Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems—cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse “real-life” situations like the climate, the coherent emission of light from lasers, chemical reaction-diffusion systems, biological cellular networks, the dynamics of stock markets and of the internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series “Understanding Complex Systems” focusing on the various applications of complexity, the “Springer Series in Synergetics”, which is devoted to the quantitative theoretical and methodological foundations, and the “Springer Briefs in Complexity” which are concise and topical working reports, case studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Editorial and Programme Advisory Board

Henry Abarbanel, Institute for Nonlinear Science, University of California, San Diego, USA
Dan Braha, New England Complex Systems Institute and University of Massachusetts Dartmouth, USA
Péter Érdi, Center for Complex Systems Studies, Kalamazoo College, USA and Hungarian Academy of Sciences, Budapest, Hungary
Karl Friston, Institute of Cognitive Neuroscience, University College London, London, UK
Hermann Haken, Center of Synergetics, University of Stuttgart, Stuttgart, Germany
Viktor Jirsa, Centre National de la Recherche Scientifique (CNRS), Université de la Méditerranée, Marseille, France
Janusz Kacprzyk, System Research, Polish Academy of Sciences, Warsaw, Poland
Kunihiko Kaneko, Research Center for Complex Systems Biology, The University of Tokyo, Tokyo, Japan
Scott Kelso, Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, USA
Markus Kirkilionis, Mathematics Institute and Centre for Complex Systems, University of Warwick, Coventry, UK
Jürgen Kurths, Nonlinear Dynamics Group, University of Potsdam, Potsdam, Germany
Ronaldo Menezes, Department of Computer Science, Florida Institute of Technology, Melbourne, FL, USA
Andrzej Nowak, Department of Psychology, Warsaw University, Poland
Hassan Qudrat-Ullah, School of Administrative Studies, York University, Toronto, ON, Canada
Linda Reichl, Center for Complex Quantum Systems, University of Texas, Austin, USA
Peter Schuster, Theoretical Chemistry and Structural Biology, University of Vienna, Vienna, Austria
Frank Schweitzer, System Design, ETH Zürich, Zürich, Switzerland
Didier Sornette, Entreprenuerial Risk, ETH Zürich, Zürich, Switzerland
Stefan Thurner, Section for Science of Complex Systems, Medical University of Vienna, Vienna, Austria
Understanding Complex Systems

Founding Editor: S. Kelso

Future scientific and technological developments in many fields will necessarily depend upon coming to grips with complex systems. Such systems are complex in both their composition – typically many different kinds of components interacting simultaneously and nonlinearly with each other and their environments on multiple levels – and in the rich diversity of behavior of which they are capable.

The Springer Series in Understanding Complex Systems series (UCS) promotes new strategies and paradigms for understanding and realizing applications of complex systems research in a wide variety of fields and endeavors. UCS is explicitly transdisciplinary. It has three main goals: First, to elaborate the concepts, methods and tools of complex systems at all levels of description and in all scientific fields, especially newly emerging areas within the life, social, behavioral, economic, neuro- and cognitive sciences (and derivatives thereof); second, to encourage novel applications of these ideas in various fields of engineering and computation such as robotics, nano-technology, and informatics; third, to provide a single forum within which commonalities and differences in the workings of complex systems may be discerned, hence leading to deeper insight and understanding.

UCS will publish monographs, lecture notes, and selected edited contributions aimed at communicating new findings to a large multidisciplinary audience.

More information about this series at http://www.springer.com/series/5394
Scientific Metrics: Towards Analytical and Quantitative Sciences
Preface

This is a selected collection of my academic representative works, which concentrates on the core idea that the scientific metrics decide scientific findings, or the selections of measures determine the discovery of scientific laws. As the choice of different measures may lead to different quantitative relations, the different measures could cause and generate different scientific discoveries.

In the first section of six physical papers, the linked-measure and the linked-field are introduced by using mathematical multi-vector methodology. It is found that a linked-measure can be geometrically represented by a vortex and the linked-field can unify microparticle standard model and macrocosmos standard model, without supersymmetry and with dispelling dark matter and dark energy.

In the following six economic papers, it is revealed that economic measures can be described by complex numbers. It is found that economic equilibrium and economic stability are mastered by Cauchy–Riemann equation and Laplace equation together, while commodity function and money function are unified by complex function.

In the last group of six scientometric papers, a unified informetric model characterized by wave-heat equations is proposed and h-type metrics are developed for the applications in network analysis. Meanwhile, a mathematical theory of knowledge is contributed to approach the unified framework of data-information-knowledge-wisdom.

There are two key ideas which are stressed throughout these chapters. One is that measures link to findings, including physical measures to physical findings, economic measures to economic findings and knowledge measures to knowledge findings. The other is that equations can be the guide for the process, including Hamilton–Lagrange equations for physical process, Cauchy–Riemann equation and Laplace equation for economic process, and wave-heat equations for information process.

These three groups of scientific papers construct a logic chain on nature (physical world)-society (economic world)-knowledge (information world). It is expected to reveal the unified mechanism via scientific metrics in the natural laws
from microparticles to macrocosmos, in the economic rules of human society and in the core knowledge among huge information.

I have been thinking three issues for many years since 1978 when I was 16: (1) what drives the physical world? (2) what masters the economic society? (3) what constructs the human knowledge? The seeking for answers leads to these physical, economic and scientometric papers organized under the title “Scientific Metrics: Towards Analytical and Quantitative Sciences”.

Actually, the collection of selected eighteen papers nearly covers all the creative ideas presented in my representative works, including vortex mechanism applied in physics, complex analysis applied in economics and $h$-type metrics applied in scientometrics. In these multidisciplinary papers, around natural, social and knowledge mechanisms, I try to clarify three issues: (1) how to measure the physical world? (2) how to measure the economic process? (3) how to measure the human knowledge? In those papers collected in this book, the creative ideas include using linked-measure to physics, complex measure to economics and $h$-type measure to scientometrics. Also, a philosophical idea is attached in Appendix I.

Concerning the copyright issue of those published papers, I rewrite five scientometric papers by merging two or three articles into one mostly. Except scientometric papers, the copyright of all papers published on and distributed in open-access journals is reserved by myself. The collection integrated them together for a more effective academic communication. I hope that my efforts put into this book could inspire new insights and promote the academic progress.

Acknowledgements: it is my honour to accept the National Scientific and Technical Publishing Fund for publishing this book. I acknowledge the National Natural Science Foundation of China Grant Nos. 70773101, 7101017006, 71173187, 71673131 for continuous supports, thank Profs. H. Eugene Stanley, Loet Leydesdorff, Ronald Rousseau, Wolfgang Glänzel, Mu-huan Huang, Dar-zen Chen, Feicheng Ma, and Drs. Lutz Bornmann, Star X. Zhao, Simon S. Li, and more colleagues mentioned in each chapter, for their kind collaborations, and thank Ms. Regina Entorf, and Dr. Xuguang Li for their English proofreading, as well as editors Ms. Xue Hui and Dr. Jian Li for their helpful editing.

Fred Y. Ye, B.S., M.A., Ph.D.
US Fulbrighter, Fellow, European Academy of Sciences and Arts;
Professor, Nanjing University, Nanjing, China
Contents

Part I Physics: Linked-Metrics Generates Vortex-World Physics

1 The Linked-Measure and Linked-Field for Linking Micro-Particles to Macro-Cosmos with Dispelling Dark Matter and Dark Energy ........................................ 3
  1.1 Introduction ........................................ 3
  1.2 Present Achievements as Constrains ...................... 4
    1.2.1 Observational Evidence ........................ 4
    1.2.2 Theoretical Constrains with Characteristics ....... 5
  1.3 New Theoretical Structure ..................................... 6
    1.3.1 Linked-Measure ................................ 7
    1.3.2 Linked-Field ................................ 8
  1.4 Dark Matter and Dark Energy: The Structural Effects ...... 12
  1.5 Experiment Suggestion ................................... 12
  1.6 Discussion: Cosmological Image ............................ 13
  1.7 Concluding Remarks ...................................... 16
References .................................................. 17

2 The Physical Linked-Measure Works as Vortex with Linking to Turbulence .................. 19
  2.1 Introduction ............................................... 19
  2.2 Methodology ............................................... 20
  2.3 Vortex Dynamics and Black Hole ............................. 23
  2.4 Wingtip Vortices ......................................... 26
  2.5 Crosswind as Wavelets Leading to Ideal Turbulence ....... 28
  2.6 Discussion and Conclusion ................................ 30
References .................................................. 31
### 3 The Clifford-Finslerian Linked-Field Leads Branching Multiverse

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.2</td>
<td>The Clifford-Finslerian Construction</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Clifford Algebra as Interior Dynamics</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Finsler Geometry and Catastrophic Theory Leads Space-Time Branching</td>
</tr>
<tr>
<td>3.3</td>
<td>Analytical Structure and Wave-Particle Duality</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Field Equations</td>
</tr>
<tr>
<td>3.3.2</td>
<td>An Interpretation of Wave-Particle Duality</td>
</tr>
<tr>
<td>3.4</td>
<td>Branching Multiverse</td>
</tr>
<tr>
<td>3.5</td>
<td>Discussion and Criticism</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusion</td>
</tr>
</tbody>
</table>

### 4 A Clifford-Finslerian Physical Unification and Fractal Dynamics

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>4.2</td>
<td>Mathematical Methods</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Algebraic Structure</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Analytical Structure</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Geometrical Structure</td>
</tr>
<tr>
<td>4.3</td>
<td>Physical Principles</td>
</tr>
<tr>
<td>4.3.1</td>
<td>The First Physical Principle</td>
</tr>
<tr>
<td>4.3.2</td>
<td>The Second Physical Principle</td>
</tr>
<tr>
<td>4.3.3</td>
<td>The Third Physical Principle</td>
</tr>
<tr>
<td>4.4</td>
<td>Clifford–Finslerian Physics</td>
</tr>
<tr>
<td>4.5</td>
<td>Finslerian Metric and Fractal Dynamics</td>
</tr>
<tr>
<td>4.6</td>
<td>Discussion</td>
</tr>
<tr>
<td>4.7</td>
<td>Conclusion</td>
</tr>
</tbody>
</table>

### 5 A Vortex Mechanism Linking Micro-Particle to Macro-Galaxy Without Supersymmetry

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>5.2</td>
<td>Phenomenological Foundations and Theoretical Constrains</td>
</tr>
<tr>
<td>5.3</td>
<td>Theoretical Approaches by Vortex Mechanism</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Linked-Measure Describes Vortex, Leading to Linked-Field</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Electromagnetic Field</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Strong and Electro-Weak Fields</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Gravitational Field</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion: The Vortex-World</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>8.4 Some Special Cases</td>
<td>99</td>
</tr>
<tr>
<td>8.5 Complex Economics: Using Complex Numbers as Unified Marginal-Multiplier Measures</td>
<td>100</td>
</tr>
<tr>
<td>8.6 Discussion: A Conjecture on Linking Marginal Effect to Multiplier Effect</td>
<td>101</td>
</tr>
<tr>
<td>8.7 More Analysis: Relations Among the Rates of Inflation, Interest and Exchange</td>
<td>103</td>
</tr>
<tr>
<td>8.8 Concluding Remarks</td>
<td>104</td>
</tr>
<tr>
<td>References</td>
<td>105</td>
</tr>
<tr>
<td>9 Economic Complex Analysis for Approaching Economic Equilibrium and Economic Stability</td>
<td>107</td>
</tr>
<tr>
<td>9.1 Introduction</td>
<td>107</td>
</tr>
<tr>
<td>9.2 Complex Economics as Theoretical Framework</td>
<td>108</td>
</tr>
<tr>
<td>9.3 Analytical Solutions</td>
<td>109</td>
</tr>
<tr>
<td>9.4 Numerical Simulation</td>
<td>110</td>
</tr>
<tr>
<td>9.5 Discussion</td>
<td>112</td>
</tr>
<tr>
<td>9.6 Concluding Remarks</td>
<td>114</td>
</tr>
<tr>
<td>References</td>
<td>114</td>
</tr>
<tr>
<td>10 Complex Economic Metrics Linking to Scaling Money Supply</td>
<td>117</td>
</tr>
<tr>
<td>10.1 Introduction</td>
<td>117</td>
</tr>
<tr>
<td>10.2 Theoretical Framework for the Upper-Limitations and Lower-Limitations of Money Supply</td>
<td>118</td>
</tr>
<tr>
<td>10.2.1 Three Ways</td>
<td>118</td>
</tr>
<tr>
<td>10.2.2 Linking to Complex Metrics</td>
<td>120</td>
</tr>
<tr>
<td>10.3 Numerical Estimation</td>
<td>121</td>
</tr>
<tr>
<td>10.4 Discussion: Money Supply and Money Policy</td>
<td>122</td>
</tr>
<tr>
<td>10.5 Conclusion</td>
<td>124</td>
</tr>
<tr>
<td>References</td>
<td>125</td>
</tr>
<tr>
<td>11 A Synthetic Macro-Economic Model Integrating Interest, Exchange and Tax Rates</td>
<td>127</td>
</tr>
<tr>
<td>11.1 Introduction</td>
<td>127</td>
</tr>
<tr>
<td>11.2 Concepts and Relations of Three Rates</td>
<td>128</td>
</tr>
<tr>
<td>11.3 Synthetic Model</td>
<td>130</td>
</tr>
<tr>
<td>11.4 Policy Analysis</td>
<td>132</td>
</tr>
<tr>
<td>11.5 Possible Extension</td>
<td>135</td>
</tr>
<tr>
<td>11.6 Discussion: Quantitative Easing and Three Ratios</td>
<td>136</td>
</tr>
<tr>
<td>11.7 More Discussion: Social Equality and Economic Efficiency</td>
<td>137</td>
</tr>
<tr>
<td>11.8 Conclusion</td>
<td>139</td>
</tr>
<tr>
<td>References</td>
<td>139</td>
</tr>
</tbody>
</table>
12 A Cubic Integrated Economic Model for Macro-Economic Analysis ................................................ 141
12.1 Introduction ........................................ 141
12.2 A Brief Review of the Analytical Foundations .............. 142
12.3 The Cubic Integrated Model ............................ 145
12.4 Discussion I: Total Economic Quantities and Relative Economic Rates .............................. 147
12.5 Discussion II: Possible Applications and Expected Extensions ............................................... 149
12.6 Conclusion ......................................... 149
References............................................... 150

Part III Scientometrics: Knowledge Metrics and $h$-type Metrics

13 Measuring Knowledge: A Quantitative Approach to Knowledge Theory ..................................... 155
13.1 Introduction ........................................ 155
13.2 DIKW Chain and Knowledge Metrics .................... 156
13.3 Measuring Explicit Knowledge and Tacit Knowledge ........ 158
13.4 Analysis and Discussion ................................ 160
13.5 Conclusion ......................................... 161
References............................................... 162

14 A Theoretical Unification of Informetric Models by Wave-Heat Equations .................................. 163
14.1 Introduction ........................................ 163
14.2 Theoretical Methodology .............................. 164
14.2.1 Shifted Power Function Fits Wave-Heat Equations . 164
14.2.2 Fundamental Solutions and the Function F(x, T) as a Fundamental Solution ................ 165
14.2.3 Space-Type Distributions and Their Size-Rank Transformation ........................................ 166
14.2.4 Time-Type Distributions ........................ 173
14.3 Theoretical Extension ................................ 174
14.4 Discussion ......................................... 177
14.5 Conclusion ......................................... 178
References............................................... 179

15 The Empirical Investigation and Theoretical Unification of Mathematical Models for the $h$-Index .............. 181
15.1 Introduction ........................................ 181
15.2 Methodology ........................................ 182
15.2.1 Method ..................................... 182
15.2.2 Data. ........................................ 183
15.3 Results ............................................ 183
## 16 The $h$-Type Core Structure in Single-Layer and Multi-layer Weighted Information Networks

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.1</td>
<td>Introduction</td>
<td>191</td>
</tr>
<tr>
<td>16.2</td>
<td>Methodology</td>
<td>192</td>
</tr>
<tr>
<td>16.2.1</td>
<td>$h$-Degree, Leading to Network $h$-Core</td>
<td>193</td>
</tr>
<tr>
<td>16.2.2</td>
<td>$h$-Strength, Leading to $h$-Subnet and $h$-Bone</td>
<td>195</td>
</tr>
<tr>
<td>16.2.3</td>
<td>$h$-Crystal</td>
<td>197</td>
</tr>
<tr>
<td>16.2.4</td>
<td>Algorithm and Procedure</td>
<td>198</td>
</tr>
<tr>
<td>16.2.5</td>
<td>Data Sets and Experiments</td>
<td>200</td>
</tr>
<tr>
<td>16.3</td>
<td>Results</td>
<td>201</td>
</tr>
<tr>
<td>16.3.1</td>
<td>The $h$-Subnet of Co-citation Network on “$h$-Index”</td>
<td>201</td>
</tr>
<tr>
<td>16.3.2</td>
<td>The $h$-Crystal of the “$h$-Set”</td>
<td>202</td>
</tr>
<tr>
<td>16.4</td>
<td>Analysis and Discussion</td>
<td>203</td>
</tr>
<tr>
<td>16.4.1</td>
<td>$h$-Centrality Measures for Nodes</td>
<td>203</td>
</tr>
<tr>
<td>16.4.2</td>
<td>$h$-Centralization for Whole Networks</td>
<td>204</td>
</tr>
<tr>
<td>16.4.3</td>
<td>Some Theoretical Properties of $h$-Strength</td>
<td>205</td>
</tr>
<tr>
<td>16.4.4</td>
<td>Some Theoretical Properties of $h$-Crystal</td>
<td>206</td>
</tr>
<tr>
<td>16.4.5</td>
<td>Limitations</td>
<td>208</td>
</tr>
<tr>
<td>16.5</td>
<td>Conclusion</td>
<td>208</td>
</tr>
<tr>
<td>16.6</td>
<td>Appendix Pseudo-codes of the $h$-Type Core Algorithm</td>
<td>209</td>
</tr>
</tbody>
</table>

## 17 The $h$-Core and $h$-Tail Distribution with Dynamic Metrics

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.1</td>
<td>Introduction</td>
<td>215</td>
</tr>
<tr>
<td>17.2</td>
<td>Methodology</td>
<td>216</td>
</tr>
<tr>
<td>17.2.1</td>
<td>Dynamic Core-Tail Measures</td>
<td>217</td>
</tr>
<tr>
<td>17.2.2</td>
<td>Distributed Core-Tail Measures</td>
<td>218</td>
</tr>
<tr>
<td>17.3</td>
<td>Empirical Study</td>
<td>221</td>
</tr>
<tr>
<td>17.4</td>
<td>Analysis and Discussion</td>
<td>226</td>
</tr>
<tr>
<td>17.4.1</td>
<td>The Evolution of $h$-Core</td>
<td>227</td>
</tr>
<tr>
<td>17.4.2</td>
<td>The Evolution of $h$-Tail</td>
<td>228</td>
</tr>
<tr>
<td>17.5</td>
<td>Conclusion</td>
<td>229</td>
</tr>
</tbody>
</table>

## 18 A Quantitative Relationship Between Per Capita GDP and Scientometric Criteria

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.1</td>
<td>Introduction</td>
<td>233</td>
</tr>
<tr>
<td>18.2</td>
<td>Methodology</td>
<td>234</td>
</tr>
<tr>
<td>18.2.1</td>
<td>Subject</td>
<td>234</td>
</tr>
</tbody>
</table>
About the Author

**Dr. Fred Y. Ye** is now a full professor in the School of Information Management at Nanjing University, a fellow of European Academy of Sciences and Arts, a lifelong member of ISSI (International Society for Scientometrics and Informetrics) and a member of the Editorial Committees of *Journal of Data and Information Science* and so on. During September 2000 to September 2001, Dr. Ye was a Fulbright research scholar in the University of Arizona and the University of Michigan in the USA.

Professor Dr. Ye’s research interest focusses on the scientific metrics. His publications include about 100 articles published in *Scientometrics, Journal of Informetrics, Journal of the Association for Information Science and Technology, Physical Journal*, etc. He also published three monographs and three textbooks. He has proposed vortex-world physics, complex economics and knowledge metrics. His present contributions concentrate on the following ideas: (1) vortex mechanism linking microparticle to macrogalaxy; (2) complex economic mechanism mastered by Cauchy–Riemann equation and Laplace equation for penetrating microeconomics and macroeconomics; (3) unified informetric mechanism described by wave-heat equations.
Part I

Physics: Linked-Metrics Generates Vortex-World Physics
Chapter 1
The Linked-Measure and Linked-Field for Linking Micro-Particles to Macro-Cosmos with Dispelling Dark Matter and Dark Energy

A mathematical multi-vector consists of a complex scalar, a complex vector and a bi-vector, which constructs a physical linked-measure, yielding a linked-field. When the linked-measure is applied as the world measure, its strong symmetric links generate electromagnetic field and its strong micro-inner links do strong field, while its weak micro-inner symmetric links synthesize electro-weak field. With adding outer space-time metric, the linked-field leads to gravitational field with a new understanding of dark matter and dark energy. In the linked-field, the micro-particle standard model and the macro-cosmos standard model are unified, where double dynamic sources drive the universe, in which one initialized big bang and another pushed rotation. Dark matter and dark energy are structural effects of the whole universe, caused respectively by rotation and big bang. A rotated vacuum explosive experiment is suggested to verify the hypothesis.

1.1 Introduction

In the 20th century, two great physical theories, quantum theory and relativity, led to two standard models respectively, the micro-particle standard model characterized by $SU_C(3) \times SU_L(2) \times U_Y(1)$ gauge field based on quantum theory and the macro-cosmos standard model characterized by big bang cosmology based on general relativity (Einstein 1916), which construct theoretical core of contemporary physics (Beringer and Navas 2012). However, the two standard models are so different that they can not be unified within a harmony theory, which causes a theoretical poser called quantum gravity. On the other side, the physical experiments and observations strongly support the two standard models, so that any revised unified theory has to include the two standard models.

Meanwhile, the issues of dark matter and dark energy perplexed scientists (Drees and Gerbier 2014; Mortonson et al. 2014). If the problems are kept in two standard models, it means that standard models mismatch the real world or there exist logic faults in those theories.
Following a brief review of present achievements in physics, the linked-measure and linked-field approaching to the unified theory of micro-particle and the macro-cosmos are suggested in this chapter, using the mathematical methodology of multi-vector (Hestenes 1975; Doran and Lasenby 2003).

1.2 Present Achievements as Constrains

We know that physics has reached a bright level with shining achievements, where we especially mention quantum theory and relativity, with linking via the Hamilton principle (the least action principle, which provided a cornerstone of analytical physics). Moreover, in micro-fields, physicists have probed into “basic particles” that construct the natural world, where we especially mention the principle of particle-wave duality, which introduced the quantum mechanism. In macro-fields, physicists have explored the universe, where we especially mention the principle of gravity-curvature equivalence, which constructs general relativity as well as scientific cosmology. Certainly, all the theoretical principles match physical observations and experiments very well. More generally, the principles of action, connection and construction have been suggested (Ye 2009, 2013).

Facing the great research object of physics, we have to push forward of the physical achievements and inherit the glorious tradition in future. At present, we have strong evidence that two standard models are qualified physical theories for approaching the world. The basic facts focus on that all substances consist of micro-particles, while the cosmic observations strongly support that the universe is expanding, with acceleration.

1.2.1 Observational Evidence

It is well known that both standard models have their solid foundations of observational evidence. During 2012 to 2013, two large-scale important experimental observations bring us new physical updates. One is Higgs that is found by LHC (large hadron collider) and the other is CMB (cosmic microwave background) that is exactly verified by Planck satellite.

In the micro-particle standard model, the greatest success is the prediction of the Higgs boson, which have been experimentally discovered in 2012 and verified in 2013 by LHC at energy 125.6 GeV·c$^{-2}$ (The ATLAS Collaboration 2012; The CMS Collaboration 2012). It was well known that the Higgs mechanism describes how the weak SU(2) gauge symmetry is broken and how fundamental particles obtain mass, which was also the last particle predicted by the micro-particle standard model to be observed, although efforts to confirm that it has all of the properties predicted by the standard model are ongoing.
Other great successes of the micro-particle standard model included the prediction of the $W$ boson and $Z$ boson, the gluon and the top and the charm quarks. However, the worst disadvantage in the micro-particle standard model was the complete absence of gravity, which predicts neutrinos to be massless while the observed evidence of neutrino oscillations implies that neutrinos have tiny mass (Olive et al. 2014).

