

# TERG

Discussion Paper No.401

『道德感情論』における良俗の一般的諸規則形成の進化  
経済学的再解釈

菅 隆彦

発行 2019 年 3 月 19 日

TOHOKU ECONOMICS RESEARCH GROUP  
Discussion Paper

---

GRADUATE SCHOOL OF ECONOMICS AND  
MANAGEMENT TOHOKU UNIVERSITY  
27-1 KAWAUCHI, Aoba-KU, SENDAI,  
980-8576 JAPAN

# 『道徳感情論』における良俗の一般的諸規則形成の進化経済学的再解釈

## 1 はじめに

アダム・スミス著『道徳感情論』<sup>1</sup>は、経済学の様々な研究成果の観点から再解釈され、その現代における意義が複数の側面から見出されてきた。ゲーム理論を用いて同書第3部における自己規制 (self-command) を再解釈するものや (Meardon & Ortmann 1996)、同書が行動経済学の研究成果を予見していることを示すもの (Ashraf *et al.* 2005)、同書を制度派経済学の観点から再解釈するもの (Tajima 2007)、行動経済学的なモデルとの差異を明らかにしつつ、効用関数を用いて同書における主体の行動を定式化するもの (Bréban 2012)、合理的選択理論の観点から再解釈するもの (Khalil 2017)、等が存在する。

しかし、『道徳感情論』には経済学の諸成果を用いて再解釈される余地のある、重要な概念が未だに存在する。本報告は、『道徳感情論』における「良俗の一般的諸規則」(general rules of morality) (以下、一般的諸規則) の形成を、進化ゲームモデルの1つである、レプリケータダイナミクスを用いて再解釈する<sup>2</sup>。一般的諸規則は、主体達が他者を繰り返し観察することによって形成されるが、この形成過程は試行錯誤学習の1種とみなされる。このような主体の学習過程は、レプリケータダイナミクスを含む進化ゲームモデルにより定式化されてきた。

『道徳感情論』における主体達は、互いの言動を継続的に観察しあうことによって相互作用し、自ら一般的諸規則を形成する。そして、この諸規則に従って自身達の言動を制御する。このようなスミスの社会秩序観は、理神論的であるとみなされてきた。しかし、同書においては、道徳感情の腐敗に代表される、非理神論的と言える記述も存在する。本報告は、これら2つの状況の混在についての、再解釈を提示する。

本報告が提示する混在状況についての解釈は、道徳感情の腐敗に代表される、非理神論的状況についての議論に、寄与する。道徳感情の腐敗論は、同書の社会秩序形成論に矛盾すると考えられてきた。しかし、本定式化によって生じた解釈からすれば、理神論的状況と非理神論的状況は矛盾しないとも、考えられる。

本報告は以下のように構成される。第2節において一般的諸規則について説明する。第3節において、進化ゲームモデルを構築し、このモデルを分析する。第4節において、本

---

<sup>1</sup> 『道徳感情論』を引用する際には、(TMS : 「グラスゴウ版のパラグラフ番号」) の形式で、対応する箇所を示す。

<sup>2</sup> スミスの著作を進化的な視点から考察するのが Evensky(1998)である。本稿では、主体の判断の進化を考察対象とするが、Evenskyは社会の進化を対象とする。

モデルを用いて一般的諸規則の形成過程を再解釈する。最終節において、本報告を総括し含意について述べる。

## 2 良俗の一般的諸規則

### 2.1 中立的な観察者

本節では、中立的な観察者について、その要点を簡潔に記す<sup>3</sup>。

『道徳感情論』における基礎的な概念が、「同感」(sympathy)である。同感とは、何らかの事柄に際して他者が抱く感情に、主体が接する際に生じる。同感とは、主体が想像の中で当事者の境遇に身を置き、当事者と同じ感情を共有することである。

「中立的な観察者」(impartial spectator)とは、主体が自身の言動を制御する際に用いられる、想像上の観察者である。主体が自身の言動の正当性を判断するには、言動をとる自分と、その言動を精査する自分の、「ふたりの人間に分割」しなくてはならない(TMS : III.1.6)。後者の役割を果たすのが中立的な観察者であって、この観察者は、言動が取られる場面・状況を理解したうえで、当事者から離れた客観的立場に立ち、主体の言動を判断する。

中立的な観察者は、主体達が社会に出て他者を観察することによって、形成される。他者は主体にとって鏡のような存在であり、主体の言動がどのような時に是認されるのか、あるいは否認されるのかを、主体に示す。自身の言動が他者に是認された場合に主体は喜び、否認された場合には主体は気持ちを落とす。そのため、主体は是認の喜びを得るために他者を注意深く観察するようになる。

『道徳感情論』における主体は、心中に中立的な観察者を形成し、この観察者に是認されるように自身の言動を制御する。しかし、人間は、たとえ中立的な観察者の見方を認識していたとしても、感情が高まると、その見方を無視してしまう弱さを持つ。このような状況を、スミスは「自己欺瞞」(self-deceit)と呼び、「この致命的な弱点は、人間生活の混乱の半分の源である」とした(TMS : III.4.6)。

### 2.2 良俗の一般的諸規則

本節では、良俗の一般的諸規則について要点を簡潔に記す<sup>4</sup>。

自己欺瞞に対処するために有用なのが、「一般的諸規則」(general rules of morality)と

---

<sup>3</sup> 東北大学経済学研究科の TERG399 で詳説している。

<sup>4</sup> 東北大学経済学研究科の TERG399 で詳説している。

いう複数の規則である。一般的諸規則は、中立的な観察者の判断と整合する規則であり、社会で共有される。一般的諸規則と中立的な観察者は、それぞれが下す判断が一致する一方で、実際に主体の言動を制御可能か否かという点で異なる。中立的な観察者が主体の心中の判断基準に過ぎないのに対し、一般的諸規則は規則であって、しかも社会で共有されている。

中立的な観察者と一般的諸規則の違いを生み出す要因が、形成過程における、他者の観察の継続性と言えよう。一般的諸規則は、ある言動に対する他者の判断を「継続的な観察」(TMS : III.4.7) によって学習することで形成されるのであって、主体の一定以上の観察経験に基づく。

一般的諸規則には、①非難される値うちがある行動についての諸規則と、②称賛される値うちがある行動についての諸規則の、2つの類型が存在する。非難される値うちがある行動とは、中立的な観察者に同感されないような行動である。称賛される値うちがある行動とは、中立的な観察者に同感されるような行動である。ここでは、①の非難される値うちがある行動について主に説明する。2種類の形成過程の差異は、後に定式化の段階で考慮される。

一般的諸規則は、集団の外部から与えられる規則ではなく、人々の相互作用によって形づくられる。一般的諸規則の形成過程は他者を観察することから始まる (TMS : III.4.7)。他者の行動を繰り返し観察する中で、主体はある種の行動から衝撃を受ける。この種の行動とは、非難される値うちがある行動のことである。主体はこのような行動を見苦しいと感じる。そして、主体はその行動に対して周りの皆が自身と同様の嫌悪感を抱いているのを知ることとなる。他者が自身と嫌悪感を共有することを知った主体は、自身の感情が正当だという思いを強くする。この経験が繰り返されることで、一般的諸規則は形成される。一般的諸規則は「継続的な観察」(TMS : III.4.7) によって形成されるのであり、その形成は主体の「経験にもとづいている」(TMS : III.4.8)。一般的諸規則が形成されるには、主体が自身の感情の正当性を確信するに至る必要があるため、少ない回数での観察では不十分である。諸規則が形成されるためには、「継続的な観察」が行われ、多くの他者との感情共有が確認されなければならない。

