

TERG

Discussion Paper No.391

『道德感情論』における道德感情の腐敗論の進化経済学的再
解釈

菅 隆彦

2018/05/30

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』における道徳感情の腐敗論の進化経済学的再解釈

菅 隆彦

1 はじめに

アダム・スミス著『道徳感情論』は、純粋な経済思想史研究の対象となるだけでなく、経済学の様々な研究成果を用いて再解釈され、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三部における「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), 制度派経済学の観点から再解釈するもの (Tajima, 2008), 行動経済学的なモデルとの差異を明らかにしつつ、効用関数を用いて定式化するもの (Bréban, 2012), 等が存在する。菅(2018)は、上記の先行研究に続いて、進化経済学のモデルである、レプリケータダイナミクスを用いて、同書における、「良俗の一般的諸規則(general rules of morality)」(以下、一般的諸規則)の形成を再解釈した。

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。

本稿は、『道徳感情論』第一部における、所謂「道徳感情の腐敗論」を、レプリケータダイナミクスを用いて定式化し再解釈する。スミスは、第六版において道徳感情の腐敗論を展開した。本稿は菅(2018)のモデルを応用して、道徳感情の腐敗論における、中立的な観察者 (impartial spectator) の見方が形成される過程を定式化する。中立的な観察者と一般的諸規則は、形成過程がほとんど同様であるため、後者についてのモデルを前者に対して応用することが可能である。本稿においては、菅(2018)のモデルとは異なり、複数種類の利得表の存在が仮定される。

本モデルの含意として、『道徳感情論』における基礎概念である、同感概念の孕む問題が指摘される。同感とは、けして社会秩序形成と相性の良い概念とは言えない。

本稿は以下のように構成される。第2節において、道徳感情の腐敗論について、中立的な観察者及び良俗の一般的諸規則との関係性に言及しつつ述べる。第3節では、中立的な観察者の形成過程を、レプリケータダイナミクスを用いて定式化する。第4節で定式化の含意を述べ、最終節で本稿を総括する。

2 道徳感情の腐敗論

2-1 道徳感情の腐敗論

所謂、道徳感情の腐敗論は、『道徳感情論』の第六版において付け加えられた。田中(2000,123), 田島(2003, 265)によれば、スミスの、第六版改訂の最大の動機は、道徳感情の腐敗論である。

スミスによれば、「富裕な人びと、有力な人びとに感嘆し、ほとんど崇拜し、そして、貧乏でいやしい状態にある人びと、軽蔑し、すくなくとも無視するという、この性向は、……道徳感情の腐敗の、大きな、そしてもっとも普遍的な、原因である¹⁾」。

スミスは、徳の道 (the road to virtue) と財産の道 (that [the road] to fortune) とい

¹ 『道徳感情論』 I .iii.3.1.

う、2つの異なる倫理基準を区別する。徳の道に進むことは、「英知の研究と徳の実行」によって尊敬を得ようとすることを意味し、財産の道に進むことは「富と地位の獲得」によって尊敬を得ようとすることを意味する²。財産の道は、しばしば徳の道とは整合しない、悪徳と愚行を伴う。

両者は明確に異なる倫理基準であるが、混同されてしまうことが多い³。中流及び下流の生活階級においては、両者はほとんど一致する一方で、上位の生活階級においては、両者は全く異なったものになる⁴。上位の生活階級においては、「成功と昇進は、理解力があり豊富な知識をもった同等者たちの評価にではなく、無知高慢で誇りの高い上位者たちの、気まぐれでばかげた好意に依存する⁵」。

人々は上位の生活階級に憧れを持ち、上位者の言動を模倣する。このため、上位者における道徳感情の腐敗は、中流及び下流においても進行する⁶。上位の生活階級は服装のみならず、悪徳もまた流行させてしまう。人々は上位の生活階級に上り詰めるために、しばしば徳の道を軽んじ財産の道へ走る。野心的な人間は、一度上位階級に上り詰めてしまえば、それに至る以前の悪徳は覆い隠されるだろうと考える。法律の上に立つ地位にたどり着けば、過去の悪徳は障害にはならない。しかし、その試みは失敗に終わることが多い。また、仮に成功したとしても、他の人から見れば彼の榮譽は汚れたものであり、彼自身からしても内心では自分の悪徳行為は恥ずべきものであって、彼を苦しめ続ける。

2—2 道徳感情の腐敗論と、中立的な観察者の複数性

柴田(2010,13-14)が主張するように、第六版で道徳感情の腐敗論が追加されたことによって、『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考えることができ」る。それは、道徳感情の腐敗論が、中立的な観察者の見方が複数存在することを、意味するからである。

中立的な観察者とは、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。中立的な観察者は、主体が言動を取る際の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。『道徳感情論』においては、各主体が中立的な観察者に是認されるような言動をとることにより、社会秩序が形成される。

中立的な観察者の「影響と権威は、あらゆるばあいにかわめて大きい。そして、この内部の裁判官に相談することによってのみ、われわれは、自分自身に関連するどんなものごとでも、その本来の形と大きさにおいて、みることができ⁷」る。

スミスが述べるような社会秩序が成り立つとすれば、中立的な観察者の見方が複数存在し、かつ互いに整合しない、ということはある得ない。『道徳感情論』における社会秩序は、厳密に守られるべき徳である、正義の徳によって成り立っている⁸。