In the macro-cosmos standard model, it is frequently referred to the big bang cosmology, since it is the simplest model that provides a reasonably good account of the following observational evidence of the cosmos: (1) the existence of the CMB and the large-scale structural distribution of galaxies; (2) the abundances of hydrogen (including deuterium), helium and lithium, as famous BBN (big bang nucleosynthesis) theory of elements synthesis explains (Sarkar 1996; Cyburt et al. 2005; Fields and Sarkar 2001); (3) the accelerating expansion of the universe observed in the light from distant galaxies and supernovae (Riess et al. 1998). According to the Planck satellite’s newest observation report (Planck Collaboration 2014a, b), we understand that there are 68.3% dark energy, 26.8% dark matter and 4.9% matter-energy in the universe.

Also, there are challenges to the macro-cosmos standard model, in which we mention particularly dark matter and dark energy. Extensive searches for dark matter particles have so far shown credible detection and dark energy almost becomes impossible to be detected in any laboratory.

Observational evidence and great successes above mean that we have to consider theoretical constrains or restrictions when we try to extend the two standard models or create new theories.

### 1.2.2 Theoretical Constrains with Characteristics

While the particle standard model supposes that gauge field theory is the best choice, the cosmos standard model assumes that general relativity is the correct theory of gravity on cosmological scales.

Because those two standard models are so successful, a complete unified theory should abide some theoretical constrains with the following characteristics.

1. The theory should unify all interactions, including $\text{SU}_C(3) \times \text{SU}_L(2) \times \text{UY}(1)$ symmetry at micro-level and accelerating expansion at macro-level.

2. The theory should accord with analytical principles of physics, particularly the Hamiltonian principle, with incorporating quantum theory and relativity. Therefore it is a quantum field theory, where local symmetries are described by Abelian and non-Abelian gauge theories.

3. The particle-wave duality should be maintained at micro-particle level and the gravity-curvature equivalence at macro-cosmos level.

Generally, a perfect unified theory should be consistent with both micro-particle standard model and macro-cosmos standard model, abiding the principles of the least action, particle-wave duality and gravity-curvature equivalence, with interpreting dark matter and dark energy. So the unified theory is demanded with high standard,
not only unifying quantum theory and gravity, but also matching experimental data and observational phenomena well.

Also, a good theoretical framework should keep math-physical balance, which means that, when we set up an equation, if the left side of the equation reflects mathematical structure, its right side should reveal physical essence, i.e., mathematical structure = physical essence. For approaching the object, space-time multi-vector (Hestenes 1975; Hestenes and Sobczyck 1984; Lasenby et al. 2004) provides a good methodology and the Clifford-type algebra can be united with Finsler-type geometry.

1.3 New Theoretical Structure

Let’s consider space-time at the beginning. For each space-time point \(x\) on Dirac frame \(\{\gamma^\mu, \mu = 0, 1, 2, 3\}\), there exists

\[x = x_\mu \gamma^\mu, \quad x_\mu = \gamma^\mu x\]  (1.1)

The coordinates’ transformation will be \(x_\mu \to x'_\mu = \alpha^\nu_\mu x_\nu; \gamma^\mu \to \gamma'^\mu = \alpha^\nu_\mu \gamma^\nu\) with \(\alpha^\mu_\nu \alpha^\nu_\lambda = \delta^\mu_\lambda\) where four Dirac matrices are no longer viewed as four matrix-valued components of a single space vector, but as four orthonormal basis vectors for real space-time,

\[\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & -\sigma^k \\ \sigma^k & 0 \end{pmatrix}\]  (1.2)

where \(\gamma^0\) is time-like vector and \(\gamma^k (k = 1, 2, 3)\) space-like vectors. Similarly, the three Pauli matrices \(\sigma^k = (\sigma^1, \sigma^2, \sigma^3)\) are no longer viewed as three matrix-valued components of a single space vector, but as three orthonormal basis vectors for three-dimensional Euclidean space.

\[\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\]  (1.3)

Both \(\{\gamma^k\}\) and \(\{\sigma^k\}\) are to be interpreted geometrically as meaningful space-time bi-vectors and not as operators.

Using multi-vector \(M_k (k = 0, 1, 2, 3, 4)\) as the world measure where \(M_k\) is a multi-vector of grade \(k\), \(k = 0\) corresponds to scalar, \(k = 1\) to vector, \(k = 2\) to bi-vector, \(k = 3\) to pseudo-vector and \(k = 4\) to pseudo-scalar and the Clifford bases of four-dimensional space-time are generated by four orthonormal vectors \(\{\gamma^\mu, \mu = 0, 1, 2, 3\}\) and spanned by \(1\) (1 scalar at grade 0), \(\{\gamma^\mu\}\) (4 vectors at grade 1), \(\{\sigma^k, i\sigma^k\}\) (6 bi-vectors at grade 2), \(\{i\gamma^\mu\}\) (4 pseudo-vectors at grade 3) and \(i\) (1 pseudo-scalar at grade 4) orderly. Now we write
\[ M = M_0 + M_1 + M_2 + M_3 + M_4 = \varphi + V + B + iU + i\theta = \psi + A + B = (\psi, A, B) \]  

(1.4)

where \( \psi = \varphi + i\theta \) constructs a complex scalar function (mass-like function), while \( A = V + iU \) forms a complex vector function (potential-like function) and \( B = (1/2) B_{\mu\nu} \gamma^\mu \wedge \gamma^\nu \) as a unique bi-vector (spin-like function).

The revision (revised conjugation) is denoted as \( \tilde{M} \).

\[ \tilde{M} = \tilde{M}_0 + \tilde{M}_1 + \tilde{M}_2 + \tilde{M}_3 + \tilde{M}_4 = \varphi + V - B - iU + i\theta = (\psi, \tilde{A}, -B) \]  

(1.5)

And there are space-time conjugation \( M \), space conjugation \( M^* \) and Hermitian conjugation \( M^\dagger \) as follows.

\[ M = -iM = M_0 - M_1 + M_2 - M_3 + M_4 = \psi - A + B = (\psi, -A, B) \]  

(1.6)

\[ M^* = \gamma^0 M \gamma^0; \quad M^\dagger = \gamma^0 \tilde{M} \gamma^0 \]  

(1.7)

Then \( M^2 \) can be defined as

\[ M^2 = M \tilde{M} \]  

(1.8)

Meanwhile, \( M \) can be divided as even part \( M_+ \) and odd part \( M_- \) as follows.

\[ M_+ = \varphi + B + i\theta = \psi + B \]  

(1.9)

\[ M_- = V + iU = A \]  

(1.10)

That is the origin of spiral and asymmetry in the world if we apply \( M \) as the world’s physical measure.

### 1.3.1 Linked-Measure

Mathematically, a multi-vector consists of a complex scalar, a complex vector and a bi-vector, where one scalar \( \psi \), one vector \( A \) and one bi-vector \( B \) contain rich structural information.

Physically, a multi-vector can be applied as linked-measure which may link scalar static mass \( m \), vector moving potential \( V \) and bi-vector rotating strength \( S \) together.

So a mathematical multi-vector corresponds to a physical linked-measure. That is to say, a multi-vector constructs a linked-measure.

The differential operators of one order derivatives can be introduced and defined as

\[ \partial_\mu = \frac{\partial}{\partial x_\mu}; \quad \nabla = \gamma^\mu \partial_\mu \]  

(1.11)
Its covariant derivative and the differential operators of two order derivatives can be introduced and defined respectively as

\[ \mathbb{D}_\mu = (\partial_\mu - \omega_\mu); \ \nabla^2 = g^{\mu\nu} \partial_\mu \partial_\nu; \ g^{\mu\nu} = \gamma^\mu \cdot \gamma^\nu \] (1.12)

where we see that space-time metric \( g^{\mu\nu} \) is naturally generated.

For keeping gauge invariance, we suppose the transformations as

\[ \psi \rightarrow \psi' = e^{i\omega} \psi; \ \overline{\psi} \rightarrow \overline{\psi'} e^{-i\omega} \]  

\[ A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \omega; \ \overline{A}_\mu \rightarrow \overline{A'}_\mu = \overline{A}_\mu - \partial_\mu \omega \] (1.14)

So the linked-measures construct a complete math-physical system, which leads to linked-field.

When we define linked-energy \( E \) and linked-momentum \( p_\mu \) with linking Hamilton function \( H \) and Lagrangian function \( L \), we obtain math-physical equations following Hamilton principle as follows.

\[ L = p_\mu x_\mu - H; \ p_\mu = \frac{\partial L}{\partial x_\mu}; \ \delta \int_s d^4 x L = 0 \] (1.15)

\[ H = p_\mu x_\mu - L = E(s, t) = \int_s e(x) dx_\mu \]  

\[ \frac{\partial H}{\partial p_i} = \frac{ds_i}{dt}; \ \frac{\partial H}{\partial s_i} = -\frac{dp_i}{dt} \] (1.17)

where \( e(x) \) means density function of linked-measure in space and \( s_i \) marks space variables. While Greek subscripts \( \mu, \nu \) denote 1, 2, 3, 4, and Latin subscripts \( i, j \) do 1, 2, 3.

Equation (1.16) is the integral form and Eq. (1.17) the differential form of multivector linked-field.

### 1.3.2 Linked-Field

When \( M(\psi, A, B) \) is viewed as physical linked-measure, fields are naturally generated in space-time.

If there exist strong links among \( \psi, A \) and \( B \) as follows,

\[ \nabla A = \psi + \partial \varphi; \ \ B_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \] (1.18)

the Maxwell’s equations under gauge group \( U(1) \) will be obtained as
\[ \nabla B = \nabla \cdot B + \nabla \wedge B = J \quad \text{or} \quad \partial_\mu B^{\mu\nu} = J^\nu \tag{1.19} \]

where \( J = \nabla \cdot B \) is current and bi-vector \( B \) includes both electrical field \( E \) and magnetic field \( H \) as \( B = E + iH \), which can also be derived by applying Hamilton principle to the QED (quantum electrodynamics) Lagrangian.

\[ \delta S = \delta \int d^4x L = 0 \tag{1.20} \]

\[ L_{\text{QED}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - J^\mu A_\mu \tag{1.21} \]

Generally, electromagnetic field is both macro-phenomenon and micro-phenomenon. At micro-level, when the complex scalar field (i.e., Higgs field) introduces mass \( m \) following Lagrangian,

\[ L_{\text{Higgs}} = (D_\mu \psi^*)_*(D^\mu \psi) - V(\psi^*\psi) \tag{1.22} \]

then QED Lagrangian can be modified as

\[ L_{\text{QED}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \overline{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{1.23} \]

If there exist strong micro-inner links as strong field, including 8 duplicates \( a \) in 3 generations, the field strength upgrades to inner higher symmetry

\[ B_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = B^{a}_{\mu\nu\kappa}; a = 1, 2, \ldots, 8 \tag{1.24} \]

and the linked-field is extended with introducing new field \( F \) as

\[ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - g f_{abc} A^a_\mu A^b_\nu; f_{abc} t^c = [t^a, t^b] \tag{1.25} \]

When \( A^a_\mu \) corresponds to the gluon fields \( a = 1, 2, \ldots, 8 \), as there are eight kinds of gluon) and the \( \psi_{q,c} \) is quark-field spinor for a quark of flavor \( q \) and mass \( m_q \) with a color-index \( c \) \((c = 1, 2, 3)\), where the \( t^a_{\mu\nu} \) corresponds to eight \( 3 \times 3 \) matrices and is the generator of the SU\( (3) \) group, the Lagrangian becomes

\[ L_{\text{QCD}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \overline{\psi}_{q}(i\gamma^\mu D_\mu - m_q)\psi_q \tag{1.26} \]

which leads to QCD (quantum chromodynamics) under SU\( (3) \) invariance, where the first item describes gluons and the second item quarks. So the strong interactions are included.

When the main field strengths break micro-inner strong symmetry into micro-inner weak symmetry, from \( F \) returning to \( B \), the interactions are stemmed by group SU\( (2) \), replacing SU\( (3) \), with color-index \( c \) changing to \( b \):
\[ B_{\mu\nu} = [\partial_\mu + (ig/2)B_\mu]B_\nu - [\partial_\nu + (ig/2)B_\nu]B_\mu = B_{\mu\nu}^b \gamma_b; \ b = 1, 2, 3 \] (1.27)

where \( i \) is imaginary unit and \( g \) is coupling constant.

Combining Lagrangian

\[ L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4} B_{\mu\nu}^b B^{b\mu\nu} \] (1.28)

with QED Lagrangian equation (1.21), the electro-weak Lagrangian is obtained under \( SU_L(2) \times U_Y(1) \) symmetry as follows.

\[ L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4} B_{\mu\nu}^b B^{b\mu\nu} - J^\mu A_\mu \] (1.29)

It is feasible to attach the outer space-time metric \( g_{\mu\nu} \) to the linked-field, with introducing Riemann-Christoffel symbol

\[ \Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho} (\partial_\nu g_{\rho\mu} + \partial_\mu g_{\rho\nu} - \partial_\rho g_{\mu\nu}) = g^\lambda (D_\mu g_\nu) \] (1.30)

Then the Riemann curvature tensor is kept as

\[ R^\lambda_{\mu\rho\kappa} = \partial_\rho \Gamma^\lambda_{\mu\kappa} - \partial_\kappa \Gamma^\lambda_{\mu\rho} + \Gamma^\rho_{\mu\nu} \Gamma^\lambda_{\nu\kappa} - \Gamma^\rho_{\mu\kappa} \Gamma^\lambda_{\nu\rho} \] (1.31)

where scalar curvature \( R \) is defined by Ricci tensor \( R = g^{\mu\nu} R_{\mu\nu} \).

The Bianchi identity is also written as

\[ D_\mu R_{\mu\lambda} + D_\lambda R_{\mu\nu} + D_\nu R_{\lambda\mu} = 0 \] (1.32)

Let’s recall gravity-curvature equivalence, we know

\[ \text{World’s physical measure} = \text{World’s mathematical curvature} \] (1.33)

or

\[ \text{World’s measure tensor} \ G_{\mu\nu} = \text{World’s curvature tensor} \ \Omega_{\mu\nu} \] (1.34)

Since

\[ M \bar{M} = (\psi + A + B)(\bar{\psi} - A + B) \] (1.35)

whereas scalar and vector have no out products \( (M \wedge M = B \wedge B) \), we expect

\[ L = \frac{1}{2} (\partial_\mu \psi \cdot \partial^\nu \bar{\psi} + \partial_\mu A \cdot \partial^\nu \bar{A} + \partial_\mu B \cdot \partial^\nu \bar{B}) + \text{c.t.} - \nabla \wedge B \] (1.36)

where c.t. denotes coupling terms.
For fitting general relativity, geometrical curvature should be

$$\Omega_{\mu\nu} = k \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right)$$  \hspace{1cm} (1.37)

With applying Hamilton principle to Einstein-Hilbert-type action (Weinberg 1972)

$$S = \int d^4x \sqrt{-g} f(M, \bar{M}, R)$$  \hspace{1cm} (1.38)

it is expected that varying $S$ gives the field equation

$$T_{\mu\nu} + p(\psi) - q(A) = G_{\mu\nu} = \Omega_{\mu\nu} = k \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right)$$  \hspace{1cm} (1.39)

where $G_{\mu\nu}$ is Einstein tensor and $T_{\mu\nu}$ denotes the total observational tensor (e.g., energy-momentum tensor), positive $p(\psi)$ integrates all positive items of Eq. (1.34) and negative $q(A)$ integrates all negative items of Eq. (1.35), where the left side denotes physical effects while the right side indicates mathematical means.

Based on Eq. (1.39), we see total observational tensor becomes

$$T_{\mu\nu} = G_{\mu\nu} - p(\psi) + q(A) = k \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) - p(\psi) + q(A)$$  \hspace{1cm} (1.40)

so that we can provide a theoretical interpretation of dark matter and dark energy (Matarrese et al. 2011), where $p(\psi)$ item describes dark energy and $q(A)$ item denotes dark matter. The dark matter and dark energy are structural effects in the physical unification of whole cosmos by linked-field.

Thus we see that all physical elements of two standard models can be concluded in multi-vector linked-field, where electromagnetic field originates from bi-vector field with strong links among scalar $\psi$, vector $A$ and bi-vector $B$ at both macro-level and micro-level; strong field acts as strong micro-inner symmetry at micro-level; electro-weak field obtains from breaking strong micro-inner symmetry into weak micro-inner symmetry at micro-level. Then, after attaching outer space-time metric to linked-field, the gravity is realized via curvature at macro-level.

Generally, electro-weak and strong fields are described by construction effects of the linked-field and gravity (phenomenological structural effects contain dark matter and dark energy) is caused by structural effects of the linked-field. The linked-field is a unified field of both micro-particles and macro-cosmos.
1.4 Dark Matter and Dark Energy: The Structural Effects

For understanding the universe, a qualitative model called “cosmos ball” model can be dynamically shown in Fig. 1.1.

In Fig. 1.1, the “cosmos ball” will synchronously rotate and expand in both inside structure and outside surface, so that the whole synchronous rotation will be seen everywhere in the ball. In Newton’s theory, he paid all attention to the center of masses, so that he found gravitation between masses. In Einstein’s theory, local space-time curvature was linked with gravity, so that the gravity was equivalent to the curvature of space-time. Now the whole universe is the rotated “cosmos ball”, so that all are included in it, where space-time curvature underwent from large to small.

For driving the “cosmos ball”, double dynamic sources are demanded: one initialized big bang and another pushed rotation. The model implies that dark matter and dark energy are structural effects, where dark matter is rotated kinetic effect and dark energy is vacuum expanded dynamic effect, in which dark matter is caused by the whole synchronous rotation and dark energy is the vacuum expansion originated by the beginning of the “big bang”. While the big bang is considered as dynamic source led to dark energy, the rotation as dynamic source caused dark matter. So dark matter and dark energy are deducted to structural effects originated by the double dynamic sources. The model supplies two images which can be checked by observation or experiment: (1) there is changeable curvature evolution of time-space; (2) there is a rotated center in the universe.

1.5 Experiment Suggestion

Any theory needs observation or experiment for double check and support as

Theoretical derivatives ⇔ Observational evidence
For verifying the hypothesis above, I suggest to design a new experiment of vacuum explosion, for simulating cosmos evolution, as shown in Fig. 1.2.

After explosive is set in the center, vacuum pump can start to produce vacuum. When explosive starts to explode in vacuum, computer system may record the process of the vacuum explosion, as well as check and analyze the results.

In the experiment, rotation and exploration are human-controlled, so that it could help us to understand the double dynamic sources of the universe, in which it is expected to find the rotated distribution (vortex) and changeable curvature of the explosive.

1.6 Discussion: Cosmological Image

Historically, as a key prediction of the big bang cosmology, CMB was discovered in 1965. From that point on, the big bang cosmology was generally accepted, where the universe started in a hot, dense state and had been expanding over time. The rate of expansion depends on the types of matter and energy present in the universe, and in particular, whether the total density is above or below the so-called critical density. During the 1970s, most attention was focused on pure-baryonic models, but there were serious challenges explaining the formation of galaxies, given the small anisotropies in the CMB (upper limits at that time). In the early 1980s, it was realized that this could be resolved if CDM (cold dark matter) dominated over the baryons and the theory of cosmic inflation motivated models with critical density. During the 1980s, most researches focused on CDM with critical density in matter, around 95% CDM and 5% baryons: those showed success at forming galaxies and clusters of galaxies, but problems remained. Notably, the model required a Hubble constant lower than preferred by observations and the model under-predicted observed
large-scale galaxy clustering. Those difficulties were sharpened with the discovery of CMB anisotropy by COBE satellite in 1992 and several alternatives including ΛCDM (Lambda cold dark matter) and mixed cold plus hot dark matter came under active consideration. The ΛCDM model can be extended by adding cosmological inflation, quintessence and other elements that are current areas of speculation and research in cosmology and the model has become the standard following the observations of accelerating expansion since 1998. Much more precise measurements of the microwave background from Wilkinson microwave anisotropy probe (WMAP) in 2003–2010 have continued to support and refine the model.

Comparison of the ΛCDM model with observations is very successful on large scales (larger than galaxies, up to the observable horizon), but may have some problems on sub-galaxy scales, possibly predicting too many dwarf galaxies and too much dark matter in the innermost regions of galaxies. Those small scales are harder to resolve in computer simulations, so it is not yet clear whether the problem is the simulations, non-standard properties of dark matter or a more radical error in the model.

Now the cosmological standard model is mastered by three parts:

1. When the universe is supposed as perfect fluid, Einstein equations become

\[ T_{\mu\nu} = -pg_{\mu\nu} + (p + \rho)u_{\mu}u_{\nu} \]  

where \( T_{\mu\nu} \) is the energy and momentum tensor; \( g_{\mu\nu} \) is the metric tensor; \( p \) is the isotropic pressure; \( \rho \) is the energy density and \( u = (1, 0, 0, 0) \) is the velocity vector for the isotropic fluid in comoving coordinates, leading to Friedmann-Lemaitre equations

\[ H^2 = \left( \frac{\dot{R}}{R} \right)^2 = \frac{8\pi G \rho}{3} - \frac{k}{R^2} + \frac{\Lambda}{3} \]  

\[ \frac{\ddot{R}}{R} = \frac{\Lambda}{3} - \frac{4\pi G}{3} (\rho + 3p) \]  

where \( G \) is gravitational constant and \( \Lambda \) is cosmological constant.

2. Cosmological state equation, if the cosmos possesses total matter (energy) density \( \rho \) and isotropic pressure \( p \) are linked by

\[ \frac{d\rho}{dt} = -3H(\rho + p) = -3H\rho(1 + w) \]  

the cosmic state is determined by \( w = p/\rho \).

3. Robertson-Walker metric of space-time, which determined the cosmic structure

\[ ds^2 = dr^2 - R^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\varphi^2) \right] \]
where \( (t, r, \theta, \varphi) \) is comoving coordinate system and \( R(t) \) is cosmic scale factor, while \( k \) is curvature index and \( K = k/R^2 \) as space curvature, with \( k = 1, 0 \) or \(-1\) corresponding to closed \((K > 0)\), flat \((K = 0)\) or open \((K < 0)\) spatial geometries.

Equations (1.42)–(1.44) construct complete equations for variables \( R, \rho \) and \( p \).

After defining the critical density (when \( k = 0 \) and \( \Lambda = 0 \)) as
\[
\rho_c \equiv \frac{3H^2}{8\pi G} = 1.88 \times 10^{-26} h^2 (\text{kg} \cdot \text{m}^{-3}) = 1.05 \times 10^{-5} h^2 (\text{GeV} \cdot \text{cm}^{-3})
\]  
(1.46)
and scaling Hubble parameter \( h \) as
\[
H \equiv 100h(\text{km} \cdot \text{s}^{-1} \cdot \text{Mpc}^{-1})
\]  
(1.47)
via Eq. (1.43), the (1.42) becomes
\[
\rho_c = \rho - \frac{3k}{8\pi G R^2} + \frac{\Lambda}{8\pi G}
\]  
(1.48)
Introducing dimensionless density parameters for pressureless matter \( \Omega_m \) and vacuum \( \Omega_v \),
\[
\Omega_m = \frac{\rho}{\rho_c}, \quad \Omega_v = \frac{|\rho\Lambda|}{\rho_c}
\]  
(1.49)
where \( |\rho\Lambda| = \Lambda/(8\pi G) \) denotes vacuum energy density, i.e., dark energy. Then we obtain Friedmann equation as follows.
\[
\Omega_m + \Omega_v - 1 = \frac{k}{R^2 H^2}
\]  
(1.50)
Combining Eq. (1.43) with Eq. (1.44), we can also introduce a deceleration parameter as
\[
q_0 \equiv - \frac{\ddot{R}}{R^2} = - \frac{\ddot{R}}{RH^2} = \frac{1}{2} \frac{\Omega_m}{\Omega_r} + \Omega_r + \frac{(1 + 3w)}{2} \Omega_v
\]  
(1.51)
where \( \Omega_r \) is the density parameter of relativistic particles.

Since the universe is evolutional and develops from high temperature and high pressure to low temperature and low pressure, it is essential to probe into the cosmological state equation, which determines the cosmic dynamic evolution.

Comparing Eq. (1.40) with Eqs. (1.41)–(1.43), we see that the effect of dark energy is caused by scalar field \( p(\psi) \) and the effect of dark matter by vector field \( q(A) \), which just match that the \( p(\psi) \) is corresponding to a negative pressure and the \( q(A) \) is similar to a positive force.

If Riemann space-time is extended to Finsler space-time, the Clifford-Finslerian physical unification could be introduced and would produce more rich and colorful results in physical theories.
Although the universe is made out of mostly matter, the standard model predicts that matter and antimatter should have been created in (almost) equal amounts if the initial conditions of the universe did not involve disproportionate matter relative to antimatter. Yet, there is no mechanism that is sufficient to explain that asymmetry exists in micro-particle standard model. Now, Eqs. (1.9) and (1.10) imply the possibility while the linked-measure leads the symmetric or asymmetric linked-field, which provides a unified interpretation of micro-particles and macro-cosmos. As a hypothesized unification, the vacuum experiment is suggested to verify the hypothesis.

