) = \sum_{i=0}^d \lambda_i t^i$ ($\lambda_i \in \mathbb{R}, \lambda_d \neq 0$) とする. このとき, 原点の近傍で定義された関数 $Q(t, \mathbf{x})$ と $r_i(\mathbf{x})$ ($0 \leq i \leq d-1$) が存在して,

$$f(t, \mathbf{x}) = Q(t, \mathbf{x})P(t) + \sum_{i=0}^{d-1} r_i(\mathbf{x})t^i.$$

写像の特異点の分類は通常 \mathcal{A} -同値 (あるいは右左同値とも呼ばれる) にて行われる. 写像芽 $f, g: (\mathbb{R}^m, 0) \rightarrow (\mathbb{R}^n, 0)$ が \mathcal{A} -同値であるとは, ソースとターゲットにおける微分同相写像芽 σ と τ が存在して, $g = \tau \circ f \circ \sigma^{-1}$ が成り立つときをいう. ルジャンドル写像には以下の典型的な特異点が現れる. 写像芽 $h: (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^{n+1}, 0)$ がカस्प辺であるとは, h が原点における写像芽 $(x_1, \dots, x_n) \mapsto (x_1^3, x_1^2, x_2, \dots, x_n)$ と \mathcal{A} -同値であるときをいう. また, h がツバメの尾であるとは, h が原点における写像芽

$$(x_1, \dots, x_n) \mapsto (3x_1^4 + x_1^2 x_2, 2x_1^2 + x_1 x_2, x_2, \dots, x_n)$$

と \mathcal{A} -同値であるときをいう. これら原点における特異点は余次元 1 である. すなわち, $\text{rank } dh_0 = n - 1$. また写像芽の代表元を原点の近傍 U と $h: U \rightarrow \mathbb{R}^{n+1}$ によって表すと, いずれの特異点の場合でも, 特異点集合 $S(h) = \{p \in U \mid \text{rank } dh_p \leq n - 1\}$ は明らかに余次元 1 の部分多様体になる. 我々はこれらの特異点に注目し標準形の導出を行う.

局所的な議論を行うため, 局所モデル $L \subset \mathbb{R}^{2n+1}$ をとり, h を概ヘッセ計量, (E, Φ, ∇^E) と $(E', \Phi', \nabla^{E'})$ を接続接束とする. 点 $\bar{p} = (\bar{\mathbf{x}}, \bar{\mathbf{p}}, z) \in L$ をとる. 以下では \bar{p} の周りの L に付随した m -波面に対する結果を述べるが, $(\Phi', \nabla^{E'})$ と (Φ, ∇^E) を入れ替えて議論することによって e -波面についても同様の結果が成り立つことに注意する.

m -ラグランジュ写像

$$\pi_1^m: L \rightarrow \mathbb{R}_p^n$$

に対し, $d\pi_1^m$ が \bar{p} において退化することを仮定する. 特に余次元 1 の場合を考える, i.e.,

$$\dim \ker d\pi_1^m(\bar{p}) (= \dim \ker \Phi'(\bar{p})) = 1.$$

このとき, \bar{p} の近傍において特異点集合 $S(\pi_1^m) \subset L$ は, (L の局所座標系をとって) 特異点識別子 ([35])

$$\lambda(q) := \det[d\pi_1^m(q)] \quad (q \in L)$$

を用いて記述される： $S(\pi_1^m) = \{\lambda = 0\}$. \bar{p} の周りで定義される非ゼロな L 上のベクトル場 X であって、 $\ker d\pi_{\bar{p}}^m$ を張るものをとる.

泉屋・佐治の判定法 [24] によると、 m -ルジャンドル写像 $\pi^m = (\pi_1^m, z') : L \rightarrow \mathbb{R}^{n+1} = \mathbb{R}_p^n \times \mathbb{R}_{z'}$ が \bar{p} においてカस्प辺であることは、次が成り立つことと同値である：

$$X\lambda(\bar{p}) \neq 0.$$

この特異点は情報幾何学には現れないものである（フィッシャー・ラオ計量は半正定値である）. 我々は、上記の条件を L に付随する概ヘッセ構造に関する条件に書き換え、 m -波面に対する多価（双対）ポテンシャルの標準形をアフィン座標系において与える. これはエクランによる結果 [15, 命題 2.7] を一般化するものである. また、エクランによるこの命題の証明は不完全なものであり、我々の証明はこれを補完する.

点 $\bar{p} = (\bar{x}, \bar{p}, \bar{z}) \in L$ を固定する. ここで $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n) \in \mathbb{R}_x^n$, $\bar{p} = (\bar{p}_1, \dots, \bar{p}_n) \in \mathbb{R}_p^n$ である（バーは変数ではなく固定点であることを意味する）.

定理 4.2.2 ([32]) $\text{rank } \Phi'(\bar{p}) = n - 1$ とする. また、 \bar{p} において $\ker \Phi'(\bar{p})$ を張る L 上のベクトル場 $X \neq 0$ に対し、

$$\tau(\Phi(X)(\bar{p}), (\nabla_X^E \Phi')(X)(\bar{p})) \neq 0 \quad (4.3)$$

が成り立つとする. このとき π^m は \bar{p} においてカस्प辺型の特異点を持つ. すなわち、 \mathbb{R}_p^n のアフィン座標系 $\mathbf{p} = (p_1, \dots, p_n)$ とその双対アフィン座標系である \mathbb{R}_x^n の $\mathbf{x} = (x_1, \dots, x_n)$ を取り直すことにより、以下が成り立つ：ある近傍 $\bar{p} \in V \subset L$, $0 \in \mathcal{U}_1 \subset \mathbb{R}$, $(\bar{p}_2, \dots, \bar{p}_n) \in \mathcal{U}_2 \subset \mathbb{R}^{n-1}$ と関数

$$\varphi \in C^\infty(\mathcal{U}_1 \times \mathcal{U}_2), \quad k_1, k_2 \in C^\infty(\mathcal{U}_2)$$

が存在して m -波面 $\pi^m(V)$ の 2 値双対ポテンシャル関数が

$$z' = z'_\pm(\mathbf{p}) = k_2(p_2, \dots, p_n) + (p_1 - k_1)\varphi(\pm\sqrt{p_1 - k_1}, p_2, \dots, p_n), \quad (4.4)$$

$p_1 \geq k_1(p_2, \dots, p_n)$ で与えられる. このときさらに、 φ の第一変数に関する偏導関数は非ゼロである.

証明： $\text{rank } \Phi'(\bar{p}) = n - 1$ であることから、 \mathbb{R}_x^n のアフィン変換を取ること、すなわち、 \mathbb{R}^{2n+1} のアフィンルジャンドル同値を取ることによって $\ker \Phi'(\bar{p})$ は $(\frac{\partial}{\partial x_1})_{\bar{p}} + \bar{p}_1(\frac{\partial}{\partial z})_{\bar{p}}$ によって張られていると仮定して良い. このとき (x_1, p_2, \dots, p_n) は \bar{p} 周りの L の局所座標系である. この座標系において、 m -ラグランジュ写像の特異点集合は $S(\pi_1^m) = \{\frac{\partial p_1}{\partial x_1} = 0\}$ とかける.

$g(x_1, p_2, \dots, p_n) = g(x_1, \mathbf{p}_J)$ を \bar{p} 周りの L の母関数とする ($I = \{1\}, J = \{2, \dots, n\}$). このとき

$$p_1 = \frac{\partial g}{\partial x_1}(x_1, \mathbf{p}_J), \quad z' = x_1 \frac{\partial g}{\partial x_1}(x_1, \mathbf{p}_J) - g(x_1, \mathbf{p}_J). \quad (4.5)$$

特異点集合 $S(\pi_1^m)$ 上で $\frac{\partial z'}{\partial x_1} = x_1 \frac{\partial p_1}{\partial x_1} = 0$ であることに注意する. ベクトル場 X であって, \bar{p} 周りで $X = \frac{\partial}{\partial x_1}$ なるものをとる (これは $\ker \Phi'(\bar{p})$ を張っている). また, (4.1) のように E と E' の平坦切断 s_i と s_j^* をとる. このとき補題 4.1.26 の証明にあるように, \bar{p} 周りの E' の任意の局所切断 $\zeta' = \sum b_j s_j^*$ に対し $\nabla_X^{E'} \zeta' = \sum X(b_j) s_j^*$ となることに注意する. 直接計算することで次を得る:

$$\begin{aligned} \Phi(X(\bar{p})) &= s_1 - \sum_{j=2}^n \frac{\partial^2 g}{\partial x_1 \partial p_j}(\bar{p}) s_j, \\ \Phi'(X(\bar{p})) &= \frac{\partial^2 g}{\partial x_1^2}(\bar{p}) s_1^* = \frac{\partial p_1}{\partial x_1}(\bar{p}) s_1^*, \\ (\nabla_X^{E'} \Phi'(X))(\bar{p}) &= \frac{\partial^3 g}{\partial x_1^3}(\bar{p}) s_1^* = \frac{\partial^2 p_1}{\partial x_1^2}(\bar{p}) s_1^*. \end{aligned}$$

従って,

$$\tau(\Phi(X(\bar{p})), (\nabla_X^{E'} \Phi'(X))(\bar{p})) = \frac{1}{2} \frac{\partial^3 g}{\partial x_1^3}(\bar{p}) = \frac{1}{2} \frac{\partial^2 p_1}{\partial x_1^2}(\bar{p}). \quad (4.6)$$

仮定 (4.3) より, $\frac{\partial^2 p_1}{\partial x_1^2}(\bar{p}) \neq 0$ を得る. このようにして, \bar{p} のある開近傍における特異点集合 $S(\pi_1^m)$ 上で次が成り立つことが分かる:

$$\frac{\partial p_1}{\partial x_1} = 0, \quad \frac{\partial^2 p_1}{\partial x_1^2} \neq 0, \quad \frac{\partial z'}{\partial x_1} = 0. \quad (4.7)$$

(4.7) の一つ目の式に陰関数定理を用いる; $(\bar{p}_2, \dots, \bar{p}_n) \in \mathbb{R}^{n-1}$ の近傍 \mathcal{U}_2 と関数 $f \in C^\infty(\mathcal{U}_2)$ が存在して, \bar{p} の周辺で特異点集合 $S(\pi_1^m)$ がグラフ $x_1 = f(\mathbf{p}_J)$ としてパラメータづけされる.