² 『道徳感情論』 I .iii.3.2。

³ 『道徳感情論』 I .iii.3.3。

⁴ 『道徳感情論』 I .iii.3.5 及び 6。

⁵ 『道徳感情論』 I .iii.3.6。

⁶ 『道徳感情論』 I .iii.3.7。

⁷ 『道徳感情論』 III.3.1。

⁸ 詳細については、菅(2018)の第2節を参照。

しかし、道德感情の腐敗論からすると、社会において複数の倫理基準が存在しており、中立的な観察者の見方が複数種類形成されている。よって、道德感情の腐敗の過程を論じるにあたっては、中立的な観察者の形成過程を考慮する必要がある。

2—3 中立的な観察者の形成

中立的な観察者は、主体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の「鏡」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⁹。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、主体は他者を注意深く観察するようになる。

スミスは、中立的な観察者の形成過程を、美醜についての観念の形成に例える¹⁰。美醜についての観念もまた、中立的な観察者と同様に、他者の判断を観察することによって形成される。他者から美しいと称賛されることによって主体は喜び、逆に非難されることで主体は失望する。主体は自身が他者からどのように見られているのかを、注意深く観察するようになる。

他者を観察して彼らの判断を考察することによって、主体は、自身の言動が他者からどのように判断されるのか、心中で想像できるようになる。これは、「われわれが、ある程度他人の目をもって、われわれ自身の行動の適宜性を熟視することができる、唯一の鏡である¹¹」。この鏡が中立的な観察者に他ならない。

2—4 一般的諸規則との類似性

中立的な観察者の形成過程は、スミス自身は本文中で直接に言及していないものの、良俗の一般的諸規則（general rules of morality, 以下、一般的諸規則）の形成過程とほとんど同一である¹²。一般的諸規則は、中立的な観察者が是認する行動をとるべきとし、中立的な観察者が否認する行動をとるべきとしない、規則である。一般的諸規則と、中立的な観察者の判断は一致する。

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。①は中立的な観察者が否認する行動に対応し、②は是認する行動に対応する。

両者の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、両者は異なる。一般的諸規則は、中立的な観察者が主体の行動を制御不可能となる、自己欺瞞という問題に対処するための規則である¹³。

しかし、形成過程については、両者はほとんど同一である。両者とも、他者の言動を観察することにより形成される。本稿は、①の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について述べる。②の形成過程は①の形成過程とほとんど同じであるため、

⁹ 『道德感情論』Ⅲ.1.3。

¹⁰ 『道德感情論』Ⅲ.1.4。

¹¹ 『道德感情論』Ⅲ.1.5。

¹² 一般的諸規則の詳細については、菅(2018)を参照。以下に述べる一般的諸規則についての説明は、菅から抜粋し編集したものである。

¹³ 『道德感情論』Ⅲ.4.7。

②の形成過程を省略する。

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。他者の行動を繰り返し観察している中で、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¹⁴。この経験を繰り返すことで、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¹⁵」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¹⁶」。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、数回の観察では不十分である。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、ひと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て、同様の規則を形成する。

2—5 一般的諸規則との相違点

以上に述べた一般的諸規則の形成過程は、中立的な観察者の形成過程と同様に、他者が何を是認しまた否認するのかという、他者の道徳的判断を観察することによって進行する。前述したように、実際に主体の言動を制御可能か否かという点で、両者は異なるが、倫理基準を確立する過程は同一である。

一般的諸規則の形成についての記述においては、継続的观察という表現が見られ、観察の継続性が強調されている。対して、中立的な観察者の形成についての記述においては、そのような強調は見られない。この点もまた両者の相違点である。

しかし、継続の長さに違いがあったとしても、両者とも、他者の道徳的判断を観察することによって、形成されることには変わりがない。スミス自身は言及していないものの、一般的諸規則は、中立的な観察者の形成過程が継続性に繰り返されることによって形成される、発展的な概念であるとみなすことも可能であろう。

3 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

¹⁴ スミス自身は言及していないものの、この感情の共有もまた同感の一つとみなされる。

¹⁵ 『道徳感情論』Ⅲ.4.7。

¹⁶ 『道徳感情論』Ⅲ.4.8。

なされる。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、他者が自身と感情を共有すると、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共有し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。この試行錯誤学習の過程を、レプリケータダイナミクス (replicator dynamics) を用いて定式化することが可能である。菅(2018)は、一般的諸規則の形成過程を、レプリケータダイナミクスを用いて定式化した。

第2節で述べたように、一般的諸規則と中立的な観察者の形成過程はほとんど同一であるから、菅のモデルによって後者の形成過程を定式化することもまた可能である。

レプリケータダイナミクスは、進化ゲーム理論モデルの一つである¹⁷。このモデルは、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したうえで、ゲームの繰り返しに伴う集団の状態の変化を分析する。本稿は、集団を個人、集団の状態を個人の混合戦略とみなした場合の、レプリケータダイナミクスを定式化に用いる。

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。戦略の集合を $S_i = \{A, B\}$ とし、 $x_i = (x_{iA}, x_{iB})$ をプレイヤー $i \in I$ の混合戦略とする。 $x = (x_1, x_2, \dots, x_n)$ を集団の混合戦略プロファイルとし、 θ を混合戦略プロファイルの集合とする。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。

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。

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、中立的な観察者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。

各プレイヤーは利得最大化を目的として選択を行う。『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。同書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¹⁸。一般的諸規則の形成における他者との感情共有の確認もまた、その他者が直接主体に同感するわけではないとしても、他者が自身と感情を共有することには違いない。よって、同様の喜びを主体にもたらすはずである。主体の利得関数が、他者との同感による喜びの程度を表すとすれば、主体が利得最大化を目的とする、と仮定