1.7 Concluding Remarks

Conclusively, the linked-field theory above is characterized as follows.

(1) A mathematical multi-vector consists of a complex scalar, a complex vector and a bi-vector, which constructs a physical linked-measure, where scalar mass, vector potential and bi-vector strength together generate linked-field.

(2) In the linked-field, strong symmetric links generate electromagnetic field and strong micro-inner links generate strong field, while weak micro-inner symmetric links synthesize electro-weak field. With adding outer space-time metric, the linked-field leads to gravitational field and a new understanding of dark matter and dark energy.

(3) Both dark matter and dark energy are structural effects of geometrical dynamics in the whole universe, which are equivalent to double dynamic sources driving the universe, in which one initialized the big bang (leads to dark energy) and the other pushed rotation (leads to dark matter).

With combining quantum field theory and general relativity, multi-vector linked-field did unify those two standard models together, which provided another way to consider physical unification, where physical achievements are kept and two standard models are saved. The linked-field model looks consistent, on which it is expected that the essence of the physical world can be revealed and the unified physics can be approached.

Furthermore, the choice of measures determines scientific relations, and thus the different measures would lead to different discoveries of scientific laws. When linked-measure based on the mathematical multi-vector is introduced, the linked-field for physical unification is generated, in which the micro-particle standard model and the macro-cosmos standard model can be combined and unified, and the phenomena of dark matter and dark energy are naturally interpreted. The linked-field leads to both mathematically and physically simple unified theory, which might stimulate further studies.

Acknowledgements This chapter is a revision of the original version published at Physical Journal, 2015, 1(2): 89–96.
References


Ye, F.Y.: Curvature mass inside hadrons: linking gravity to QCD. Natural Sci. 5(2), 182–186 (2013)
Chapter 2
The Physical Linked-Measure Works as Vortex with Linking to Turbulence

A physical linked-measure is mathematically consisted of a complex scalar, a complex vector and a bi-vector and is geometrically equivalent to a vortex. When the complex scalar means mass, the complex vector implies directed momentum and the bi-vector rotates angular momentum, with using the least action principle to the linked-measure, yielding energy-mass-momentum-angular momentum joint conservation. Hamilton equation and Lagrange equation are kept as the core of physics, leading fluid dynamics of energy way. As any periodic function can be expressed as a Fourier series, energy spectrum is suggested to be an analytical method. Combining the vortex dynamics with relativity and thermodynamics, Bekenstein-Hawking entropy and “no-hair theorem” of black hole are naturally derived. Applying to wingtip vortices, with adding wavelets, simplified ideal turbulence is described.

2.1 Introduction

During its glorious history, physics had contributed much knowledge of the human understanding to the universe. However, there are still some problems perplexing scientists, such as turbulence, which has been a difficult issue for more than a hundred years. It is well known that Navier-Stokes equations master the fluids including turbulence (Lesieur 1997; Ecke 2005) and the incompressible Navier-Stokes equations with conservative external field are the fundamental equations of fluid mechanics (Landau and Lifshits 1987; Davidson 2015). However, the Navier-Stokes equations are also kept as math-physical puzzle, while the equations act on the flow velocity in the case of incompressible homogeneous flows. Although the equations seem useful because they describe the physics of many models across weather change, ocean currents, water flow and air flow around the world, the solving Navier-Stokes equations still meet some difficult problems, whether in their full and simplified forms. There is an outstanding open problem in mathematics: how to determine an initial condition of the velocity field. There exists, in some senses, a unique solution of the Navier-Stokes equations starting with that initial condition and valid for all later
times (Rosa 2006). In the studies of turbulence, from a mathematical perspective, it is fundamental to develop a rigorous background upon which to study the physical quantities of a turbulent flow, as the mathematical theory is related to the deterministic nature of chaotic systems assumed in dynamical system theory and is believed to hold in turbulence.

Meanwhile, vortex (Saffman 1992; Nitsche 2006) is an interesting path to approach turbulence (Chow et al. 1997), since wingtip vortices could become wingtip turbulence and cause damages in practice. With the wavelet applications, the vortex and turbulence may be linked (Farge 1992; Farge and Schneider 2006).

Since physics focuses on the fundamental mechanism of the nature and the universe, physical theories have to explain all the realities of the world (Penrose 2004). In order to simplify the problems and find solutions, here I introduce the physical linked-measure for characterizing natural vortex with the aim of leading to a solution of ideal turbulence theoretically, using the mathematical multi-vector method (Hestenes 2003; Doran and Lasenby 2003; Lasenby et al. 2004).

### 2.2 Methodology

Using multi-vector $M_k$ ($k = 0, 1, 2, 3, 4$) as the total measure, where $M_k$ is a multi-vector of grade $k$. $k = 0$ corresponds to scalar, $k = 1$ to vector, $k = 2$ to bi-vector, $k = 3$ to pseudo-vector and $k = 4$ to pseudo-scalar and the Clifford bases of four-dimensional space-time are generated by four orthonormal vectors $\{\gamma^\mu, \mu = 0, 1, 2, 3\}$ and spanned by 1 (1 scalar at grade 0), $\{\gamma^\mu\}$ (4 vectors at grade 1), $\{\sigma^k, i\sigma^k\}$ (6 bi-vectors at grade 2), $\{i\gamma^\mu\}$ (4 pseudo-vectors at grade 3) and $i$ (1 pseudo-scalar at grade 4) orderly. Now we write the physical linked-measure as

$$M = M_0 + M_1 + M_2 + M_3 + M_4 = \varphi + V + B + iU + i\theta = \psi + A + B$$

(2.1)

where $\psi = \varphi + i\theta$ constructs a complex scalar function (massive function), while $A = V + iU$ forms a complex vector function (potential function) and $B = (1/2)B_{\mu\nu}\gamma^\mu \wedge \gamma^\nu$ as a unique bi-vector.

A physical linked-measure (Ye 2015) mathematically consists of a complex scalar, a complex vector and a bi-vector, which contain rich structural information. A linked-measure can be geometrically represented by a vortex, as shown in Fig. 2.1 (here is right-rotated system. Similarly, left-rotated system can be defined), where one scalar $\psi$, one vector $A$ and one bi-vector $B$ construct just one vortex.

Each $M$ has its space-time conjugation, denoted by $\overline{M}$, as follows (revised conjugation is $\overline{M} = (\overline{\psi}, \overline{A}, -\overline{B})$)

$$\overline{M} = -iM_i = M_0 - M_1 + M_2 - M_3 + M_4 = \psi - A + B = (\psi, -A, B)$$

(2.2)
2.2 Methodology

Fig. 2.1 Geometric representation of a linked-measure as vortex (*right-rotated* system)

![Diagram](image1.png)

Fig. 2.2 The linked-measure of $\overline{MM}$

When $M$ and $\overline{M}$ act to be $M\overline{M}$, leading to define $M^2 = |M\overline{M}|$, in which the two opposite directions offset so that scalar $\psi$ and bi-vector $B$ become main measure of $M^2$, as shown in Fig. 2.2.

In Fig. 2.2, if the right side denotes $M$ and left side means $\overline{M}$, $M$ and $\overline{M}$ construct right-rotated system, while $M$ and $\overline{M}$ construct left-rotated system when $A$ or $B$ indicates inverse direction in $M$ and $\overline{M}$. When time is put into the linked-measure, $M(s, t)$ is a fluid-type measure and a physical linked-measure $M(s, t)$ is related to space-time distribution of a complex scalar, a complex vector and a bi-vector, which is just suitable to apply into fluid dynamics.

The differential operators of one order derivatives can be introduced and defined as (Greek sub-indices $\mu, \nu$ denote 1, 2, 3, 4)

$$\partial_{\mu} = \frac{\partial}{\partial x_{\mu}}; \quad \nabla = \gamma^{\mu} \partial_{\mu} \quad (2.3)$$

Then we keep the concepts of energy and momentum and define linked-energy $E$ and linked-momentum $p_{\mu}$ with linking Hamilton function $H$ and Lagrangian function $L$ as follows.

$$H = \frac{\partial L}{\partial \dot{x}_{\mu}} \cdot x_{\mu} - L = p_{\mu} x_{\mu} - L = E(s, t) = \int_{s}^{e(t)} e(t) dt \quad (2.4)$$

$$L = p_{\mu} x_{\mu} - H; \quad p_{\mu} = \frac{\partial L}{\partial \dot{x}_{\mu}} \quad (2.5)$$
where $e(t)$ means density function of linked-measure in space. Now the energy-mass conservation is extended to energy-mass-momentum-angular momentum joint conservation.

So we obtain math-physical equations following Hamilton principle (Latin subindices $i, j$ denote 1, 2, 3)

$$\frac{\partial H}{\partial p_i} = \frac{ds_i}{dt}; \quad \frac{\partial H}{\partial s_i} = -\frac{dp_i}{dt}$$ \hspace{1cm} (2.6)

$$\delta \int L dt = 0$$ \hspace{1cm} (2.7)

The Eq. (2.6) is the differential form and Eq. (2.7) is the integral form of the linked-field equations, which fit analytical tradition of physics.

Corresponding to the vorticity $\omega$ of a flow field with velocity distribution $v$, $\omega = \nabla \times v$, the Helmholtz vortex laws ($\nabla \cdot \omega = 0$) can be simplified to approach as

$$\partial \mu M \to 0$$ \hspace{1cm} (2.8)

Differentiating from Navier-Stokes equations, where flow velocity is mainly concerned, the linked-measure focuses on energy, leading directly to fluid dynamics of energy way, while it also bypasses the difficulties of solving the Navier-Stokes equations.

Recalling $\psi = \varphi + i\theta$ and $A = V + iU$, both $\psi$ and $A$ are complex functions, which can be rewritten as

$$\psi = \varphi + i\theta = re^{i\alpha} = r(\cos \alpha + i \sin \alpha)$$ \hspace{1cm} (2.9)

$$A = V + iU = Re^{i\beta} = R(\cos \beta + i \sin \beta)$$ \hspace{1cm} (2.10)

where ($r, \alpha$) and ($R, \beta$) are polar coordinates.

Let’s mention the complex potential $A$, which satisfies the Cauchy-Riemann equations

$$\frac{\partial V}{\partial x} = \frac{dU}{dy} = v_1$$ \hspace{1cm} (2.11)

$$\frac{\partial V}{\partial y} = -\frac{dU}{dx} = v_2$$ \hspace{1cm} (2.12)

where $v_1$ and $v_2$ can be understood as velocities, so that $A$ is just the velocity potential.

The method above also supplies a generalized methodology of mathematical physics (Ye 2009).
2.2 Methodology

In general, any periodic function \( g \) with a period of \( 2l \), i.e., \( g(x) = g(x + 2l) \), can be expressed as a Fourier series as

\[
g(x) = \frac{1}{2}a_0 + \sum_{n=0}^{\infty} (a_n \cos k_n x + b_n \sin k_n x) \quad (2.13)
\]

where \( x \) is a spatial coordinate and \( k_n = n \pi / l \) has been called the wavenumber. The Fourier coefficients are given by

\[
a_n = \frac{1}{l} \int_{-l}^{l} g(x) \cos k_n x \, dx \quad (2.14)
\]

\[
b_n = \frac{1}{l} \int_{-l}^{l} g(x) \sin k_n x \, dx \quad (2.15)
\]

and Parseval identity holds

\[
\int_{-l}^{l} g^2(x) \, dx = \frac{l}{2} a_0^2 + l \sum_{n=0}^{\infty} (a_n^2 + b_n^2) \quad (2.16)
\]

Then the total energy can be obtained by integrating over the whole wavenumber space

\[
\int_{0}^{\infty} E(k) \, dk = l \sum_{n} g^2(k_n) \quad (2.17)
\]

It is shown that \( E(k) \) characterizes the energy, so that it is called as the energy spectrum, which provides the methodology of energy spectrum analysis.

2.3 Vortex Dynamics and Black Hole

Since a vortex can be geometrically represented by a linked-measure, the linked-measure matches following dynamic equation via Hamilton principle

\[
\frac{\partial H}{\partial s_i} = \partial_{\mu} M = - \frac{\partial p_i}{\partial t} \quad (2.18)
\]

Considering that Hamiltonian energy \( H = E = kT \) in thermodynamics, where \( k \) is Boltzmann constant and \( T \) denotes temperature, it yields

\[
\frac{\partial E}{\partial s_i} = k \frac{\partial T}{\partial s_i} = - \frac{\partial p_i}{\partial t} \quad (2.19)
\]
Replacing flow-velocity way characterized by Navier-Stokes equations, Eqs. (2.18) and (2.19) provide the fluid dynamics of energy way, which could introduce another effective analysis for vortex movement, such as black hole, as shown in Fig. 2.3.

Visually, the process of a vortex rotates into the black hole can be simply viewed as rotated energy measure, in which $B$ approaches to 0 but the total energy measure $M$ keeps conservation when $A$ and mass concentrate to the center of the black hole. Consequently, the center of the black hole forms a strong energy field, absorbing everything. However, all observed black holes in astronomy belong to universal vortices, not time-space singular points.

The theory of general relativity predicts that a sufficiently compact mass can deform space-time to form a black hole and the quantum field theory in curved space-time predicts that event horizons emit Hawking radiation, with the same spectrum as a black body of a temperature inversely proportional to its mass. The simplest static black holes have mass but neither electric charge nor angular momentum, which just fits the linked-measure with $B \to 0$.

However, following the no-hair theorem, a black hole achieves a stable condition after formation and it has only three independent physical properties: mass, charge and angular momentum. Any two black holes that share the same values of those properties, or parameters, are indistinguishable according to classical (non-quantum) mechanics. Those properties are special because they are visible outside a black hole.

Recalling the energy-mass-momentum relation in relativity

$$E^2 = m^2 c^4 + p^2 c^2$$

we mention that the total energy should include the angular momentum part, so that we should have the form

$$E^2 = \psi^2 (m) + A^2 (p) + B^2 (J)$$
According to “the first law of black hole mechanics” (Frolov and Novikov 1997; Ashtekar 2006; Papantonopoulos 2009),

$$dM = \frac{\kappa}{8\pi G} dA + \phi dQ + \omega dJ$$ (2.22)

where $\kappa$ is the surface gravity of the black hole and $Q$ is electric charge.

That is identical to the first law of thermodynamics

$$dE = \theta dS + \phi dQ + \omega dJ$$ (2.23)

where $\theta$, $\phi$ and $\omega$ are black hole parameters, $\phi$ being the electrostatic potential at the horizon, and the thermodynamic entropy $S$ is introduced as

$$dS = \frac{dE}{T}$$ (2.24)

Since $M = E$, combining Eqs. (2.22) and (2.23) yields the Bekenstein-Hawking entropy when $\theta = \kappa/(8\pi G) = c^3 k/(4G\hbar)$

$$S = \frac{c^3 k}{4G\hbar} A$$ (2.25)

where $A$ is the horizon area of the black hole and constants are the speed of light $c$, the Boltzmann constant $k$, Newton’s constant $G$ and the reduced Planck constant $\hbar$.

When $\psi(m) = \theta m$, $A(p) = \phi p$ and $B(J) = \omega J$, we derive

$$E^2 = (\theta m)^2 + (\phi p)^2 + (\omega J)^2$$ (2.26)

and

$$dE = \theta dm + \phi dp + \omega dJ$$ (2.27)

Considering the potential, $A(p) = \phi p$, if $\phi$ comes from electromagnetic interaction $\phi-Q$, we approach Eq. (2.28) in Planck units

$$Q^2 + \left(\frac{J}{m}\right)^2 + 1 = E^2 = M^2$$ (2.28)

When $m \approx M$, it leads to famous “no-hair theorem”

$$Q^2 + \left(\frac{J}{M}\right)^2 \approx M^2$$ (2.29)

where the black hole mass $M$ equals to its total energy $E$. Black holes saturating that equation are called extremal. Solutions of Einstein’s equations that violate that equation exist, but they do not possess an event horizon. Those solutions have so-
called naked singularities that can be observed from the outside, and hence are
deemed unphysical. The cosmic censorship hypothesis rules out the formation of
such singularities, when they are created through the gravitational collapse of realistic
matter.

As there are many vortices in the universe, including atmospheric vortices, water
vortices and galaxies vortices, it is believed that the vortex dynamics simplified by
linked-measure will be useful in physics.

### 2.4 Wingtip Vortices

Wingtip vortices are sometimes named contrails or lift-induced vortices because they
also occur at points other than at the wingtips. Indeed, vorticity is trailed at any point
on the wing where the lift varies span-wise (a fact described and quantified by the
lifting-line theory) and it eventually rolls up into large vortices near the wingtip,
at the edge of flap devices, or at other abrupt changes in wing planform. Three-
dimensional lift and the occurrence of wingtip vortices can be approached with the
concept of horseshoe vortex and described accurately with the Lanchester-Prandtl
theory (Anderson 2001). In that point of view, the trailing vortex is a continuation of
the wing-bound vortex inherent to the lift generation.

When viewed from the tail of the aircraft, looking forward in the direction of
flight, there is one wingtip vortex trailing from the left-hand wing and circulating
clockwise (M1), and the other trailing from the right-hand wing and circulating anti-
clockwise (M2). The result is a region of downwash behind the aircraft, between
those two vortices, as shown in Fig. 2.4.

Wingtip vortices are associated with induced drag, the imparting of downwash,
and are a fundamental consequence of three-dimensional lift generation. Careful
selection of wing geometry (in particular, aspect ratio), as well as of cruise conditions,
is one of the most important designing and operational methods to minimize induced
drag.

As the aircraft will keep symmetric mass $m$ during the flight, we must have
$M1=M2$. Supposing $E_1(k)$ and $E_2(k)$ as the energy spectra of M1 and M2 respectively,
using the method of energy spectrum characterized by Eq. (2.17), the equilibrium
condition of those two wingtip vortices becomes

$$
\int_0^\infty E_1(k)dk = \int_0^\infty E_2(k)dk = l \sum_n g^2(k_n)
$$  \hspace{1cm} (2.30)

where $g(x)$ is a periodic function that can be expressed by Fourier series Eq. (2.13),
which describes a stable scattering and circulating wingtip vortices in the tail of the
aircraft.
If there is no crosswind, those two wingtip vortices do not merge because they are circulating in opposite directions. They dissipate slowly and linger in the atmosphere long after the airplane has passed, approaching to laminar flow at last.

However, if there is crosswind, they are hazards to other aircrafts, known as wake turbulence, i.e., the wingtip vortices form the primary component of wake turbulence.

The wingtip turbulence occurrence is related to some conditions of aerodynamics and thermodynamics. Depending on ambient atmospheric humidity as well as the geometry and wing loading of aircraft, water may condense or freeze in the core of the vortices, making the vortices visible. When a wing generates aerodynamic lift, the air on the top surface has lower pressure relatively to the bottom surface. Air flows below the wing and out around the tip to the top of the wing in a circular fashion, and the pressure on the top of the wing is lower than that on the bottom, causing air to move around the edge of the wing from the bottom surface to the top, so that aerodynamic issues meet thermodynamic solutions.

In the situation of low pressure, vortex cores are regions. As a vortex core begins to form, the water in the air (i.e., in the region that is about to become the core) is in vapor phase, which means that the local temperature is above the local dew point. After the vortex core forms, the pressure inside it has decreased from the ambient value and so the local dew point $T_c$ has dropped from the ambient value. Thus, in and of itself, a drop in pressure would tend to keep water in vapor form, where the initial dew point was already below the ambient air temperature and the formation of the vortex has made the local dew point even lower. However, as the vortex core forms, its pressure (and its dew point) is not the only property that is dropping. In other words, the vortex-core temperature is also dropping and in fact, it can drop by much more than the dew point does. Approximately, the formation of vortex cores is thermodynamically an adiabatic process. To put it simply, there is no exchange of heat. In such a process, the drop in pressure is accompanied by a drop in temperature, according to the equation (Green 1995)

$$\frac{T_f}{T_i} = \left(\frac{P_f}{P_i}\right)^{\gamma - 1}$$

(2.31)

where $T_i$ and $P_i$ are the absolute temperature and pressure at the beginning of the process (here, equal to the ambient air temperature and pressure); $T_f$ and $P_f$ are the
absolute temperature and pressure in the vortex core (which is the end result of the process) and the constant $\gamma$ is $7/5 = 1.4$ for air.

Thus, even though the local dew point inside the vortex cores is even lower than that in the ambient air, the water vapor may nevertheless condense, if the formation of the vortex brings the local temperature below the new local dew point.

Noticing that there are four most common Maxwell relations among temperature $T$, pressure $P$, volume $V$ and entropy $S$ as

$$ \left( \frac{\partial T}{\partial V} \right)_S = - \left( \frac{\partial P}{\partial S} \right)_V ; \left( \frac{\partial T}{\partial V} \right)_P = - \left( \frac{\partial P}{\partial S} \right)_T ; $$

$$ \left( \frac{\partial T}{\partial p} \right)_S = \left( \frac{\partial V}{\partial S} \right)_P ; \left( \frac{\partial T}{\partial p} \right)_V = \left( \frac{\partial V}{\partial S} \right)_T $$

(2.32)

We can also replace $P$ with $S$ in same space

$$ \frac{T_f}{T_i} = \left( \frac{S_f}{S_i} \right)^{\frac{\gamma - 1}{\gamma}} $$

(2.33)

So the temperature, energy and entropy are linked together. That is an issue concerning fluid mechanics and thermodynamics and that is only a simple preliminary exploration. In the case of wingtip vortices leading turbulence, it is important that the local temperature would be lower than the local dew point, so that the water vapor inside the vortices would indeed condense. In that circumstance, the local temperature in vortex cores may drop below the local freezing point, in which case ice particles will form inside the vortex cores.

### 2.5 Crosswind as Wavelets Leading to Ideal Turbulence

Wingtip turbulence indicates that there is a relation between vortex and turbulence (Pullin and Saffman 1998), within a fuzzy layer mixing vortex and turbulence, which can be simulated by introducing wavelets (Speziale 1998).

A wavelet is a wave-like oscillation with changed amplitude that begins at zero, increases and then, decreases back to zero. As a mathematical tool, wavelets can be used to extract information from many different kinds of data, including audio signals and video images. Sets of wavelets are generally needed to analyze data fully. A set of “complementary” wavelets will decompose data without gaps or overlap so that the decomposition process is mathematically reversible. Thus, sets of complementary wavelets are useful in wavelet based compression/decompression algorithms where it is desirable to recover the original information with minimal loss.

Instead of solving the Navier-Stokes equations over a highly refined mesh, we can use the wavelet decomposition of a low-resolution simulation to determine the loca-
2.5 Crosswind as Wavelets Leading to Ideal Turbulence

![Fig. 2.5](image)

Vortex meets wavelet leading to turbulence

We then synthesize those missing components using a novel incompressible turbulence function and provide a method to maintain the temporal coherence of the resulting structures.

Visually, the crosswind acts onto bi-vector $B$, so that $B$ could be changed by wavelets and the energy is kept in $\psi$ and $A$. However, a part of the energy is scattered and disappeared by crosswind (wavelet), so that the total energy may be dissipated, then the energy spectrum is changed by wavelets, as shown in Fig. 2.5.

If the crosswind takes less energy than vortex itself, the wavelet vorticity $\omega = \nabla \times v$ looks weak, and much smaller than $B$, which produces stable fluid.

If the crosswind takes more energy as much as vortex itself, the wavelet vorticity $\omega = \nabla \times v$ shows the same as, even stronger than $B$, which leads to turbulence.

Generally, if the wavelets are so strong that cause the energy change of the vortex ($\nabla \times v \geq B$), the turbulence will occur. If the wavelets are weak ($\nabla \times v \ll B$), the vortex will keep stable dissipation and then merge into laminar flow.

Supposing the vortex has energy density $g(x)$, the results lead to

$$g(x) = \int_{\mathbb{R}^n} \tilde{g}(x)f^*(x)dx \quad (2.34)$$

The inner products $g(x)$ and $f^*(x)$ give wavelet coefficients

$$\tilde{g}(x) = \int_{\mathbb{R}^n} \hat{g}(k)f^*(k)dk \quad (2.35)$$

where $\mathbb{R}^n$ marks the real space.