主体が自身の感情の正当性を確信するに至ると、その次の段階として、自身がその行動をとった場面を想像する。自身が見苦しいと確信しているのだから、当然、自分がその行動をとった場合には、他者から見苦しいとみなされると主体は判断する。他者から見苦しいとみなされることを避けるため、主体はその行動はとるまいと決意する。かくして主体は、非難される値うちがある行動を取ってはならないとする、1つの一般的規則を形成する。この規則が形成される過程で、ある行動に対する、他者との感情の一致が確認された。他の主体も同様の過程を経るのであり、同様の一般的規則を形成する。

上の諸規則の形成過程では、非難される値うちがある行動に対する嫌悪感が、一般的諸規則の形成につながった。この嫌悪感を、称賛される値うちがある行動に対する好意に読

み替えることで、②称賛される値うちがある行動についての諸規則の形成過程は説明される。

### 2.3 良俗の一般的諸規則と理神論

スミスは、人間がすべての場面で一般的諸規則に厳密に従うとはしていないし、従うべきだともしていない。とはいえ、スミスは、社会秩序を形成するうえで不可欠な役割を果たす、正義の徳が求める行動の一般的諸規則に関しては、人間が厳密に従うべきだとした<sup>5</sup>。「正義の諸規則にしっかり固執することには、杓子定規というものは存在しない」(TMS : III.6.10)。正義の徳とは、「否認されるのが自然な動機から、ある特定の人びとにたいして、現実的で積極的な害をなす」(TMS : II.ii.1.5) ことを、回避することである。正義の徳は、直接的に同感できないような、あるいは間接的に同感できないような、他者へ有害な影響を及ぼす言動を取ることを、禁止する。正義の徳の一般的諸規則は、ある行動をとるべきでないとする、諸規則に該当する。

正義の徳は、その遵守が社会秩序の維持に不可欠である。社会の秩序は「たがいに害をあたえ侵害しようと、いつでも待ちかまえている人びとのあいだには、存立しえない」(TMS : II.ii.3.3)。正義の徳は、「大建築物の全体を支持する主柱」であり、「もしそれが除去されるならば、人間社会の偉大で巨大な組織は、一瞬に崩壊して」しまう (TMS : II.ii.3.4)。それゆえ、正義の徳を守ることは、「われわれ自身の意志の自由にまかされず、力づくで強制されてもよく、その侵犯は憤慨の、したがって処罰の的となる」(TMS : II.ii.1.3)。スミスは、『道徳感情論』において、「正義の原理論と処罰の原理論とを完全に一体化し、正義と不正をいかに判断するのかと言う問題と処罰に値する罪悪をいかに判断するのかという問題を一括して」論じた (新村 1994 : 182)。

正義の一般的諸規則が人々の相互作用によって形成され、厳守されることで、社会秩序は形成される。このようにして社会秩序が形成されるように、人間社会は、「最高存在」(Deity)、すなわち自然の創造主に設計された。スミスの著作に見られる、「神聖な設計者」「偉大な技師」「宇宙の管理者」「全知の存在」「自然の創造者」等の用語は、創造主としての神を意味していた (山口 2014 : 109)。スミスの「基本的世界観によれば、神は『全知・全善』という属性を有し、またこの世を想像した創造主」である (村越 1999 : 40)。一般的諸規則は、上述のように、「あくまでも人間の現世における個々の経験の集積を通じてつくり出されたもの」(石沢 1990 : 37) だが、同時にそれは、最高存在の諸法であるともみなされ人々から尊敬される (TMS : III.5.3)。一般的諸規則は最高存在が「われわれの内面

---

<sup>5</sup> 正義の徳とは対照的に、値うちの徳は強制されるものではなく、その欠如が正当な処罰の対象にはならないとされた。値うちの徳は「たとえて言うと建物の装飾であり、その存在は快適なものであるが、社会の存続にとって不可欠というわけではない」(中谷 1996 : 13)。また、我々は値うちの徳が義務感からなされるのを望まない (TMS : III.6.4)。

に設定した代理人たちによって布告された、かれの諸命令および諸法律とみなされるべきである」(TMS : III.5.6)。しかし、人間にとって、最高存在の設計の目的を意識して一般的諸規則を遵守することは稀である (TMS : II.ii.3.9)。人間が主に関心を持つのは、社会の維持への顧慮ではなく、限定された範囲の個別的顧慮に過ぎない (TMS : II.ii.3.10)。人間が、正義の一般的諸規則を遵守するのは、侵害を受ける個人への関心からであったり、侵害に対しての来世における処罰への関心であったりする (TMS : III.3.12)。このように、スミスは目的原因と作用原因とを明確に区別し、人間による正義の一般的諸規則の遵守を作用原因によって説明し、この遵守が、神の目的である社会秩序を形成するとした。

このスミスの社会秩序観は、理神論的であるとみなされてきた<sup>6</sup>。「理神論」(deism)とは、神が世界を創造したとみなす一方で、人間の活動への神の干渉を否定する、宗教観である<sup>7</sup> (Morris 1883 : 10, Corfe 2007 : 54)。理神論という宗教観は、「神の発動性を宇宙の創造に限定し、創造されたのち、宇宙は神から独立的に自己展開の力を有するとし、あたかもネジを巻かれた時計のように、自動的に自らの委ねられたコースを進んでいくと見るものである」(山口 1983 : 88)。また、理神論は「奇蹟や啓示を否認し、宗教を合理的倫理学に解消する」(田中 1993 : 61)。理神論は、「西ヨーロッパ市民革命期の思想的主流であって、人間の理性による真理の認識とそれにもとづく人類の進歩の無限性に対する確信を基調とした啓蒙思想に立つ宗教観である」(石沢 1990 : 24)。一般的諸規則は、神から与えられるものではなく、人間が自ら形成するものであるから、理神論の考えた方に整合すると言える。

しかしながら、理神論的な一般的諸規則の形成とは、整合しないと思われる記述が、『道徳感情論』中に存在する。典型的には、第6版で追加された、「道徳感情の腐敗」(corruption of our moral sentiments) についての記述がそれに該当する<sup>8</sup>。道徳感情の腐敗論は、徳の道 (road to virtue) と財産の道 (that [road] to fortune) という、2つの異なる倫理基準についての議論であり、財産の道が社会で拡大することにより、一般的諸規則とは異質の、腐敗した倫理基準が確立されてしまうとした (TMS : I.iii.3.全体)。腐敗論の加筆は、スミスが第6版を出版した最大の動機とも言われている (田中 2000 : 123, 田島 2003 : 265)。スミスが腐敗論を加筆したのは、現実の経済世界の認識を深め、制度改革では解決されえない、人間本性に根付く問題を自覚したことによる<sup>9</sup> (田中 2000 : 122-123)。道徳感情の

---

<sup>6</sup> スミス自身が理神論者であったか否かは、1つの論点である。石沢(1990)や山口(2014)はスミスを理神論者とみなす。理神論者でないとする立場の1つが、田中(1993)であり、スミスは理神論ではなく自然宗教の伝統に立脚しているとした。田中の主張については村越(1999 : 38-39)もまた参照。無神論者であるとの見方もある (Minowitz 1993)。

<sup>7</sup> Hefelbower(1920), Harrison (1990 : 62)もまた参照。

<sup>8</sup> 腐敗論は、『道徳感情論』第1部第3篇第3章において展開されており、本稿はこの章を議論の対象とする。一方で、篠原(1998 : 106-107)は、道徳感情の腐敗論の「真実の対象は…第3部の増補部分に潜伏している議論のなかにみられる」との見解を持つ。