マルグラングジュの割り算定理より (必要であれば \mathcal{U}_2 を小さくすることで), 関数 $\varphi_1, \varphi_2, k_1, k_2, k'_1, k'_2$ が存在して,

$$\begin{aligned} p_1(x_1, \mathbf{p}_J) &= (x_1 - f)^2 \varphi_1(x_1, \mathbf{p}_J) + (x_1 - f) k'_1(\mathbf{p}_J) + k_1(\mathbf{p}_J), \\ z'(x_1, \mathbf{p}_J) &= (x_1 - f)^2 \varphi_2(x_1, \mathbf{p}_J) + (x_1 - f) k'_2(\mathbf{p}_J) + k_2(\mathbf{p}_J). \end{aligned}$$

このとき, (4.7) より, $k'_1 = k'_2 = 0$ かつ特異点集合上で $\varphi_1 \neq 0$ となることが分かる.

$\varphi_1 > 0$ を仮定する ($\varphi_1 < 0$ なら p_1 を $-p_1$ にすればよい). 座標変換 $\Psi: (x_1, \mathbf{p}_J) \mapsto (y_1, \mathbf{p}_J)$, $y_1 := (x_1 - f)\sqrt{\varphi_1}$ を考え, $\varphi := (\varphi_2/\varphi_1) \circ \Psi^{-1}$ とおく. このとき, $p_1 - k_1 =$

$y_1^2 \geq 0$ であり

$$z' - k_2 = y_1^2 \varphi(y_1, \mathbf{p}_J) = (p_1 - k_1) \varphi(\pm \sqrt{p_1 - k_1}, \mathbf{p}_J).$$

\mathcal{U}_1 を Ψ の像が $\mathcal{U}_1 \times \mathcal{U}_2$ となるようにとり, $V \subset L$ を \bar{p} の対応する近傍としてとる. このようにして, 標準形 (4.4) を得る.

上式より, $\frac{\partial^3 z'}{\partial y_1^3}(\bar{p}) = 6 \frac{\partial \varphi}{\partial y_1}(\bar{p})$ を得る. また, (4.5) より,

$$z'(y_1, \mathbf{p}_J) = x_1 p_1 - \tilde{g}(y_1, \mathbf{p}_J) = x_1 (y_1^2 + k_1) - \tilde{g}(y_1, \mathbf{p}_J)$$

であることから, $\frac{\partial^3 z'}{\partial y_1^3} = 2y_1 \frac{\partial^2 x_1}{\partial y_1^2} + 4 \frac{\partial x_1}{\partial y_1}$ を得る. ここで, $\tilde{g} := g \circ \Psi^{-1}$ である. 従って, Ψ が座標変換であることと $y_1(\bar{p}) = 0$ であることから,

$$\frac{\partial \varphi}{\partial y_1}(\bar{p}) = \frac{1}{6} \frac{\partial^3 z'}{\partial y_1^3}(\bar{p}) = \frac{2}{3} \frac{\partial x_1}{\partial y_1}(\bar{p}) \neq 0.$$

であることが分かる. □

例 4.2.3 (カスプ辺) $n = 2$ のときを考え, 母関数が次で与えられるものとする:

$$g(x_1, p_2) = \frac{1}{3} x_1^3 - \frac{1}{2} p_2^2.$$

このとき, e -波面 $W_e(L) \subset \mathbb{R}_x \times \mathbb{R}_z$ は滑らかな部分多様体であり, m -波面 $W_m(L)$ は $p_1 = 0$ で分岐する 2 値双対ポテンシャル

$$z' = x_1 p_1 - g = \pm \frac{2}{3} p_1^{3/2} + \frac{1}{2} p_2^2, \quad p_1 \geq 0$$

のグラフである (cf. 例 4.1.8).

注意 4.2.4 上記証明中の式 (4.6) より,

$$\tau(\Phi(X(\bar{p})), (\nabla_X^{E'} \nabla_X^{E'} \Phi'(X))(\bar{p})) = \frac{1}{2} \frac{\partial^4 g}{\partial x_1^4}(\bar{p}) = \frac{1}{2} (X X \lambda)(\bar{p}), \quad (4.8)$$

を得る. ここで, $X = \frac{\partial}{\partial x_1}$ および $\lambda = \frac{\partial p_1}{\partial x_1}$ である. 従って泉屋・佐治による判定法 [24] を用いることで, \bar{p} における m -ルジャンドル写像 π^m の写像芽がツバメの尾と \mathcal{A} -同値であることと以下の式が \bar{p} において成り立つことは必要十分であることが分かる: $d\lambda(\bar{p}) \neq 0$ および

$$\begin{aligned} \tau(\Phi(X)(\bar{p}), (\nabla_X^{E'} \Phi'(X))(\bar{p})) &= 0, \\ \tau(\Phi(X)(\bar{p}), (\nabla_X^{E'} \nabla_X^{E'} \Phi'(X))(\bar{p})) &\neq 0. \end{aligned}$$

この場合においてもアファイン座標系における多価ポテンシャルの標準形を書き下すことは可能であるが、複雑になるためここでは行わない。例えば、母関数

$$g(x_1, p_2) = \frac{1}{4}x_1^4 + \frac{1}{2}p_2x_1^2$$

を考えると、 m -波面はツバメの尾をなす (cf. [6]).

一方、情報幾何学における応用を考える上では、概ヘッセ計量が半正定値であると仮定することは自然であろう (例えばフィッシャー・ラオ計量は半正定値である)。このような場合において最も単純な場合を考える。概ヘッセ計量 h は半正定値かつ $\text{rank } \Phi'(\bar{p}) = n - 1$ であると仮定する。このとき、 π_1^m に対する特異点識別子 λ の微分は特異点集合上でゼロになることが分かる。実際、補題 4.1.16 より、母関数 $g(x_1, \mathbf{p}_J)$ について概ヘッセ計量は

$$h = \begin{bmatrix} \frac{\partial^2 g}{\partial x_1^2} & 0 \\ 0 & -g_{JJ} \end{bmatrix}$$

と表現される。ここで、 $J = \{2, \dots, n\}$, $g_{JJ} = [\frac{\partial^2 g}{\partial p_i \partial p_j}]_{i,j \in J}$ である。 h は半正定値であるため、 $\frac{\partial^2 g}{\partial x_1^2} = \lambda \geq 0$ である。従って、特異点集合 $\{\lambda = 0\}$ 上で $d\lambda = 0$ である。これは $\lambda = 0$ に対する陰関数定理が成り立たないことを意味する。このような状況における最も単純な場合として、特異点集合が余次元 1 の部分多様体であると仮定することは妥当であろう。次の定理はこのような場合に現れる標準形をアファイン座標系において記述したものであり、これはまた、ヴァシリエフ (V. A. Vasil'ev) による関数の極小点の分類の文脈としての特徴づけを行うものである [42].

定理 4.2.5 ([32]) $\text{rank } \Phi'(\bar{p}) = n - 1$ とする。また、 \bar{p} 周りにおいて、特異点集合 $S(\pi^m)$ の任意の点 q で $\ker \Phi'(q)$ を張る L 上のベクトル場 $X \neq 0$ に対し、

$$\tau(\Phi(X)(q), (\nabla_X^{E'} \Phi'(X))(q)) = 0, \quad (4.9)$$

$$\tau(\Phi(X)(\bar{p}), (\nabla_X^{E'} \nabla_X^{E'} \Phi'(X))(\bar{p})) \neq 0 \quad (4.10)$$

が成り立つとする。さらに、 \bar{p} の周辺で $S(\pi^m)$ は L における余次元 1 の部分多様体であると仮定する。このとき、アファイン座標系 \mathbf{p} と \mathbf{x} を取り直すことにより、以下が成り立つ：ある近傍 $\bar{p} \in V \subset L$, $0 \in \mathcal{U}_1 \subset \mathbb{R}$, $(\bar{p}_2, \dots, \bar{p}_n) \in \mathcal{U}_2 \subset \mathbb{R}^{n-1}$ と関数

$$\varphi \in C^\infty(\mathcal{U}_1 \times \mathcal{U}_2), \quad k_1, k_2 \in C^\infty(\mathcal{U}_2)$$

が存在して m -波面 $\pi^m(V)$ に対する双対ポテンシャルが

$$z' = z'(\mathbf{p}) = k_2(p_2, \dots, p_n) + (p_1 - k_1)\varphi((p_1 - k_1)^{1/3}, p_2, \dots, p_n) \quad (4.11)$$

で与えられる。このときさらに、 φ の第一変数に関する偏導関数は非ゼロである。

注意 4.2.6 先ほど見たように概ヘッセ計量 h が半正定値であるときには, π_1^m に対する特異点識別子の微分は特異点集合上でゼロになるため, 上記の条件 (4.9) は自動で満たされる.

証明: 定理 4.2.2 の証明における設定と記法を用いる. このときベクトル場 $\frac{\partial}{\partial x_1}$ は, \bar{p} 周りにおける特異点集合 $S(\pi_1^m) = \{\frac{\partial p_1}{\partial x_1} = 0\}$ の任意の点 q において $\ker \Phi'(q)$ を張っていることに注意する. 定理 4.2.2 の証明と同様にして, 条件 (4.9) と (4.10) より, \bar{p} のある開近傍における特異点集合 $S(\pi_1^m)$ 上で次が成り立つことが分かる:

$$\frac{\partial p_1}{\partial x_1} = 0, \quad \frac{\partial^2 p_1}{\partial x_1^2} = 0, \quad \frac{\partial^3 p_1}{\partial x_1^3} \neq 0, \quad \frac{\partial z'}{\partial x_1} = 0, \quad \frac{\partial^2 z'}{\partial x_1^2} = 0 \quad (4.12)$$

上記 (4.12) の二つ目の方程式に陰関数定理を用いる; $(\bar{p}_2, \dots, \bar{p}_n)$ の近傍 $\mathcal{U}_2 \subset \mathbb{R}^{n-1}$ と関数 $f \in C^\infty(\mathcal{U}_2)$ が存在し, \mathcal{U}_2 上で $\frac{\partial^2 p_1}{\partial x_1^2}(f(\mathbf{p}_J), \mathbf{p}_J) = 0$ を満たす. $S(\pi_1^m)$ が L の余次元 1 の部分多様体であるため, このグラフは \bar{p} の周辺で $S(\pi_1^m)$ と一致する.