Now the original energy spectrum $E(k)$ of $g(x)$ is changed by wavelets’ one. Supposing $f^*(x)$ has energy spectrum $E^*(k)$. If $E^*(k)$ keeps much weaker corresponding to $E(k)$, the stable vortices can be kept and turbulence will not happen. However, if $E^*(k)$ has near equivalent strength at the same level of $E(k)$, the equilibrium will be destroyed and the turbulence will occur. So, we give the conditions of introducing the ideal turbulence as

$$E^*(k) \sim E(k); \quad \omega = \nabla \times v \geq B \quad (2.36)$$

Equation (2.36) means that wavelet energy spectrum $E^*(k)$ approaches the same level as vortex energy spectrum $E(k)$ and the wavelet vorticity $\omega = \nabla \times v$ shows dominance to vortex angular momentum measure $B$. In those conditions, we meet an ideal turbulence, which causes changes of the energy spectrum and structure of
the vortex or vortices, leading to turbulence. Then the energy (or power) spectrum analysis provides a simplified methodology for studying fluids and turbulence, which is expected to introduce a new idea and develop useful analysis.

2.6 Discussion and Conclusion

In fluid mechanics or physics, there is still no perfect definition of turbulence. However, three main characteristic features can be listed as follows.

(1) Nonlinearity or irregularity. Differentiating from linear laminar flow, turbulence is nonlinear and random flow, which consists of a spectrum of different scales (i.e., irregularity). However, although turbulence seems chaotic, it is deterministic and can be studied via the Navier-Stokes equations or the energy spectra.

(2) Dissipation and diffusivity. Turbulent dissipation means that kinetic energy in the small dissipative vorticities is transformed into thermal energy. The small vorticities receive the kinetic energy from slightly larger vorticities and the slightly larger vorticities receive their energy from even larger vorticities and vice versa. The largest vorticities extract their energy from the mean flow. That process of transferring energy from the largest turbulent scales to the smallest is called the cascade process, in which the diffusivity increases. The increased diffusivity also increases the resistance and heat transfer in internal flows.

(3) Three-dimensional vorticity. Turbulence is also characterized by apparently random and chaotic three-dimensional vorticity, which usually dominates all other phenomena and results in increased energy dissipation and heat transfer and also increases the exchange of momentum in boundary layers.

The Reynolds number \( (Re) \) is not mentioned in this chapter. As the fluid dynamics of energy way and the energy spectrum analysis are not directly related to flow velocity, the \( Re \) that links to flow velocity is ignored. However, it keeps effective that turbulent flow occurs at high \( Re \), such as the transition to turbulent flow in pipes occurs at \( Re \simeq 2,300 \) and in boundary layers at \( Re \simeq 500,000 \).

When a physical linked-measure is geometrically equivalent to a vortex, the methodology concentrates on the energy and the energy-mass conservation is extended to energy-mass-momentum-angular momentum joint conservation. By using the least action principle to the linked-measure, Hamilton equation and Lagrange equation are kept as the core of physics, leading fluid dynamics of energy way. Meanwhile, the energy (or power) spectrum is suggested to be the main analytical method, which looks useful for probing into the unified mechanism of fluids. Adding wavelets, simplified ideal turbulence can be described, which could introduce a new solution to approach turbulent flow.

Conclusively, the linked-measure as a simplified methodology may benefit for understanding vortex as well as turbulence, with focusing on energy construction and energy spectrum analysis, which could stimulate further considerations and studies.

Acknowledgements This chapter is a revision of the original version published at Physical Journal, 2015, 1(3): 209–215
References

Chapter 3
The Clifford-Finslerian Linked-Field Leads Branching Multiverse

Focusing on the issue of multiple universes (multiverse), the physical linked-measure contributes the Clifford-Finslerian linked-field, which generates branching multiverse. While Clifford algebra supplies interior dynamics for generating inner branching structure, Finsler geometry provides catastrophic branches of space-time metric and curvature, as exterior dynamics do. The branching multiverse is different from countless multiverse, as the branches are deterministic and based on mainstream, where all branches integrate and produce limited multiverse. This chapter makes contributions by enabling the multiverse to get rid of puzzles from a non-scientific term and then return to the correct path for approaching real physics. Meanwhile, being in line with the tradition of analytical mechanics, the wave-particle duality is clearly interpreted at both micro-level and macro-level and the cosmological model is suggested to verify its curvature change.

3.1 Introduction

After Everett contributed his relative state (Everett 1957) and DeWitt subsequently introduced the “many-worlds interpretation (MWI)” (DeWitt and Graham 1973) to describe a complete measurement history of an observer, the idea of multiverse had gradually developed into a popular concept in the academic world (Steinhardt 2014), though it never became real physics. Physically, Everett’s relative state formulation makes two assumptions. The first one is that the wave function is not a simple description of the object’s state, but it actually is entirely equivalent to the object. That claim has much in common with some other interpretations. The second one is that observation or measurement has no special laws or mechanics, unlike the Copenhagen interpretation that considers the wave function which collapses as a special kind of event occurring as a result of observation. Instead, measurement in the relative state formulation is the consequence of a configuration change in the memory of an observer described by the same basic wave physics as the object
being modeled. When one models an isolated quantum system subject to external observation, one could mathematically model an object as well as its observers as purely physical systems.

Following Everett’s original work, there have appeared a number of similar formalisms as the many-worlds or multiverse interpretation, where one had referred to the combined observer-object system as being split by an observation, each split corresponding to the different or multiple possible outcomes of an observation. The multiverse had been hypothesized in astronomy, cosmology, physics, philosophy, religion, as well as in science fiction and even fantasy, with using various terms such as “parallel universes” “alternate universes” “quantum universes” “parallel worlds” “parallel dimensions” “alternate realities” “alternate timelines” “interpenetrating dimensions” “dimensional planes” and so on. Actually, the multiverse is only a hypothesis, in which there are a set of infinite or finite possible universes that together comprise everything that exists.

In representative multiverse “models”, it is necessary to mention Tegmark’s four levels (Tegmark 2009) and Greene’s nine types (Greene 2011).

The four-level “model” gives a taxonomy of universes beyond the familiar observable universe, where the four levels are arranged so that subsequent levels can be understood to encompass and expand upon previous levels, including level I (where a generic prediction of chaotic inflation is an infinite ergodic universe, which, being infinite, must contain Hubble volumes realizing all initial conditions), level II (where there exist “bubble universes” and each bubble universe has its special physical constants. Those different bubbles cover Universe 1 to Universe 6 and they have different physical constants. Our universe is just one of the bubbles), level III (where multiverse does not have more possibilities in the Hubble volume than a level I-II multiverse and “multiverse equals quantum many-worlds”) and level IV (which is based on a mathematical universe hypothesis. That level considers equally all real universes that can be described by different mathematical structures as ultimate ensemble).

The nine-type “model” informs us there are nine parallel universes: (1) Quilted multiverse, which works only in an infinite universe. In that universe, every possible event may occur an infinite number of times, with an infinite amount of space. However, the speed of light prevents us from being aware of those other identical areas. (2) Inflationary multiverse, which is composed of various pockets where inflation fields collapse and new universes form. (3) Brane multiverse, which follows $M$-theory and states that our universe is a three-dimensional brane that exists with many others on a higher-dimensional brane or “bulk”. Particles are bound to their respective branes except for gravity. (4) Cyclic multiverse, which has multiple branes (each is a universe) that collided, causing the big bang. The universes bounce back and pass through time, until they are pulled back together and again collide, destroying the old contents and creating them anew, via the ekpyrotic scenario. (5) Landscape multiverse, which relies on string theory’s Calabi-Yau manifolds. Quantum fluctuations drop the manifolds to a lower energy level, creating a pocket with a different set of laws from the surrounding space. (6) Quantum multiverse,
3.1 Introduction

which creates a new universe when a diversion in event occurs, as in the many-worlds interpretation of quantum mechanics. (7) Holographic multiverse, which is derived from the theory that the surface area of a space can simulate the volume of the region. (8) Simulated multiverse, which exists on complex computer systems that simulate the entire universe. (9) Ultimate multiverse, which contains every mathematically possible universe under different laws of physics. Certainly, all things in different universes never meet each other.

However, the physics community continues to fiercely debate the multiverse hypothesis. Serious concerns have been raised about whether attempts to exempt the multiverse from experimental verification may erode public confidence in science and ultimately damage the nature of fundamental physics. Although there are some active proponents such as Greene and Tegmark, there are also many skeptics including Penrose (2004) and Weinberg (2007), who disagree about whether the multiverse may exist and whether it is even a legitimate topic of scientific inquiry.

Now, under the framework of multi-vector methodology (Hestenes 2003; Doran and Lasenby 2003), the multiverse would become deterministic and could be characterized by branching, if the Clifford algebraic structure and Finsler geometry provided the real description of the physical universe via the linked-measure and the linked-field (Ye 2009, 2015b), leading to branching construction (Ye 2013). Here we formulate a methodological approach for characterizing the branching multiverse.

3.2 The Clifford-Finslerian Construction

At first, let’s consider Clifford algebraic branching structure and Finsler geometric catastrophic structure respectively.

3.2.1 Clifford Algebra as Interior Dynamics

Supposing the space-time point $x$ be based on Dirac frame $\{\gamma^\mu, \mu = 0, 1, 2, 3\}$, there are

$$x = x_\mu \gamma^\mu, \quad x_\mu = \gamma^\mu x$$  \hspace{1cm} (3.1)

The coordinates’ transformation will be

$$x_\mu \rightarrow x'_\mu = \alpha^\nu_\mu x_\nu; \quad \gamma^\mu \rightarrow \gamma'^\mu = \alpha^\nu_\mu \gamma^\nu$$  \hspace{1cm} (3.2)

where four Dirac matrices are viewed as four orthonormal basis vectors for four-dimensional Riemann-Finslerian space-time and $\alpha^\nu_\mu \alpha^\nu_\lambda = \delta^\mu_\lambda$, where $\gamma^0$ is time-like vector and $\gamma^k (k = 1, 2, 3)$ space-like vectors. Similarly, the three Pauli matrices...
σ^k = (σ^1, σ^2, σ^3) are viewed as three orthonormal basis vectors for three-dimensional Euclidean space.

For a multi-vector \( M_k \) \((k = 0, 1, 2, 3, 4)\), where \( M_k \) is a multi-vector of grade \( k \). \( k = 0 \) corresponds to scalar, \( k = 1 \) to vector, \( k = 2 \) to bi-vector, \( k = 3 \) to pseudo-vector and \( k = 4 \) to pseudo-scalar and the bases of four-dimensional space-time is generated by four orthonormal vectors \{\( \gamma^\mu \), \( \mu = 0, 1, 2, 3 \}\) and spanned by \( 1(1 \text{ scalar at grade } 0) \), \{\( \gamma^\mu \)\} \( (4 \text{ vectors at grade } 1) \), \{\( \sigma^k \), \( i\sigma^k \)\} \( (6 \text{ bi-vectors at grade } 2) \), \{\( i\gamma^\mu \)\} \( (4 \text{ pseudo-vectors at grade } 3) \) and \( i(1 \text{ pseudo-scalar at grade } 4) \) orderly.

Now we can write a multi-vector as

\[
M = M_0 + M_1 + M_2 + M_3 + M_4 = \varphi + V + B + iU + i\theta = \psi + A + B = (\psi, A, B) \quad (3.3)
\]

where \( \psi = \varphi + i\theta \) constructs a complex scalar function (mass-like function), while \( A = V + iU \) forms a complex vector function (particle-like function) and \( B = (1/2)B_{\mu\nu}\gamma^\mu \wedge \gamma^\nu \) as a unique bi-vector (span-like function).

The space-time conjugation of \( M \) is \( \bar{M} \).

\[
\bar{M} = -iMi = M_0 - M_1 + M_2 - M_3 + M_4 = \psi - A + B = (\psi, -A, B) \quad (3.4)
\]

The revised conjugation is denoted as \( \tilde{M} \).

\[
\tilde{M} = \tilde{M}_0 + \tilde{M}_1 + \tilde{M}_2 + \tilde{M}_3 + \tilde{M}_4 = \varphi + V - B - iU + i\theta = (\psi, \bar{A}, -B) \quad (3.5)
\]

And \( M \) can be divided as even part \( M_+ \) and odd part \( M_- \) as follows.

\[
M_+ = \varphi + B + i\theta = \psi + B = (\psi, B) \quad (3.6)
\]

\[
M_- = V + iU = A \quad (3.7)
\]

Those give asymmetry algebraic structure. If \( M \) is divided by following two parts

\[
Q_+ = \varphi + A = (\varphi, A) \quad (3.8)
\]

\[
Q_- = B + i\theta = (B, \theta) \quad (3.9)
\]

Those provide another asymmetry structure, where \( A \) represents a three-dimensional vector and the set of four-unit multi-vectors become \( \{1, i, j, k\} \), with \( i^2 = j^2 = k^2 = ijk = -1 \) and \( ij = -ji = k, jk = -kj = i, ki = -ik = j \), making up a basis for \( (\varphi, A) \). The sub-algebra \( Q_+ \) constructs a linear space of four dimensions and its elements just are quaternions, and quaternions are spinors. Quaternions and spinors have equivalent algebraic properties as well as the same geometric significance, which all belong to multi-vectors \( M \).
When we view the $M$ as a root, $M_+, M_-, Q_+$ and $Q_-$ look like main stems, which may lead to interesting algebraic branch structure and produce branching multiverse.

Here, it can be supposed that $R$, $C$ and $H$ denote respectively real, complex and quaternion fields. For a complex $Z = (X, X')$ with $X \in R$, record $\Re(Z) = X$ and $\Im(Z) = X'$, called the real part and imaginary part respectively, so that a quaternion $Q$ has its dual complex construction as

$$Q = (Z, Z') = q_0 + q_1i + q_2j + q_3k \subset M$$

with $Z \in C$, record $\Sa(Q) = Z$ and $\Pu(Q) = Z'$,

As scalar $q_0 = \varphi$ and the vector $(q_1, q_2, q_3) = A$, forming the scalar part and vector or pure quaternion part respectively, a quaternion $Q$ has also its scalar-vector construction $(\varphi, A)$, called Hamilton presentation, as follows.

$$Q = (\varphi, A) = \varphi + A \subset M$$

And its conjugation is

$$\overline{Q} = (\overline{Z}, -Z') = (\varphi, -A) \subset \overline{M}$$

with $\overline{Q} = Q$ and norm $Q\overline{Q} = \overline{Q}Q = |Q|^2 = \sum_{i=0}^{3} q_i^2$.

Meanwhile, a quaternion also has an equivalent scalar-vector construction which is called the Pauli representation as

$$Q = (\varphi, -A) = \varphi^0\sigma_0 - A^1\sigma_1 - A^2\sigma_2 - A^3\sigma_3$$

We see that Eq. (3.13) is the same as Eq. (3.12), the conjugation of Eq. (3.11), as $Q\sigma_i\overline{Q} = \sigma_i$, which means that the Pauli representation and Hamilton representation become conjugations each other. The algebraic structure shows that the Hamilton representation and Pauli representation exist naturally for a quaternion, which constructs a conjugation pair.

Then we have

$$\Sa(Q) = \frac{1}{2}(Q + \overline{Q}) \in R$$

and

$$\Pu(Q) = \frac{1}{2}(Q - \overline{Q}) \in P$$

where $H = R \oplus P$ and $P$ constructs the 3-dimensional Euclidean vector space.
When we do the multiplication and analysis of quaternion (Deavours 1973), there will generate a Cayley-Dickson branch and a Euclid-Grassmann branch further, in which the Cayley-Dickson branch is generated by the multiplication of quaternion \( Q_i \) and quaternion \( Q_j \) with the form of the dual complex function form as

\[
Q_i Q_j = (Z_i, Z'_i)(Z_j, Z'_j) = (Z_i Z_j - Z'_i Z'_j, Z_i Z'_j + Z'_i Z_j)
\]  

(3.16)

And the Euclid-Grassmann branch is produced by the multiplication of quaternion \( Q_i \) and quaternion \( Q_j \) with the form of scalar-vector form as

\[
Q_i Q_j = (\varphi_i, A_i)(\varphi_j, A_j) = (\varphi_i \varphi_j - A_i \cdot A_j, \varphi_i A_j + A_i \varphi_j + A_i \wedge A_j)
\]  

(3.17)

or

\[
Q_i * Q_j = (\varphi_i, A_i) * (\varphi_j, A_j) = (\varphi_i \varphi_j + A_i \cdot A_j, \varphi_i A_j - A_i \varphi_j - A_i \wedge A_j)
\]  

(3.18)

where \( \wedge \) is Wedge product and \( a \wedge b = -b \wedge a \).

Since Eqs. (3.17) and (3.18) co-exist, we have

\[
Q_i Q_j = (Z_i Z_j - Z'_i Z'_j, Z_i Z'_j + Z'_i Z_j) = (\varphi_i \varphi_j - A_i \cdot A_j, \varphi_i A_j + A_i \varphi_j - A_i \wedge A_j)
\]  

(3.19)

Both the Cayley-Dickson branch and the Euclid-Grassmann branch belong to the Clifford algebraic branches, which are driven by interior sources, so that the Clifford algebra supplies inner branching structure as interior dynamics, which looks like generic action and establishes the foundations of branching.

### 3.2.2 Finsler Geometry and Catastrophic Theory Leads Space-Time Branching

In Finsler geometry, according to Chern’s analysis (Chern 1996), there is a Chern connection form \( \omega^j_i \) in Finsler bundle \( p^*TM \rightarrow PTM \) as the unique solution of structural equations

\[
d\omega^j = \omega^j \wedge \omega^j
\]  

(3.20)

and

\[
\omega_{ij} + \omega_{ji} = -2A_{ijk} \omega^k_m
\]  

(3.21)

where \( \omega_{ij} = \omega^k_i \delta_{kj} \), and \( A = A_{ijk} \omega^j \otimes \omega^j \otimes \omega^k \) is a Cartan tensor. The Finsler structure becomes Riemann structure if \( A = 0 \). And the curvature of the Chern connection is
\[ \Omega^i_k = \frac{1}{2} R^i_{kjl} \omega^j \wedge \omega^l + P^i_{kjn} \omega^j \wedge \omega^n_m = \Omega(R) + \Omega(P) \]  
\[ (3.22) \]

where \( R \)-part is horizon-horizon (1st Chern) curvature and \( P \)-part is horizon-vertical (2nd Chern) curvature. When there is no \( P \)-part, a Finsler curvature becomes a Riemann curvature.

For a function \( f \), there is its Hessian matrix as

\[ H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_i \partial y_j} \end{bmatrix} \]  
\[ (3.23) \]

If \( \det H = 0 \), there is \( \nabla f = 0 \). If \( \det H \neq 0 \), there is \( f = M^2_i \), i.e., Morse saddle.

Referring catastrophic theory (Zeeman 1977; Thom 1989), Finsler space-time metric function can be defined as

\[ f(t, x, y, z) = (ct)^4 + x^4 + y^4 + z^4 - 2(ct)^2 x^2 + 2y^2 z^2 \]  
\[ (3.24) \]

In its critical point, there is

\[ f = \frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0 \]  
\[ (3.25) \]

which produces two branches \( x = \pm ct \) and \( t = x = y = x = 0 \), leading to following two parts of the stable Hessian sub-matrices with

\[ \det H(t, x) = -48[(ct)^4 + x^4 - 2(ct)^2 x^2] = 0 \]  
\[ (3.26) \]

\[ \det H(y, z) = 48(y^2 + z^2)^2 \neq 0 \]  
\[ (3.27) \]

so that the space-time concludes four parts: \( (ct)^2 - x = 0 \), which denotes critical plane; \( (ct)^2 - x > 0 \), which indicates time-like area; \( (ct)^2 - x < 0 \), which marks space-like area; \( t = x = 0 \), which is zero point.

Then the Finslerian space-time metric \( ds \) is characterized by the Lorentz-Cao formula (Cao 2001)

\[ ds^4 = c^4(1 - \beta^2)^2 dt^4, \quad \beta = v/c \]  
\[ (3.28) \]

where we see that there is a fractal factor within a typical catastrophic branching

\[ B = \lambda(\alpha^2 - \beta^2)^2 \]  
\[ (3.29) \]

If \( \lambda \neq 0 \), Eq. (3.29) also includes four branches as

\[ (\alpha^2 - \beta^2)^2 = (\alpha - \beta)^2(\alpha + \beta)^2 \]  
\[ (3.30) \]
Mathematically, the catastrophe theory as a branch of bifurcation theory in dynamical systems just gives a special case of more general singularity theory in geometry, which constructs geometric bifurcation, leading to space-time branching.

The Clifford-Finslerian branching structure provides a general mathematical structure for multiverse, where it seems that Clifford algebraic branches engine interior branching dynamics and Finsler geometric branches determine exterior branching dynamics.

### 3.3 Analytical Structure and Wave-Particle Duality

Mathematically, a multi-vector contains a complex scalar, a complex vector and a bi-vector, which provides rich structural information. Physically, a multi-vector can be applied as linked-measure (Ye 2015a, b).

The differential operators of the derivatives can be introduced and defined as

\[
\partial_\mu = \frac{\partial}{\partial x_\mu}; \quad D_\mu = (\partial_\mu - \omega_\mu)
\]  

\[
\nabla = \gamma^\mu \partial_\mu; \quad \nabla^2 = g^{\mu\nu} \partial_\mu \partial_\nu; \quad g^{\mu\nu} = \gamma^\mu \cdot \gamma^\nu
\]

where space-time metric \(g^{\mu\nu}\) is naturally generated.

Then the linked-energy \(E\) and linked-momentum \(p_\mu\) with linking Hamilton function \(H\) and Lagrangian function \(L\) can be defined as follows.

\[
L = p_\mu x_\mu - H; \quad p_\mu = \frac{\partial L}{\partial x_\mu}; \quad \delta \int_s d^4x L = 0
\]  

\[
H = p_\mu x_\mu - L = E(s, t) = \int_s (M/V) dx_\mu\]

\[
\frac{\partial H}{\partial p_i} = \frac{ds_i}{dt}; \quad \frac{\partial H}{\partial s_i} = -\frac{dp_i}{dt}
\]

where \(V\) is volume of space so that \(M/V\) means density function of linked-measure in the space. So the energy-mass conservation is extended to energy-mass-momentum-angular momentum joint conservation, and Hamilton principle keeps in the analytical center of physics.
3.3 Analytical Structure and Wave-Particle Duality

3.3.1 Field Equations

For keeping gauge invariance, we suppose the transformations as

\[ \psi \to \psi' = e^{i\omega} \psi; \quad \overline{\psi} \to \overline{\psi}'e^{-i\omega} \]  \hspace{1cm} (3.36)

\[ A_\mu \to A'_\mu = A_\mu + \partial_\mu \omega; \quad \overline{A}_\mu \to \overline{A}'_\mu = \overline{A}_\mu - \partial_\mu \omega \]  \hspace{1cm} (3.37)

The spin caused by self-mass and self-potential is

\[ B = \nabla(\psi A) \]  \hspace{1cm} (3.38)

So the linked-measures construct the linked-field. The Riemann-Einstein space-time provides the following field equation (Ye 2015a)

\[ T_{\mu\nu} + p(\psi) - q(A) = G_{\mu\nu} = \Omega_{\mu\nu} = k \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \]  \hspace{1cm} (3.39)

where \( G_{\mu\nu} \) is Einstein tensor, \( \Omega_{\mu\nu} \) is curvature tensor and \( T_{\mu\nu} \) denotes the total observational tensor (e.g., energy-momentum tensor), positive \( p(\psi) \) integrates all positive items of the linked-field and negative \( q(A) \) integrates all negative items of the linked-field, where the left side denotes physical effects while the right side indicates mathematical means.

When Finslerian curvature Eq. (3.22) is applied, the total observational tensor equals to the total curvature as

\[ T_{\mu\nu} = \Omega(R)_{\mu\nu} + \Omega(P)_{\mu\nu} = k \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) + q(A) - p(\psi) \]  \hspace{1cm} (3.40)

Equation (3.40) shows that \( q(A) - p(\psi) \) contributes the \( P \)-part of space-time curvature, while the \( R \)-part supplies Riemann curvature of space-time, which is just the results of general relativity: when the \( P \)-part of Finslerian curvature is ignored, the field equation returns to Einstein equation.

\[ T_{\mu\nu} = G_{\mu\nu} = k \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \]  \hspace{1cm} (3.41)

Thus we see the analytical equation matches general relativity and standard cosmology.
3.3.2 An Interpretation of Wave-Particle Duality

Generally, micro-particle and macro-cosmos can be jointly measured by measure $M = (\psi, A, B)$, where $\psi = \varphi + i\theta$ is just a complex scalar particle-bias (massive) function and $A = V + iU$ looks like a complex vector wave-bias (potential) function while $B$ characterizes spin (rotation). Meanwhile, $M$ is also the measure of total energy. When the energy distributes bias $\psi$, the particle looks like real particle and there is in $M$

$$\psi \gg A \subset M$$ (3.42)

When the energy distributes bias $A$, the particle looks like wave and there is in $M$

$$\psi \ll A \subset M$$ (3.43)

Thus that the even part $M_+$ and the odd part $M_-$ just describe particle-bias distribution and wave-bias distribution respectively.