¹⁷ レプリケータダイナミクス等の、進化ゲーム理論の解説書として、スミス(1985)、ウェイブル(1999)、大浦(2008)が存在する。

¹⁸ 「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、もっともつくやしがらせる、もっとも不快なものとしたのである」『道徳感情論』III.2.6。

することは不自然ではない。

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 $s = (s_1, \dots, s_n) \in S$ が使われる確率を、 $x(s) = \prod_{i=1}^n x_{is_i} \in [0,1]$ とする。

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

$$u_i(x) = x(s)\pi_i(s).$$

関数 u_i は混合戦略 $x \in \Theta$ がプレイされる時の期待利得である。プレイヤー*i*の純粋戦略 k を e_i^k とする。プレイヤー*i*が e_i^k をとる際の、プレイヤー*i*の期待利得を、 $u_i(e_i^k, x_{-i})$ と表す。

x の挙動を知るには、 $x_A = (x_{1A}, \dots, x_{nA})$ 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

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

とする。

$u_i(e_i^A, x_{-i})$ 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。 $u_i(x)$ は定義より、各純粋戦略をとった場合の期待値、 $x_{iA} \cdot u_i(e_i^A, x_{-i}) + x_{iB} \cdot u_i(e_i^B, x_{-i})$ 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、同感の喜びが大きい戦略ほど採用確率の上昇が大きくなる。

x_i の挙動は、各プレイヤーが形成する、中立的な観察者の判断基準の挙動を表す。ゲームが繰り返されたときに、各人の中立的な観察者の判断がどのような点に収束するのかを分析するのが、本モデルの目的である。 $x = (x_1, x_2, \dots, x_n)$ は、全員の成分が一致する点に収束するかもしれないし、各人でまったく成分が異なる点に収束するかもしれない。前者の場合には、全員共通の中立的な観察者が形成される点に x が収束すると言える。あるいは、 x はどこにも収束しないかもしれない。この場合には、中立的な観察者の基準が各期で変動する。

プレイヤー*i*の利得は表 1 のように決定されるとする。

表 1 : プレイヤー*i*の利得

$i /$ 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

3-2 2プレイヤーゲーム

最も単純な場合である、 $n = 2$ としたときの混合戦略プロファイル、 $x = (x_1, x_2)$ の挙動を考察する。 x の挙動を知るには、 $x_A = (x_{1A}, x_{2A})$ の挙動を調べれば十分である。

ダイナミクスは以下のようなになる¹⁹。

$$\begin{aligned}\dot{x}_{1A} &= \{(a_{AA}^1 - a_{BA}^1 - a_{AB}^1 + a_{BB}^1)x_{2A} + a_{AB}^1 - a_{BB}^1\}x_{1A}(1 - x_{1A}), \\ \dot{x}_{2A} &= \{(a_{AA}^2 - a_{BA}^2 - a_{AB}^2 + a_{BB}^2)x_{1A} + a_{AB}^2 - a_{BB}^2\}x_{2A}(1 - x_{2A}).\end{aligned}$$

ここで、 $\alpha_i \equiv a_{AA}^i - a_{BA}^i$, $\beta_i \equiv a_{AB}^i - a_{BB}^i$, とすれば、

$$\begin{aligned}\dot{x}_{1A} &= \{(\alpha_1 - \beta_1)x_{2A} + \beta_1\}x_{1A}(1 - x_{1A}), \\ \dot{x}_{2A} &= \{(\alpha_2 - \beta_2)x_{1A} + \beta_2\}x_{2A}(1 - x_{2A}).\end{aligned}$$

全ての端点が平衡点である。 x_{1A} の変化の方向は、 α_1 と β_1 に依存して決まる。 x_{2A} の変化についても、 α_2 と β_2 に依存して決まる。点 $(\frac{\beta_2}{\beta_2 - \alpha_2}, \frac{\beta_1}{\beta_1 - \alpha_1})$ が定義域に含まれる場合には、この点において \dot{x}_{1A} と \dot{x}_{2A} が同時に0となり平衡点となる。

3—3 利得表の場合分け

菅(2018)では、一般的諸規則についての記述を根拠として、仮定1を課した²⁰。

仮定1

プレイヤー*i*について、

$$a_{AB}^i < a_{AA}^i, \quad a_{BA}^i < a_{AA}^i, \quad a_{AB}^i < a_{BB}^i, \quad a_{BA}^i < a_{BB}^i.$$

仮定1は、マッチングされた2プレイヤーの戦略が一致した場合に、一致しない場合に比べて、利得が大きくなることを意味する²¹。対戦相手と戦略が一致するということは、相手と感情を共有することを意味しており、共有しない場合に比べて同感の喜びが高くなる。この仮定は、感情の共有が形成される一般的諸規則についての記述に対応する。

仮定1を満たすプレイヤー*i*は、 $\frac{\beta_i}{\beta_i - \alpha_i} \in (0, 1)$ となる。2人ゲームの場合には、相手の戦略Aをとる確率がこの値を超える際に、 x_{iA} が増加する。この値を下回る際には、 x_{iA} が減少する。

しかし、道徳感情の腐敗についての記述を考慮すると、必ずしも仮定1が成立するとはみなされない。ここで、戦略Aを徳の道に該当する戦略とし、戦略Bを財産の道に該当する戦略であると定義する。戦略Aを徳戦略、戦略Bを財戦略と呼ぶ。

¹⁹ $u_i(e^A, x_{-i}) = a_{AA}^i x_{2A} + a_{AB}^i (1 - x_{2A})$, $u_i(e^B, x_{-i}) = a_{BA}^i x_{2A} + a_{BB}^i (1 - x_{2A})$, を以下の方程式に代入すると導出される。 $\dot{x}_{iA} = [u_i(e^A, x_{-i}) - u_i(x)]x_{iA} = (1 - x_{iA}) \cdot u_i(e^A, x_{-i}) - x_{iB} \cdot u_i(e^B, x_{-i})$ 。