<sup>9</sup> 田中は、先行する『国富論』第3版においても同様の影響が見られ、両書に連続性が見られるとする。一方で、田島(2003 : 259)は両書間に「明らかな緊張」を見る。

腐敗論を含む第6版における改訂は、スミス自身の「社会科学体系の道徳性に対する疑問の増大に基づく」ものであった(田中 1993 : 176-177)。柴田(2010 : 13-14)が述べるように、道徳感情の腐敗論の加筆によって、『道徳感情論』の「理神論的世界観にもほころびが生じている」とも考えられる。さらには、「スミスの道徳哲学体系を破綻させる」ともみなされている(田島 2003 : 264)。道徳感情の腐敗論あるいはその他の根拠から、スミスは理神論者を装っていたとする立場がある<sup>10</sup> (Lindgren 1973 : 148, Kristol 1980 : 204, ラファエル 1986 : 41-42, Kennedy 2011 : 400, Peach 2014 : 80)。

### 3 良俗の一般的諸規則形成の定式化

#### 3.1 モデル

前節で述べた、一般的諸規則の形成過程は、主体の試行錯誤学習によって進行するとみなされる。①非難される値うちがある行動についての、諸規則について述べる。主体はまず、他者の見苦しい行動を観察し嫌悪感を持つ。その上で、他者が自身と同様の感情を抱くことを知ると、自身の感情が正当だという思いを強くする。これが継続的に繰り返されれば、一般的諸規則が形成される。一方で、他者が自身と同じ感情を抱かないことが観察された場合には、感情の正当性が確認されることはない。主体は、他者と感情を共有するか否かを判断基準として、自身の感情の正当性を試行錯誤的に学習し、一般的諸規則を形成する<sup>11</sup>。

この、試行錯誤学習による感情の正当性の上昇過程を、各主体がその感情を持つべきとみなす確率の、上昇過程とみなし、定式化する。本モデルにおけるプレイヤーは、2つの戦略の中から1つの戦略を確率的に選択する。プレイヤーの集合を  $I = \{1, 2, \dots, n\}$  とし、 $n = 2m (m \in \mathbb{N})$  とする。 $S_i = \{A, B\}$  をプレイヤー  $i \in I$  の純粋戦略の集合とし、 $x_i = (x_{iA}, x_{iB}) \in [0, 1]^2$  をプレイヤー  $i \in I$  の混合戦略とする。 $x = (x_1, x_2, \dots, x_n) \in [0, 1]^{2 \times n}$  を集団の混合戦略プロファイルとし、 $\theta$  を混合戦略プロファイル  $x$  の母集合とする。プレイヤー  $i \in I$  は、 $t \in \mathbb{N}$  期に戦略 A を確率  $x_{iA}(t) \in [0, 1]$  で選択する。同様に、戦略 B を確率  $x_{iB}(t) \in [0, 1]$  で選択する。

2つの戦略は、例えば、ある特定の行動に対して、A がそれを非難するという意味

---

<sup>10</sup> 理神論についての前注でも述べた通り、スミスの真の宗教的態度がいかなるものであったのかは、未解決の論争となっている。典型的な争点は「見えざる手 (invisible hand)」の解釈である。Hill(2001), Harrison(2011), Oslington(2012)を参照。

<sup>11</sup> 社会的学習による規範の形成は、その存在の証拠が、心理学の研究によって与えられている。子供は、他者を観察・模倣することにより、社会的行動を自発的に学習する (Bandura 1977, Rosenthal & Zimmerman 2014)。関連文献については、Henrich *et al.*(2005 : 813)を参照。

し、Bが非難しないということを意味する、というように解釈される。本モデルでは、プレイヤーが戦略Aをとるということを、プレイヤーがある状況下において、戦略Aをとるべきだと考えていることとみなす。戦略Bについても同様である。プレイヤー*i*は、*t*期に戦略Aを確率 $x_{iA}(t)$ で選択するべきと考えていることになる。プレイヤーの確率的な選択を仮定するのは、一般的諸規則が形成される前には、どの言動をとるべきかについて主体が不確実な判断を行うからである。一般的諸規則が形成される前の段階にあり、自己欺瞞に陥っている主体は、中立的な観察者の判断を無視してしまう可能性がある。主体は中立的な観察者と同じ判断をある確率で行うし、そうでない判断もある確率で行う。このような状態の主体ではあるが、他者の観察を繰り返すことで、中立的な観察者の判断の正当性についての確信度が上昇して行く。確信度が上昇するにつれて、主体が中立的な観察者と同様の判断を行う確率が増加する。そして、一般的諸規則が形成されるに至れば、確実に判断を行うことが可能になる。

試行錯誤学習を含む学習の過程は、進化ゲームモデルによって定式化されてきた。進化ゲームモデルには、レプリケータダイナミクス、模倣ダイナミクス、試行錯誤ダイナミクス、最適反応ダイナミクス等が存在する（大浦 2008）。本報告は様々な進化ゲームモデルの中から、レプリケータダイナミクスを採用する。後述するように、本モデルは他者との感情共有の喜びをゲームの利得と解釈し、この喜びが大きい戦略ほど採用確率が増加し、最終的に一般的諸規則となる過程を定式化する。本モデルにおいては、各戦略の採用確率はもっぱら期待利得の大小に依存して変動する。他の進化ゲームモデルにおいては、期待利得の大小以外の要素がダイナミクスに影響する。そのような仮定を課すことは、実証研究との整合性という面では望ましいと思われるが、一般的諸規則についての、『道徳感情論』中の記述からは直接導かれない<sup>12</sup>。他者との感情共有が一般的諸規則の形成につながるという、同書中の記述を忠実に再現するために、本報告はレプリケータダイナミクスを採用する。なお、一般的なレプリケータダイナミクスは、集団の状態が、各プレイヤー間のゲームの結果に依存して変化すると想定されるが、このモデルでは、プレイヤーの各戦略の採用確率の変化をモデル上に表現できない<sup>13</sup>。よって、本報告では一般的なレプリケータダイナミクスではなく、各プレイヤーの混合戦略の変化を考察対象とする、特殊なレプリケータダイナミクスを採用する。

本モデルにおいては、每期、*n*人のプレイヤー集合の中で、全プレイヤーが2人1組ずつランダムマッチングされる。マッチングされた2プレイヤーは、対戦において、相手の選んだ戦略を知る。この過程が一般的諸規則の形成過程における、他者を観察する過程に相当する。このマッチングと対戦が連続的に幾度も繰り返され、継続的な観察が行われると

---

<sup>12</sup> 実証研究に整合するモデルの構築を課題とするのが、例えば、Roth & Erev(1995), Erev & Roth(1998)である。これらは、忘却や誤認等を導入した、強化学習モデルを用いた。

<sup>13</sup> レプリケータダイナミクスの一般的な解説としては、スミス(1985), Weibull(1997), 大浦(2008)等を参照。

想定する。

各プレイヤーは利得最大化を目的として選択を行う<sup>14</sup>。『道徳感情論』における主体が利得最大化を目的として行動するのは、一見して不適當だとみなされかねない。しかし、このように想定することは、スミス自身の記述と整合する。本モデルにおける利得は、一般的諸規則の形成過程における、他者との感情共有の喜びを意味する。この感情共有は、将来的な、他者からの自身に対する、同感の喜びを主体に想像させるだろう。主体は、他者の言動の観察結果から、将来に観察対象と同様の言動を自身がとった際に、他者から同感されるだろうことを、想像する。『道徳感情論』における主体にとって、他者に同感されることは喜びである<sup>15</sup>。利得の最大化は、想像される同感の喜びの最大化を意味するのであり、これを仮定することは自然であろう。