マルグラングの割り算定理より, 関数 $\varphi_1, \varphi_2, k_1, k_2, k'_1, k'_2, k''_1, k''_2$ が存在して,

$$\begin{aligned} p_1(x_1, \mathbf{p}_J) &= (x_1 - f)^3 \varphi_1(x_1, \mathbf{p}_J) + (x_1 - f)^2 k'_1(\mathbf{p}_J) + (x_1 - f) k''_1(\mathbf{p}_J) + k_1(\mathbf{p}_J), \\ z'(x_1, \mathbf{p}_J) &= (x_1 - f)^3 \varphi_2(x_1, \mathbf{p}_J) + (x_1 - f)^2 k'_2(\mathbf{p}_J) + (x_1 - f) k''_2(\mathbf{p}_J) + k_2(\mathbf{p}_J). \end{aligned}$$

このとき (4.12) より, $k'_1 = k''_1 = k'_2 = k''_2 = 0$ と特異点集合上で $\varphi_1 \neq 0$ であることが分かる. すなわち,

$$\begin{aligned} p_1(x_1, \mathbf{p}_J) &= (x_1 - f)^3 \varphi_1(x_1, \mathbf{p}_J) + k_1(\mathbf{p}_J), \\ z'(x_1, \mathbf{p}_J) &= (x_1 - f)^3 \varphi_2(x_1, \mathbf{p}_J) + k_2(\mathbf{p}_J). \end{aligned}$$

座標変換 $\Psi : (x_1, \mathbf{p}_J) \mapsto (y_1, \mathbf{p}_J)$, $y_1 := (x_1 - f)(\varphi_1)^{1/3}$ を考え, $\varphi := (\varphi_2/\varphi_1) \circ \Psi^{-1}$ とおく. このとき

$$z' - k_2 = y_1^3 \varphi(y_1, \mathbf{p}_J) = (p_1 - k_1) \varphi((p_1 - k_1)^{1/3}, \mathbf{p}_J).$$

ここで \mathcal{U}_1 と $V \subset L$ を適切にとることにより, 標準形 (4.11) を得る.

また, $\frac{\partial \varphi}{\partial y_1}(\bar{p}) \neq 0$ であることも分かる. 実際, 上式より, $\frac{\partial^4 z'}{\partial y_1^4}(\bar{p}) = 24 \frac{\partial \varphi}{\partial y_1}(\bar{p})$ である. さらに, $g(x_1, \mathbf{p}_J)$ を母関数とすることで

$$z'(y_1, \mathbf{p}_J) = x_1 p_1 - \tilde{g}(y_1, \mathbf{p}_J) = x_1 (y_1^3 + k_1) - \tilde{g}(y_1, \mathbf{p}_J)$$

を得る. ここで $\tilde{g} := g \circ \Psi^{-1}$ である. このとき $\frac{\partial^4 z'}{\partial y_1^4} = 3y_1^2 \frac{\partial^3 x_1}{\partial y_1^3} + 18y_1 \frac{\partial^2 x_1}{\partial y_1^2} + 18 \frac{\partial x_1}{\partial y_1}$ が従う. 以上より, Ψ が座標変換であることと $y_1(\bar{p}) = 0$ であることから,

$$\frac{\partial \varphi}{\partial y_1}(\bar{p}) = \frac{1}{24} \frac{\partial^4 z'}{\partial y_1^4}(\bar{p}) = \frac{3}{4} \frac{\partial x_1}{\partial y_1}(\bar{p}) \neq 0$$

が分かる. □

例 4.2.7 (A_3 -特異点 [42]) 母関数

$$g(x_1, p_2) = \frac{1}{4}x_1^4 - \frac{1}{2}p_2^2$$

を考える. このとき, e -波面 $W_e(L)$ は滑らかな部分多様体であり, m -波面 $W_m(L)$ は双対ポテンシャル

$$z' = x_1 p_1 - g = \frac{3}{4}p_1^{4/3} + \frac{1}{2}p_2^2$$

のグラフである. また, e -コースティクスは空である一方で, m -コースティクスは $p_1 = 0$ に沿って現れる (cf. 例 4.1.9).

4.2.2 判定法の内在的表現

§4.1 でみたように, 概ヘッセ多様体 M には概ヘッセ計量 h と標準 3 次テンソル C が備わっており, これらはそれぞれ, 局所的に母関数の 2 階および 3 階偏導関数によって表現されるものである. 本節では, これらのようなテンソルと定理 4.2.2 と 4.2.5 の関係について議論する.

定理 4.2.2 の条件式 (4.3) に現れる量 $\tau(\Phi(X), \nabla_X^{E'} \Phi'(X))$ はテンソル h と C を用いて書き換えることができる.

命題 4.2.8 ([32]) M 上の任意のベクトル場 X に対して, 次が成り立つ:

$$\tau(\Phi(X), \nabla_X^{E'} \Phi'(X)) = \frac{1}{4}(C(X, X, X) + Xh(X, X)).$$

証明: 定義 4.1.15 と 4.1.28 より,

$$\begin{aligned} h(X, X) &= 2\tau(\Phi(X), \Phi'(X)), \\ C(X, X, X) &= 2\tau(\Phi(X), \nabla_X^{E'} \Phi'(X)) - 2\tau(\nabla_X^E \Phi(X), \Phi'(X)). \end{aligned}$$

を得る. 補題 4.1.26 より, $Xh(X, X) = 2\tau(\nabla_X^E \Phi(X), \Phi'(X)) + 2\tau(\Phi(X), \nabla_X^{E'} \Phi'(X))$ であるため, 定理の等式が従う. \square

命題 4.2.8 のように, 定理 4.2.5 の条件式 (4.10) に現れる量 $\tau(\Phi(X), \nabla_X^{E'} \nabla_X^{E'} \Phi'(X))$ もテンソル h と C を用いて表現できることを期待する. しかしながらこれは不可能であることが分かる. 素朴に考えると, 二つの量

$$\begin{aligned} XC(X, X, X) &= 2\{\tau(\Phi(X), \nabla_X^{E'} \nabla_X^{E'} \Phi'(X)) - \tau(\nabla_X^E \nabla_X^E \Phi(X), \Phi'(X))\}, \\ XXh(X, X) &= 2\{\tau(\Phi(X), \nabla_X^{E'} \nabla_X^{E'} \Phi'(X)) + \tau(\nabla_X^E \nabla_X^E \Phi(X), \Phi'(X))\} \\ &\quad + 4\tau(\nabla_X^E \Phi(X), \nabla_X^{E'} \Phi'(X)), \end{aligned}$$

を足し合わせてできる量が候補になるが, $\tau(\nabla_X^E \Phi(X), \nabla_X^{E'} \Phi'(X))$ は一般には消えない. そうであっても, M 上の 4 次テンソルであって, h や C のように母関数の 4 階導関数を局所的な表示に持つ Q が存在するかもしれない, このとき h, C および Q によって表示できることを期待する. 以下ではこのようなテンソルの存在について, 統計多様体の幾何学あるいはコントラスト関数の理論の観点から議論する.

§2.2.2 でみたように, 一般に多様体 N 上にコントラスト関数 $\rho: N \times N \rightarrow \mathbb{R}$ が与えられたとき, ベクトル場 $X_1, \dots, X_k, Y_1, \dots, Y_l$ に対して定まる関数 $\rho[X_1 \cdots X_k | Y_1 \cdots Y_l]: N \rightarrow \mathbb{R}$,

$$\rho[X_1 \cdots X_k | Y_1 \cdots Y_l](r) = (X_1)_p \cdots (X_k)_p (Y_1)_q \cdots (Y_l)_q (\rho(p, q))|_{p=q=r}$$

を組み合わせることで, 擬リーマン計量や 3 次テンソルを始め N 上にさまざまなテンソルが導入される.

近年 [13] において, 統計多様体 N に対して上記の意味でコントラスト関数 ρ を組み合わせてできる 4 次テンソルは次の二つしかないことが報告されている:

$$\begin{aligned} Q_1(X, Y, Z, W) = & \rho[XW | YZ] - \rho[YZ | XW] + \rho[X | YZW] - \rho[YZW | X] \\ & + \rho[XZW | T] - \rho[Y | XZW] + \rho[XZ | YW] - \rho[TW | XZ], \quad (4.13) \end{aligned}$$

$$\begin{aligned} Q_2(X, Y, Z, W) = & \rho[XYZW | \cdot] - \rho[\cdot | XYZW] + \rho[XYZ | W] - \rho[W | XYZ] \\ & + \rho[XYW | Z] - \rho[Z | XYW] + \rho[XY | ZW] - \rho[ZW | XY]. \quad (4.14) \end{aligned}$$

ここで X, Y, Z, W は N 上のベクトル場である. 双対平坦多様体 (M, h, ∇, ∇^*) の場合には, そのダイバージェンス \mathcal{D} に関する上記のテンソルはいずれも消えてしまう. 実際, $\mathbf{x} = (x_1, \dots, x_n)$ を ∇ -アファイン座標系, $f(\mathbf{x})$ をポテンシャルとし, $\partial_i := \frac{\partial}{\partial x_i}$ ($1 \leq i \leq n$) とおくと, 任意の $1 \leq i, j, k, l \leq n$ に対して次が分かる:

$$\begin{aligned} \mathcal{D}[\partial_i \partial_j \partial_k \partial_l \cdot] &= \partial_i \partial_j \partial_k \partial_l f, & \mathcal{D}[\partial_i \partial_j \partial_k | \partial_l] &= 0, & \mathcal{D}[\partial_i \partial_j | \partial_k \partial_l] &= 0, \\ \mathcal{D}[\partial_i | \partial_j \partial_k \partial_l] &= -\partial_i \partial_j \partial_k \partial_l f, & \mathcal{D}[\cdot | \partial_i \partial_j \partial_k \partial_l] &= 3\partial_i \partial_j \partial_k \partial_l f. \end{aligned}$$

これらを式 (4.13) および (4.14) に代入することで, $Q_1 = Q_2 = 0$ を得る. 従って, 双対平坦多様体上には母関数の 4 次導関数による局所的な表示を持つ非自明な 4 次テンソルは構成できないことが分かる. 次の 5 章でみるように, 概ヘッセ多様体上には, 双対平坦多様体におけるダイバージェンスの自然な一般化である正準ダイバージェンスが定義される (定義 5.1.6); これは概ヘッセ計量がいたるところ非退化であるときダイバージェンスを復元する. 従って概ヘッセ多様体上にも上記の性質を持つ非自明な 4 次テンソルは構成できないことが分かる.