Since $\psi$ and $A$ are particle-bias function and wave-bias function respectively, the $M$ will look like particle if $\psi \gg A$ and the $M$ will look like wave if $\psi \ll A$. As $M = (\psi, A, B)$ becomes unified linked-measure for both micro-particle and macro-cosmos, all things have wave-particle duality. For micro-particle, the energy changes easily between $\psi$ and $A$, so that the micro-particle looks like obvious wave-particle duality. For macro-things, the mass-energy concentrates and keeps mostly in $\psi$, so that there is no obviously observed wave-particle duality. When vortex (Ye 2015b) becomes general methodological approach for both micro-thing and macro-thing, the world can be naturally and jointly interpreted.

3.4 Branching Multiverse

Synthesizing Clifford algebra and Finsler geometry, branching space-time is naturally derived.

Rooted by the linked-measure $M$, four types of main stems, $M_+, M_-, Q_+, Q_-$, can be generated. Stemmed by $Q_+$, the Cayley-Dickson branch and the Euclid-Grassmann branch can be produced, as shown in Fig. 3.1, based on the mainstream.

Therefore, there are only limited branches in the branching multiverse, where there are 1 ($M$), 2($M_+, M_-$ or $Q_+, Q_-), 3(M, M_+, M_- or M, Q_+, Q_-), 4(M_+, M_-, Q_+, Q_-), 5(M_+, M_-, Q_+, Q_- and $Q_+$ divides two sub-branches) or 6 ($M, M_+, M_-, Q_+, Q_- and $Q_+$ divides two sub-branches) basic branches (more branches will be their multiples). Figure 3.1 looks like a universe organ, on which it is expected to verify that our fingers are not occasional in the universe. That is a real universal holography.

Meanwhile, the branching multiverse actually belongs to a unified universe, where all branches have the same “root”. In different branches, there may be
different physical parameters. However, there are unified physical laws in the “root” universe. Otherwise, from the physical perspective, we can not understand our universe and branching multiverse physically.

In the double-dynamic cosmos (Ye 2015a) and the branching universe, possible evolution of cosmological curvature should be modified from big to small and each branch might have its unique curvature, which could provide experimental ways for verifying the suggested cosmological model.

3.5 Discussion and Criticism

There were two artificial “principles” for interpreting the multiverse, holographic principle (Susskind 1995) and anthropic principle (Susskind 2003).

The holographic principle was inspired by black hole thermodynamics, which conjectures that the maximal entropy in any region scales with the radius squared and not cubed as might be expected. In the case of a black hole, the insight was that the informational content of all the objects that have fallen into the black hole might be entirely contained in surface fluctuations of the event horizon. The holographic principle resolves the black hole information paradox within the framework of string theory or $M$-theory (Becker et al. 2007), where a multiverse of a somewhat different kind has been envisaged within string theory and its higher-dimensional extension. Those theories require the presence of 10 or 11 space-time dimensions, in which extra 6 or 7 dimensions may either be compactified on a very small scale and our universe may simply be located on a dynamical $(3 + 1)$-dimensional object, a $D$-brane. As there exist classical solutions to the Einstein equations that allow values of the entropy larger than those allowed by an area law, hence in principle larger than those of a black hole. The holographic principle is a property of string theories and a supposed property of quantum gravity that states that the description of a volume
of space can be considered as encoded on a boundary to the region, preferably a
light-like boundary like a gravitational horizon.

The anthropic principle is the philosophical consideration that observations of the
universe must be compatible with the conscious and sapient life which observes
it. While the strong anthropic principle (SAP) states that this is all the case
because the universe is compelled to eventually have conscious and sapient life
emerge within it, the weak anthropic principle (WAP) states that the universe’s osten-
ssible fine tuning is the result of selection bias. That is to say, only when a universe is
capable of eventually supporting life will there be living beings capable of observing
and reflecting upon fine tuning. The key issue of anthropic principle concerns that
other universes have been proposed to explain how our own universe appears to be
fine-tuned for conscious life as we experience it. If there were a number of universes,
each with possibly different physical laws or different fundamental physical con-
stants, some of those universes, even if just few, would have the combination of laws
and fundamental parameters that are suitable for the development of matter, astro-
nomical structures, elemental diversity, stars and planets that can exist long enough
for life to emerge and evolve. The weak anthropic principle could then be applied to
make such a conclusion: we, as conscious beings, would only exist in one of these
few universes that happened to be finely tuned, permitting the existence of life with
developed consciousness.

However, those interpretations above belong to philosophical views, unless we
can find physical evidence. A data analysis from WMAP claimed to find preliminary
evidence suggesting that our universe collided with other parallel universes in the
distant past. But, a more thorough analysis of data from the Planck satellite (Planck
Collaboration 2014), which has a resolution of 3 times higher than WMAP, failed
to find any statistically significant evidence of such a bubble universe collision. In
addition, there is no evidence of any gravitational pull of other universes on ours.

Actually, the multiverse is philosophical rather than scientific because it lacks
falsifiability and the ability to disprove a theory by means of scientific experiment has
always been a part of the accepted scientific method. Not only there is no experiment
that can rule out a theory if it provides all possible outcomes, but also there exist
too many possibilities (Linde and Vanchurin 2010; de Simone et al. 2010; Guth and
Nomura 2012).

According to the image of branching multiverse, physical even biological branches
are generally natural phenomena in the universe, so that we see the rich
branching phenomena in the world. Because there are different exterior environ-
ments, complete symmetry seldom happens in nature, even if the same “genes” drive
as interior dynamics. Different genes and different environments will introduce dif-
ferent kinds of branching and various branches can be produced and form fractals
at the ends, via the fundamental mathematical mechanism of the Clifford-Finslerian
branching.

The situation shows that the multiverse is possible. However, the multiverse is
branching from 1 to 6 basic types (others belong to their combination), either arbitrary
or chaos. So, branching multiverse is different from the countless multiverse based
on anthropic principle.
3.5 Discussion and Criticism

It is expected that this contribution can get rid of puzzle that multiverse is only a non-scientific image, then helps the physical study return to the correct path for approaching real physics (Olive et al. 2014).

3.6 Conclusion

To conclude, the Clifford-Finslerian linked-field generates branching multiverse. While Clifford algebra supplies interior dynamics for generating inner branching structure, Finsler geometry provides catastrophic branches of space-time metric and curvature, as exterior dynamics. Based on dual-complex representation, Hamilton representation and Pauli representation of quaternion, the multiplication of quaternions generates Cayley-Dickson branch and Euclid-Grassmann branch, which causes limited branching multiverse.

The branching multiverse is different from countless multiverse, as the branches are deterministic and based on mainstream, where all branches integrate and produce limited multiverse. Although there might be different physical parameters in different branches, it still should follow unified physical laws in the “root” universe. Meanwhile, with keeping the tradition of analytical mechanics, the wave-particle duality can be clearly interpreted at both micro-level and macro-level. The branching multiverse seems to be a reasonable physical choice of multiverse rather than philosophical one, which provides a feasible theoretical framework for interpreting physical even biological branching phenomena in the universe. Furthermore, as the Clifford-Finslerian linked-field leads to branching multiverse, the multiverse can solve the puzzle as a non-scientific term and then enable the relevant study to return to the correct path for exploring real physics. In the physical framework, deeper and wider issues could be further investigated and it is expected to find more interesting results.

Acknowledgements This chapter is a revision of the original version published at Physical Journal, 2015, 1(3): 375–381

References


A Clifford–Finslerian physical unification is proposed based on Clifford–Finslerian mathematical structures and three physical principles. In the Clifford–Finslerian mathematical structure, spontaneous symmetry breaking is automatically embedded in fractal branches. With the action principle, connection principle and construction principle, physics can be unified, in which the Riemman-Einstein system and Euclid-Newton system are naturally included when quaternion are reduced to complex and real phases.

4.1 Introduction

Some scientists believe that supersymmetry and string theory are biased towards mathematics rather than physics (Penrose 2004; Smolin ?). I share the similar belief with them. Let us go back to physical tradition and recall experimental and empirical foundations and the main theories of physics. Various string theories and M theory have shown us many beautiful situations in theoretical physics, but they are only mathematical images. According to our knowledge, real physics that relies on experiments ended up at two standard models: (1) standard model of cosmos (big bang) based on general relativity; (2) standard model of particle physics based on gauge field theory (Amsler et al. 2008). Thus, those theories are our starting point for reflection and further consideration.

Since the 1980s, some mathematicians and physicists have probed into quaternionic physics (Alder 1986, 1995; de Leo 1996) and Clifford–Finslerian structures (Vargas and Torr ?, Vacaru ?), while the others developed string theory or M theory (Witten ?; Seiberg and Witten 1999; Smolin ?), loop quantum gravity (Ashtekar and Lewandowski 2004; Rovelli 2004), spinor and twistor theory (Penrose and Rindler 1984, 1986; Penrose 1999), causal dynamical triangulations (Ambjorn et al. 2006; Jurkiewicz et al. 2008) and so on El Naschie (2008). In this chapter, some basic principles (Ye ?) are developed for Clifford–Finslerian physical unification with incorporation of fractal dynamics.
4.2 Mathematical Methods

A quaternion consists of two complex or four real factors. In mathematics, a quaternion has both Cayley–Dickson construction (Baez 2002) and scalar-vector construction (Sweetser ?).

4.2.1 Algebraic Structure

According to the Cayley–Dickson construction, we can construct logically complex or quaternion with real number or function $X$, in which a complex is expressed by two real parameters or functions and a quaternion by four real parameters or functions.

For real $X$, we define its conjugation $X = X$. And two real functions $X_1$ and $X_2$ multiply as

$$X_1X_2 = X_2X_1$$ (4.1)

For complex $Z = (X, X')$ with $\text{Re}(Z) = X$ and $\text{Im}(Z) = X'$, we have its conjugation

$$\overline{Z} = (\overline{X}, -X')$$ (4.2)

and the multiplication for two complexes $Z_1$ and $Z_2$ becomes

$$Z_1Z_2 = (X_1, X'_1)(X_2, X'_2) = (X_1X_2 - X'_2X'_1, X_1X'_2 + X'_1X_2)$$ (4.3)

For quaternion $Q = (Z, Z')$ with dual complex $Z$ and $Z'$, we have its conjugation

$$\overline{Q} = (\overline{Z}, -Z')$$ (4.4)

and the multiplication for quaternion $Q_1$ and quaternion $Q_2$ becomes

$$Q_1Q_2 = (Z_1, Z'_1)(Z_2, Z'_2) = (Z_1Z_2 - Z'_2\overline{Z'_1}, \overline{Z'_1}Z'_2 + Z'_1Z_2)$$ (4.5)

And according to scalar-vector construction, a quaternion $Q = q_0 + q_1i + q_2j + q_3k = (\varphi, A)$ with scalar $q_0 = \varphi$ and vector $(q_1, q_2, q_3) = A$ has its conjugate

$$\overline{Q} = (\varphi, -A)$$ (4.6)

and the multiplication for quaternion $Q_1$ and quaternion $Q_2$ becomes

$$Q_1Q_2 = (\varphi_1, A_1)(\varphi_2, A_2) = (\varphi_1\varphi_2 - A_1 \cdot A_2, \varphi_1A_2 + A_1\varphi_2 + A_1 \wedge A_2)$$ (4.7)

where $\wedge$ is Wedge product and $a \wedge b = -b \wedge a$. 
As Eqs. (4.5) and (4.7) co-exist, we have

\[ Q_1 Q_2 = (Z_1 Z_2 - Z'_{Z_1}, Z_1 Z'_{Z_2} + Z'_{Z_1} Z_2) = (\varphi_1 \varphi_2 - A_1 \cdot A_2, \varphi_1 A_2 + A_1 \varphi_2 - A_2 \wedge A_1) \] (4.8)

We see that the quaternion has two algebraic branches, a Cayley–Dickson branch and a scalar-vector branch.

### 4.2.2 Analytical Structure

When we consider mathematical analysis, the differential operator \(d\) and integral operator \(\int\) should be defined as a whole operator:

\[
d = (\partial_Q, \partial_Q), \partial_Q = \frac{1}{2}(\partial_Z, -\partial_Z), \partial_Q = \frac{1}{2}(\partial_Z, \partial_Z)
\] (4.9)

\[
d = (\partial, \nabla); \quad d^2 = (\partial, \nabla)(\partial, \nabla)
\] (4.10)

where Eq. (4.9) acts on structure \((Z, Z')\) and Eq. (4.10) on \((\varphi, A)\), which will result in different branches.

With the same reason, the integral operator \(\int\) acts on differential form \(\omega\) in a Cayley–Dickson construction or scalar-vector construction as different branches.

\[
\int \omega = \left( \int_Z \omega, \int_{Z'} \omega' \right) = \left( \int_\varphi \omega, \int_A \omega \right)
\] (4.11)

All the algebraic and analytical branches formulate fractal branches of the quaternionic world.

### 4.2.3 Geometrical Structure

According to Chern’s analysis (Bao et al. 2000), there is a Chern connection form \(\omega^i_j\) in Finsler bundle \(p^*TM \rightarrow PTM\) as the unique solution of structural equations

\[
d \omega^i_j = \omega^i_j \wedge \omega^j_i
\] (4.12)

and

\[
\omega_{ij} + \omega_{ji} = -2A_{ijk} \omega^k_m
\] (4.13)
where $\omega_{ij} = \omega^k_i \delta_{kj}$ and $A = A_{ijk} \omega^j \otimes \omega^j \otimes \omega^k$ is a Cartan tensor. The Finsler structure becomes Riemann structure if $A = 0$. And the curvature of the Chern connection is
\[
\Omega_k^i = \frac{1}{2} R_{kij}^l \omega^l \wedge \omega^j + P_{kijn}^l \omega^n \wedge \omega_m^i
\]
(4.14) where $R$-part is horizon-horizon (1st Chern) curvature and $P$-part is horizon-vertical (2nd Chern) curvature. When there is no $P$-part, a Finsler curvature becomes a Riemann curvature.

The Clifford–Finslerian structure with fractal branches above can be a general mathematical structure for physics.

### 4.3 Physical Principles

For formulating a physical system, we need some basic physical principles with some requisite physical quantities. Based on physical tradition, time $t$ and space $S$ are needed, so are energy $E$ and momentum $p$, which just construct two scalar-vector quaternion pairs $X = (t, -S)$ and $F = (E, -p)$. While scalar time $t$ is one dimension, vector space $S$ naturally is three dimensions.

In synthesized traditional and modern physics, the following three principles should be introduced and maintained.

#### 4.3.1 The First Physical Principle

The first physical principle is the action principle, which links dynamic mechanism and symmetry as
\[
\delta \int dt L(t, S, \dot{S}) = \delta \int dt' L'(t', S', \dot{S}') = 0; \quad L' = L + \frac{d}{dr} G
\]
(4.15) where $L$ is the Lagrange function of a physical system and $G$ the transformation group, which can be quaternion, complex or real.

The first physical principle originates from the Lagrange-Hamilton principle and the Nöther theorem and determines the kinematical and dynamic mechanism of a physical system, including Newton system, Einstein system and Yang-Mills system.
4.3 Physical Principles

4.3.2 The Second Physical Principle

The second physical principle is the connection principle, which links physical potential and mathematical connection, physical field strength and mathematical curvature together, with the following mathematical structure

$$ P = k_1 \omega; \quad F = k_2 \Omega; \quad \Omega = d\omega - \omega \wedge \omega \quad (4.16) $$

where $P$ is potential (quaternion, complex or real); $\omega$ connection (quaternion, complex or real); $k_1$ a constant (real scalar); $F$ field strength (quaternion, complex or real); $\Omega$ curvature (quaternion, complex or real); $k_2$ a constant (real scalar).

The second physical principle originates from kinematics, Newton’s 2nd law and Einstein’s equivalence principle, which determines the kinematical characteristics and the dynamic structure of a physical system. That system covers Newtonian theory, Einstein’s general relativity and Yang-Mills gauge field theory. While every potential has its relative connection and each connection determines its potential, every field strength links its relative curvature and each curvature determines its field strength.

If a relation between connection $\omega$ and time-space $X$ with $\omega = dX$ and energy $E$ (scalar) and momentum $p$ (vector) constructing with $F = (E, -p) = dP$ become physical laws, matching curvature tensor $\Omega = d^2 X$, a physics will be established in the time-space.

Those two principles above link physical structures with mathematical structures together.

4.3.3 The Third Physical Principle

The third physical principle is the construction principle, which links inner geometric structure and outer topological index together as

$$ \int_M \Omega = \chi; \quad \int_M d\omega = \int_{\partial M} \omega \quad (4.17) $$

where $\Omega$ is curvature and $\omega$ is connection, while $\chi$ is characteristic index and $M$ is manifold on real, complex or quaternion.

The third physical principle originates from Gauss-Bonnet and Stokes theorems, which are mathematical laws, to help probe into inner structure with surface parameters for constructing a computational system for physics.

In those three principles, each consists of both mathematical and physical structures. While the mathematical algebra-geometry structures are respectively real-Euclid, complex-Riemann and quaternion-Finsler, the physical systems may become Newton system, Einstein system and a new one. We can apply those three principles above to any physical system as a general approach of physics.
4.4 Clifford–Finslerian Physics

Applying those mathematical structures and physical principles above to any physical system, we see the following quaternionic Finslerian physics in quaternionic Finslerian time-space \( X = (t, S) \).

Cayley–Dickson branch → scalar-vector branch

\[
P = (P_Z, P_{Z'}) = k_1 \omega = k_1 (\partial, \nabla) (t, S) \quad (4.18)
\]

\[
F = (F_Z, F_{Z'}) = k_2 \Omega = k_2 (d \omega - \omega \wedge \omega) = k_2 (\partial, \nabla)^2 (t, S) [(1 - (t, S)] \quad (4.19)
\]

If we know a system with physical laws as

\[
F = k_3 dP \quad (4.20)
\]

\[
L = \int k F \overline{F} - \omega \wedge d \omega \quad (4.21)
\]

where \( k_3 \) and \( k = k(k_1, k_2, k_3) \) are constants (real scalars), a quaternionic physics set up.

In a real-Euclidean structure, Eqs. (4.18)–(4.21) embed Newtonian mechanics and Maxwellian equations. In a complex-Riemannian structure, Eqs. (4.18)–(4.21) support Einstein’s general relativity and Dirac’s theory, when energy-momentum tensor \((E, p)\) is applied. And in a quaternion-Finslerian structure, Eqs. (4.18)–(4.21) provide a new system. As a quaternion-Finslerian structure contains double complex-Riemannian structures, it constructs two branches of physics in complex-Riemannian time-space. As one complex-Riemannian structure includes two real-Euclidean structures, a quaternion-Finslerian structure can be divided into four real-Euclidean structures, which means that those four branches co-exist and may match four basic interactions.

When the quaternion-Finslerian structure is reduced to the complex-Riemannian structure, its algebra is complex (it can be divided as two real parts) and its geometry is Riemannian. When it is reduced further to the real-Euclidean structure, its algebra is real and its geometry is Euclidean.

4.5 Finslerian Metric and Fractal Dynamics

If we want to measure the geometry of space-time, we need a key element, which is metric \( ds \) of time-space. In a Clifford–Finslerian structure, the metric \( ds \) should apply the Lorentz-Cao transformation (Cao 2001), which is

\[
ds^4 = g_{\mu \nu \rho \sigma} dx^\mu dx^\nu dx^\rho dx^\sigma = c^4 (1 - \beta^2)^2 dt^4, \quad \beta = v/c \quad (4.22)
\]
4.5 Finslerian Metric and Fractal Dynamics

In time-space $X = (t, S)$, the Lorentz-Cao invariance remains as
\[ X^T G X = X'^T G X' ; \quad G = \det(g_{\mu\nu\rho\sigma}) \] (4.23)

In Eq. (4.22), we see that there is a fractal factor with a typical catastrophic pattern like symmetry breaking
\[ B = \lambda (\alpha^2 - \beta^2)^2 \] (4.24)

where spontaneous symmetry breaking embedded without Higgs mechanism and will arise from the time-space fractal branches. If $\lambda \neq 0$, Eq. (4.24) includes four branches as
\[ (\alpha^2 - \beta^2)^2 = (\alpha - \beta)^2 (\alpha + \beta)^2 \] (4.25)

And Eq. (4.21) embeds a mechanism of fractal dynamics with the following geodesic line
\[ l = \int ds = \int (g_{\mu\nu\rho\sigma} dx^\mu dx^\nu dx^\rho dx^\sigma)^{1/4}, \quad \nabla g = 0 \] (4.26)

where $g = g_{\mu\nu\rho\sigma} dx^\mu dx^\nu dx^\rho dx^\sigma$.

When it is Riemann space, Eq. (4.26) reduces into Riemann geodesic line
\[ l = \int ds = \int (g_{\mu\nu} dx^\mu dx^\nu)^{1/2}, \quad \nabla g = 0 \] (4.27)

where $g = g_{\mu\nu} dx^\mu dx^\nu$.

When Einstein’s view of space as warped was confirmed, we knew that Riemann’s geometry fits Einstein’s physics or Einstein’s physics is based on Riemann’s geometry. While we logically expand algebraic structures from real, complex to quaternion, we have to develop matching geometric structures from Euclid, Riemann to Finsler. The developments will introduce some new approaches to the world and some new ideas for physics.

4.6 Discussion

In the Clifford–Finslerian unification above, we see three constants of $k_1$, $k_2$ and $k_3$. Those three constants may be sources of basic physical constants. If minimum quantum space is the unit of reduced Planck constant $\hbar$, we can have $\min(dS) = \hbar$. If we accept special relativity, we know $\max(dS/dt) = c$, where $c$ means the velocity of light. Therefore, we have $\min(dr/dS) = 1/c$. Then, if minimum two order derivative of space-time is gravitational constant $g$, we have $\min(d^2 S/dr^2) = g$. That is summed up as quantum mechanism.
\[
\min(dS) = \hbar; \quad \min\left(\frac{dt}{dS}\right) = \frac{1}{c}; \quad \min\left(\frac{d^2S}{dt^2}\right) = g
\] (4.28)

If we think that time-space is always the combination of minimum quantum units and all constants relate to the minimum quantum units, we can introduce the physical constants by

\[
k_1 = k_1(\hbar, c, g); \quad k_2 = k_2(\hbar, c, g); \quad k_3 = k_3(\hbar, c, g)
\] (4.29)

When we choose a suitable unit system, we may set \( \hbar = c = g = 1 \) or \( k_1 = k_2 = k_3 = 1 \).

In the new system, we can maintain experimental foundations and theoretical traditions of physics, while general relativity and standard models can be saved. Beyond standard models, string theory and \( M \)-theory have developed so far from phenomena, as well as other unified theories of quantum gravity (Connes 1995; Markopoulou and Smolin 2004; Amelino-Camelia et al. 2004; Bilson-Thompson et al. 2007). That is a new investigation.

## 4.7 Conclusion

The physical unification above is based on Clifford–Finslerian mathematical structures and three physical principles. The first physical principle is the action principle, which links dynamic mechanism and symmetry and determines the kinematical and dynamic mechanism of a physical system. The second physical principle is the connection principle, which links physical potential and mathematical connection, physical field strength and mathematical curvature and determines the kinematical characteristics and dynamic structure of a physical system. The third physical principle is the construction principle, which links together inner geometric structure to outer topological index and constructs a computational system. Those three principles (action, connection and construction, ACC) can be applied to physical systems with real, complex and quaternion algebraic phases and Euclidean, Riemmanian, Finslerian geometric structures, which has produced traditional approaches developed by Newton and Einstein and can inspire new approaches to the world and new physics.

The Clifford–Finslerian physical unification is a new idea based on Clifford–Finslerian mathematical structures and three physical principles. In the Clifford–Finslerian mathematical structure, spontaneous symmetry breaking is automatically embedded in fractal branches. With the action principle, connection principle and construction principle, physics can be unified, in which the Riemman-Einstein system and Euclid-Newton system are naturally included when quaternion are reduced to complex and real phases. While Eqs. (4.15)–(4.17) construct an abstract framework,
4.7 Conclusion

Eqs. (4.20)–(4.23) and (4.26) will provide the concrete physics when they are applied. Although that is only a preliminary idea, we hope that it can pave the way for fruitful development.

Acknowledgements This chapter is a revision of the original version published at Chaos, Solitons and Fractals, 2009, 41(5): 2301–2305.

References

Vacaru, S.L.: Nonholonomic clifford structures and noncommutative riemann-finsler geometry. 

Vargas, J.G., Torr, D.G.: Marriage of Clifford algebra and Finsler geometry: a lineage for unification?