各プレイヤー $i$ の $x_{iA}$ は、各期のプレイヤー $i$ の期待利得に応じて変化する。 $s_i \in S_i$ をプレイヤー $i$ の純粋戦略とし、 $s = (s_1, \dots, s_n)$ を純粋戦略プロファイルとする。純粋戦略プロファイルの集合を、 $S = \times_{i=1}^n S_i$ とする。任意の純粋戦略プロファイル $s \in S$ について、プレイヤー $i$ の利得を、純粋戦略利得関数 $\pi_i: S \rightarrow \mathbb{R}$ によって定義する。

混合戦略プロファイル $x \in \Theta$ がプレイされる時、純粋戦略 $s = (s_1, \dots, s_n) \in S$ が採用される確率を、 $x(s) = \prod_{i=1}^n x_{is_i} \in [0,1]$ とする。

関数 $u_i: \Theta^{2 \times n} \rightarrow \mathbb{R}$ を以下のように定義する。

$$u_i(x) = \sum_{s \in S} x(s) \pi_i(s)。$$

関数 $u_i$ は混合戦略 $x \in \Theta$ がプレイされる時の、プレイヤー $i$ の期待利得である。プレイヤー $i$ が純粋戦略 $s_i$ をとる際の、混合戦略を $e_i^{s_i}$ と表す。その際の期待利得は、 $u_i(x_1, \dots, e_i^{s_i}, \dots, x_n)$ となる。

$x$ の挙動を知るには、 $x_A = (x_{1A}, \dots, x_{nA}) \in [0,1]^n$ の挙動を知るだけで十分である。 $x_A$ を社会状態と呼ぶ。 $x_{iA}$ のダイナミクスを

$$\dot{x}_{iA} = [u_i(x_1, \dots, e_i^A, \dots, x_n) - u_i(x)] x_{iA},$$

とする。これは、ウェイブル(1999 : 215)の方程式である。このダイナミクスは、ウェイブル(1999 : 93)におけるダイナミクスと、変数は異なるものの、同様の方法で導出される。集団 $i$ において、戦略 $K$ をとる個体数を $p_{iK}$ としたときに、 $\beta$ を基礎適応度、 $\delta$ を死亡率として、 $p_{iK} = [\beta + u_i(x_1, \dots, e_i^K, \dots, x_n) - \delta] p_{iK}$ となる。これと、 $x_{iK} = p_{iK} / \sum_L p_{iL}$ から、 $\dot{x}_{iK}$ が導出される。 $u_i(x_1, \dots, e_i^A, \dots, x_n)$ は純粋戦略 $A$ をとった場合の、プレイヤー $i$ の期待利得である。

<sup>14</sup> ここでの利得最大化とは、進化経済学で仮定される、限定合理性の下での利得最大化である。

<sup>15</sup> 「自然はかれ [人間] に、かれら [他者] の好意的な顧慮に喜びを感じ、好意的でない顧慮に苦痛を感じるように、教えた。自然は、かれらの明確な是認をそれ自体で、かれにとってもっとも嬉しがらせるもの、最も快適なものとし、かれらの明確な否認を、もっともくやしがるさせる、もっとも不快なものとしたのである」(TMS : III.2.6)。

$u_i(x)$  は、混合戦略  $x$  がプレイされる場合の、プレイヤー  $i$  の期待利得であり、 $x_{iA} \cdot u_i(x_1, \dots, e_i^A, \dots, x_n) + x_{iB} \cdot u_i(x_1, \dots, e_i^B, \dots, x_n)$  となる。 $x_{iB} = 1 - x_{iA}$  より、 $x_{iA}$  のダイナミクスから  $x_{iB}$  のダイナミクスが得られる。このダイナミクスにおいては、期待利得の高い戦略ほど採用確率の上昇が大きくなる。言い換えれば、感情共有の喜びが大きい戦略ほど採用確率の上昇が大きくなる。

プレイヤー  $i$  の利得は下の表のように決定されるとする。1 行目の利得はプレイヤー  $i$  が戦略 A をとるときに得る、利得である。2 行目の利得はプレイヤー  $i$  が戦略 B をとるときに得る、利得である。

$i$ / 相手	戦略 A	戦略 B
戦略 A	$a_{AA}^i$	$a_{AB}^i$
戦略 B	$a_{BA}^i$	$a_{BB}^i$

表：プレイヤー  $i$  の利得

ここで、 $\alpha_i \equiv a_{AA}^i - a_{BA}^i$ ,  $\beta_i \equiv a_{AB}^i - a_{BB}^i$ , とする。本ダイナミクスの連立微分方程式は、以下の通りになる。 $u_i(x_1, \dots, e_i^A, \dots, x_n) = a_{AA}^i (\sum_{j \neq i} x_{jA} / n - 1) + a_{AB}^i (1 - (\sum_{j \neq i} x_{jA} / n - 1))$ ,  $u_i(x_1, \dots, e_i^B, \dots, x_n) = a_{BA}^i (\sum_{j \neq i} x_{jA} / n - 1) + a_{BB}^i (1 - (\sum_{j \neq i} x_{jA} / n - 1))$  から、 $\dot{x}_{iA}$  が導出される。

$$\begin{aligned} \dot{x}_{1A} &= \left( (\alpha_1 - \beta_1) (\sum_{j \neq 1} x_{jA} / n - 1) + \beta_1 \right) x_{1A} (1 - x_{1A}), \\ &\vdots \\ \dot{x}_{iA} &= \left( (\alpha_i - \beta_i) (\sum_{j \neq i} x_{jA} / n - 1) + \beta_i \right) x_{iA} (1 - x_{iA}), \\ &\vdots \\ \dot{x}_{nA} &= \left( (\alpha_n - \beta_n) (\sum_{j \neq n} x_{jA} / n - 1) + \beta_n \right) x_{nA} (1 - x_{nA}). \end{aligned}$$

この連立微分方程式より、全ての端点が平衡点であることがわかる。

本報告は 3 種類の利得表の存在を仮定し、それに応じて 3 種類のプレイヤーを定義する<sup>16</sup>。

<sup>16</sup> 本モデルにおいては、各プレイヤーの利得表が異なるという意味で、プレイヤー間に

プレイヤー*i*の利得が、 $\alpha_i > 0$ ,  $\beta_i < 0$ ,  $a_{AB}^i < a_{AA}^i$ ,  $a_{BA}^i < a_{BB}^i$ を満たすとき、このプレイヤーを、**同調タイプ**とする。同調タイプは、自身の取る戦略を固定した時に、対戦相手と戦略が一致した場合の方が、一致しない場合に比べて利得が大きくなる。対戦相手と戦略が一致するという事は、相手と感情を共有することを意味する。同調タイプのプレイヤーは、相手と感情を共有したときに、そうでないときに比べて喜びが大きくなる。

プレイヤー*i*の利得が、 $\alpha_i > 0$ ,  $\beta_i > 0$ を満たすとき、このプレイヤーを、**Aタイプ**とする。Aタイプのプレイヤー*i*は、対戦相手の戦略がいずれの場合にも、戦略Aのもたらす利得が、戦略Bのもたらす利得より大きくなる。このタイプのプレイヤー*i*は、戦略の変化が相手プレイヤーの戦略に依存せず、定義域内では平衡点を除いて常に $\dot{x}_{iA} > 0$ が成り立つ。