テンソルを用いる代わりに正準ダイバージェンスを用いた表現を考えることができる. これは §5.3 にて議論する.

第5章

ダイバージェンス

この章では、双対平坦多様体におけるダイバージェンスの一般化である正準ダイバージェンスを概ヘッセ多様体上に導入し、概ヘッセ多様体における拡張ピタゴラスの定理と射影定理を述べる。さらに、これまで導入した概念と江口によるコントラスト関数の理論との関係性について述べる。

5.1 拡張ピタゴラスの定理と射影定理

5.1.1 e/m -曲線

$(M, h, (E, \Phi, \nabla^E), (E', \Phi', \nabla^{E'}))$ を概ヘッセ多様体とし、 $c : I \rightarrow M$ を曲線とする。ここで $I (\neq \emptyset) \subset \mathbb{R}$ は开区間であり、速度ベクトル場を $\dot{c}(t) := \frac{d}{dt}c(t) \in T_{c(t)}M$ ($t \in I$) とする。

定義 5.1.1 ([33]) 曲線 $c : I \rightarrow M$ が m -曲線であるとは、はめ込み ($\dot{c}(t) \neq 0$) であり、任意の $t \in I$ に対し、 $E'_{c(t)}$ のベクトル

$$\Phi' \circ \dot{c}(t), \nabla_{\dot{c}}^{E'}(\Phi' \circ \dot{c})(t), (\nabla_{\dot{c}}^{E'})^2(\Phi' \circ \dot{c})(t), \dots$$

が同時にはゼロにならず、任意の2つが線形従属であるときをいう。 e -曲線についても、 Φ' と E' をそれぞれ Φ と E に取り換えることによって同様に定義する。

m -曲線 $c_m : I \rightarrow L_\alpha$ が局所モデル L_α 内で与えられているとし、

$$\mathbf{p}(t) := \pi_1^m \circ c_m(t) \in \mathbb{R}_p^n$$

によって m -ラグランジュ写像 $\pi_1^m : L_\alpha \rightarrow \mathbb{R}_p^n$ による像を表すとする。 E'_p は z' -軸に沿った線形射影によって自然に \mathbb{R}_p^n と同型であることに注意する。 $\Phi' \circ \dot{c}_m(t)$ がゼロでなければ、速度ベクトル $\dot{\mathbf{p}}(t)$ は消えず、定義 5.1.1 の条件によって、加速度ベクトル $\ddot{\mathbf{p}}(t)$ は速度

と平行になる（ゼロにはなり得る）．それゆえに $\mathbf{p}(t)$ は \mathbb{R}_p^n 内に直線上を動く．すなわち， $c_m(t)$ は m -測地線（ $\nabla^* = \nabla^{E'}$ に関する測地線）の再パラメータ化である．ある t_0 において $\Phi' \circ \dot{c}_m(t_0) = 0$ になるとき，測地線とは異なる問題が起こる．このとき， $\dot{\mathbf{p}}(t_0) = 0$ であるが， m -曲線の条件より，ある高階導関数はゼロにはならない： $\frac{d^k}{dt^k} \mathbf{p}(t_0) \neq 0$ ．そしてこのとき，ベクトル $\frac{d^{k+1}}{dt^{k+1}} \mathbf{p}(t)$ は $\frac{d^k}{dt^k} \mathbf{p}(t)$ に平行であり， $\mathbf{p}(t)$ が直線上を動くことが分かる．しかし，この m -曲線は m -コースティクスに $t = t_0$ において交わる；このとき，一度止まり，引き返すかあるいは同一直線上をそのまま通過する（ k が偶数と奇数の場合におけるこれらの現象については図 5.1 を見よ）．その直線方向ベクトル \mathbf{m} をとる． e -曲線 $c_e(t)$ に対しても，全て同様であり， \mathbb{R}_e^n における対応する直線方向ベクトルを \mathbf{e} によって表す．

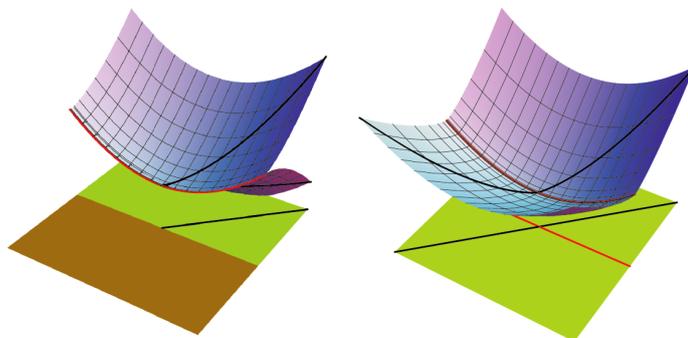


図 5.1 m -ラグランジュ写像による m -曲線の像．

注意 5.1.2 (1) 任意の 2 点を与えられたとき，それらを繋ぐ e/m -曲線は必ずしも存在しない（図 5.1 の左において，そのような m -波面上の 2 点を容易に見つけられるだろう）．任意の 2 点は（以降では用いないが）区分的 e/m -曲線によって繋ぐことができる．(2) 局所モデル L_α に対して座標系 $(\mathbf{x}_I, \mathbf{p}_J)$ をとる．任意の e/m -曲線は $h = (h_{ij})$ と $C = (C_{ijk})$ を用いた（測地方程式のような）ある微分方程式を満たすことが簡単に分かる．§4.1.5 では， α -族 $N^{(\alpha)}$ を導入した．このようにして， e/m -曲線の α -版が考えられるかもしれない；実際， $M - \Sigma$ 上では， $\nabla^{(\alpha)}$ と $\nabla^{(-\alpha)}$ に関する測地線と同じである．

次の定義は L_α や方向ベクトルの取り方にはよらない．

定義 5.1.3 ([33]) c_e , c_m と \mathbf{e} , \mathbf{m} を上記の通りとする． S を M の部分多様体とし， $q \in S$ において c_m は S に交わるとする． c_m と S が q において直交するとは，任意の $u \in T_q S$ に対し，次が成り立つときをいう：

$$\mathbf{m}^T dx(u) = 0.$$

ここで $(\mathbf{x}, \mathbf{p}, z)$ は, q を含む局所モデル L_α に対する \mathbb{R}^{2n+1} の座標系である. 同様に, c_e と S の q における直交性は, 任意の u に対して $e^T d\mathbf{p}(u) = 0$ が成り立つときをいう. さらに, c_e と c_m が q において真に直交しているとは, $e^T \mathbf{m} = 0$ が成り立つときをいう.

$q \notin \Sigma$ であるとき, c_m と S の直交性の定義は計量 h に関する直交性と同様である. 実際, q 周りの正則モデルをとると, ヘッセ行列 $H(q)$ は非退化であり, ある $k \neq 0$ に対して

$$h(\dot{c}_m(t), u) = \dot{\mathbf{x}}(t)^T H d\mathbf{x}(u) = (H \dot{\mathbf{x}}(t))^T d\mathbf{x}(u) = \dot{\mathbf{p}}(t)^T d\mathbf{x}(u) = k \mathbf{m}^T d\mathbf{x}(u).$$

一方で, $q \in \Sigma$ であるとき, 一般にはその意味は異なる. 例えば, $\dot{\mathbf{p}}(t_0) = 0$ だが $\mathbf{m} \neq 0$ なることが起き得る (この場合には, \mathbf{m} は $\mathbf{p}(t)$ のある高階導関数によって決定される). 真の直交性を定義する理由も同様である; e と \mathbf{m} は速度ベクトルによって定まらない可能性がある.

5.1.2 正準ダイバージェンスと拡張ピタゴラスの定理

L を \mathbb{R}^{2n+1} のルジャンドル部分多様体とする. 次で座標を表すとする:

$$p = (\mathbf{x}(p), \mathbf{p}(p), z(p)) \in \mathbb{R}^{2n+1}, \quad z'(p) = \mathbf{x}(p)^T \mathbf{p}(p) - z(p) \in \mathbb{R}.$$

定義 5.1.4 ([33]) 正準ダイバージェンス $\mathcal{D} = \mathcal{D}_L : L \times L \rightarrow \mathbb{R}$ を次で定義する:

$$\mathcal{D}(p, q) = z(p) + z'(q) - \mathbf{x}(p)^T \mathbf{p}(q).$$

$\mathcal{D}(p, p) = 0$ であり, 一般には非対称であることに注意する; $\mathcal{D}(p, q) \neq \mathcal{D}(q, p)$. 特に, L が正定値ヘッセ計量に関する正則モデルであるとき, これはある凸ポテンシャル $z = f(\mathbf{x})$ に対するブレグマンダイバージェンスに他ならない (cf. 式 (2.2));

$$\mathcal{D}(p, q) = f(\mathbf{x}(p)) + \varphi(\mathbf{p}(q)) - \mathbf{x}(p)^T \mathbf{p}(q).$$

ここで $z' = \varphi(\mathbf{p})$ はポテンシャルのルジャンドル変換である.