Ye, F.Y.: Approaches to the Natural World Based on Three Principles of Mathematical Physics.
Chapter 5
A Vortex Mechanism Linking Micro-Particle to Macro-Galaxy Without Supersymmetry

A vortex is mathematically equivalent to a multi-vector and physically equivalent to a linked-measure. It is suggested that vortex is a bridge to link micro-particle and macro-galaxy in phenomenological view and can naturally unify the micro-particle standard model and the macro-cosmos standard model theoretically. Approached by vortex mechanism, physics keeps Hamilton principle and balances mathematical structure and physical essence, in which the particle-wave duality and the gravity-curvature equivalence are reasonably interpreted and the behaviors of micro-particles and macro-galaxies as well as black holes are jointly understood.

5.1 Introduction

Contemporary physics reached a bright level with colorful achievements, where we especially mention quantum theory and relativity, linked by Hamilton principle. Those two great physical theories led to two physical standard models respectively, i.e., the micro-particle standard model characterized by $SU_C(3) \times SU_L(2) \times U_Y(1)$ gauge field as quantum field theory and the macro-cosmos standard model characterized by big bang cosmology based on general relativity (Einstein 1916), which construct theoretical core of contemporary physics (Beringer et al. 2012). However, those two standard models are so different that they can not be unified into a harmony theory, which causes a theoretical poser called as quantum gravity. On the other side, the physical experiments and observations strongly support those two standard models, so that any revised unified theory has to include those two standard models in it.

We see that the micro-particle standard model did successfully process all phenomena concerning micro-particles, while the macro-cosmos standard model did so for macro-cosmos. Only is there lack of link micro-particles and macro-galaxies.

In the micro-particle standard model, the greatest success was the prediction of the Higgs boson, which had been experimentally discovered in 2012 and verified in 2013 by LHC at energy 125.6 GeV·c$^{-2}$ (The ATLAS Collaboration 2012; The CMS
Collaboration 2012). It is well known that the Higgs mechanism describes how the weak SU(2) gauge symmetry is broken and how fundamental particles obtain mass, which is also the last particle predicted by the micro-particle standard model to be observed, although efforts to confirm that it has all of the properties predicted by the standard model are ongoing. Other great successes of the micro-particle standard model included the prediction of the W boson and the Z boson, the gluon, and the top and the charm quarks, before they had been observed. However, the worst disadvantage in the micro-particle standard model is the complete absence of gravity and it predicts neutrinos to be massless while the observed evidence of neutrino oscillations implies that neutrinos have tiny mass (Olive et al. 2014).

In the macro-cosmos standard model, it is frequently referred to the big bang cosmology and it is the simplest successful model that provides a reasonably good account of the following observational evidence of the cosmos: (1) the existence of the CMB and the large-scale structural distribution of galaxies; (2) the abundances of hydrogen (including deuterium), helium and lithium, as famous BBN theory of elements synthesis (Sarkar 1996; Cyburt et al. 2005; Fields et al. 2014); (3) the accelerating expansion of the universe observed in the light from distant galaxies and supernovae (Riess 1998). However, the dark energy and the dark matter are kept unknown in the model and the Planck satellite’s newest observation report (Planck Collaboration 2014a, b) shows that there are 68.3% dark energy, 26.8% dark matter and 4.9% matter-energy in the universe.

Meanwhile, in micro-fields, physicists have probed into “basic particles” that construct the natural world, where there is the principle of particle-wave duality, which introduced the quantum mechanism. In macro-fields, physicists have explored the galaxies, where there is the principle of gravity-curvature equivalence, which constructed scientific cosmology. All the theoretical principles matched physical observations and experiments very well, but lost a link mechanism. More generally, the principles of action, connection and construction had been suggested (Ye 2009).

Facing the great physics, we have to inherit the glorious achievements and develop it in future. Since there exists strong evidence that those two standard models are good physical theories for approaching the world, we’d better explore the linkage between the micro-particle and macro-galaxy.

5.2 Phenomenological Foundations and Theoretical Constrains

There are universal vortex phenomena around us, from micro-level to macro-scale, as shown in Fig. 5.1.

Figure 5.1 describes various vortices from micro-level to macro-scale, where we see, from left to right, the imaged spinning particle, real river vortex, real atmospheric vortex and the typical galaxy (NGC3370). Those phenomenological vortices mean
Fig. 5.1 Phenomenological vortices at different scales in the universe

that the vortices are popular phenomena in the universe, so that we should have their
mathematical and physical description.

Since those two standard models are so successful and the vortex is equivalent to
the linked-measure geometrically, it is feasible to apply the linked-measure as well
as linked-field to approach the physical phenomena (Ye 2015a, b).

While the micro-particle standard model supposes that gauge field theory is the
best choice, the macro-cosmos standard model assumes that general relativity is the
correct theory of gravity at cosmological scale. Because those two standard models
are so successful, a complete unified theory should abide some theoretical constrains
with following characteristics:

(1) The theory could interpret all interactions, including phenomenological
$SU_C(3) \times SU_L(2) \times U_Y(1)$ symmetry at micro-level and cosmic accelerating
expansion at macro-level.

(2) The theory should accord with analytical principles of physics, particularly the
Hamiltonian principle, with incorporating quantum theory and relativity, and
therefore it is quantum field theory, where local symmetries are described by
Abelian and non-Abelian gauge theories.

(3) The particle-wave duality should be maintained for micro-particles and the
gravity-curvature equivalence should be kept for macro-galaxies.

Generally, a suitable unified theory should be consistency of both micro-particle
standard model and macro-cosmos standard model, abiding the principles of the least
action, particle-wave duality and gravity-curvature equivalence. Here, we introduce
vortex mechanism for approaching the unification.

5.3 Theoretical Approaches by Vortex Mechanism

A good theoretical framework should keep math-physical balance, which means that,
when we set up an equation, if the left side of the equation reflects mathematical
structure, its right side should reveal physical essence, i.e., mathematical structure= physical essence. For reaching the object, space-time multi-vector (Hestenes and
Sobczyck 1984; Hestenes 2003; Doran and Lasenby 2003) provided a good mathematical method, in which Clifford algebra could be united with Riemann or Finsler geometry. The new math-physical methodology can be called linked-measure and linked-field.

### 5.3.1 Linked-Measure Describes Vortex, Leading to Linked-Field

A linked-measure \( M \) is geometrically equivalent to a vortex, as shown in Fig. 5.2. Mathematically, a vortex can be simply expressed by using a multi-vector \( M_k \) \((k = 0, 1, 2, 3, 4)\) as follows.

\[
M = M_0 + M_1 + M_2 + M_3 + M_4 = \varphi + V + B + iU + i\theta = \psi + A + B = (\psi, A, B) \tag{5.1}
\]

where \( \psi = \varphi + i\theta \) constructs a complex scalar function (mass-like function), while \( A = V + iU \) forms a complex vector function (potential-like function) and \( B = (1/2)B_{\mu\nu}\gamma^{\mu} \wedge \gamma^{\nu} \) as a unique bi-vector (spin-like function). One scalar \( \psi \), one vector \( A \) and one bi-vector \( B \) just describe a vortex and construct a linked-measure.

And its space-time conjugation is

\[
\overline{M} = -iMi = M_0 - M_1 + M_2 - M_3 + M_4 = \psi - A + B = (\psi, -A, B) \tag{5.2}
\]

In the multi-vectors, for each space-time point \( x \) on Dirac frame \( \{\gamma^\mu, \mu = 0, 1, 2, 3\} \), there exists

\[
x = x_\mu \gamma^\mu, \quad x_\mu = \gamma^\mu x \tag{5.3}
\]
The coordinates’ transformation will be

\[ x_\mu \rightarrow x_\mu' = \alpha^\nu_\mu x_\nu; \quad \gamma^\mu \rightarrow \gamma'^\mu = \alpha^\mu_\nu \gamma^\nu \quad \text{with} \quad \alpha^\mu_\nu \alpha^\nu_\lambda = \delta^\mu_\lambda \quad (5.4) \]

where Dirac matrices \( \{ \gamma^k \} \) as well as Pauli matrices \( \{ \sigma^k \} \) are to be interpreted geometrically as meaningful space-time bi-vectors.

Physically, a multi-vector can be applied as linked-measure which link scalar mass \( m \), vector potential \( V \) and bi-vector spin \( S \) together. At micro-particle level, a multi-vector \( M \) and its space-time conjugation \( \overline{M} \) just describe particle and anti-particle. At macro-cosmos level, the interaction of \( M \) and \( \overline{M} \) could generate dynamic space-time.

For keeping gauge invariance, we can suppose the transformations as

\[ \psi \rightarrow \psi' = e^{i \omega \psi}; \quad \overline{\psi} \rightarrow \overline{\psi} e^{-i \omega} \quad (5.5) \]

\[ A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \omega; \quad \overline{A}_\mu \rightarrow \overline{A}'_\mu = \overline{A}_\mu - \partial_\mu \omega \quad (5.6) \]

where the linked-measures construct linked-field, where the differential operators of one order derivatives can be introduced and defined as

\[ \partial_\mu = \frac{\partial}{\partial x_\mu}; \quad \nabla = \gamma^\mu \partial_\mu \quad (5.7) \]

Its covariant derivative and the differential operators of two order derivatives can be introduced and defined respectively as

\[ D_\mu = (\partial_\mu - \omega_\mu); \quad \nabla^2 = g^{\mu\nu} \partial_\mu \partial_\nu; \quad g^{\mu\nu} = \gamma^\mu \cdot \gamma^\nu \quad (5.8) \]

where we see that space-time metric \( g^{\mu\nu} \) is naturally generated.

For understanding and measuring physical world, basic physical measures would be kept in physics, including mass \( m \), energy \( E \) and momentum \( p \). When spin is caused by self-mass and self-potential, the measure of rotation becomes

\[ B = \text{rot}(\psi A) \quad (5.9) \]

In order to introduce quantum mechanism, there should exist canonical commutation relations in \( \psi \) and \( A \) as

\[ [\psi, A] = \psi A - A \psi = i \hbar \quad (5.10) \]

where \( i \) is the imaginary unit and \( \hbar \) is the reduced Planck’s constant (\( \hbar = \hbar/2\pi \)).

Under Lorentz invariance, it is well known that there is the relation

\[ E^2 - p^2 c^2 = m^2 c^4 \quad (5.11) \]
where \( c \) is velocity of light in vacuum. If we choose physical unit and let \( c = h = 1 \) where \( h = 2\pi\hbar \) is Planck’s constant, it is

\[
m^2 = E^2 - p^2
\]  

(5.12)

The energy \( E \) (linking to \( \psi \)) and momentum \( p \) (linking to \( A \)) construct canonical commutation relations similarly to Eq. (5.2) as

\[
[E, p] = \begin{pmatrix} E \\ p \end{pmatrix} = \begin{pmatrix} h & 0 \\ 0 & \hbar \end{pmatrix} \begin{pmatrix} \omega \\ k \end{pmatrix}
\]

(5.13)

where \( E = \hbar \omega \) and \( p = \hbar k \) are combined and unified.

Under the quantization rule of momentum \( p \) and energy \( E \), we may also have operator expression

\[
[E, p] \rightarrow \left( \frac{i\partial}{\partial t} - i\nabla \right) \sim p_\mu \rightarrow i\partial_\mu
\]

(5.14)

When mass operator is introduced

\[
m \rightarrow i\partial/\partial t - i\nabla
\]

(5.15)

there is Dirac equation

\[
(i\gamma^\mu \partial_\mu - m)\Psi = 0
\]

(5.16)

When we define linked-energy \( E \) as well as linked-momentum \( p_\mu \) with linking Hamilton function \( H \) and Lagrangian function \( L \), we obtain math-physical equations following Hamilton principle as follows.

\[
H = p_\mu x_\mu - L; \quad \frac{\partial H}{\partial p_i} = \frac{ds_i}{dt}; \quad \frac{\partial H}{\partial s_i} = -\frac{dp_i}{dt}
\]

(5.17)

\[
L = p_\mu x_\mu - H; \quad p_\mu = \frac{\partial L}{\partial x_\mu}; \quad \delta \int_s d^4x L = 0
\]

(5.18)

where Greek subscripts \( \mu, \nu \) denote 1, 2, 3, 4 and Latin subscripts \( i, j \) do 1, 2, 3. Now the energy-mass conservation is extended to energy-mass-momentum-angular momentum joint conservation and Hamilton principle is kept as core analytical principle in physics.

Equation (5.17) is the differential form and Eq.(5.18) the integral form of the linked-field. When \( M (\psi, A, B) \) is viewed as physical linked-measure, fields are naturally generated in space-time.
5.3.2 Electromagnetic Field

Spreading all the space-time, if there exist strong links among $\psi$, $A$ and $B$ as

$$A = \nabla \psi; \quad B_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu}$$  \hspace{1cm} (5.19)

the Maxwell’s equations are obtained

$$\nabla B = \nabla \cdot B + \nabla \wedge B = J \quad \text{or} \quad \partial_{\mu} B^{\mu\nu} = J^{\nu}$$  \hspace{1cm} (5.20)

where $J = \nabla \cdot B$ is current and bi-vector $B$ includes both electrical field $E$ and magnetic field $H$ as $B = E + iH$ and the QED Lagrangian becomes exactly

$$L_{\text{QED}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - J^{\mu} A_{\mu}$$  \hspace{1cm} (5.21)

Generally, electromagnetic field is both macro-phenomenon and micro-phenomenon. At micro-level, when the complex scalar field (i.e., Higgs field) introduces mass $m$ following Lagrangian

$$L_{\text{Higgs}} = (D_{\mu} \psi^*)(D^{\mu} \psi) - V(\psi^* \psi)$$  \hspace{1cm} (5.22)

Then quantum QED Lagrangian can be re-written as

$$L_{\text{QED}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \bar{\psi}(i\gamma^{\mu} D_{\mu} - m) \psi$$  \hspace{1cm} (5.23)

The Lagrangian above determined electromagnetics.

5.3.3 Strong and Electro-Weak Fields

If there exist strong inner links at micro-level, strong field would generate. Including 8 duplicates $a$ in 3 generations, the field strength upgrades to inner higher symmetry

$$B_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} = B_{\mu\nu}^{a}; \quad a = 1, 2, \ldots, 8$$  \hspace{1cm} (5.24)

and the linked-field is extended with introducing new field $F$ as

$$F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} - g f_{abc} A_{\mu}^{a} A_{\nu}^{b}; \quad f_{abc} t^{c} = [t^{a}, t^{b}]$$  \hspace{1cm} (5.25)
When $A_\mu^a$ corresponds to the gluon fields ($a = 1, \ldots, 8$, as there are eight kinds of gluon) and the $\psi_{q, c}$ is quark-field spinors for a quark of flavor $q$ and mass $m_q$ with a color-index $c$ ($c = 1, 2, 3$), where the $t_{\mu\nu}^a$ corresponds to eight $3 \times 3$ matrices and is the generators of the SU$_c$(3) group, the Lagrangian becomes

$$L_{\text{QCD}} = -\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + \bar{\psi}_q (i \gamma^\mu D_\mu - m_q) \psi_q$$  (5.26)

which leads to QCD under SU(3) invariance, where the first item describes gluons and second item quarks. So the strong interactions are included.

When the main field strengths break micro-inner strong symmetry to micro-inner weak symmetry, from $F$ returning to $B$, the interactions are stemmed by group SU(2), replacing SU(3), with color-index $c$ changing to $b$

$$B_{\mu\nu} = [\partial_\mu + (ig/2)B_\mu]B_\nu - [\partial_\nu + (ig/2)B_\nu]B_\mu = B^{b\mu\nu}_b; \quad b = 1, 2, 3$$  (5.27)

where $i$ is imaginary unit and $g$ is coupling constant.

Combining Lagrangian

$$L = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} - \frac{1}{4} B_{\mu\nu}^b B_{\mu\nu}^b$$  (5.28)

with QED Eq. (5.21), the electro-weak Lagrangian is obtained under SU$_L$(2)×UY(1) symmetry as follows.

$$L = -\frac{1}{4} F_{\mu\nu} F_{\mu\nu} - \frac{1}{4} B_{\mu\nu}^b B_{\mu\nu}^b - J^\mu A_\mu$$  (5.29)

The Lagrangian above determined micro-particle physics.

### 5.3.4 Gravitational Field

It is feasible to attach the outer space-time metric $g^{\mu\nu}$ to the linked-field, with keeping Riemann-Christoffel symbol $\Gamma^\lambda_{\mu\nu} = g^\lambda (D_\mu g_{\nu})$. Then the Riemann curvature tensor is kept as

$$R^\lambda_{\mu\nu\kappa} = \partial_\nu \Gamma^\lambda_{\mu\kappa} - \partial_\kappa \Gamma^\lambda_{\mu\nu} + \Gamma^\rho_{\mu\kappa} \Gamma^\lambda_{\rho\nu} - \Gamma^\rho_{\mu\nu} \Gamma^\lambda_{\rho\kappa}$$  (5.30)

and scalar curvature $R$ is defined by Ricci tensor $R = g^{\mu\nu}R_{\mu\nu}$.

For fitting general relativity, geometrical curvature of space-time is equivalent to energy-momentum tensor

$$T_{\mu\nu} = \frac{1}{8\pi G} (G_{\mu\nu} - \Lambda g_{\mu\nu}) = \frac{1}{8\pi G} \left[ R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} (R - 2\Lambda) \right]$$  (5.31)
5.3 Theoretical Approaches by Vortex Mechanism

Considering total action consists of scalar part $L_s$, vector part $L_v$, bi-vector part $L_b$ and matter field $L_m$

$$L = L_s + L_v + L_b + L_m$$  \hspace{1cm} (5.32)

where $L_m = -mc^2$ is the ordinary matter Lagrangian density and $L_s$ is an Einstein-Hilbert type action

$$L_s = -\frac{1}{16\pi G}(R + 2\Lambda)\sqrt{-g}$$  \hspace{1cm} (5.33)

where $R$ is the trace of the Ricci tensor; $G$ is the gravitational constant; $g$ is the determinant of the metric tensor $g_{\mu\nu}$; $\Lambda$ is the cosmological constant.

Meanwhile, $L_v$ is a Maxwell-Proca type action

$$L_v = \frac{k}{8\pi}[m^2 A_\mu A^\nu - V(\psi)]\sqrt{-g}$$  \hspace{1cm} (5.34)

where $m$ is the mass of the vector field $A$; $k$ characterizes the strength of the coupling between a new force and matter; $V$ is a self-interaction potential of vector field $A$.

And $L_b$ is bi-vector action as

$$L_b = -\frac{c}{4\pi G}[B^{\mu\nu} B_{\mu\nu} - V(\psi)]\sqrt{-g}$$  \hspace{1cm} (5.35)

where $B_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $c$ is a constant linking with three parameters $G$, $k$ and $m$.

Similarly to scalar-tensor-vector gravity (STVG) (Moffat 2006), the action integral takes following form

$$S = \int d^4x L = \int d^4x (L_s + L_v + L_b + L_m)$$  \hspace{1cm} (5.36)

which could effectively explain the galaxy rotation curves and so on Brownstein and Moffat (2006).

In the vortex-based physics, there are following predictions:

(1) The natural world follows quantum mechanism and vortex mechanism without supersymmetry. Two standard models are really approaching the reality.

(2) There are the whole rotation and the initial “big bang” of the universe. The whole rotation contributes to the gravitational effects of both matter-energy and dark matter and the initial “big bang” has been driving the expansion of the universe like dark energy.

(3) The curvature of space-time is evolutional, from high curvature to low curvature.
5.4 Discussion: The Vortex-World

The vortex-based physics will lead to vortex-world at both micro-level and macro-level. With coherence of both mathematics and physics, the vortex-world may be real reality (Penrose 2004). The vortex-based is universal, covering both micro-thing and macro-thing and acting local and global universe.

The vortex-based physics looks like a natural and harmonic “theory” with fitting quantum theory and general relativity and particles and anti-particles can be simply unified in it, so that it seems a good choice for unifying contemporary physics.

In the situation of non-existence of super-partners, the vortex-based physics is the simplest theoretical framework for unifying micro-particle and macro-galaxy. Certainly, it is only a phenomenological fitting theory, not a final perfect theory of everything.

Meanwhile, the issue of “dark matter” and “dark energy” perplexed scientists (Ellis 2003). If the problem is kept in two standard models, it means that standard models mismatch the real world or there exist logic faults in the theories.

After the starting of “big bang”, the vortex mechanism supplied a stable evolitional process, so that we see the present universe. Although changes always happen, the world and the universe are relatively stable, so that we are here, the earth keeps steadily running and the stars staying in the sky, during thousands of years.

Furthermore, some special phenomenological topics are discussed and interpreted as follows.

5.4.1 An Interpretation of Wave-Particle Duality

Generally, micro-particle and macro-cosmos can be jointly measured by measure \( M = (\psi, A, B) \), where \( \psi = \varphi + i\theta \) is just a complex scalar particle-bias (massive) function and \( A = V + iU \) looks like a complex vector wave-bias (potential) function while \( B \) characterizes spin (rotation). Meanwhile, \( M \) is also the measure of total energy. When the energy distributes bias \( \psi \), the particle looks like real particle and there is in \( M \)

\[
\psi \gg A \subset M \tag{5.37}
\]

When the energy distributes bias \( A \), the particle looks wave and there is in \( M \)

\[
\psi \ll A \subset M \tag{5.38}
\]

So that the even part \( M_+ \) and odd part \( M_- \) just describe particle-bias distribution and wave-bias distribution respectively.

Since \( \psi \) and \( A \) are particle-bias function and wave-bias function respectively, the \( M \) will look like particle if \( \psi \gg A \) and the \( M \) will look like wave if \( \psi \ll A \). As \( M = (\psi, A, B) \) becomes unified linked-measure for both micro-particle and macro-
Discussion: The Vortex-World

Cosmos, all things have wave-particle duality. For micro-particle, the energy changes easily between $\psi$ and $A$, so that the micro-particle looks like obvious wave-particle duality. For macro-things, the mass-energy concentrates and keeps mostly in $\psi$, so that there is no obviously observed wave-particle duality. When vortex becomes general methodological approach for both micro-thing and macro-thing, the world can be naturally and jointly interpreted.

5.4.2 An Interpretation for the Asymmetry of Particles and Antiparticles

As $M$ can be divided as even part $M_+$ and odd part $M_-$ as follows

$$M_+ = \varphi + B + i\theta = \psi + B = (\psi, B)$$

$$M_- = V + iU = A$$

which give asymmetry algebraic structure.

If particles show $M_+$-bias in $M$ and antiparticles do $M_-$-bias in $\bar{M}$, the asymmetric mechanism of particles and antiparticles is set up.

5.4.3 An Interpretation of Dark Matter and Dark Energy

Combining special relativity with quantum theory, referring to Qian (2014), there is a relation between Hubble constant $H$ and physical constants as

$$H = \frac{cGm_e m_n^2}{2\hbar^2}$$

where $c$ is velocity of light in vacuum; $G$ is gravitational constant; $m_e$ and $m_n$ are mass of electron and neutron respectively. The theoretical computation value $H = 2.299 \times 10^{-18}$ $\text{m} \cdot \text{s}^{-1}$ = 70.937 $\text{km} \cdot \text{s}^{-1} \cdot \text{Mpc}^{-1}$, which matches observational value well.

Recall the definition of dimensionless density parameter $\Omega_j$

$$\Omega_j = \frac{\rho_j}{\rho_c} = \frac{8\pi G \rho_j}{3H_0^2}; \quad \sum_j \Omega_j = 1$$

where the subscript $j$ is one of these following capitals: $m$ for matter; $r$ for radiation (photons plus relativistic neutrinos); $b$ for baryons; $c$ for CDM; $dm$ for dark matter, $em$ for matter-energy; $de$ or $\Lambda$ for dark energy.
If the cosmos consists of energy-mass, dark matter and dark energy, according to both observation and theory, its Friedmann equations will be

$$\Omega_{\text{em}}(t) + \Omega_{\text{dm}}(t) + \Omega_{\text{de}}(t) = 1 + \frac{k}{R^2 H^2} = 1 + \frac{K}{H^2} \quad (5.43)$$

where $K = k/R^2$ is space curvature, with $k = 1, 0$ or $-1$ corresponding to closed ($K > 0$), flat ($K = 0$) or open ($K < 0$) spatially geometries. When $k = 0$ and $\rho_t = \rho_A$, the cosmic model becomes de Sitter universe.

Now insert Eqs. (5.41) into (5.43), it derives without considering dark energy

$$K = \frac{c^2 G^2 m_e^2 m_n^4}{2\hbar^4} (\Omega_{\text{em}} + \Omega_{\text{dm}} - 1) \quad (5.44)$$

yielding

$$\Omega_{\text{dm}} = 1 - \Omega_{\text{em}} + \frac{2\hbar^4 K}{(c G m_e m_n^2)^2} \quad (5.45)$$

Equation (5.45) indicates the relation between dark matter and matter-energy, linking to curvature $K$. It means that the change of curvature reflects the evolution of dark matter, as there is matter-energy conservation. If all physical constants are stable, we can study the evolution of dark matter via the change of curvature, or inverse. If we find $H$ is not stable (for example, $H$ decreases), perhaps $G$ is changing. Or if $K$ and $H$ decrease meantime, dark matter could show stability.