²⁰ 菅(2018)では利得が各期で異なると仮定したが、本稿ではそのようには仮定しないため、仮定1から*t*を省略する。

²¹ これは、調整ゲームの利得と共通の利得構造である。

道徳感情の腐敗論においては、財産の道に進む主体は、徳の道を軽視する。この記述を利得表上で解釈するなら、 a_{AA}^i 及び a_{AB}^i が、 a_{BA}^i 及び a_{BB}^i に比べて相対的に小さくなっている状況と解釈できる。この状況においては、主体が徳戦略を取った際の利得が、財戦略を取った場合に比べて、相対的に小さくなる。

道徳感情の腐敗論における、徳の道の軽視を仮定2によって表現する。

仮定2

$$\text{任意の } i \text{ で, } a_{BA}^i > a_{AA}^i, \quad a_{AB}^i < a_{BB}^i.$$

仮定2を満たすプレイヤー*i*は、対戦相手の戦略がいずれの場合にも、財戦略のもたらす利得が、徳戦略のもたらす利得より大きくなる。仮定2を満たすプレイヤー*i*は、戦略の変化が相手プレイヤーの戦略に依存しない。プレイヤー*i*が仮定2を満たす場合、定義域内では、常に $\dot{x}_{iA} < 0$ が成り立つ。

『道徳感情論』においては、財産の道へ進む主体のような、徳の道と整合しない人間だけの存在が仮定されているわけではない。例えば、欺瞞論について述べる箇所であっても、有徳な人物の存在についての記述が見られる²²。道徳感情の腐敗論においても、徳の道に進む主体の存在が完全に否定されているわけではないだろう。

このようなプレイヤーが満たす条件を、仮定3とする。

仮定3

$$\text{任意の } i \text{ で, } a_{BA}^i < a_{AA}^i, \quad a_{AB}^i > a_{BB}^i.$$

仮定3を満たすプレイヤー*i*は、対戦相手の戦略に関係なく、徳戦略のもたらす利得が、財戦略のもたらす利得より大きくなる。仮定3を満たすプレイヤー*i*は、戦略の変化が相手プレイヤー*i*の戦略に依存しない。定義域内では、常に $\dot{x}_{iA} > 0$ が成り立つ。

本稿は、仮定1を満たすプレイヤーを「同感タイプ」のプレイヤーであると定義する。このタイプのプレイヤーは、徳戦略（戦略A）に関して、他のプレイヤーがとる確率（平均）が基準値（ $\frac{\beta_i}{\beta_i - \alpha_i}$ ）より低い場合には、自身の確率も下がり、他のプレイヤーのとる確率が基準値より高い場合には、自身の確率も上がる。

仮定2を満たすプレイヤーを「財タイプ」のプレイヤーであると定義する。

仮定3を満たすプレイヤーを「徳タイプ」のプレイヤーであると定義する。

以上のように、本稿では3種類のプレイヤーが存在する。財タイプのプレイヤーの存在に、道徳感情の腐敗論が反映されている。道徳感情の腐敗の存在によって、中立的な観察者の形成過程（*x*の挙動）は、どのような影響を受けるのだろうか。

次節において、これら3タイプのプレイヤー間の対戦を分析する。

3—4 各タイプ間の対戦

3—4—1 2プレイヤーゲーム

最も単純なケースである、2プレイヤーゲームを考察する。各タイプのプレイヤー間の

²² 『道徳感情論』VI.3.23以降。

対戦は、6通りが考えられる。なお、全ての場合において端点が平衡点である。

場合1-1：プレイヤー1：同感 プレイヤー2：同感

同感タイプの2プレイヤーが対戦した場合は、図1-1の位相図が成り立つ。これは、菅(2018)で分析されたケースである。(0,0)と(1,1)の、2点が漸近安定となる。社会状態は、2点のいずれかに収束する。

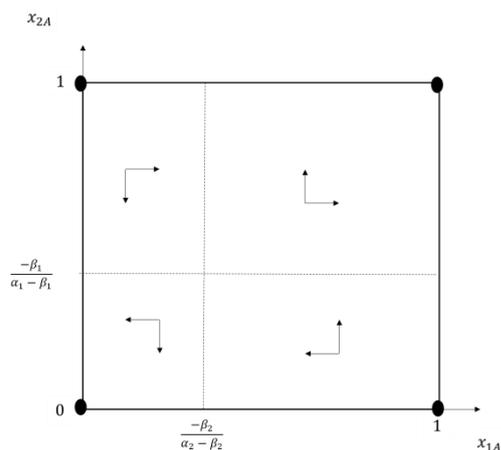


図1-1

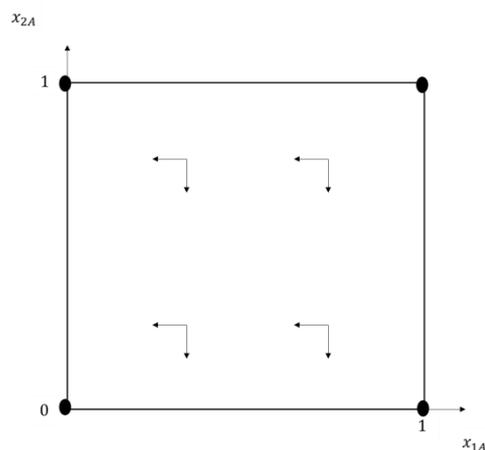


図1-2

場合1-2 (1-3)：プレイヤー1：財(徳) プレイヤー2：財(徳)