補題 5.1.5 ([33]) 正準ダイバージェンス \mathcal{D}_L はアファインルジャンドル同値で不変である. すなわち, \mathbb{R}^{2n+1} のルジャンドル部分多様体 L_1 と L_2 が \mathcal{L}_F によりアファインルジャンドル同値であるとき, $L_1 \times L_1$ 上で次が成り立つ:

$$\mathcal{D}_{L_1} = \mathcal{D}_{L_2} \circ (\mathcal{L}_F \times \mathcal{L}_F).$$

証明: $\mathcal{L}_F : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}^{2n+1}$ が次で得られるとする:

$$(\mathbf{u}, \mathbf{v}, w) = \mathcal{L}_F(\mathbf{x}, \mathbf{p}, z) = (A\mathbf{x} + \mathbf{b}, A'\mathbf{p} + \mathbf{b}', z + \mathbf{c}^T \mathbf{x} + d).$$

ここで、 $A' = (A^T)^{-1}$, $\mathbf{b}' = A'\mathbf{c}$, $\mathcal{L}_F(L_1) = L_2$ である。このとき $w' = \mathbf{v}^T \mathbf{u} - w$ とおくと、

$$\begin{aligned}
& \mathcal{D}_{L_2}(\mathcal{L}_F(p), \mathcal{L}_F(q)) \\
&= w(p) + w'(q) - \mathbf{u}(p)^T \mathbf{v}(q) \\
&= w(p) - w(q) + \mathbf{v}(q)^T (\mathbf{u}(q) - \mathbf{u}(p)) \\
&= z(p) - z(q) + \mathbf{c}^T (\mathbf{x}(p) - \mathbf{x}(q)) + (A'(\mathbf{p}(q) + \mathbf{c}))^T (A(\mathbf{x}(q) - \mathbf{x}(p))) \\
&= z(p) - z(q) + \mathbf{p}(q)^T (\mathbf{x}(q) - \mathbf{x}(p)) \\
&= z(p) + z'(q) - \mathbf{x}(p)^T \mathbf{p}(q). \\
&= \mathcal{D}_{L_1}(p, q).
\end{aligned}$$

□

$(M, \mathcal{U} = \{L_\alpha\})$ を局所モデルの張り合わせにより得られた概ヘッセ多様体とし、 $\Delta_M = \{(p, p) \in M \times M\}$ とする。 $M \times M$ の部分集合 $U(\Delta_M)$ を、ある局所モデル L_α に含まれる p, q の組み (p, q) からなる集合とする。 M には商位相が備わっていたことから、 $U(\Delta_M)$ は対角集合 Δ_M の開近傍である。

定義 5.1.6 ([33]) ある α に対する $p, q \in L_\alpha$ において、 $\mathcal{D}_M(p, q) := \mathcal{D}_{L_\alpha}(p, q)$ とおく。このとき、 $\mathcal{D}_M : U(\Delta_M) \rightarrow \mathbb{R}$ は補題 5.1.5 より well-defined である。これを M の正準ダイバージェンスと呼ぶ。

§2.1.4 でみたように、ヘッセ多様体のダイバージェンスに対して、拡張ピタゴラスの定理 (定理 2.1.17) と射影定理 (定理 2.1.18) が成り立つ。これらの定理は甘利・長岡による双対平坦多様体の理論において統計的推定や em -アルゴリズム、機械学習などへの応用の中心的な役割を担うものであり [3, 1], 我々の特異的設定において直ちに一般化される。

次の2つの定理では、 M は局所モデル (すなわち、 $M = L \subset \mathbb{R}^{2n+1}$) であるとする (とにかく正準ダイバージェンス $\mathcal{D} (= \mathcal{D}_M)$ は $M \times M$ 上で定義されているとする)。

定理 5.1.7 (拡張ピタゴラスの定理 [33]) $p, q, r \in M$ を3つの異なる点とし、 p と q は e -曲線 c_e によって繋がれ、 q と r は m -曲線 c_m によって繋がれているとする。さらに、 c_e と c_m が q において真に直交しているとする。このとき、次が成り立つ：

$$\mathcal{D}(p, r) = \mathcal{D}(p, q) + \mathcal{D}(q, r).$$

証明：直接計算により、

$$\mathcal{D}(p, r) - \mathcal{D}(p, q) - \mathcal{D}(q, r) = -(\mathbf{x}(p) - \mathbf{x}(q))^T (\mathbf{p}(r) - \mathbf{p}(q))$$

が分かる。写像 $\pi_1^e \circ c_e$ と $\pi_1^m \circ c_m$ の像はそれぞれ、方向ベクトルを \mathbf{e}, \mathbf{m} とする直線上に位置する。このとき、

$$\mathbf{x}(p) - \mathbf{x}(q) = k_0 \mathbf{e}, \quad \mathbf{p}(r) - \mathbf{p}(q) = k_1 \mathbf{m}$$

がある $k_0, k_1 \in \mathbb{R}$ において成り立つ. 仮定 $\mathbf{e}^T \mathbf{m} = 0$ により, 等式が従う. \square

定理 5.1.8 (射影定理 [33]) S を M の部分多様体とし, $c_m : [0, 1] \rightarrow M$ を $q = c_m(1) \in S$ なる m -曲線とする. $p = c_m(0) \in M$ とおく. このとき, c_m と S が q において直交することと, q が関数 $F = D(-, p) : S \rightarrow \mathbb{R}$ の臨界点になることは同値である. e -曲線 c_e と $F = D(p, -)$ に対しても同様に成り立つ.

証明: q の周りの母関数 $g(\mathbf{x}_I, \mathbf{p}_J)$ をとる. このとき, $\mathbf{p}_I = \frac{\partial g}{\partial \mathbf{x}_I}$, $\mathbf{x}_J = -\frac{\partial g}{\partial \mathbf{p}_J}$, $z = \mathbf{p}_J^T \mathbf{x}_J + g(\mathbf{x}_I, \mathbf{p}_J)$ である. $\gamma = \gamma(s)$ を $\gamma(0) = q$ なる S 上の曲線とする. この曲線上で, $\frac{d}{ds} g(\mathbf{x}_I, \mathbf{p}_J) = \mathbf{p}_I^T (\frac{d}{ds} \mathbf{x}_I) - \mathbf{x}_J^T (\frac{d}{ds} \mathbf{p}_J)$ と $\frac{dz}{ds} = (\frac{d}{ds} \mathbf{p}_J)^T \mathbf{x}_J + \mathbf{p}_J^T (\frac{d}{ds} \mathbf{x}_J) + \frac{d}{ds} g = \mathbf{p}^T \frac{d}{ds} \mathbf{x}$ を得る. 従って,

$$\begin{aligned} \frac{d(F \circ \gamma)}{ds}(s) &= \frac{d}{ds} (z(\gamma(s)) + z'(p) - \mathbf{p}(p)^T \mathbf{x}(\gamma(s))), \\ &= (\mathbf{p}(\gamma(s)) - \mathbf{p}(p))^T \frac{d(\mathbf{x} \circ \gamma)}{ds}(s) \end{aligned}$$

が分かる. $s = 0$ において, ベクトル $\mathbf{p}(q) - \mathbf{p}(p)$ は \mathbb{R}_p^n における m -曲線 c_m を射影した直線方向ベクトル \mathbf{v} のスカラー倍であり, $\frac{d}{ds} \mathbf{x} = d\pi_1^e(\frac{d\gamma}{ds}(0)) \in \mathbb{R}_x^n$ である. それゆえに, q における S と c_m の直交性は, 任意の γ に対して $\frac{d}{ds} F \circ \gamma(0) = 0$ なることと同値であり, これは q が F の臨界点であることを意味する. \square

例 5.1.9 拡張ピタゴラスの定理を例 4.1.8 における簡単な例で確かめる.

$$f(\mathbf{x}) = \frac{x_1^3}{3} + \frac{x_2^2}{2}$$

とし, L に対するアファイン座標系 $\mathbf{x} = (x_1, x_2)$ を用いる. m -ラグランジュ写像は $(x_1, x_2) \mapsto (p_1, p_2) = (x_1^2, x_2)$ であり, 退化集合 Σ は x_2 -軸である. 従って, $P := \mathbf{x}(p) = (a_1, a_2)$ と $Q := \mathbf{x}(q) = (b_1, b_2)$ に対して, 次を得る:

$$D(P, Q) = \frac{a_1^3}{3} + \frac{a_2^2}{2} + \frac{2b_1^3}{3} + \frac{b_2^2}{2} - a_1 b_1^2 - a_2 b_2.$$

\mathbb{R}_p^2 における直線 $p_2 = ap_1 + b$ (すなわち, m -曲線) は \mathbb{R}_x^2 における放物線 $x_2 = ax_1^2 + b$ に対応する. 今, 例えば, m -曲線 $c_m: x_2 = \frac{1}{2}x_1^2$ ($\mathbf{m} = (2, 1)^T$) と, それ上の 2 点 $Q := (u, \frac{u^2}{2})$ と $R := (t, \frac{t^2}{2})$ をとる. \mathbb{R}_x^2 における $\mathbf{e} = (1, -2)^T$ 方向 ($\mathbf{e}^T \mathbf{m} = 0$ となっている) で Q を通る直線 (すなわち e -曲線) 上の点 $P := (s, -2(s-u) + \frac{u^2}{2})$ をとる. このとき, $\triangle PQR$ は条件を満たし, $D(P, Q) + D(Q, R) = D(P, R)$ となることが分かる. これは点 Q が Σ 上にあるかどうかは関係なく成立するものである.

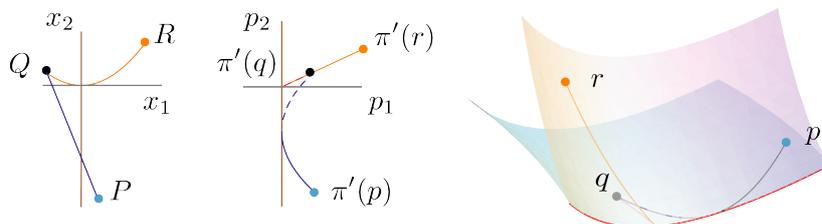


図 5.2 例 4.1.8 の L 上における三角形 $\triangle pqr$ の 2 種類の射影と m -波面上で見た三角形

5.2 弱コントラスト関数

§2.2.2 でみたように, 多様体 M に対し, 対角集合 $\Delta_M \subset M \times M$ の開近傍 U 上で定義される関数 $\rho: U \rightarrow M$ がコントラスト関数であるとは, ρ が次を満たすときをいった:

$$(1) \rho[-|-] = \rho(p, p) = 0, \quad (2) \rho[X|-] = \rho[-|X] = 0,$$

$$(3) h(X, Y) := -\rho[X|Y] \text{ は } M \text{ 上の擬リーマン計量である.}$$

ここで, (2.6) の記法を用いている. ρ が (1) と (2) のみを満たす時に, 弱コントラスト関数と呼ぶ. 与えられた統計多様体の構造を復元するコントラスト関数の存在は松本 [29] によって示されているが, そこでは計量の非退化性は用いられていないことに注意する. 従ってその結果は多様体上に与えられた (退化し得る) 対称 $(0, 2)$ -テンソルおよび対称 $(0, 3)$ -テンソルを復元する弱コントラスト関数の存在を示しているものである.