On the dark energy, at the beginning of evolution, it should reach equilibrium to the matter-energy plus dark matter, so Eq. (5.43) had equivalent form as

$$\Omega_{\text{de}} = 1 + \frac{K}{2H^2} = 1 + \frac{2\hbar^4 K}{(c G m_e m_n^2)^2} \quad (5.46)$$

According to experimental and observational evidence, there are no electromagnetic interactions among visible matter-energy, dark matter and dark energy. Since both matter-energy and dark matter expressed gravitational effects, they are induced to space-time curvature, based on Einstein’s general relativity that the gravity was equivalent to the curvature of space-time. As dark energy has driven the cosmic accelerating expansion, it is theoretically explained by the cosmological constant that linked with physical constants referring to Eq. (5.46). The way provides the simplest image mathematically and physically.
5.4 Discussion: The Vortex-World

5.4.4 Black Holes as Astronomical Vortices in The Universe

More and more black holes are discovered by astronomical observations. Not each black hole constructs an independent cosmos, only an astronomical vortex is in the universe.

Visually, the process of a black hole can be viewed as a vortex rotated into a hole, with rotating energy measure, in which $B$ approaches to 0 but the total energy measure $M$ keeps conservation when $A$ and $\psi$ (mass) concentrate and condense to the center of the black hole, so that the center of the black hole forms a strong energy field, absorbing everything.

Since a vortex can be geometrically represented by a linked-measure, the linked-measure matches following dynamic equation via Hamilton principle

$$\frac{\partial H}{\partial s_i} = \partial_\mu M = -\frac{\partial p_i}{\partial t} \quad (5.47)$$

Considering that Hamiltonian energy $H = E = kT$ in thermodynamics, where $k$ is Boltzmann constant and $T$ denotes temperature, it yields

$$\frac{\partial E}{\partial s_i} = k \frac{\partial T}{\partial s_i} = -\frac{\partial p_i}{\partial t} \quad (5.48)$$

The energy way replacing flow-velocity characterized by Navier-Stokes equations, so that Eqs. (5.47) and (5.48) provide the fluid dynamics of energy way, which could introduce another analytical model for black holes.

5.5 More Evidence and Limitations

In the standard model of particle physics, there are two basic classes of elementary particles: bosons with integer-valued spin and fermions with half-integer spin. After supersymmetric hypothesis was introduced, each fermion should have a partner boson and each boson should have a partner fermion. By that hypothesis, each particle in standard model should be associated with its super-partner and each pair of super-partners would share the same mass and internal quantum numbers besides spin. Scientists have suggested to name the super-partner of fermion as $s$-particle (which is boson) and to name the super-partner of boson as particle-info (which is fermion). For example, super-partner of electron is named as selectron and super-partner of photon is named as photino. Incorporating supersymmetry into the standard model requires doubling the number of particles since there is no way that any of the particles in the standard model can be super-partners of each other. According to the supersymmetric hypothesis, the preferred masses for squarks and gluino are estimated around 2 TeV.

However, unfortunately, there is no any direct confirmation for the existence of supersymmetry. The supersymmetry is only an unproved hypothesis. Direct confir-
mation would entail production of super-partners in collider experiments, such as the LHC. But the LHC and other major particle physics experiments never detected supersymmetric partners (as well as evidence of extra dimensions), so that “super-string theory” and “M theory” face to fall into failure or need to be significantly revised. Since experiment discovers neither supersymmetric partners nor extra spatial dimensions, it is necessary to explore new theory for interpreting the universe. Also, since there is no any supersymmetric evidence, and dark matter and dark energy show their existence, it is suggested to accept vortex-based physics without supersymmetry and dark sectors.

We may mention more advantageous and disadvantageous evidence.

The in-spiral gravitational waves look like advantageous evidence, which show that vortices exist everywhere. The Laser Interferometer Gravitational-wave Observatory (LIGO) observation verified the gravitational waves from a binary black hole merger (Abbott et al. 2016).

The cosmic rotation seems a disadvantageous evidence, where a UK group claimed that the universe is isotropic on large scales, without traces of rotations and anisotropic expansion of the universe is strongly disfavored (Saadeh et al. 2016). However, early in 1946, Gamow started the issue of rotating universe (Gamow 1946). Then the rotating angular velocity had been estimated about $10^{-13}$ rad per year (Birch 1982) or less than $10^{-9}$ rad per year (Su and Chu 2009). Also the stability of Gödel universe had been discussed (Novello and Reboucas 1978; Barrow and Tsagas 2004). The issue keeps unsolved.

Then limitations also exit. We still have some unknown things, such as the large-scale structure of the universe, which may be observed if one only uses red-shift to measure distances to galaxies. For example, galaxies behind a galaxy cluster are attracted to it and they seem slightly blue-shifted. On the near side, things are slightly red-shifted. Thus, the environment of the cluster looks a bit squashed if using red-shifts to measure distance. An opposite effect works on the galaxies already within a cluster: the galaxies have some random motion around the cluster center and when those random motions are converted to red-shifts, the cluster appears elongated.

Another issue came from the mass of neutrinos. Someone believes that neutralino is the lightest super-partner of neutrino and simplest possible supersymmetric model consistent with the standard model is the minimal supersymmetric standard model (MSSM), which can include the necessary additional new particles that are able to be super-partners of these in the standard model. However, if there would be no phenomenological super-partner, we have to abandon the supersymmetry.

5.6 Conclusion

Conclusively, vortex-based physics above is characterized as follows.

(1) A mathematical multi-vector is equivalent to a physical linked-measure, which describes a vortex geometrically. Both micro-particle and macro-galaxy can be approached by vortices.
(2) The vortex-based field (linked-field) is a unified field, where electromagnetic field, strong field and electro-weak field can be synthetically united in it. Adding outer space-time metric, gravitational field can be also united together.

(3) The particle-wave duality and the gravity-curvature equivalence can be naturally and reasonably interpreted in the framework of vortex-based physics.

Combining quantum field theory and general relativity, we applied vortex-based field to unify those two standard models, which provided another way to consider physical unification, where physical achievements are kept and two standard models are saved. The vortex mechanism did so for reaching unified physics linking micro-particles and macro-galaxies.

Generally, from tiny particle to huge cosmos, our world is a vortex-based world. Everything looks like vortex and the vortex mechanism provides a beautiful and simple image for interpreting the universe. Furthermore, the choice of measures determines scientific relations, so that the different measures would lead to difference discoveries of scientific laws. When physical vortex (linked-measure) based on the mathematical multi-vector is introduced, the vortex-based field (linked-field) could combine and unify the micro-particles and the macro-galaxies, leading to a mathematically and physically simple unified theory, which might stimulate further studies.

Acknowledgements This chapter is firstly published in this book.

References

Ye, F.Y.: The physical linked-measure works as vortex with linking to turbulence. Phys. J. 1, 209–215 (2015b)
Chapter 6
The Physical Foundation of Information
and the Unification of Physics

It is introduced that order coordinate is the physical foundation of information and suggested that the order-space-time becomes a new unified framework of physics. It is found that the basic constants c, h, k coupled to be a unified constant e, which is just electron. The substance generation hypothesis that “coupling introduces mass and waving produces particles” has been proposed. Order-space-time system as new physical foundation means the unification of gravity theory and thermodynamics, as well as relativity and quantum theory.

Early in 1984, Kovanic had discussed the relation between information and physics (Kovanic 1984). In 1990, Stonier and Wheeler probed into the problem again (Stonier 1990; Wheeler 1990). However, the problem about the physical foundation of information remains. Here, I suggest that we should find out a unique coordinate for information and provide a dimension as the physical foundation of information. The dimension is different from the 5th dimension in five-dimensional Kaluza-Klein theory (Sajko 1999) and in the $M$-theory (Brandhuber et al. 1997). If we can find the real physical meaning of the 5th dimension (Wesson 2002), we may have essential progress in physics. Therefore, the issue is probed as follows.

6.1 Order Coordinate as the Dimension of Information

About 10 years ago, I thought that the temperature might be the 5th dimension as a new coordinate corresponding to space and time. However, temperature means energy so that it may not become an independent coordinate. Due to the same reason, space-time-matter (Wesson 1999) can not become a complete system because matter is not a unique coordinate. What is the new coordinate? It is a basic dimension ignored in physics in the past. I discovered it on the basis of a comparison among mass $M$, energy $E$ and information $I$.
In physics, mass $M$ can be written as

$$M = \int \rho(s)ds$$

(6.1)

where $\rho(s)$ is the function of mass density (international unit: kg·m$^{-3}$) and $s$ is space coordinate (volume).

And energy $E$ can be expressed as

$$E = \int \varepsilon(t)dt$$

(6.2)

where $\varepsilon(t)$ is the function of energy current density (international unit: J·s$^{-1}$) and $t$ is time coordinate.

So, similarly, information $I$ can be defined as

$$I = \int \eta(r)dr$$

(6.3)

where $\eta(r)$ is the function of information density (suggested unit: bit/tet) and $r$ is a new coordinate. It can be called order coordinate or quantity coordinate, which is the basic dimension of the world. We need the coordinate for measure. We do measure anytime and nothing can be separated from quantity in the world. But we ignored the dimension of measure. Order or quantity coordinate is like space and time. Only when there is order or quantity coordinate, can space and time be measured. Thus, order-space-time is indispensable. Order-space-time constructs the cosmic coordinate system in which space is three-dimensional while order and time are one-dimensional. Besides, number relates them together as a whole (Eldon 2000).

The basic 3+2 dimension physical frame with order-space-time coordinate system is shown in Table 6.1.

On the basis of the second law of thermodynamics, we have

$$dH \geq \frac{dQ}{T}$$

(6.4)

<table>
<thead>
<tr>
<th>Coordinate</th>
<th>Constant</th>
<th>Differential</th>
<th>Integral</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order($r$)</td>
<td>$u_{\text{min}} = k$</td>
<td>$u = \frac{dr}{ds}$</td>
<td>$I = \int \eta(r)dr$</td>
<td>$D = \frac{d}{dr}(Iu)$</td>
</tr>
<tr>
<td>Space($s$)</td>
<td>$v_{\text{max}} = c$</td>
<td>$v = \frac{ds}{dt}$</td>
<td>$m = \int \rho(s)ds$</td>
<td>$F = \frac{d}{ds}(mv)$</td>
</tr>
<tr>
<td>Time($t$)</td>
<td>$w_{\text{min}} = h$</td>
<td>$w = \frac{dr}{dt}$</td>
<td>$E = \int \varepsilon(t)dt$</td>
<td>$G = \frac{d}{dt}(Ew)$</td>
</tr>
</tbody>
</table>

Note $k$ is Boltzmann constant; $h$ Planck constant; $c$ velocity of light in vacuum; $\eta$, $\rho$, $\varepsilon$ respectively density of mass, of energy and of information; $m$, $E$, $I$ mass, energy and information separately
where $H$ is entropy; $Q$ is heat; $T$ is temperature. In the system, there are two hidden ideas: information is the positive accumulation and entropy is negative accumulation of an order density function. Considering that information equals negative entropy, because $E = -Q$ and $I = -H$, we get

$$dI \leq \frac{dE}{T} \text{ or } E \geq \int T dI$$

(6.5)

Analogizing with

$$E = mc^2$$

(6.6)

We should have

$$E = ITh^2$$

(6.7)

which contains the idea of unifying macro-cosmos and micro-world and the thought of combining the theory of relativity and the quantum theory, as well as gravity theory and thermodynamics.

### 6.2 Physical Constants with Dimension Analysis

Supposing $\nu$ is frequency and $T$ is temperature, according to energy formula $E = mc^2 = h\nu = kT$, we can only choose $c, h, k$ as basic physical constants in the system. If

$$c = 2.997,92 \times 10^8 \text{ m} \cdot \text{s}^{-1}$$

(6.8)

$$h = 6.626,08 \times 10^{-34} \text{ J} \cdot \text{s}$$

(6.9)

$$k = 1.379,51 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}$$

(6.10)

we have

$$e = \frac{h}{kc} = 1.602,18 \times 10^{-19} \text{ C}$$

(6.11)

For Eq. (6.11) becoming correct in measure unit, C should equal $m^{-1} \cdot s^2 \cdot \text{K}$. Let \([L],[T],[B],[K],[I],[C]\) and \([M]\) be separately dimension of space, time, order, temperature, current, charge and mass, then \([C]=[L^{-1} \cdot T^2 \cdot \text{K}]\). As $C=A \cdot s$ in SI unit system, i.e., \([C]=[IT]\), so \([I]=[L^{-1} \cdot T \cdot \text{K}]\) or \([K]=[L \cdot T^{-1} \cdot I]\).

Considering gravity constant

$$G = \frac{1}{3} \left( \frac{4\pi}{3} \right) \frac{h}{k} = 6.706,55 \times 10^{-11} \text{ K} \cdot \text{s}$$

(6.12)

where $K$ equals $m^3 \cdot s^{-3} \cdot \text{kg}^{-1}$, i.e., \([K]=[L^3 \cdot T^{-3} \cdot M^{-1}]\) or \([M]=[L^3 \cdot T^{-3} \cdot K^{-1}]\).
If unit of information is byte or bit, the density of information will be byte/tet or bit/tet and the unit of order may use tet (trinary system). As \( E = ITh^2 \), we get \( J = \text{kg} \cdot \text{m}^2 \cdot \text{s}^{-2} = \text{bit} \cdot \text{K} \cdot \text{J}^2 \cdot \text{s}^2 \), therefore bit = kg\(^{-1}\) \cdot m\(^{-2}\) \cdot K\(^{-1}\), i.e., \([\text{B}] = [M^{-1} \cdot L^{-2} \cdot K^{-1}]\).

The vacuum is real integration of order-space-time. The vacuum is not empty. The structure of vacuum has been decided by physical constants. When we use vacuum unit system, \( c = h = k = 1\).

### 6.3 Analytical Principle and Gauge Invariance

Whether the physical system changes, we hope to keep unchanged analytical principle in physics, which is Hamilton principle or the least action principle as

\[
\delta \int_s \delta t \ L = 0 \quad (6.13)
\]

Equation (6.13) means that Lagrangian action is the key to dynamic equations. In extended order \((r)\)-space \((s)\)-time \((t)\) framework, the analytical principle should transfer as

\[
\delta \int \delta t \ L(r, s, t) = \delta \int \delta t ' L(r', s', t') = 0 \quad (6.14)
\]

where \(L\) is the Lagrange function of a physical system. For keeping gauge invariance (or Noether’s theorem), it is supposed that the transformations follow

\[
\begin{align*}
  r & \rightarrow r' = e^{\omega} r; \quad \bar{r} \rightarrow \bar{r}' e^{-i\omega} \\
  t & \rightarrow t' = t + \omega; \quad \bar{t} \rightarrow \bar{t}' = \bar{t} - \omega \\
  s_i & \rightarrow s_i' = s_i + d_i \omega; \quad \bar{s_i} \rightarrow \bar{s_i}' = \bar{s_i} - d_i \omega
\end{align*}
\]

where \(i = 1, 2, 3\) indicates the different space dimension. The metrics or line element is different from Wesson’s five-dimensional relativity (Wesson 1999) and the system is different from Fisher’s information theory (Frieden 1990; Lavis and Streater 2002).

### 6.4 Substance Generation

The generation of substances in above 3+2 dimension physics focuses on that “coupling introduces mass and waving produces particles”, an extension of wave-particle duality. A coupling case corresponding to contemporary physics should be drawn as follows.
## Table 6.2 Main particle system

<table>
<thead>
<tr>
<th>Wave/Particles</th>
<th>Particles</th>
<th>Positive particles</th>
<th>Negative particles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Genons</td>
<td>Generation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Single wave</td>
<td>Actors</td>
<td>$G$, $g$, $\gamma$</td>
<td>None</td>
</tr>
<tr>
<td>Couple wave</td>
<td>Leptons Quarks</td>
<td>$e^-$</td>
<td>$\mu^-$</td>
</tr>
<tr>
<td></td>
<td>$\nu_e$</td>
<td>$\nu_{\mu}$</td>
<td>$\nu_{\tau}$</td>
</tr>
<tr>
<td></td>
<td>$d$</td>
<td>$s$</td>
<td>$t$</td>
</tr>
<tr>
<td>Comples wave</td>
<td>Baryons Mesons</td>
<td>$p$</td>
<td>$\Sigma$, $\Xi$</td>
</tr>
<tr>
<td></td>
<td>$n$</td>
<td>$\Lambda$, $\Omega$</td>
<td>$\bar{\Delta}$</td>
</tr>
<tr>
<td></td>
<td>$\pi$</td>
<td>$K$, $\eta$</td>
<td>$J/\psi$, $D$</td>
</tr>
</tbody>
</table>

The origin vibration of order, space and time will produce massless actors. The origin vibration in order direction produces gluon $G$, space direction graviton $g$ and time direction photon $\gamma$. Coupling wave of order-space-time will produce massy particles. The basic coupling wave of order-space introduces down quark $d$; the basic coupling wave of order-time introduces up quark $u$; the basic coupling wave of space-time produces neutrino $\nu_e$ and the basic coupling wave of order-space-time produces electron $e$. All the basic coupling waves are genons or basicons (generation I). Genons are relatively stable. When it is excited by energy, genons will become excitons (generation II and generation III). Excitons are relatively unstable and will decay into genons. The system of main particles is shown in Table 6.2.

Baryons are composed of triad quarks and Mesons of dual quarks, leptons and positive-negative excitons. For example

\[
    u + u + d \rightarrow p \quad (6.18)
\]

\[
    p + e^- + \bar{\nu}_e \rightarrow n \quad (6.19)
\]

\[
    p + \pi^- \rightarrow \Lambda^0 \quad (6.20)
\]

\[
    \mu^- + \bar{\nu}_\mu \rightarrow \pi^- \quad (6.21)
\]

\[
    \mu^+ + \nu_\mu \rightarrow \pi^+ \quad (6.22)
\]

\[
    \pi^+ + \pi^- \rightarrow K^0 \quad (6.23)
\]
The decay is the inverse process of composing. The proton $p$ does not decay because of quark confinement. The mechanism of quark confinement and undivided electric charge is because order-space, order-time, space-time and order-space-time can not be separated.

Two basic ideas for quantitative analysis are: the physical constants relate to the strength of vibration and the constant coupling waves relate to the mass of particles, as well as other possible mechanisms (Witten 1996, 1997; Hawking et al. 2000).

### 6.5 Physical Information and Probable Information

The physical information in this chapter is different from probable information and Fisher information (Frieden 1990). Probable information was defined in classical Shannon information theory. As for Fisher information, it can be considered as an isolated statistical system specified by a coordinate $x$ and its probability density $p(x)$. The function of $p(x)$ called “Fisher information” can be used to measure the degree of disorder of the system due to the spread in $p(x)$.

And there are relations between physical information and probable information or Fisher information. Let $H$ be probable information and $I$ be physical information. When we measure them by unit bit, $H = k \ln 2 I$. When we measure them by unit tet, $H = k \ln 3 I$. When we measure them with unit nat, $H = k \ln e I = k I$. $k$ is Boltzmann constant. Thus, the general relation between $H$ and $I$ should be

$$H = k \ln b I$$

where $b$ is constant.

### 6.6 Cosmos Model

Under the order-space-time framework, a single black-white hole cosmos model can be updated. A new cosmos model is suggested as shown in Fig. 6.1, where huge amounts of dark matter and dark energy can be reunderstood.

The dark matter or negative substances are kept inside the cosmos. All substances fall into a black hole and spurt out from a white hole, which forms a cosmic cycle. The CPT theorem should be correct in the cosmos. All living beings exist in an expansive semi-sphere cosmos. Living cycles such as fruits and seeds are similar to the model. The holographic principle will be real and correct for all things in the cosmos. That means that everything reflects the cosmos.

If the outside radius of the cosmos is $R$ and inside radius $r$, with mass density $\rho_M$ and $\rho_m$ respectively, the total mass of matter $M$ and dark matter $m$ becomes
If $\rho_M \ll \rho_m$, we can have $m \gg M$ even if $r \ll R$. For example, when $\rho_M = 10^{-5} \rho_m$ and $r = 0.2R$, $M/m = 0.00064$, which means $m$ occupies close to 99.5% and $M$ a little more than 0.5%. Therefore, only a small cosmos channel can keep and transfer huge dark matter or dark energy. That is another explanation of the theoretical mechanism of dark matter and dark energy.

According to traditional physical theories, when an object falls into a black hole, any information about the shape of the object or distribution of charge on it is evenly distributed along the horizon of the black hole and is lost to outside observers. That phenomenon is so puzzling that it has been called the black hole information loss paradox. The behavior of the horizon in that situation is a dissipative system that is closely analogous to that of a conductive stretchy membrane with friction and electrical resistance—the membrane paradigm. Because a black hole eventually achieves a stable state with only three parameters, there is no way to avoid losing information about the initial conditions: the gravitational and electric fields of a black hole give very little information about what came in. The information that is lost includes every quantity that can not be measured far away from the black hole horizon, including approximately conserved quantum numbers such as the total
baryon number and lepton number. However, in the new cosmos model above, all information and matter could be emitted from the white hole, therefore, information-mass-energy would keep jointed conservation.

### 6.7 Conclusion

The advantages of order-space-time framework above are concise and terse. In the framework, Einstein gravitational field, Maxwell electromagnetic field and Yang-Mills gauge field could be naturally kept and unified. That is a substantial unification rather than pure geometric image (Cachazo et al. 2002). With the idea that “coupling introduces mass and waving produces particles”, we can get massy substances within wave-particle unification. The vacuum is interpreted as order-space-time integrated space in which three basic physical constants $c$, $h$ and $k$ are embedded and synthesized as electron $e$. The fact that order-space, order-time, space-time, and order-space-time can not independently exist results in quark confinement and undivided leptons and electric charge.

Certainly, all hypotheses need to be verified by experiences and experiments, matching real physical data (Eidelman et al. 2004). After we input information into fundamental physics, physics could be changed, which might generate benefit to stimulate physical progress.


**References**


Eldon, F.Y.: Triad philosophy and triad science based on triad logic. J. Zhejiang University (Science) 1(4), 467–475 (2000)


Part II
Economics: Complex Metrics
Leads to Complex Economics
Chapter 7
A Probe into the Unification of Micro-Macro-Economics: Arrow-Debreu-Mundell-Fleming Model as a Standard Model

For unifying micro-economics and macro-economics, it is proposed that demand-supply relation, marginal effect and multiplier model construct three keys. As aggregate demand (AD)-aggregate supply (AS) can be approached by multiplier model and marginal effect is also connected to multiplier one, it is revealed that micro-Arrow-Debreu (A-D) model can be combined with macro-Mundell-Fleming (M-F) model, which is suggested to become economic standard model.

7.1 Introduction

As an important field of social sciences, economics applied many models for approaching economic phenomena. Micro-economics examines how entities, forming a market structure, interact within a market to create a market system while macro-economics examines the economy as a whole to explain broad aggregates and their interactions as well as effects of monetary policy and fiscal policy (Samuelson and Nordhaus 2005).

As a mathematical description of general equilibrium of economic market, Arrow-Debreu model (Arrow and Debreu 1954) explains micro-foundations of economics, while Mundell-Fleming model (Mundell 1963; Fleming 1962) supplies macro-economic images of goods and finance (Jones and Kenen 1985; Friedman and Hahn 2000). However, those two important models ignored necessary relations in linking each other, as if macro-economics lost its micro-foundations (Weintraub 1977, 1979; Forni and Lippi 1997; Horwitz 2000). For better economic and theoretical analysis, we need a unified micro-macro-economic model for understanding unified economics. Here, I suggest a way to reach the unification via Arrow-Debreu-Mundell-Fleming model as a standard model.
7.2 Theoretical Background

Arrow-Debreu model is a mathematical economic model under certain economic assumptions (i.e., convex preferences, perfect competition and demand independence), which suggests that there must be a set of prices to allow aggregate supplies to equal AD for every commodity in economic system. The model is the central theory of general economic equilibrium and it is often used as a general reference for other micro-economic models. Arrow-Debreu model is one of the most general models of competitive economy and is a crucial part of general equilibrium theory, as it can be used to prove the existence of general equilibrium (or Walrasian equilibrium) of an economy. In general, there may be a lot of equilibria. However, with extra assumptions on consumer preferences, to be specific, their utility functions are strongly concave and twice continuously differentiable, a unique equilibrium exists.