プレイヤー*i*の利得が、 $\alpha_i < 0$ ,  $\beta_i < 0$ を満たすとき、このプレイヤーを、**Bタイプ**とする。Bタイプのプレイヤー*i*は、対戦相手の戦略がいずれの場合にも、戦略Bのもたらす利得が、戦略Aのもたらす利得より大きくなる。このタイプのプレイヤー*i*は、戦略の変化が相手プレイヤーの戦略に依存せず、定義域内では平衡点を除いて常に $\dot{x}_{iA} < 0$ が成り立つ。

以上のタイプに加えて、例えば、 $\alpha_i < 0$ ,  $\beta_i > 0$ を満たすようなプレイヤー*i*を想定することも可能であるが、本報告はこのようなプレイヤーの存在を想定しない。このようなプレイヤーは、戦略Aと戦略Bのどちらを取ったときにも、対戦相手と戦略が異なる場合に、そうでない場合に比べて利得が増える。これは、感情共有の喜びが、その感情の正当性を高め、対応する戦略の採用確率を高めることと、整合しない。AタイプとBタイプについては、どちらかの戦略で、対戦相手と採用戦略が一致したときに、そうでないときに比べて利得が大きくなる。

試行錯誤学習と本ダイナミクスの整合性について、補足する。本ダイナミクスには、試行錯誤学習と一見整合しないような、場合が存在する。全プレイヤーがAタイプの場合を想定する。この場合には、平衡点以外で任意のプレイヤー*i*で $\dot{x}_{iA} > 0$ が成り立つから、試行錯誤の要素がなく、戦略Aの採用確率が上昇するように一見思われる。しかし、このケースの背後には、試行錯誤学習が存在する。本ダイナミクスが表しているのは、 $x_A$ の変化の期待値であり、 $x_A$ の平均的な変化である。仮に、期待値を用いず、シミュレーションを行って本ゲームを考察することを考える。この場合には、毎ゲームに各プレイヤーが選ぶ戦略は1つだけだから、 $x_A$ は期待値通りには変化しない。変化の仕方は相手の戦略に依存するため、4通りがある。この4通りの中には、全プレイヤーがAタイプであっても、戦略Aの採用確率が減るような場合が含まれる。平均的には、試行錯誤の要素がなく戦略Aの採用確率が上昇するが、個々のゲームにおいては、試行錯誤的な変化（戦略Aの採用確率の上昇と下降）が生じる。確率的モデルを分析するにおいて、期待値を用いるのは標準的

---

異質性がある。本モデルにおいては、異質なプレイヤー達が、互いに選択を観察しあうという意味で相互作用し、自身の判断（混合戦略）を変化させる。このような、異質な主体間の相互作用は、社会動学において、確率的動学システム理論やシミュレーションを用いて考察されてきた (Durlauf & Young 2001)。本モデルは、簡素な構造を持つ社会動学モデルの1つと位置付けられる。

な手法である。この手法を採用するため、本ダイナミクスでは、試行錯誤学習と一見整合しないようなケースが生じる<sup>17</sup>。

$\alpha_i$ と $\beta_i$ を変化させることによって、2つの種類の一般的諸規則の形成過程の相違が、モデル上に表現されうる。2つの種類とは、①非難される値うちがある行動についての諸規則の形成過程と、②称賛される値うちがある行動についての諸規則の形成過程である。①の過程における各プレイヤーの $\alpha_i$ と $\beta_i$ と、②の過程におけるそれを区別することを、考える。 $\alpha_i$ と $\beta_i$ の絶対値は、他者と戦略を共有した際の喜びと、そうでない場合の喜びとの差である。この絶対値は、他者と戦略が共有されない場合の、落胆の程度を表す数値と解釈できる。落胆の程度は、①に該当する行動と②に該当する行動では異なると思われる。①の行動の代表例は正義の侵害であり、②の行動の代表例は慈恵である。スミスによれば、慈恵は無償であって、その欠如は憤慨の自然な対象とはならないが、正義の侵害は憤慨の自然な対象となる(TMS : II.ii.1.3-5)。よって、他者と戦略が共有されない場合の落胆は、憤慨に伴う、①の行動の方が大きいと思われる。

### 3.2 ダイナミクスの分析

一般的諸規則を定義する。

#### 定義 一般的諸規則

社会状態 $(0,0, \dots, 0)$ が実現しているとき、あるいは、社会状態 $(1,1, \dots, 1)$ が実現しているとき、一般的諸規則が存在している。

一般的諸規則が存在するときには、各主体が自身の戦略の正当性を確信するため、各プレイヤーが戦略 A を取る確率は、0あるいは1となる。加えて、一般的諸規則は社会で共有されるから、全プレイヤーの混合戦略が一致する。

これより、本ダイナミクスを分析し、その性質を調べる。命題1から命題3は、一般的諸規則の漸近安定性についての命題であり、命題4は内点平衡点についての命題である。漸近安定点は、その近傍から出発する軌道が、いずれその点自体に収束するような点である。漸近安定点は微小な変化に対して安定な平衡点であり、モデル外の要因による変動に対して頑強と言える。進化ゲームモデルにおいては、漸近安定点が、ダイナミクスの収束先の候補とされている。複数の漸近安定点が存在する場合には、ダイナミクスの具体的な収束先は、初期点とモデル外の変動に依存して、決定される。

---

<sup>17</sup> 試行錯誤学習の代表的進化ゲームモデルと言える、Erev&Roth(1998)のモデル(2019/1/02 時点、google scholar 上で1996回の被引用)は、シミュレーションモデルである。このモデルであっても、ダイナミクスの期待値に着目すれば、本モデルのような、試行錯誤学習と一見整合しない場合が生じうる。大浦(2008)の第4.3章を参照。

本ダイナミクスの連立微分方程式の、 $i$ 番目の方程式を $f_i(t)$ とする。この連立方程式のヤコビ行列は、 $i$ 行目の対角成分において、

$$\frac{\partial f_i(t)}{\partial x_{iA}} = \left( (\alpha_i - \beta_i) \left( \sum_{j \neq i} x_{jA} / (n-1) \right) + \beta_i \right) (1 - 2x_{iA}).$$

また、ヤコビ行列は非対角成分 $j \neq i$ において、 $\frac{\alpha_i - \beta_i}{n-1} x_{iA} (1 - x_{iA})$ 、となるから、端点平衡点においては非対角成分が0となる。ある端点平衡点が漸近安定となる必要十分条件は、ヤコビ行列が安定となることである。全ての非対角成分が0となるので、ヤコビ行列が安定となる必要十分条件は、全ての対角成分が負となることである。

**命題 1** : いずれか1つの一般的諸規則が漸近安定となるための必要十分条件は、1) あるいは2) の、いずれか1つが成り立つことである。

1) (1,1, ..., 1)が漸近安定 : プレイヤー集合が、A タイプのみで構成されるか、A タイプと同調タイプの2タイプで構成される、

2) (0,0, ..., 0)が漸近安定 : プレイヤー集合が、B タイプのみで構成されるか、B タイプと同調タイプの2タイプで構成される。

証明

社会状態(1,1, ..., 1)が漸近安定となる必要十分条件は、対応するヤコビ行列の対角成分がすべて負となることである。ヤコビ行列の対角成分は第 $i$ 行で $-\alpha_i$ となるから、全て負となるためには、各プレイヤーがAタイプあるいは同調タイプでなければならない。