ヘッセ多様体上の幾何構造とダイバージェンスの関係 (定理 2.2.8) は概ヘッセ多様体とその上の正準ダイバージェンスの場合に拡張される.

定理 5.2.1 ([33]) 概ヘッセ多様体 M に対し, 正準ダイバージェンス \mathcal{D}_M は弱コントラスト関数であり, それは概ヘッセ計量と標準 3 次テンソルを次のように復元する:

$$\begin{aligned} h(X, Y) &= -\mathcal{D}_M[X|Y], \\ C(X, Y, Z) &= -\mathcal{D}_M[Z|XY] + \mathcal{D}_M[XY|Z]. \end{aligned}$$

証明: 局所的性質であるため, 局所モデル $L_\alpha \subset \mathbb{R}^{2n+1}$ をとる. $g(\mathbf{x}_I, \mathbf{p}_J)$ を $p \in L_\alpha$ の周りの L_α に対する母関数とする. このとき, $(\mathbf{x}_I, \mathbf{p}_J)$ は p の周りの L_α に対する局所座標系であり, p に近い $q \in L_\alpha$ に対して $x_j(q) = -\frac{\partial g}{\partial p_j}(q)$, $p_i(q) = \frac{\partial g}{\partial x_i}(q)$, $z(q) = \mathbf{p}_J(q)^T \mathbf{x}_J(q) + g(q)$ が成り立つ. 従って,

$$\begin{aligned} \mathcal{D}_M(p, q) &= z(p) - z(q) + \mathbf{p}(q)^T (\mathbf{x}(q) - \mathbf{x}(p)) \\ &= g(p) - g(q) + \mathbf{x}_J(p)^T (\mathbf{p}_J(p) - \mathbf{p}_J(q)) + \mathbf{p}_I(q)^T (\mathbf{x}_I(q) - \mathbf{x}_I(p)). \end{aligned}$$

簡単のために, ∂_k で, $k \in I$ なるとき $\frac{\partial}{\partial x_k}$ を, $k \in J$ なるとき $\frac{\partial}{\partial p_k}$ を表すとする. このとき,

$$\begin{aligned} (\partial_k)_p \mathcal{D}_M(p, q) &= \epsilon(k)(\partial_k g(p) - \partial_k g(q)) + \partial_k \mathbf{x}_J(p)^T (\mathbf{p}_J(p) - \mathbf{p}_J(q)), \\ (\partial_k)_q \mathcal{D}_M(p, q) &= (1 - \epsilon(k))(\partial_k g(p) - \partial_k g(q)) + \partial_k \mathbf{p}_I(q)^T (\mathbf{x}_I(q) - \mathbf{x}_I(p)). \end{aligned}$$

ここで, $\epsilon(k)$ は $k \in I$ なるとき 1 を, $k \in J$ なるとき 0 を表す. これより直ちに次を得る:

$$\mathcal{D}_M[-|-] = 0, \quad \mathcal{D}_M[\partial_k|-] = \mathcal{D}_M[-|\partial_k] = 0.$$

従って, 正準ダイバージェンスが弱コントラスト関数であることが分かる. また

$$\epsilon(k, l) = \begin{cases} 1 & (k, l \in I) \\ -1 & (k, l \in J) \\ 0 & (\text{その他}). \end{cases}$$

とおくと, 単純な計算で次が分かる:

$$\begin{aligned} (\partial_l)_p (\partial_k)_p \mathcal{D}_M(p, q) &= \epsilon(k, l) \partial_l \partial_k g(p) + \partial_l \partial_k \mathbf{x}_J(p)^T (\mathbf{p}_J(p) - \mathbf{p}_J(q)), \\ (\partial_l)_q (\partial_k)_q \mathcal{D}_M(p, q) &= \epsilon(k, l) \partial_l \partial_k g(q) + \partial_l \partial_k \mathbf{p}_I(q)^T (\mathbf{x}_I(q) - \mathbf{x}_I(p)). \end{aligned}$$

よって, $\mathcal{D}_M[\partial_k \partial_l|-] = \mathcal{D}_M[-|\partial_k \partial_l] = h(\partial_k, \partial_l)$ が従う. さらに,

$$\begin{aligned} (\partial_l)_p (\partial_k)_p (\partial_m)_q \mathcal{D}_M(p, q) &= -\partial_l \partial_k \mathbf{x}_J(p)^T \partial_m \mathbf{p}_J(q), \\ (\partial_l)_q (\partial_k)_q (\partial_m)_p \mathcal{D}_M(p, q) &= -\partial_l \partial_k \mathbf{p}_I(q)^T \partial_m \mathbf{x}_I(p) \end{aligned}$$

であるため,

$$-\mathcal{D}_M[\partial_m|\partial_k \partial_l] + \mathcal{D}_M[\partial_k \partial_l|\partial_m] = \partial_k \partial_l \partial_m g$$

が任意の k, l, m に対して成り立ち, 命題 4.1.30 により, これは二つ目の等式の成立を意味する. \square

5.3 e/m -波面の特異点の判定法と正準ダイバージェンス

本節では, 正準ダイバージェンスと定理 4.2.2 および定理 4.2.5 の関係について議論する. §4.2.2 では, 概ヘッセ多様体 $(M, h, (E, \Phi, \nabla^E), (E', \Phi', \nabla^{E'}))$ 上のテンソルと上記の定理の条件式に現れる量 $\tau(\Phi(X), \nabla_X^{E'} \Phi'(X))$ および $\tau(\Phi(X), \nabla_X^{E'} \nabla_X^{E'} \Phi'(X))$ との関係について議論した. 特に二つ目の量をテンソルによって表現することは叶わない一方で, これらの量を正準ダイバージェンスによって表現することができる. このことは次の定理から直ちに従う.

定理 5.3.1 ([32]) X, Y, Z, W を M 上の任意のベクトル場とする. このとき次の等式が成り立つ:

$$\tau(\Phi(Z), \nabla_Y^{E'} \Phi'(W)) = -\frac{1}{2} D_M[Z|YW], \quad (5.1)$$

$$\tau(\Phi(Z), \nabla_X^{E'} \nabla_Y^{E'} \Phi'(W)) = -\frac{1}{2} D_M[Z|XYW]. \quad (5.2)$$

証明: M の局所モデル $L \subset \mathbb{R}^{2n+1}$ をとり, $g(\mathbf{x}_I, \mathbf{p}_J)$ を L の母関数とする. ここで, $(\mathbf{x}_I, \mathbf{p}_J)$ は $p \in L$ 周りの局所座標系である. このとき p に十分近い $q \in L$ に対して

$$\mathbf{x}_J(q) = -\frac{\partial g}{\partial \mathbf{p}_J}(q), \quad \mathbf{p}_I(q) = \frac{\partial g}{\partial \mathbf{x}_I}(q), \quad z(q) = \mathbf{x}_J(q)^T \mathbf{p}_J(q) + g(\mathbf{x}_I(q), \mathbf{p}_J(q))$$

および

$$D_M(p, q) = g(p) - g(q) + \mathbf{x}_J(p)^T (\mathbf{p}_J(p) - \mathbf{p}_J(q)) + \mathbf{p}_I(q)^T (\mathbf{x}_I(q) - \mathbf{x}_I(p)) \quad (5.3)$$

が成り立つ. ここで $g(q) := g(\mathbf{x}_I(q), \mathbf{p}_J(q))$ である. 任意の $1 \leq i \leq n$ に対して ∂_i は, $i \in I$ であるとき $\frac{\partial}{\partial x_i}$, $i \in J$ であるとき $\frac{\partial}{\partial p_i}$ を表すものとする.

定理 5.2.1 より $h(Z, W) = -D_M[Z|W]$ であるため, 等式 (5.1) の成立は $Y = \partial_t, Z = \partial_u, W = \partial_v$ のときに確かめれば十分である. E' の平坦切断 s_i^* ($1 \leq i \leq n$) をとると,

$$\nabla_{\partial_t}^{E'} \Phi'(\partial_v) = \sum_{i \in I} \partial_t \partial_v \partial_i g s_i^*.$$

従って,

$$2\tau(\Phi(\partial_u), \nabla_{\partial_t}^{E'} \Phi'(\partial_v)) = \begin{cases} \partial_t \partial_u \partial_v g & (u \in I), \\ 0 & (u \in J). \end{cases}$$

一方, 式 (5.3) の右辺を微分することで次を得る (定理 5.2.1 の証明参照):

$$(\partial_u)_p (\partial_t)_q (\partial_v)_q D_M(p, q) = -\partial_t \partial_v \mathbf{p}_I(q)^T (\partial_u \mathbf{x}_I(p)).$$

従って,

$$D_M[(\partial_u)|(\partial_t)(\partial_v)] = \begin{cases} -\partial_t \partial_u \partial_v g & (u \in I), \\ 0 & (u \in J). \end{cases}$$

以上より等式 (5.1) の成立がわかる. $X = \partial_s$ とする. 同様にして,

$$2\tau(\Phi(\partial_u), \nabla_{\partial_s}^{E'} \nabla_{\partial_t}^{E'} \Phi'(\partial_v)) = \begin{cases} \partial_s \partial_t \partial_u \partial_v g & (u \in I), \\ 0 & (u \in J). \end{cases}$$

および

$$D_M[(\partial_u)|(\partial_s)(\partial_t)(\partial_v)] = \begin{cases} -\partial_s \partial_t \partial_u \partial_v g & (u \in I), \\ 0 & (u \in J). \end{cases}$$

すなわち, 等式 (5.2) が成立する. □

系 5.3.2 ([32]) M 上の任意のベクトル場 X に対して, 以下が成り立つ:

$$\tau(\Phi(X), \nabla_X^{E'} \Phi'(X)) = -\frac{1}{2} D_M[X|XX], \quad (5.4)$$

$$\tau(\Phi(X), \nabla_X^{E'} \nabla_X^{E'} \Phi'(X)) = -\frac{1}{2} D_M[X|XXX]. \quad (5.5)$$

統計多様体の幾何学から見ると, 双対平坦多様体上ではダイバージェンスから誘導される (双対) バートレットテンソルは消えており (定理 2.2.9), 双対平坦多様体の幾何学はこれに伴った制約を受けることがわかる. 概ヘッセ多様体上に直接的な方法でそれらのテンソルを構成することはできないが, 定理 5.3 の式 (5.2) は双対平坦多様体における上記の事実を反映したものであり, 概ヘッセ多様体も類似した幾何構造を持つと考えられる.