両プレイヤーが財タイプの場合には、(0,0)が唯一の漸近安定点となる。この点において連立微分方程式を線形近似すると、

$$\begin{pmatrix} \dot{x}_{1A} \\ \dot{x}_{2A} \end{pmatrix} = \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} \begin{pmatrix} x_{1A} \\ x_{2A} \end{pmatrix}。$$

各時点 t において、ヤコビ行列の固有値は β_1 と β_2 である。仮定2より、 β_1 と β_2 は常に負であり固有値が負となるから、ヤコビ行列は安定である。よって、平衡点(0,0)は漸近安定と判定される。

両プレイヤーが徳タイプの場合には、(1,1)が唯一の漸近安定点となる。

この点において連立微分方程式を線形近似すると、

$$\begin{pmatrix} \dot{x}_{1A} \\ \dot{x}_{2A} \end{pmatrix} = \begin{pmatrix} -\alpha_1 & 0 \\ 0 & -\alpha_2 \end{pmatrix} \begin{pmatrix} x_{1A} \\ x_{2A} \end{pmatrix}。$$

仮定3より、 α_1 と α_2 は常に正である。前の場合と同様の論理から、平衡点(1,1)は漸近安定と判定される。

場合1-4：プレイヤー1：同感 プレイヤー2：財

(0,0)が唯一の漸近安定点となる。この点において連立微分方程式を線形近似すると、

$$\begin{pmatrix} \dot{x}_{1A} \\ \dot{x}_{2A} \end{pmatrix} = \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} \begin{pmatrix} x_{1A} \\ x_{2A} \end{pmatrix}，$$

となる。仮定1及び仮定2より、前の場合と同様の論理から、平衡点(0,0)は漸近安定と判定される。

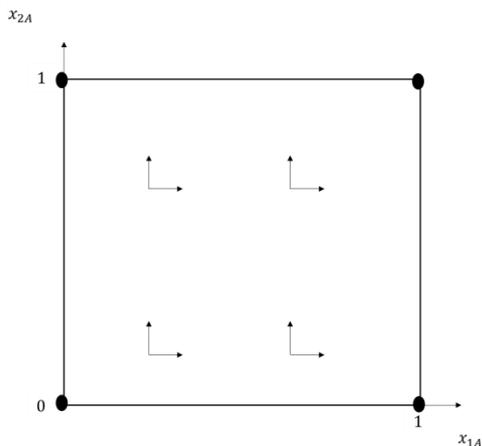


図 1 - 3

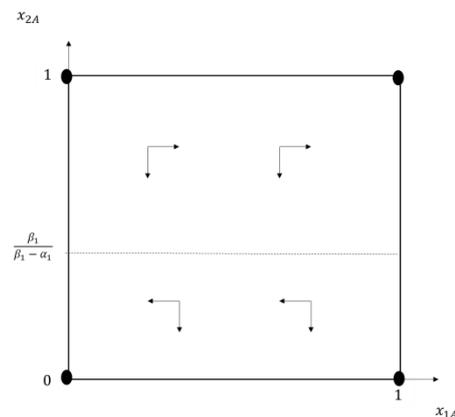


図 1 - 4

場合 1 - 5 : プレイヤー 1 : 同感 プレイヤー 2 : 徳

(1,1)が唯一の漸近安定点となる。この点において連立微分方程式を線形近似すると、

$$\begin{pmatrix} \dot{x}_{1A} \\ \dot{x}_{2A} \end{pmatrix} = \begin{pmatrix} -\alpha_1 & 0 \\ 0 & -\alpha_2 \end{pmatrix} \begin{pmatrix} x_{1A} \\ x_{2A} \end{pmatrix},$$

となる。仮定 1 及び仮定 3 より、 α_1 と α_2 は常に正となる。前の場合と同様の論理から、平衡点(1,1)は漸近安定と判定される。

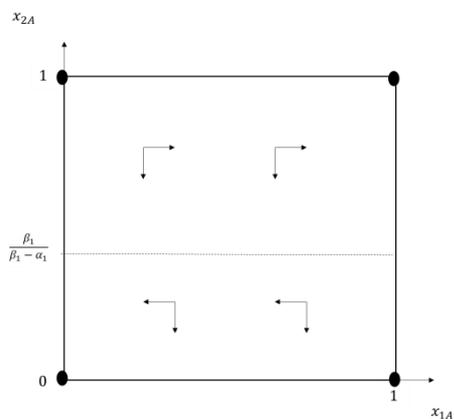


図 1 - 5

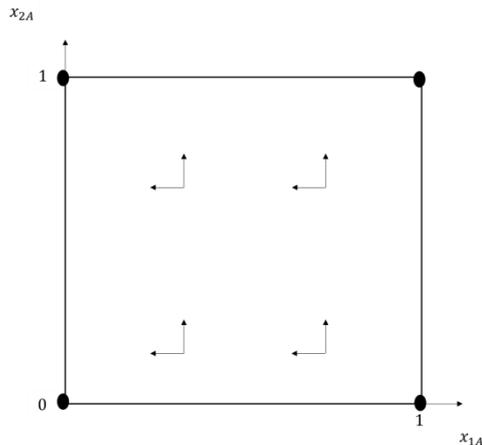


図 1 - 6

場合 1 - 6 : プレイヤー 1 : 財 プレイヤー 2 : 徳

(0,1)が唯一の漸近安定点となる。この点において連立微分方程式を線形近似すると、

$$\begin{pmatrix} \dot{x}_{1A} \\ \dot{x}_{2A} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ 0 & -\beta_2 \end{pmatrix} \begin{pmatrix} x_{1A} \\ x_{2A} \end{pmatrix}.$$