社会状態(0,0, ..., 0)が漸近安定となる必要十分条件は、対応するヤコビ行列の対角成分がすべて負となることである。ヤコビ行列の対角成分は第 $i$ 行で $\beta_i$ となるから、全て負となるためには、各プレイヤーがBタイプあるいは同調タイプでなければならない。

以上から、プレイヤー集合が同調タイプのみで構成されるときには、両方の一般的諸規則が漸近安定となる。

したがって、2つの一般的諸規則の内、(1,1, ..., 1)のみが漸近安定である必要十分条件は、1) プレイヤー集合が、A タイプのみで構成されるか、A タイプと同調タイプの2タイプで構成されることである。また、2つの一般的諸規則の内、(0,0, ..., 0)のみが漸近安定である必要十分条件は、2) プレイヤー集合が、B タイプのみで構成されるか、B タイプと同調タイプの2タイプで構成されることである。

証明終わり

命題 1 より、いずれか 1 つの一般的諸規則が漸近安定となるためには、1) か 2) のいずれか 1 つが満たされなければならない。この条件が成り立つときには、一般的諸規則のうち 1 つが収束先の候補となる。なお、他の平衡点の漸近安定性については、この命題では何も述べられていない。

**命題 2 :** 命題 2 : 一般的諸規則が、各プレイヤーの  $\alpha_i$  と  $\beta_i$  の大小に関わらず漸近安定となる、唯一の点となる必要十分条件は、1) あるいは 2) の、いずれか 1 つが成り立つことである。

- 1)  $(1,1, \dots, 1)$  が漸近安定 : プレイヤー集合が、A タイプのみで構成される、
- 2)  $(0,0, \dots, 0)$  が漸近安定 : プレイヤー集合が、B タイプのみで構成される。

証明

社会状態  $(1,1, \dots, 1)$  が、 $\alpha_i$  と  $\beta_i$  の大小に関わらず漸近安定となる、唯一の点となる、必要十分条件が、1) であることを示す。社会状態  $(1,1, \dots, 1)$  が唯一の漸近安定点であるのに、B タイプか同調タイプの、プレイヤーが存在したとする。B タイプの場合、 $(1,1, \dots, 1)$  の近傍において、このプレイヤーに該当する成分で、 $\dot{x}_{iA} < 0$  が成り立つ。これは  $(1,1, \dots, 1)$  が漸近安定でないことを意味し、矛盾である。同調タイプの場合、 $\alpha_i$  と  $\beta_i$  の大小によっては、他の点が漸近安定となりうるから矛盾する。同調タイプのプレイヤーの人数を、 $E$  とする。同調タイプの各プレイヤーに対応する成分が 0 で、残りの A タイプの各プレイヤーに対応する成分が 1 となる、社会状態を考える。同調タイプのプレイヤーの 1 人を、 $j$  とする。この社会状態のヤコビ行列の、第  $j$  対角成分は  $(\alpha_j - \beta_j)(\frac{n-E}{n-1}) + \beta_j$  である。A タイプのプレイヤーの 1 人を、 $i$  とする。第  $i$  対角成分は、 $-(\alpha_i - \beta_i)(\frac{n-E-1}{n-1}) + \beta_i$  である。ヤコビ行列のすべての対角成分が負であれば、この社会状態は漸近安定となる。プレイヤー  $j$  が同調タイプであるから、第  $j$  対角成分は負となりうる。プレイヤー  $i$  が A タイプであるから、第  $i$  対角成分は常に負である。よって、同調タイプの各プレイヤーに対応する成分が 0 で、残りの成分が 1 となる社会状態は、漸近安定となりうる。これは、 $(1,1, \dots, 1)$  が、 $\alpha_i$  と  $\beta_i$  の大小に関わらず漸近安定となる、唯一の点となることと、矛盾する。

全プレイヤーが、A タイプであるとき、社会状態  $(1,1, \dots, 1)$  以外の平衡点が存在し、 $\alpha_i$  と  $\beta_i$  の大小に関わらず漸近安定だとする。この平衡点には 0 の成分が必ず存在する。漸近安定性より、この平衡点の近傍では、0 の成分に該当するプレイヤー  $i$  について  $\dot{x}_{iA} < 0$  が成り立つ。一方で、仮定より各プレイヤーが A タイプであるから、全ての平衡点の近

傍で、プレイヤー*i*について $\dot{x}_{iA} > 0$ となる。これは矛盾である。

社会状態 $(0,0, \dots, 0)$ が唯一の漸近安定点であるための必要十分条件が2)であることを示す。社会状態 $(0,0, \dots, 0)$ が唯一の漸近安定点であるのに、Aタイプか同調タイプのプレイヤーが存在したとする。Aタイプの場合、 $(0,0, \dots, 0)$ の近傍において、このプレイヤーに該当する成分で $\dot{x}_{iA} > 0$ が成り立つ。これは $(0,0, \dots, 0)$ が漸近安定でないことを意味し、矛盾である。同調タイプの場合、 $\alpha_i$ と $\beta_i$ の大小によっては、他の点が漸近安定となりうるから矛盾する。同調タイプのプレイヤーの人数を、 $E$ とする。同調タイプの各プレイヤーに対応する成分が1で、残りのBタイプの各プレイヤーに対応する成分が0となる、社会状態を考える。同調タイプのプレイヤーの1人を、 $j$ とする。この社会状態のヤコビ行列の、第*j*対角成分は $-[(\alpha_j - \beta_j)(E - 1/n - 1) + \beta_j]$ である。Bタイプのプレイヤーの1人を、 $i$ とする。第*i*対角成分は、 $(\alpha_i - \beta_i)(E/n - 1) + \beta_i$ である。ヤコビ行列のすべての対角成分が負であれば、この社会状態は漸近安定となる。プレイヤー*j*が同調タイプであるから、第*j*対角成分は負となりうる。プレイヤー*i*がBタイプであるから、第*i*対角成分は常に負である。よって、同調タイプの各プレイヤーに対応する成分が1で、残りの成分が0となる社会状態は、漸近安定となりうる。これは、 $(0,0, \dots, 0)$ が、 $\alpha_i$ と $\beta_i$ の大小に関わらず漸近安定となる、唯一の点となることと、矛盾する。

全プレイヤーが、Bタイプであるとき、社会状態 $(0,0, \dots, 0)$ 以外の平衡点が存在し漸近安定だとする。この平衡点には1の成分が必ず存在する。漸近安定性より、この平衡点の近傍では、1の成分に該当するプレイヤー*i*について $\dot{x}_{iA} > 0$ が成り立つ。一方で、仮定より全プレイヤーがBタイプであるから、全ての平衡点の近傍で、プレイヤー*i*について $\dot{x}_{iA} < 0$ となる。これは矛盾である。

証明終わり

全平衡点の中で、一般的諸規則のみが各プレイヤーの $\alpha_i$ と $\beta_i$ の大小に関わらず漸近安定であるときには、全プレイヤーが同じタイプでなければならない。このとき、一般的諸規則のうち1つのみが漸近安定である。