第6章

応用可能性の検討

概ヘッセ多様体上では、計量が退化していてもなお甘利・長岡による拡張ピタゴラスの定理と射影定理が依然として有効であることが示された。従って、これらの定理の帰結として得られる情報幾何学的な応用は特異的設定においても正当化される。この章では、いくつかの話題について我々の理論の応用可能性を検討する。

6.1 深層学習と統計的推定

深層ニューラルネットワークは（回帰問題の設定において）次の典型的な方法により統計モデルとみなされる：入力を $\mathbf{x} \in \mathbb{R}^n$ ，重みパラメータを $\mathbf{w} \in \mathbb{R}^k$ とする出力が1次元である深層ニューラルネットワークを $f_{\mathbf{w}}(\mathbf{x}) := f(\mathbf{x}, \mathbf{w}) \in \mathbb{R}$ とする；平均0，分散1であるガウスノイズ ε で出力 $f_{\mathbf{w}}(\mathbf{x})$ を摂動して得られる確率変数 $y := f_{\mathbf{w}}(\mathbf{x}) + \varepsilon$ の従う条件付き密度関数

$$p(y|\mathbf{x}, \mathbf{w}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y - f_{\mathbf{w}}(\mathbf{x}))^2\right)$$

からなる族 $\{p(y|\mathbf{x}, \mathbf{w})\}_{\mathbf{w}}$ を考える。このとき重みパラメータの空間 $M = \mathbb{R}^k$ 上では三次テンソル T が恒等的に消え， M は退化したフィッシャー・ラオ計量 h を持つ自己双対なリーマン多様体 $(M, h, \nabla = \nabla^*)$ とみなされる [1]。実際， $\partial_i := \frac{\partial}{\partial w_i}$ とおくと

$$T(\partial_i, \partial_j, \partial_k) = \partial_i f_{\mathbf{w}}(\mathbf{x}) \partial_j f_{\mathbf{w}}(\mathbf{x}) \partial_k f_{\mathbf{w}}(\mathbf{x}) \int (y - f_{\mathbf{w}}(\mathbf{x}))^3 p(y|\mathbf{x}, \mathbf{w}) dy = 0.$$

このような観点に基づいた深層学習の解析として、甘利による自然勾配学習法が提案されている [4]。しかしながら、理論的には、上記の枠組みでは固定されていた分散 $s > 0$ をパラメータに含めることは自然であろう (e.g., [11, 19])。この場合、

$$T\left(\frac{\partial}{\partial s}, \frac{\partial}{\partial s}, \frac{\partial}{\partial s}\right) = \frac{1}{s^3} \neq 0$$

となり、三次テンソル T は恒等的にはゼロにならない。従って、真に特異な計量を許す双対平坦構造、すなわち概ヘッセ構造からの解析が必要となる。もちろん実用的な場面では重みパラメータの推定が興味の対象になることが多い一方で、このような場合であっても分散パラメータを含めた推定は、例えば、局外母数を含んだパラメータの推定理論やハイパーパラメータの推定理論などの枠組みから捉えられる可能性があり、いずれにしても双対平坦構造に基づいた解析は有効であると考えられる [1].

統計的推論においては以下のような概ヘッセ多様体の典型的な例が現れることが分かっている。 (M, h, ∇, ∇^*) を指数型分布族とすると、確率密度関数の定義に現れる正規化関数 ψ はポテンシャルを与えていた (例 2.2.3)。 $\mathcal{D} : M \times M \rightarrow \mathbb{R}$ を ψ から定まるブレグマンダイバージェンスとすると、これはよく知られているカルバック・ライブラーダイバージェンス \mathcal{D}_{KL} と一致する：

$$\mathcal{D}_{KL}(q, p) = \int q(u) \log \frac{q(u)}{p(u)} du \quad (p, q \in M).$$

統計モデルとして曲指数型分布族 S を考える。すなわち、 $S \subset M$ を部分多様体とする。統計的推論では、データ $\{u_i\}$ が与えられたとき、得られる観測点 $\hat{p} \in M$ に対し \hat{p} を最も良く近似する $q_0 \in S$ を見つけることが問題意識である。情報幾何学は特に最尤推定に対して見通しの良い幾何学的な解釈を与える [3, 1]。すなわち、最尤推定量 \hat{p} に対して $q_0 \in S$ は $\mathcal{D}(-, \hat{p}) : S \rightarrow \mathbb{R}$ の最小値を与える点として解釈され、同様であるがこれはまた、 q_0 において S と直交する m -測地線 (m -曲線) 沿って \hat{p} から射影された結果であると解釈される。このような解釈はフィッシャー・ラオ計量が退化するような M においても有効である (定理 5.1.8)。

\hat{p} が S に十分近い場合や、 Σ から十分離れている場合には推定の漸近理論が展開される。しかしながら、実際の応用では \hat{p} がそのような場合であるかは必ずしも分からない。例えば、尤度関数が極値を複数持つ、すなわち尤度方程式が複数解を持つことはしばしば起きる。このとき、データの更新などで \hat{p} の S に対する相対的位置が変わるに従いカラストロフ現象 (最大値や最小値の個数が変化する) が起き得る。

このような分岐現象については、解選択の曖昧さが推定に与える影響の数値実験的分析 (e.g., [38, §4]), スコア統計量と曲指数型分布の曲率との関係を調べた研究 [14, 40], 推定した根の分布を特徴づける研究 [39] などがある一方で、いずれの研究も各論的でありまた数値実験的側面が強い。我々の理論は、この問題に対し理論的な視点から扱うための統一的な枠組みを提供し情報幾何学的なアプローチを可能にするものであると考えられる。

$$F : S \times M \rightarrow \mathbb{R} \quad F(q, p) := \mathcal{D}(\iota(q), p)$$

を定める。ここで $\iota : S \rightarrow M$ は包含写像である。ここで、 F を母関数族として捉える [6,

p.323]. このとき, F は $T^*M \times \mathbb{R}$ の次のルジャンドル部分多様体を定める:

$$L_S := \left\{ (p, \eta, z) \mid \exists q \in S, \frac{\partial F}{\partial q}(q, p) = 0, \eta = \frac{\partial F}{\partial p}(q, p), z = F(p, \eta) \right\}.$$

ここで, $\frac{\partial F}{\partial q}, \frac{\partial F}{\partial p}$ はそれぞれ S, M 上での微分を意味するものとする. この L_S は概ヘッセ多様体の典型的な例である. ラグランジュ写像 $\pi: L_S \rightarrow M$ ($\pi(p, \eta, z) \mapsto p$) の特異値集合は, S に直交する m -曲線の族から定まる M 上の包絡線である. これを S から定まる m -コースティクスと呼ぶことにする. S が ∇ -平坦ではないとき, m -コースティクスは一般に現れる. データ多様体 D が S から定まる m -コースティクスと交差するような際に, 上に述べたカタストロフ現象が起きることが分かる. 逆に, データ多様体 D が与えられたとき, D の $M \times D$ への制限を考えることで, D から定まる e -コースティクスが同様に定義される. これら 2 つの e/m -コースティクスの間の相互作用が EM アルゴリズムの性能に影響を与える可能性がある (cf. [1, 11 章]). 原理的には, これまで述べたことは任意の統計モデルとダイバージェンスに適用可能であることに注意する. 詳細な解析については今後行っていく.

6.2 非凸最適化問題

凸最適化問題とは, 凸領域 $M \subset \mathbb{R}^n$ を実行可能領域とし, その上の凸関数 $c: M \rightarrow \mathbb{R}$ を目的関数とする最適化問題である. 凸最適では障壁関数と呼ばれる目的関数 c に依存する凸関数が定まり, この関数をポテンシャルとする M 上の双対平坦構造 (h, ∇, ∇^*) は様々な内点法アルゴリズムに見通しの良い解釈を与える. 実際, リーマン計量 h を考慮した勾配降下法 (自然勾配) によって得られるパラメータの軌跡は ∇^* -測地線として理解され [1, 41], また, 小原・土谷・柿原は主双対内点法アルゴリズムの複雑さ (計算回数の多さ) を, 接続を用いた軌跡の曲率により評価している [25, 26].

一般には非凸関数を目的関数とする非凸最適化問題があらゆる場面で現れる. 非凸ポテンシャルを許容する我々の理論は非凸最適化問題におけるアルゴリズムの幾何学的特徴づけを与え得る. 実際, 上記の枠組みの範囲で考えても, ∇^* -測地線は m -曲線として捉えられ, また双対座標系における解析は連接接束上の接続によって行われる. 凸の場合と異なり, 一般に軌跡は特異点と接触する可能性があるが, このときには §4.2.1 で与えた e/m -波面の標準形が特異点周りでのアルゴリズムの解析を可能にするだろう. このように非凸最適化問題においても凸最適化の場合と同様に, 豊富な議論が行われると考えられる.