仮定 2 及び仮定 3 より、 α_1 は常に負、 β_2 は常に正である。前の場合と同様の論理から、平衡点(0,1)は漸近安定と判定される。

3—4—2 n プレイヤーゲーム

人数を一般化した n プレイヤーゲームを考察する。3タイプの組み合わせは7つに場合わけされ、以下に結果が示される。命題1から命題3の証明は、本稿末の数学補注に記載してある。

n 人ゲームにおける連立微分方程式は、以下の通りである²³。

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

なお、 θ の内部にリューヴィルの公式を適用可能なことから、 $x \in \theta$ の内点は漸近安定とはならないことが、ウェイブル(1999)によって証明されている²⁴。よって、漸近安定点の候補となる平衡点は端点のみである。

場合2—1：全員が同感タイプ

菅(2018)の命題1より、社会状態 $(0,0, \dots, 0)$ 及び社会状態 $(1,1, \dots, 1)$ の2点が漸近安定となる。他の点の漸近安定性は、必ずしも漸近安定ではない²⁵。

場合2—2・2—3：全員が財タイプ・全員が徳タイプ

全員が財タイプの場合2—2では、直感的に明らかのように、社会状態 $(0,0, \dots, 0)$ が唯一の漸近安定点となる。各プレイヤー i の x_{iA} は、平衡点を除いて、他プレイヤーの戦略に依存せずに常に減少する。

全員が徳タイプの場合2—3では、やはり直感的に明らかのように、社会状態 $(1,1, \dots, 1)$ が唯一の漸近安定点となる。各プレイヤー i の x_{iA} は、平衡点を除いて、他プレイヤーの戦略に依存せずに常に増加する。

²³ $u_i(e^A, x_{-i}) = a_{AA}^i (\sum_{j \neq i} x_{jA} / n - 1) + a_{AB}^i (1 - (\sum_{j \neq i} x_{jA} / n - 1))$,

$u_i(e^B, x_{-i}) = a_{BA}^i (\sum_{j \neq i} x_{jA} / n - 1) + a_{BB}^i (1 - (\sum_{j \neq i} x_{jA} / n - 1))$ から、 \dot{x}_{iA} が導出される。

²⁴ ウェイブル(1999)の219頁及び、第6章6節を参照。

²⁵ この結果は、後に述べる命題3の証明より得られる。 $p = 0$, $q = 0$ を仮定して、命題3の証明を読み替えると、他点が必ずしも漸近安定とならないことが導かれる。

場合 2-4 : 同感タイプと財タイプ

財タイプのプレイヤーがいるので、社会状態 $(1, \dots, 1)$ は明らかに漸近安定ではない。

命題 1 : プレイヤーには財タイプと同感タイプがいるとする。社会状態 $(0, 0, \dots, 0)$ は漸近安定である。それ以外の平衡点は必ずしも漸近安定ではない。

場合 2-5 : 同感タイプと徳タイプ

徳タイプのプレイヤーがいるので、社会状態 $(0, \dots, 0)$ は明らかに漸近安定ではない。

命題 2 : プレイヤーには徳タイプと同感タイプがいるとする。社会状態 $(1, 1, \dots, 1)$ は漸近安定である。それ以外の平衡点は、必ずしも漸近安定ではない。

場合 2-6 : 財タイプと徳タイプ

一般性を失わずに、プレイヤー1から p までが、財タイプであるとする。第1から、第 p 成分までが0、それ以降の成分が1となる社会状態が唯一の漸近安定点である。

これも直感的に明らかであろう。財タイプの、プレイヤー i の x_{iA} は他プレイヤーの戦略に依存せずに、常に減少する。また徳タイプの、プレイヤー i の x_{iA} は他プレイヤーの戦略に依存せずに、常に増加する。

場合 2-7 : 財タイプと徳タイプと同感タイプが混在する場合

命題 3 : プレイヤーには財タイプと徳タイプと同感タイプがいるとする。必ず漸近安定となる社会状態は存在しない。また、漸近安定な社会状態は必ずしも存在しない。

全てのタイプが混在する、最も一般的といえる状態においては、必ず漸近安定となる社会状態は存在しない。また、漸近安定な社会状態が存在しない場合もある。数学補注に具体例を示してある。

4 含意

4-1 場合 1-1・2-1

場合 1-1 は、同感タイプのプレイヤーだけが存在する場合には、財戦略の平衡点 $(0, 0)$ か徳戦略の平衡点 $(1, 1)$ の、いずれにも社会状態が収束しうることを表している。場合 2-1 はこの一般化であり、 $(0, \dots, 0)$ と $(1, \dots, 1)$ 以外は必ずしも安定とならないことを示す。

同感タイプの戦略は、対戦相手の戦略に依存して変化する。全員が同感タイプなのだから、社会状態は2つの正反対の点に収束しうる。1つ目の社会状態 $(0, \dots, 0)$ は全員が財戦略を取る点であり、反対に、もう1つの社会状態 $(1, \dots, 1)$ は、全員が徳戦略を取る点である。中立的な観察者は、この2点において全員一致となる。

4-2 場合1-4・2-4

場合1-4においては、同感タイプと財タイプの組み合わせでは、財戦略の平衡点(0,0)が漸近安定となることが示された。同感タイプどうしの場合1-1では、平衡点(0,0)と(1,1)の2つが漸近安定となるが、この場合には、平衡点(0,0)のみが漸近安定となる。同感タイプの戦略は、財タイプに影響されて財戦略に収束する。