**命題3**：2つの一般的諸規則が同時に漸近安定となるための必要十分条件は、全プレイヤーが同調タイプであることである。

証明

命題1の証明より、全プレイヤーが同調タイプであるとき、2つの一般的諸規則は同時に漸近安定となる。

2つの一般的諸規則が同時に漸近安定であるときに、あるプレイヤー $j$ が存在し、同調タイプでないとする。社会状態 $(0,0, \dots, 0)$ に対応するヤコビ行列の、 $j$ 行目の対角成分は $\beta_j$ である。社会状態 $(1,1, \dots, 1)$ に対応するヤコビ行列の、 $j$ 行目の対角成分は $-\alpha_j$ である。これらの対角成分は、プレイヤー $j$ が同調タイプではないから、同時には負とならない。よって、一般的諸規則の片方は漸近安定ではない。したがって、同調タイプでないプレイヤーの存在を仮定するのは誤りであり、全員が同調タイプである。

証明終わり

命題3より、2つの一般的諸規則が同時に漸近安定となるための必要十分条件は、全プレイヤーが同調タイプであることである。これが成り立つ場合には、2つの一般的諸規則が収束先の候補となる。

**命題4**：内点平衡点が存在するための必要十分条件は、全プレイヤーが同調タイプであることである。また、内点平衡点は漸近安定でない。

証明

プレイヤー $i$ について、端点以外で平衡点が存在したとすれば、それは点 $\frac{\beta_i}{\beta_i - \alpha_i}$ のみである。この点が内点となる必要十分条件は、 $0 < \frac{\beta_i}{\beta_i - \alpha_i} < 1$ であるが、3タイプのプレイヤーのうち、これが成り立つのは同調タイプのみである。

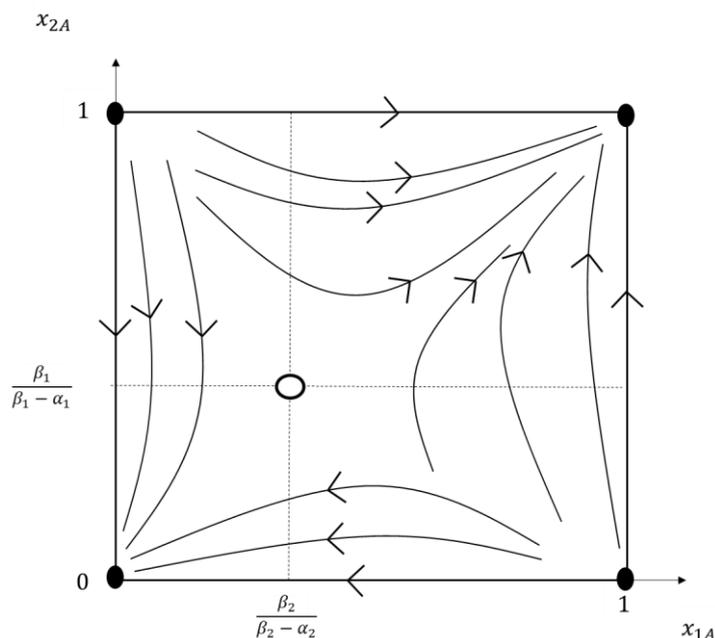
漸近安定な内点が存在しないことを、ウェイブル(1999 : 219)に基づいて証明する。連立微分方程式は一般に、それと同様の変数を持つ正関数で除しても、解軌道が影響を受けず、移動速度のみが変わる。本モデルの連立微分方程式の各式を $\prod_{j=1}^n x_{jA}(1 - x_{jA})$ で除すことにより、新たな連立微分方程式を作る。すべての内点において、任意の $j$ で $x_{jA}(1 - x_{jA}) > 0$ が成り立つから、この除算は正関数による除算である。除算後の新たな連立微分方程式は、各式で、右辺に左辺と同じ変数を含まない。よって、すべての内点において、ヤコビ行列の全対角成分は0となり、トレースは0となる。新たな連立微分方程式は連続で微分可能だから、ウェイブル(1999)の命題6.6より、すべての内点は非漸近安定である。元の連立微分方程式の解軌道は、新たな連立微分方程式のそれと同じであるから、同様に、すべての内点が非漸近安定である。

証明終わり

命題 4 より，内点平衡点は，全プレイヤーが同調タイプであれば存在するが，漸近安定ではなく，収束先の候補とはならない。社会状態が内点平衡点に一時的に留まる可能性はあるが，この点は微小な変動に対して安定ではない。内点平衡点が存在するときには，全プレイヤーが同調タイプであるから，命題 3 より，社会状態はどちらの一般的諸規則にも収束しうる。

#### 4 考察

本節では，理神論的な一般的諸規則の形成過程を，本モデル上で再解釈を試みる。次に，道德感情の腐敗論に代表される，理神論と整合しない状況の再解釈を試みる。そして，以上の 2 つが混在する状況の再解釈を試みる。



図：2人ゲームにおいて内点平衡点が存在する場合

理神論的な一般的諸規則の形成過程においては，一般的諸規則が人々の相互作用によって形成され，厳守される。この状況は，ゲームが進行した後に，社会状態が一般的諸規則に，すなわち  $(0,0, \dots, 0)$  か  $(1,1, \dots, 1)$  に，収束した状態と解釈できる。よって，少なくとも 1

つの一般的諸規則が漸近安定でなければならない。これが満たされるためには、命題1あるいは命題3で述べた条件が満たされる必要がある。命題1の条件が満たされたときには、漸近安定な一般的諸規則はどちらか1つであるが、命題3の条件が満たされたときには両方になる(図)。この点で2つの命題の条件は異なるが、一般的諸規則が漸近安定であることには変わりがない。命題1・命題3で述べた条件は、プレイヤー集合が、限定されたタイプから構成されることを要求する。これが成り立たない場合には、一般的諸規則は漸近安定にならない。しかし、これらの命題の条件が成り立っていても、一般的諸規則以外に漸近安定点が存在する可能性があり、社会状態は一般的諸規則以外に収束しうる。この可能性を排除し、一般的諸規則が確実に形成されるものだとみなすならば、命題2の条件が必要である。プレイヤー集合は、Aタイプのみで構成されるか、Bタイプのみで構成されなければならない。

スミス自身が言及した、道德感情の腐敗のような、理神論と整合しない状況はどのように解釈されるだろうか。この状況においては、一般的諸規則は当然実現していない。この状況は、大きく3通りに解釈が可能である。1つ目は、一般的諸規則が漸近安定でない場合である。この場合に、一般的諸規則は収束先の候補とならないから、社会状態が他の点に収束していると考えられる。

2つ目は、一般的諸規則が漸近安定であるが、それ以外にも漸近安定点が存在する場合である。この場合には、一般的諸規則だけでなく他の平衡点も収束先の候補となるから、一般的諸規則とは別の点に社会状態が収束しうる。

3つ目は、ダイナミクスが未だに収束していない場合である。考えられる原因は、そもそも漸近安定点が存在しないこと、ダイナミクスの進行度合いが遅いこと、内点平衡点に社会状態が止まっていること、である。ダイナミクスの進行度合いが遅ければ、たとえ一般的諸規則が漸近安定であっても、そこに社会状態が到達していない可能性がある。ダイナミクスの進行度合いは、ゲームの繰り返し回数と、各利得表の値に依存する。また、内点平衡点が存在する場合には、社会状態がこの点に到達したまま止まる可能性がある(図)。命題4より、内点平衡点は漸近安定ではないが、仮に微小な変動が起こらなければ、社会状態はこの点に止まる。