第7章

結論

本研究は接触幾何学や特異点理論に基づく情報幾何学の新しい方向性を探るものであり、本論文では、双対平坦多様体の理論のリーマン計量が退化する場合における一般化として概ヘッセ多様体の理論を構築した。技術的なアイデアは、双対平坦多様体の局所的な幾何学をダブルファイブレーション構造から導入されるルジャンドル部分多様体の幾何学として捉えることであり、そこではアファイン構造に基づいて導入された接続接束が重要な役割を果たしていた。この理論では、計量が退化していてもなお概ヘッセ多様体上に甘利・チェンツォフの3次テンソルや正準ダイバージェンスが一般化され、拡張ピタゴラスの定理と射影定理の成立が示された。さらに、弱コントラスト関数の観点から概ヘッセ多様体の理論と江口によるコントラスト関数の理論との整合性も示された。これらのことは我々の理論がポテンシャル関数の凸性に依拠しない新しい理論であることを意味するものである。また、概ヘッセ計量の退化性は e/m -ルジャンドル写像の特異点として特徴づけられるが、その典型的な特異点型を概ヘッセ多様体上の幾何学的諸量によって捉え、 e/m -波面のアファイン座標系における標準形（すなわち特異点を持ったポテンシャル関数と双対ポテンシャル関数）を導出した。これにより特異点周りのアファイン微分幾何学的（情報幾何学的）な解析が可能になる。

本論文で導入した概ヘッセ多様体は様々な領域の話題と関連するものである。拡張ピタゴラスの定理と射影定理は特異点設定においても成立するため、従来の情報幾何学に基づく統計的推論や機械学習、最適化問題等に対する応用的考察は空間に特異点があってもなお正当化される。また、数学の話題としてはアファイン微分幾何学やケーラー幾何学、ポアソン多様体の理論等と関連する。佐治・梅原・山田らによる波面のリーマン幾何学が豊富な理論であるように、波面のアファイン微分幾何学もまた充実しているに違いない。本論文で導入したアファイン構造に基づいた接続接束や波面のアファイン座標系における標準形はその発展の基礎になるだろう。また、ケーラー幾何学やポアソン多様体の文脈から反変擬ヘッセ多様体という概念が導入されており [10]、これは我々の退化ポテンシャルを

持つ概ヘッセ多様体と密接に関わるものである。今後はこのような関連をもとにし、概ヘッセ多様体の理論の応用と深化を行っていく。

謝辞

本論文の作成にあたり前指導教員の大本亨先生に深く感謝いたします。先生には学部4年生のゼミからお世話になっており、当時専門が決まっていなかった私に微分トポロジーの勉強を勧めてくださいました。その後大学院では、数学のみならず機械学習等の応用領域にも強い興味がある私に、(あまりにも漠然とした興味の方向性であるにも関わらず)大本先生はこの博士論文のテーマを挙げてくださいました。研究を行うにあたっては微分幾何学や接触幾何学、特異点理論など幅広い知識を勉強する必要がある中で、理解が遅い私にも痺れを切らさずに親身に議論やアドバイスをこなってくださいました。また、奨学金や日本学術振興会特別研究員の申請書の作成指導でもお世話になり、これらを修士時代から見通して研究計画を立てていただいたおかげで、私は博士課程への進学を決めることができました。このように数えられないほど大本先生には研究のみならず大学院生活全般においてお世話になり、元々博士課程への進学を考えていなかった私にとってこの博士論文は先生の存在なくしては文字通り書くことができないものでした。重ねて感謝申し上げます。

また、主査を担当してくださいました沼田泰英先生と副査を担当してくださいました今井幸幸先生に感謝いたします。沼田先生は着任されてまもなくの勝手がわからない中にも関わらず学位申請の手続きを引き受けてくださいました。ありがとうございました。

函館工業高等専門学校の友人や先生方、また編入生同期の友人や研究室の同期だった中野陽太郎氏は、卒業をして遠くにいてもいつも気にかけて、励ましてくださいました。皆の存在は北海道大学での生活をする上での大きな精神的支えでした。心より感謝申し上げます。

いつも見守り応援し続けてくれた家族に感謝いたします。何をしているのか分からない大学院生活では不安にさせてしまっていたかもしれませんが、それでも好きなことをやるようにと全力で応援してくれたことで、自信を持って研究生活を送ることができました。何度感謝しても足りないかもしれません。ありがとうございました。

最後に、私の研究生活を支えてくださいました全ての関係者の皆様に心より感謝申し上げます。

参考文献

- [1] S. Amari, *Information Geometry and Its Application*, Appl. Math. Sci. **194**, Springer, Tokyo, 2016.
- [2] S. Amari and H. Nagaoka, *Differential geometry of smooth families of probability distributions*, METR, **82** (1982), 49–98.
- [3] S. Amari and H. Nagaoka, *Methods of information geometry*, Transl. Math. Monogr. **191**, A.M.S. and Oxford Univ. Press, New York, 2000.
- [4] S. Amari and T. Ozeki, *Differential and Algebraic Geometry of Multilayer Perceptrons*, IEICE Trans., **84** (2001), 31–38.
- [5] V.I. Arnol'd, *Mathematical Methods of Classical Mechanics*, 2nd Edition, Grad. Texts Math. **60**, Springer-Verlag (1989).
- [6] V.I. Arnol'd, S.M. Gusein-Zade and A.N. Varchenko, *Singularities of Differentiable Maps I*, Monographs in Math. **82**, Birkhäuser, Boston, 1985.
- [7] V.I. Arnol'd, *Singularities of Caustics and Wave Fronts*, Kluwer Acad. Publ. (1990).
- [8] V.I. Arnol'd, *Catastrophe Theory*, 3rd Edition, Springer-Verlag (1992).
- [9] N. Ay, J. Jost, H. V. Lê and L. Schwachhöfer, *information geometry*, Springer, Cham (2017).
- [10] S. Benayadi and M. Boucetta, *On para-Kähler Lie algebroids and contravariant pseudo - Hessian structures*, Math. Nachrichten **292** (2019), 1418–1443.
- [11] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer (2006).
- [12] N. Chentsov, *Statistical decision rules and optimal inference*, Translation of Math. Monograph **53**, AMS, Providence (1982).
- [13] F. M. Ciaglia, G. Marmo and J. M. Pérez-Pardo, *Generalized potential functions in differential geometry and information geometry*, Int. J. Geom. Methods Mod. Phys. **16** (2019), 1940002.
- [14] B. Efron, *The Geometry of Exponential Families*, Ann. Statist. **6** (1978), 362–

- 376.
- [15] I. Ekeland, *Legendre duality in nonconvex optimization and calculus of variations*, SIAM J. Control Optim. **15** (1977), 905–934.
 - [16] S. Eguchi, *Geometry of minimum contrast*, Hiroshima Math. J. **22** (1992), 631–647.
 - [17] 福田拓生, 野口広, 復刊 初等カタストロフイー, 共立出版 (2002).
 - [18] 福水健次, 栗木哲, 竹内啓, 赤平昌文, 特異モデルの統計学, 統計科学のフロンティア 7, 岩波書店 (2004).
 - [19] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press (2016).
 - [20] G. Ishikawa, *Parametrization of a singular Lagrangian variety*, Trans. Amer. Math. Soc. **331** (1992), 787–798.
 - [21] 泉屋周一, 波面の伝播と特異性, 共立出版 (2018).
 - [22] 泉屋周一, 石川剛郎, 応用特異点論, 共立出版 (1998).
 - [23] S. Izumiya, M.C. Romero-Fuster, M.A.S. Ruas and F. Tari, *Differential Geometry from a Singularity Theory Viewpoint*, World Scientific (2015).
 - [24] S. Izumiya and K. Saji, *The mandala of Legendrian dualities for pseudospheres in Lorentz-Minkowski space and “flat” spacelike surfaces*, J. Singul. **2** (2010), 92–127.
 - [25] S. Kakihara, A. Ohara and T. Tsuchiya, *Information Geometry and Interior-Point Algorithms in Semidefinite Programs and Symmetric Cone Programs*, J. Optim. Theory Appl., **157** (2013), 749–780.
 - [26] S. Kakihara, A. Ohara and T. Tsuchiya, *Curvature integrals and iteration complexities in SDP and symmetric cone programs*, Comput Optim Appl **57** (2014), 623–665.
 - [27] M. Kokubu, W. Rossman, K. Saji, M. Umehara and K. Yamada, *Singularities of flat fronts in hyperbolic 3-space*, Pacific J. Math. **221** (2005), 303–351.
 - [28] S. L. Lauritzen, *Statistical manifolds*, IMS Lecture Notes - Monograph Series, (1987), 96–163.
 - [29] T. Matumoto, *Any statistical manifold has a contrast function – On the C^3 -functions taking the minimum at the diagonal of the product manifold*, Hiroshima Math. J. **23** (1993), 327–332.
 - [30] H. Matsuzoe, *Geometry of contrast functions and conformal geometry*, Hiroshima Math. J. **29** (1999), 175–191.
 - [31] H. Matsuzoe, *Statistical manifolds and affine differential geometry*, Advanced Stud. Pure Math. **57** (2010), 303–321.

-
- [32] N. Nakajima, *Local normal forms of em-wavefronts in affine coordinates*, Kodai Math. J. **45** (2022), 388–403.
- [33] N. Nakajima and T. Ohmoto, *The dually flat structure for singular models*, Inf. Geom. **4** (2021), 31–64.
- [34] T. Poston, I. Stewart, *Catastrophe Theory and Its Applications*, Pitman Publ. Ltd. (1978).
- [35] K. Saji, M. Umehara, K. Yamada, *The geometry of fronts*, Annals of Math., **169** (2009), 491–529.
- [36] K. Saji, M. Umehara, K. Yamada, *Differential geometry of curves and surfaces with singularities*, World Scientific (2021).
- [37] H. Shima, *The geometry of Hessian Structures*, World Scientific (2007).
- [38] C. Small and W. Jinfang, *Numerical Methods for Nonlinear Estimating Equations*, Oxford Univ. Press (2003).
- [39] C. Small and Z. Yang, *Multiple roots of estimating functions*, The Canadian Journal of Statistics, **27** (1999), 585–598.
- [40] R. Sundberg, *Statistical Modelling by Exponential Families*, Cambridge University Press (2019).
- [41] K. Tanabe, *A geometric method in nonlinear programming*, J Optim Theory Appl **30** (1980), 181–210.
- [42] V.A. Vasil’ev, *Asymptotic exponential integrals, Newton’s diagram, and the classification of minimal points*, Funct. Anal. Appl. **11** (1977), 163–172.
- [43] S. Watanabe, *Algebraic Geometry and Statistical Learning Theory*, Cambridge Univ. Press (2008).