場合2-4では、このことがより一般的に示されている。命題1より、財タイプのプレイヤーが1人でもいれば、(0, ..., 0)以外の平衡点は必ずしも漸近安定ではない。同感タイプが1人でも徳戦略をとれば、その平衡点は漸近安定にはならない。同感タイプと財タイプの組み合わせにおいては、財タイプの存在の影響が大きいと言える。この組み合わせにおいては、1人でも財タイプのプレイヤーが存在すれば、全員が財戦略をとる点が必ず漸近安定になる。道徳感情の腐敗が存在し、残りが同感タイプであるとき、(0, 0, ..., 0)は必ず漸近安定となる。この点においては、各人の中立的な観察者が全員一致となり、財戦略をとるべきであると判定する。

4-3 場合1-5・2-5

場合1-5においては、場合1-4における財タイプが、徳タイプに変更されている。その結果は、財タイプを徳タイプに置き換えれば、場合1-4と同様である。場合2-5では、このことがより一般的に示されている。命題2より、徳タイプのプレイヤーが1人でもいれば、(1, ..., 1)以外の平衡点は、必ずしも漸近安定ではない。同感タイプが1人でも財戦略をとれば、その平衡点は必ずしも漸近安定ではない。

徳タイプが1人でも存在すれば、(1, ..., 1)が唯一の漸近安定点となる。同感タイプと徳タイプの組み合わせにおいては、徳タイプの存在の影響が大きいと言える。同感タイプと徳タイプの組み合わせにおいては、1人でも徳タイプのプレイヤーが存在すれば、全員が徳戦略をとる点が必ず漸近安定になる。収束先においては、中立的な観察者が全員一致となり、徳戦略をとるべきであると判定する。

4-4 場合2-7

場合2-7は、全タイプが存在するという意味で、最も一般的な場合である。場合2-7においては、必ず漸近安定となる点は存在しない。また、各プレイヤーの利得が特定の条件を満たす場合には、漸近安定点が存在しないことがある。

どの社会状態も一般的には漸近安定と限らないのは、同感タイプのプレイヤーが、道徳感情の腐敗により財タイプから影響を受け、かつ徳タイプからも影響を受けるからである。

場合2-4・場合2-5においては、同感タイプが片方のタイプのみから影響を受ける状況であったので、必ず漸近安定点が存在した。

しかし、場合2-7においては、財・徳の2タイプのプレイヤーが存在するので、同感タイプは両者から影響を受ける²⁶。そのため、場合2-4等とは異なり、必ず漸近安定と

²⁶ 同感タイプのプレイヤーが、財・徳の両タイプのプレイヤーの両者から影響を受けるという点は、命題3の証明における $\sum_{j \neq i} x_{jA}$ に反映されている。財・徳の両タイプが存在するために、 $\sum_{j \neq i} x_{jA}$ が δ 、必ずしも漸近安定点を生じさせないような値となる。

なる点は存在しない。

したがって、ゲームが繰り返された後に、各人の中立的な観察者の判断がどのような判断をとるのかは不確かである。また、 $(0, \dots, 0)$ と $(1, \dots, 1)$ のいずれもが漸近安定とはならず、収束先において、各人の中立的な観察者の判断が一致することはない。

4—5 同感と社会秩序

以上から明らかになったのは、『道徳感情論』における基本概念である、同感の持つ性質である。同感タイプのプレイヤーは、他のプレイヤーに依存して、財戦略をとりうるのと同時に、徳戦略をとりうる。

場合2—4においては、道徳感情の腐敗に影響されて、同感タイプの全員が財戦略をとる点が、必ず漸近安定となる。対照的に、場合2—5においては、徳タイプの存在に影響されて、同感タイプの全員が徳戦略をとる点が、必ず漸近安定となる。場合2—7においては、同感タイプが財・徳両タイプのプレイヤーに影響され、必ず漸近安定となる点は存在しない。場合2—1においては、同感タイプの全員が財戦略をとる点と、同感タイプの全員が徳戦略をとる点の、2点が漸近安定となる。

スミスは『道徳感情論』において、同感という概念を基礎とする、社会秩序の形成を論じた。しかし、同感という概念は本稿で見てきたように、けして社会秩序形成と相性の良い概念とは言えない。徳戦略のような向社会的と言える言動が同感されうるのと同時に、財戦略のような反社会的と言える言動もまた、同感されうる。このため、各人の形成する中立的な観察者は様々な状態に収束しうる。よって、中立的な観察者の働きによって形成される社会秩序の内容もまた様々となる。

5 おわりに

本稿は、『道徳感情論』第一部における、所謂、道徳感情の腐敗論を、レプリケータダイナミクスを用いて再解釈した。

本稿で考察対象とした、道徳感情の腐敗論と同様に、中立的な観察者が複数種類存在する状況が、『道徳感情論』中に現れる。それは、同書の第六部における、適宜性の第一基準と第二基準の区別である²⁷。適宜性が複数種類存在するということは、中立的な観察者が複数種類存在することを意味する。よって、本稿の定式化を用いて、適宜性の第一基準と第二基準の区別を、考察する余地があると言える。