『道德感情論』には、理神論的な一般的諸規則の形成過程についての記述が存在すると同時に、理神論と整合しない状況についての記述も存在する。それぞれについての、本モデルにおける再解釈を、上に提示したわけであるが、この相反する記述の混在は、如何にして整合的に解釈可能であろうか。1つの方法は、非理神論的状況を、一般的諸規則形成の途中の一時的な状況と捉える、ことである。現状は一時的に非理神論的状況にあるが、最終的には一般的諸規則が実現されると想定する。この想定は、社会状態が既に収束している場合には成り立たない。しかし、ダイナミクスの進行度合いが遅い場合、あるいは、内点平衡点に社会状態が止まっている場合(図)には、この想定が成り立つ。本解釈が成り立つとすれば、非理神論的状況の解決シナリオは、ゲーム回数の増加か、モデル外の要

困による変動になる<sup>18</sup>。

## 5 おわりに

本報告は『道徳感情論』における一般的諸規則の形成過程を、主体達が他者を繰り返し観察することによって進む、試行錯誤学習の過程とみなし、レプリケータダイナミクスを用いて定式化した。そして、このモデルを用いて、同書における理神論的な一般的諸規則についての記述、及び、道徳感情の腐敗に代表される非理神論的な記述を、再解釈した。加えて、両者が混在する状況についての解釈を提示した。

本報告が提示した混在状況についての解釈は、道徳感情の腐敗に代表される非理神論的状况についての議論に、寄与する。道徳感情の腐敗論は、同書の社会秩序形成論に矛盾すると考えられ、『道徳感情論』第6版改訂の主要な原因であるとされてきた。しかし、本定式化によって生じた解釈からすれば、理神論的状况と非理神論的状况は矛盾しないとも、考えられる。非理神論的状况を、一般的諸規則形成の途中の、一時的な状況と捉えるならば、2つの状況は矛盾しない。この解釈は、動学による定式化を採用したゆえに生じており、本モデルの利点の1つと言える。

## 参考文献

### 英文

- Ashraf, N., Camerer, C. F., & Loewenstein, G. (2005). Adam Smith, behavioral economist. *Journal of Economic Perspectives*, 19(3), 131-145.
- Bandura, A. (1977). *Social learning theory*. Prentice Hall.
- Bréban, L. (2012). Sensitivity to prosperity and adversity: What would a Smithian function of happiness look like?. *The European Journal of the History of Economic Thought*, 19(4), 551-586.
- Corfe, R. (2007). *Deism and Social Ethics: The Role of Religion in the Third Millennium*. Arena books.
- Den Uyl, D. J. (2016). Impartial spectating and the price analogy. *Econ Journal*
- Durlauf, S. N., & Young, H. P. (2001). The new social economics. In Durlauf, S. N., &

---

<sup>18</sup> このシナリオは、『道徳感情論』の文脈において、如何ようにも解釈しうる。例えば、第6版の改訂箇所のみを考慮するとすれば、良心論の改訂（「称賛にあたいすること」等々）は、モデル外の要因に働きかけるための、スミスの策と解釈できるのではないか。

Young, H. P. (Eds.). (2001). *Social dynamics*. Mit Press, 1-14.

- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American economic review*, 88(4), 848-881.
- Evensky, J. (1998). Adam Smith's moral philosophy: The role of religion and its relationship to philosophy and ethics in the evolution of society. *History of Political Economy*, 30, 17-42.
- Harrison, P. (1990). *Religion and the Religions in the English Enlightenment*. Cambridge University Press.
- Harrison, P. (2011). Adam Smith and the history of the invisible hand. *Journal of the History of Ideas*, 72(1), 29-49.
- Hefelbower, S. G. (1920). Deism Historically Defined. *The American Journal of Theology*, 24(2), 217-223.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., ... & Henrich, N. S. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and brain sciences*, 28(6), 795-815.
- Hill, L. (2001). The Hidden Theology of Adam Smith. *European Journal of the History of Economic Thought*, 8(1), 1-29.
- Kennedy, G. (2011). The hidden Adam Smith in his alleged theology. *Journal of the History of Economic Thought*, 33(3), 385-402.
- Khalil, E. L. (2017). Socialized view of man vs. rational choice theory: What does smith's sympathy have to say?. *Journal of Economic Behavior & Organization*, 143, 223-240.
- Kristol, I. (1980). Rationalism in economics. *The Public Interest*, 61, 201-18.
- Lindgren, J.R (1973). *The Social Philosophy of Adam Smith*. The Hague: Martinus Nijhoff.
- Meardon, S. J., & Ortmann, A. (1996). SELF-COMMAND IN ADAM SMITH'S THEORY OF MORAL SENTIMENTS A GAME-THEORETIC REINTERPRETATION. *Rationality and Society*, 8(1), 57-80.
- Minowitz, P. (1993). *Profits, Priests, and Princes: Adam Smith's Emancipation of Economics from Politics and Religion*. Stanford University Press.
- Morris, G. S. (1883). *Philosophy and Christianity*, Robert Carter & Brothers.
- Oslington, P. (2012). God and the market: Adam Smith's invisible hand. *Journal of Business Ethics*, 108(4), 429-438.
- Peach, T. (2014). Adam Smith's "Optimistic Deism," the Invisible Hand of Providence, and the Unhappiness of Nations. *History of Political Economy*, 46(1),

55-83.

- Rosenthal, T. L., & Zimmerman, B. J. (2014). *Social learning and cognition*. Academic Press.
- Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1), 164-212.
- Smith, A. (1976a). *An inquiry into the nature and causes of the wealth of nations*. Ed. R.H. Campbell & A.S. Skinner, OXFORD: CLARENDON PRESS. 大内兵衛・松川訳(1959)『諸国民の富』岩波書店。
- Smith, A. (1976b). *The theory of moral sentiments*. Ed. D.D. Raphael & A.L. Macfie, OXFORD: CLARENDON PRESS. 水田洋訳(2003)『道徳感情論』岩波書店。
- Tajima, K. (2007). The theory of institutions and collective action in Adam Smith's Theory of Moral Sentiments. *The Journal of Socio-Economics*, 36(4), 578-594.
- Weibull, J. W. (1997). *Evolutionary game theory*. MIT press. 大和瀬監訳(1999)『進化ゲームの理論』オフィスカノウチ。

## 和文

- 石沢芳次郎(1990)『経済学における神』財団法人産業経済研究協会。
- 大浦宏邦(2008)『社会科学者のための進化ゲーム理論：基礎から応用まで』勁草書房。
- 篠原久(1998)「スミス道徳論における『歓喜』と『悲哀』—『道徳感情の腐敗』をめぐって—」『経済学論究』, 52(1), 85-107。
- 柴田徳太郎(2010)『『見えざる手』と『コンヴェンション』—スミスとヒュームの秩序生成論』『経済学論集』, 75(4), 2-22。
- スミス・メイナード著／寺本英・梯正之訳(1985)『進化とゲームの理論』産業図書。
- 田島慶吾(2003)『アダム・スミスの制度主義経済学』ミネルヴァ書房。
- 田中正司(1993)『アダム・スミスの自然神学』御茶の水書房。
- 田中正司(2000)『アダム・スミスと現代』御茶の水書房。
- 中谷武雄(1996)『スミス経済学の国家と財政』ナカニシヤ出版。
- 新村聡(1994)『経済学の成立—アダム・スミスと近代自然法学—』御茶の水書房。
- 山口正春(1983)「アダム・スミスと理神論」『政経研究』, 19(3), 510-533。
- 山口正春(2014)『アダム・スミスとその周辺 思想・経済・社会』三恵社。
- 村越好男(1999)「アダム・スミス自然神学と経済生活の視座」『日本の神学』, 38, 34-59。
- ラファエル・D・D 著／久保芳和訳(1986)『アダム・スミスの哲学思考』雄松堂出版。