数学補注

n プレイヤーゲームにおける、連立微分方程式のヤコビ行列は、 i 行目の対角成分において、

²⁷ 『道徳感情論』VI.3.23以降。

対角成分の正負は必ずしも負にはならない。

以上 i, ii より, $(0,0,\dots,0)$ 以外の平衡点は必ずしも漸近安定ではない。

証明終わり

命題 3

命題 3 : プレイヤーには財タイプと徳タイプと同感タイプがいるとする。必ず漸近安定となる社会状態は存在しない。また, 漸近安定な社会状態は必ずしも存在しない。

証明

一般性を失わずに, プレイヤー 1 から p 人が財タイプ, プレイヤー $p+1$ 以降の q 人が徳タイプ, 残りが同感タイプであるとする。

i 財タイプの p 人に該当する成分

第 p 成分までに 1 が含まれる平衡点は, 財タイプのプレイヤー i については, 仮定 2 より, 他のプレイヤーの戦略に依存せず, 平衡点を除いて常に $\dot{x}_{iA} < 0$ が成り立つから, 明らかに漸近安定ではない。

ii 徳タイプの q 人に該当する成分

また, 第 q 成分までに 0 が含まれる平衡点は, 徳タイプのプレイヤー i については, 仮定 3 より, 他のプレイヤーの戦略に依存せず, 平衡点を除いて常に $\dot{x}_{iA} > 0$ が成り立つから, 明らかに漸近安定ではない。

iii 同感タイプに該当する成分

ヤコビ行列の対角成分は, $((\alpha_i - \beta_i)(\sum_{j \neq i} x_{jA}/n - 1) + \beta_i)(1 - 2x_{iA})$ である。このプレイヤーに該当する成分に 1 を割り当てた時, 連立微分方程式を線形近似した際の, 対角成分は $-(\alpha_i - \beta_i)(\sum_{j \neq i} x_{jA}/n - 1) + \beta_i$ である。 $q \leq \sum_{j \neq i} x_{jA} \leq n - 2$ 。仮定 1 より, 対角成分は必ずしも負にはならない。

このプレイヤーに該当する成分に 0 を割り当てた時, 連立微分方程式を線形近似した際の, 対角成分は $((\alpha_i - \beta_i)(\sum_{j \neq i} x_{jA}/n - 1) + \beta_i)$ である。 $q \leq \sum_{j \neq i} x_{jA} \leq n - 2$ と仮定 1 より, 対角成分は必ずしも負にならない。

以上 i, ii, iii より, 必ず漸近安定となる社会状態は存在しない。

次に, 漸近安定点が存在しない場合の存在を示す。

ある p と q が与えられたときに, 同感タイプのプレイヤーが 1 人を除いて, 連立微分方程式を線形近似したときの, 該当するヤコビ行列の対角成分が, 成分に 1 を割り当てないと, 負にならないとする。この状況は, α_i が β_i に比べて相対的に, 一定水準以上に大きいときに起こり得る。ヤコビ行列の対角成分は, $((\alpha_i - \beta_i)(\sum_{j \neq i} x_{jA}/n - 1) + \beta_i)(1 - 2x_{iA})$ であり, この成分が, 所与の $p \geq 1$ で必ず正となるような, α_i と β_i の組み合わせが存在する。

残り 1 人のプレイヤーの成分に, 0 と 1 のどちらを割り当てべきか検討する。このプレイヤー以外に割り当て成分は, すべて決定済である。どちらを割り当てても, 連立微分方程式を線形近似したときの, 該当するヤコビ行列の成分が 0 になってしまう場合がある。

中間値の定理より $(\alpha_i - \beta_i)(\sum_{j \neq i} x_{jA}/n - 1) + \beta_i = 0$ となるような, $\sum_{j \neq i} x_{jA}$ が存在する。この場合には, 0と1のどちらかを割り当てても, ヤコビ行列の成分は0となる。他に漸近安定となる候補の点は存在せず, 漸近安定となる社会状態は存在しない。

証明終わり

例 1 : $n = 4, p = 1, q = 1$ としたときの, 命題 3

財タイプのプレイヤーに該当する成分は0, 徳タイプのプレイヤーに該当する成分は1となる社会状態以外は, 漸近安定とはならない。漸近安定点を求めるにおいて, 残り 2 人の同感タイプのプレイヤーに該当する成分を0とするか1にするべきかは, 各人の $k_i \equiv \frac{\beta_i}{\beta_i - \alpha_i}$ に依存する。同感タイプなので両プレイヤーで, $0 < k_i < 1$ が成り立つ。

$k_3 < 2/3$ かつ $k_4 < 2/3$ のとき, (0,1,1,1)が漸近安定となる。

$k_3 > 1/3$ かつ $k_4 > 1/3$ のとき, (0,1,0,0)が漸近安定となる。

$k_3 < 1/3$ かつ $k_4 > 2/3$ のとき, (0,1,1,0)が漸近安定となる。

$k_3 > 2/3$ かつ $k_4 < 1/3$ のとき, (0,1,0,1)が漸近安定となる。

$k_3 \leq 1/3$ かつ $k_4 = 2/3$ のとき, $k_3 = 1/3$ かつ $k_4 \geq 2/3$ のとき, $k_3 = 2/3$ かつ $k_4 \leq 1/3$ のとき, $k_3 \geq 2/3$ かつ $k_4 = 1/3$ のときには, 漸近安定点は存在しない。

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.
- Bréban, L. (2014). Smith on happiness: towards a gravitational theory. *The European Journal of the History of Economic Thought*, 21(3), 359-391.
- 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.
- Smith, A. (1976). *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』(上)・(下) 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.

和文

- ウェイブル・J・W 著／大和瀬監訳(1999)『進化ゲームの理論』オフィスカノウチ。
- 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。
- 菅隆彦(2018) 『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」 *TERG Discussion Papers*, No.384.
- 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(2000)『アダム・スミスと現代』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。