Meanwhile, Mundell-Fleming model is described in terms of an IS-LM-BP graph with the domestic interest rate plotted vertically and real gross domestic product (GDP) plotted horizontally, where IS curve represents the equilibrium of product market, LM curve means the equilibrium of money market and BP addresses the balance of payments or international income-payment. Geometrically, the IS curve is downward sloped and the LM curve is upward sloped, while the BP curve is upward sloped unless there is perfect capital mobility, in which case it is horizontal at the level of the world interest rate.

In the IS-LM-BP graph, under less than perfect capital mobility, the positions of both the IS curve and the BP curve depend on the exchange rate, since the IS-LM graph is actually a two-dimensional cross-section of a three-dimensional space involving all of the interest rate, the income and the exchange rate. Under perfect capital mobility, the BP curve is simply horizontal, and the domestic interest rate equals to that of the world interest rate. In pure IS-LM model, the domestic interest rate is a key component for keeping both the money market and the goods market in equilibrium. However, differentiated from the pure IS-LM model, Mundell-Fleming model adds international financial elements, for fitting the open economy assumption.

7.3 The Keys to Link Micro-Economics with Macro-Economics

According to Weintraub’s conclusion, Arrow-Debreu general equilibrium model (Walrasian equilibrium model) has the following characteristics.\(^1\)

1. There is a class of agents, called consumers, who have preferences over different bundles of final goods.
2. The consumer preferences are sufficiently regular so that preferences can be represented by utility indicators.

(3) Consumers’ income comes from sales of factor services and distributed profits of firms.
(4) Members of another class of agents, called firms, have preferences over output configurations, which lead to profits.
(5) Consumers, taking product and factor prices as given, attempt to maximize utility subject to their income constraint.
(6) Firms, taking product and factor prices as given, attempt to maximize profits subject to technology constraint.
Those mean that we have the first key to link micro-economy with macro-economy, which is the relation of demand and supply.

7.3.1 The First Key: From Micro-Demand-Supply to Macro-AD-AS

The consumers’ demand $D$ and firms’ supply $S$ at individual level as well as AD and AS at national level are basic micro-macro-relations in economics. From micro-demand-supply $D$-$S$ model to macro-AD-AS model, the form is very similar, i.e., demand curve is a downward-sloping curve and supply curve is an upward-sloping curve, as shown in Fig. 7.1, which gives strong sign to unify micro-economics and macro-economics.

Figure 7.1 suggests us that micro-D-S (demand-supply) system can be integrated to become aggregate macro-AD-AS system, while macro-AD-AS system can also reduce to discrete micro-D-S system. With using logarithms of the real output $Y$, nominal money $M$, nominal prices $P$ and nominal wages $W$, the AD-AS framework can be expressed as

\[
Y = a(M - P), \quad a > 0 \tag{7.1}
\]

\[
Y = b(M - W), \quad b < 0 \tag{7.2}
\]

When nominal wages are fixed at level $W^*$, the framework above plus $W = W^*$ constructs Keynesian general theory, where the increase in $M$ will increase both output $Y$ and the price level $P$. However, the linkages of discrete and aggregate processes need another two keys, namely marginal effect and multiplier model.

7.3.2 The Second Key: Marginal Effects in Both Micro-Economics and Macro-Economics

The utility-maximizing consumers and the profit-maximizing firms are considered to be important rules in economics, which can be expressed by marginal effect, as follows.
(1) The ratios of marginal utilities of goods for all consumers are equal to the relative prices of those goods.

(2) The ratios of marginal costs of goods produced by firms are equal to the relative prices of those goods.

(3) The relative marginal revenue products of all inputs are equal for all firms and all goods and are equal to those inputs’ relative prices.

In mathematical analysis of general equilibrium (Samuelson and Nordhaus 2005), the conditions for competitive general equilibrium are classified into two categories: consumer equilibrium and producer equilibrium. In consumer equilibrium, the ratio of the marginal utilities of two goods \((\text{MU}_1/\text{MU}_2)\) is equal to the ratio of their prices \((P_1/P_2)\). In producer equilibrium, the ratio of the marginal costs of two final products \((\text{MC}_1/\text{MC}_2)\) is equal to the ratio of their prices \((P_1/P_2)\). Therefore, in the market, when final products become goods, combining consumer equilibrium and producer equilibrium, the general equilibrium means

\[
\frac{\text{MC}_1}{\text{MC}_2} = \frac{\text{MU}_1}{\text{MU}_2} \tag{7.3}
\]

Equation (7.3) reveals that one of the essences of general equilibrium or Arrow-Debreu mode focuses on marginal effect of goods.

Besides, macro-economics marks that the marginal propensity to consumption (MPC) and marginal propensity to save (MPS) related like mirror images, i.e.,

\[
\text{MPC} + \text{MPS} = 1 \tag{7.4}
\]

Equation (7.4) reveals one of characteristics of national consumption behaviors and the \(1/\text{MPS} = 1/(1-\text{MPC})\) is just the expenditure multiplier or investment multiplier.
7.3.3 The Third Key: AD-AS Model Can Be Approached by Multiplier Model

The existence of a multiplier effect was initially proposed by Keynes’ student Richard Kahn in the 1930s. Although some scholars reject or downplay the importance of multiplier effects, particularly in terms of the long run, the multiplier model has been used as an argument for the efficacy of government spending or taxation relief to stimulate AD.

A multiplier is a factor of proportionality that measures how much an endogenous variable changes in response to a change in some exogenous variables in macro-economics. Including both money multipliers and fiscal multipliers, multiplier model causes macro-economic effects. While a fiscal multiplier is a ratio of a change in national income to the change in government spending, it causes that a money multiplier is one of various closely related ratios of commercial bank money to central bank money under a fractional-reserve banking system. More generally, the exogenous spending multiplier is the ratio of a change in national income to any autonomous change in spending (including private investment spending, consumer spending, government spending, or spending by foreigners on the country’s exports). A money multiplier measures the maximum amount of commercial bank money that can be created by a given unit of central bank money. That is to say, in a fractional-reserve banking system, the total amount of loans that commercial banks allowed to extend (i.e., the commercial bank money that they can legally create) is a multiple of reserves; that multiple is the reciprocal of the reserve ratio and it is an economic multiplier.

One of the key assumptions in the multiplier analysis is that prices and wages are fixed in the short run. Although that is an oversimplification (as many prices adjust quickly in the real world), that assumption captures the point that some of the adjustments to AD shifts will come through output adjustments if some wages and prices are sticky. So, Keynesian economists calculate multipliers that measure the effect on AD only. Then, American economist Alvin Hansen and Paul Samuelson extended it to get the multiplier-accelerator model, i.e., Hansen-Samuelson model.

As multiplier model can approach AD-AS model (Samuelson and Nordhaus 2005), we can combine marginal effect with multiplier model then penetrate demand and supply for unification. Thus, we can find a unified model via three keys to link micro-economics and macro-economics, in which demand and supply, marginal effect and multiplier model are interrelated to each other.

7.4 Macro-Economic Model: More Details

If we apply the following variables: $Y$ as GDP, $C$ as consumption, $I$ as physical investment, $G$ as government spending (an exogenous variable), $T$ as taxes, $X$ as net exports, $M$ as the nominal money supply, $\gamma$ as the nominal interest rate, $P$ as the
price level and $L$ as liquidity preference (real money demand), the Mundell-Fleming model is based on these three following equations.

1. The IS curve
   \[ Y = C + I + G + X \]  
   where $X$ is net exports. That is the GDP equation in demand (payment) view and another equivalent supply (income) view as $Y = C + S + T + M$, where $S$ is save and $M$ is imports.

2. The LM curve
   \[ \frac{M}{P} = L(\gamma, Y) \]  
   Equation (7.6) means that a higher interest rate or a lower income (GDP) level leads to lower money demand.

3. The BP curve
   \[ BP = CA + KA \]  
   where BP is the balance of payments surplus; CA is the current account surplus; KA is the capital account surplus.

Different from micro-economy, there are other two important factors that affect macro-economy, namely inflation $\lambda$ and exchange rate $\pi$.

As consumption and investment expenditure together account for a large percentage (about 80%) of GNP in major nations (Abel 2000), we can focus on the consumption $C$ and investment $I$. In IS components, there is

\[ C = C(Y - T(Y), \gamma - E(\lambda)) \]  
where $E(\lambda)$ is the expected rate of inflation. Higher disposable income or a lower real interest rate (nominal interest rate minus expected inflation) leads to higher consumption spending.

\[ I = I(\gamma - E(\lambda), Y_{-1}) \]  
where $Y_{-1}$ is GDP in the previous period. Higher lagged income or a lower real interest rate leads to higher investment spending. Then

\[ X = X(\pi, Y, Y^*) \]  
where $\pi$ is the nominal exchange rate (the price of domestic currency in terms of units of the foreign currency); $Y$ is GDP and $Y^*$ is the combined GDP of countries that are foreign trading partners. Higher GDP leads to more spending on imports and hence lower net exports. Higher foreign income leads to higher spending by foreigners on the country’s exports and thus higher net exports. A higher $\pi$ leads to lower net exports.
Then the BP components became \( CA = X \), i.e., the current account is viewed as solely consisting of imports and exports. So we have

\[
KA = z(\gamma, \gamma^*) - k
\]  

(7.11)

where \( \gamma^* \) is the foreign interest rate; \( k \) is the exogenous component of financial capital flows; \( z \) is the interest-sensitive component of capital flows; the derivative of the function \( z \) is the degree of capital mobility (the effect of differences between domestic and foreign interest rates upon capital flows \( KA \)).

One of the assumptions about the Mundell-Fleming model is that domestic and foreign securities are perfect substitutes. If the world interest rate \( \gamma^* \) was given, the model can predict that the domestic rate would become the same level of the world rate by arbitrage in money markets.

Under flexible exchange rates, the exchange rate is the third endogenous variable while \( BP \) is set equal to zero. In contrast, under fixed exchange rates, \( \pi \) is exogenous and the BP surplus is determined by the model.

Under both types of exchange rate regime, the nominal domestic money supply \( M \) is exogenous, but for different reasons. Under flexible exchange rates, the nominal money supply is completely under the control of the central bank. However, under fixed exchange rates, the money supply in the short run (i.e., at a given point in time) is fixed based on past international money flows, while as the economy evolves over time these international flows cause future points in time to inherit higher or lower (but pre-determined) values of the money supply. Thus, the changes of money supply and the global interest rate would become the most important factors affecting macro-economy, via inflation \( \lambda \) and exchange rate \( \pi \).

The Mundell-Fleming model under a fixed exchange rate regime also has completely different implications from the closed economy IS-LM model. In the closed economy model, if the central bank expands the money supply the LM curve shifts out. As a result, income goes up and the domestic interest rate goes down. But in the Mundell-Fleming open economy model with perfect capital mobility, monetary policy becomes ineffective. An expansionary monetary policy resulting in an incipient outward shift of the LM curve would make capital flow out of the economy. The central bank under a fixed exchange rate system would have to instantaneously intervene by selling foreign money in exchange for domestic money to maintain the exchange rate. The accommodated monetary outflows exactly offset the intended rise in the domestic money supply, completely offsetting the tendency of the LM curve to shift to the right and the interest rate remains equal to the world rate of interest.

Under the Mundell-Fleming framework of a small economy facing perfect capital mobility, the domestic interest rate is fixed and equilibrium in both markets can only be maintained by adjustments of the nominal exchange rate or the money supply (i.e., by international funds flows). Results for a large open economy, however, can be consistent with those predicted by the IS-LM model. The reason is that a large open economy has the characteristics of both an autarky and a small open economy. In particular, it may not face perfect capital mobility, thus allowing internal policy...
measures to affect the domestic interest rate and it may be able to sterilize balance-of-payments-induced changes in the money supply. Overall, the complexity of macro-economic system is mostly caused by money, so that we need to pay more attention to monetary issues in unification of micro-macro-economics.

7.5 More Discussion: Other Models

Except Arrow-Debreu model and Mundell-Fleming model, there is another important model that links micro-economy with macro-economy, which is Heckscher-Ohlin (H-O) model. H-O model is characterized by the features as relative endowments of the factors of production (land, labor and capital) that determine a country’s comparative advantage. Countries have comparative advantages in those goods for which the required factors of production are relatively abundant locally. That is because the profitability of goods is determined by input costs. Goods that require inputs that are locally abundant will be cheaper to produce than these goods that require inputs that are locally scarce. Is it another bridge across micro-economics and macro-economics?

The standard H-O model assumes that the production functions are identical for all countries concerned. That means that all countries are in the same level of production and have the same technology, yet that is highly unrealistic. Technological gap between developed countries and developing countries is the main concern for the development of poor countries. The standard H-O model ignores all those vital factors when one wants to consider development of less developed countries in the international context. Even between developed countries, technology differs from industry to industry and firm to firm. Indeed, that is the very basis of the competition between firms, inside the country and across the country. Nevertheless, the standard H-O theory assumes the same production function for all countries. That implies that all firms are identical. The negative theoretical consequence is that there is no room for firms in the H-O model, thus the H-O model is not a good bridge across micro-economics and macro-economics.

Although the H-O model can be extended to the Heckscher-Ohlin-Vanek (H-O-V) model, they all perform poorly in their predictive power. H-O theory is also badly adapted to analyze the South-North trade problems. Meanwhile, modern econometric estimates have shown that the H-O model performs poorly, particularly, in ignoring the roles of finance and bank (Lengwiler 2004; Mishkin 2004; Ueda 2013). Therefore, the H-O model can not be put into standard micro-macro-economic unification.

There are many other economic models, such as Cobb-Douglas model of production, Solow-Swan model of economic growth, Black-Scholes model of option pricing and so on. However, those concern special economic issues, without global micro-macro-economic modeling, thus they are excluded to standard model. Moreover, Lucas model (Lucas 1972, 1973), Fischer model (1977) and Taylor model (1979) could be other non-standard models, for the imperfect information and imperfect competition cases.
7.6 Conclusion and Suggestion

While Arrow-Debreu model describes general micro-economic equilibrium, Mundell-Fleming model expresses general macro-economic one. There are three keys to link Arrow-Debreu model and Mundell-Fleming model together: (1) demand-supply relation, which penetrates micro-economy and macro-economy with similar pattern; (2) marginal effect, which reveals same principles in micro-economy and macro-economy and constructs expenditure multiplier or investment multiplier; (3) multiplier model, which approaches to AD-AS. As AD-AS is approached by multiplier model and linkage of expenditure multiplier or investment multiplier to marginal effects, it is revealed that micro-Arrow-Debreu model can be combined with macro-Mundell-Fleming model. Indeed, a macro-economic model is much more complicated than micro-economic one, because there are complex money interactions, as well as global interest rates and exchange rates. Fortunately, demand-supply relation, marginal effect and multiplier model become key bridges across them. Thus, we can understand unified micro-macro-economics. For the unified understanding, it is suggested that the Arrow-Debreu-Mundell-Fleming model become standard economic model, which is characterized by a geometrical analytical system constructed by AD-AS-IS-LM curves. The unified system could promote further studies.

Acknowledgements This chapter is a revision of the original version published at Euro-Asian Journal of Economics and Finance, 2015, 3(1): 1–8.

References

Dornbusch, R.: Exchange rate expectations and monetary policy. J. Int. Econ. 6(3), 231–244 (1976)
Lucas, R.E.: Expectations and the neutrality of money. J. Econ. Theory 4, 103–124 (1972)
Chapter 8

The commodity function \( X \) and the money function \( Y \) are introduced to construct a unified micro-macro-economic analytical framework. In the commodity-money analytical framework, deflation and inflation are naturally interpreted. By considering that commodity is real and money looks imaginary that reflects the value of commodity, complex economics could be set up, where the module \( |M| = |Y/X| \) of economic measure \( M = X + iY \) is applied as unified marginal-multiplier for linking micro-economics and macro-economics. That could benefit economic analysis by providing a new perspective framework.

8.1 Introduction

Since economics became an independent science, it has evolved into many branches including micro-economics, macro-economics, international economics and so on. Although there are different economic thoughts in different economic branches, one would expect to have unified economic models for understanding economics and benefiting economic analysis.

At the beginning, a commodity is defined as a substantially fungible marketable item produced to satisfy needs, including goods and services. After money is introduced to intermediate the exchange of goods and services, it performs a function as a special commodity or a medium of exchange, acting as a standard measure and common denomination of trade, with currency as a universal equivalent. Then, commodity and money interact with each other, constructing complicated economic system.

In both micro-economics and macro-economics, the demand and supply relation is a basic relation for economic analysis, where micro-D-S model at individual level (including firm, family and person) and macro-AD-AS model at national level look similar, as demand curve is a downward-sloping curve and supply curve is an upward-sloping curve. The relevant knowledge can be found in any standard economic textbook (Samuelson and Nordhaus 2005).
Methodologically (Blaug 1992), there is another basic relation, which could be applied in economic analysis. That is the commodity-money relation: commodity is real and money seems imaginary, but money reflects the value of commodity. As changes in money will lead to changes in AD, the commodity-money relation could provide a complement to demand-supply analytical framework. Meanwhile, marginal effect is a famous theory to explain the discrepancy in the value of commodity and services by reference to their marginal utility in micro-economics, while a multiplier is a factor of proportionality that measures how much an endogenous variable changes in response to a change in some exogenous variable in macro-economics. When I combine the ideas above together, a new economic analytical framework can be introduced and mathematical methods of complex numbers can be applied, for approaching unified economics (Ye 2015).

8.2 Methodology: The Commodity-Money Analytical Framework

In order to use symbols in a clear way, I represent micro-economic measures with small Latin letters and macro-economic measures with capital Latin letters, as well as record rates and ratios as small Greek ones.

Let’s consider that there are two basic functions in economic system: commodity function (micro, \(x\); macro, \(X\)) and money function (micro \(y\); macro \(Y\)), which are determined by two kinds of economic variables (micro \(p\), macro \(P\); micro \(q\), macro \(Q\)). \(P\), \(Q\) or \(p\), \(q\) are general symbols, which can be economic factors or measures (capital, consumption, income, investment, labor, land, price, etc.), in which the values of all factors or measures are calculated with money unit (such as dollar). We write micro-commodity function

\[
x = x(p, q)
\]

micro money function

\[
y = y(p, q)
\]

macro commodity function

\[
X = X(P, Q)
\]

macro money function

\[
Y = Y(P, Q)
\]

Consider \(x\) or \(X\) is real and \(y\) or \(Y\) reflects the value of \(x\) or \(X\). Theoretically, increasing (or decreasing) commodity will cause money increase (or decrease) and increasing (or decreasing) money will affect commodity increase (or decrease), so that the commodity curve and money curve will spread upward-sloping, keeping in
the same direction. If there is no special limitation, we have two basic patterns of commodity-money relations, type I and type II, as shown in Figs. 8.1 and 8.2 respectively, where $P(p)$ denotes nominal prices and $Q(q)$ marks quantity of commodity (products) or money (income) and their values are uniformly measured by money units (such as dollar). IS-LM model keeps normal meaning.

Since IS curve represents the equilibrium of product (commodity) market and LM curve means the equilibrium of money market, the cross point of IS and LM curves is the equilibrium point of commodity and money, which is just the $E$ point in both Figs. 8.1 and 8.2.

In Fig. 8.1 (model I), the commodity curve starts from $O$ point to upward-sloping direction and money curve spreads up more quickly, which get equilibrium at the point $e$ or $E$, where commodity are equal to money. In Fig. 8.2 (model II), the money curve starts from $O$ point to upward-sloping direction and money curve spreads up more quickly, which get equilibrium at the point $e$ or $E$, where money is equal to commodity. Obviously, there is a common characteristic in both model I and model II: commodity and money need simultaneous increase or decrease for equilibrium. If we raise the commodity production, we have to add the money supply.

![Fig. 8.1 Basic commodity-money model I](image1)

![Fig. 8.2 Basic commodity-money model II](image2)
Both in Figs. 8.1 and 8.2, there are three structural parts: equilibrium point \( E \) or \( e \); area \( A \), where money is less than commodity, which means commodity is rich; area \( B \), where commodity is less than money, which means money is much more.

In IS-LM model or more generalized Mundell-Fleming model (Mundell 1963a; Fleming 1962), commodity equilibrium and money equilibrium are fixed in IS line and LM line respectively. In the commodity-money framework, commodity and money curves extend to same upward-sloping with different slopes, thus there produce new analytical spaces. In addition, different from the demand-supply framework, the commodity-money framework could provide another perspective for understanding economic phenomena (especially for inflation), which could enrich economic analysis.

### 8.3 An Interpretation of Inflation and Deflation

The commodity-money analytical framework is useful in analyzing inflation and deflation. Although there are Sidrauski or Sidrauski-type model (Sidrauski 1967) for explaining inflation (Mundell 1963b; Friedman 1963; Hahn 1983; Friedman and Hahn 1990) and Keynesian model for interpreting deflation or depression (Tobin 1975), the commodity-money analytical framework provides a simple visual understanding.

The inflation could happen in area \( B \), where we can see that there exist two types of inflation: type I in model I and type II in model II respectively. To cope with inflation type I, the governance strategy should diminish commodity production and money supply for returning to equilibrium (let commodity decrease less than money), because there are excesses of commodity and money. As for inflation type II, the governance strategy should promote commodity production and money supply to reach equilibrium (let commodity increase more than money), because there is shortage of commodity and money.

It is widely accepted that different types of inflation need different policies of governance. For example, as type I in Fig. 8.1, the suggested governance policies are shown in Fig. 8.3.

If the commodity (product) is in shortage, diminishing money could cause deflation (Fig. 8.3a), so that the governance policies should stimulate commodity production and increase money supply. If the money is excessive, raising commodity and causing inflation (Fig. 8.3b), the governance policies should diminish commodity production and decrease money supply.
8.4 Some Special Cases

Any analytical framework may meet some special cases. For example, in macro-case of Fig. 8.1, there may be following special cases as shown in Fig. 8.4.

In Fig. 8.4, some special cases are concluded. In Fig. 8.4a, it is just a local view of Fig. 8.1, in area B, so that there is no independent meaning. In Fig. 8.4b, money supply maintains rigid while commodity supply changes, which means money keeps stable to commodity supply, so that the commodity-money equilibrium is replaced by money stable status. In Fig. 8.4c, it is indicated that excessive quantity could cause price decreasing so that both commodity production and money supply might go down, in which double equilibrium points may exist, where the left one ($E_1$) may be stagflation point and the right one ($E_2$) marks the real effective equilibrium. Different from basic model I (Fig. 8.1), perhaps Fig. 8.4c indicates more generalized cases, where commodity production and money supply go down when their quantities

Fig. 8.4 Some special cases of model I
become excessive over $E_2$. Area $C$ means another kind of deflation. For model II, there are similar results.

Generally, economic complexity is mainly caused by money. Thus, the financial economics, particularly international financial economics (Mishkin 2004), becomes an important branch in economics. Certainly, if economic growth and microfoundations are synthetically considered (Tobin 1965; Sidrauski 1967; Dornbusch 1976; Weintraub 1979; Forni and Lippi 1997), the complexity could dramatically increase and a simplified methodology should be explored.

### 8.5 Complex Economics: Using Complex Numbers as Unified Marginal-Multiplier Measures

As commodity is real and money looks imaginary (but money reflects the value of commodity), we can introduce complex number as unified measurement as follows.

**Micro-economic measure**

$$m(z) = x(p, q) + iy(p, q)$$  \hspace{1cm} (8.5)

**Macro-economic measure**

$$M(Z) = X(P, Q) + iY(P, Q)$$  \hspace{1cm} (8.6)

Now, the module $|M| = |Y/X|$ of complex measure $M$ constructs potential indicator for judging inflation: if $|M| > 1$, there is potential risk of the inflation type I; if $|M| < 1$, there is potential risk of the inflation type II.

Using calculation of complex numbers, we define general marginal effect as

$$0 < M_e = \frac{\partial m}{\partial z} = \frac{1}{2} \left( \frac{\partial m}{\partial p} - i \frac{\partial m}{\partial q} \right) < 1$$  \hspace{1cm} (8.7)

$$0 < M_E = \frac{\partial M}{\partial Z} = \frac{1}{2} \left( \frac{\partial M}{\partial P} - i \frac{\partial M}{\partial Q} \right) < 1$$  \hspace{1cm} (8.8)

If $m$ or $M$ is analytical function, according to Cauchy-Riemann condition, the complex functions $m$ and $M$ will fit

$$\frac{\partial x}{\partial p} = \frac{\partial y}{\partial q}$$  \hspace{1cm} (8.9)

$$\frac{\partial x}{\partial q} = -\frac{\partial y}{\partial p}$$  \hspace{1cm} (8.10)
\[
\frac{\partial X}{\partial P} = \frac{\partial Y}{\partial Q} \tag{8.11}
\]

\[
\frac{\partial X}{\partial Q} = -\frac{\partial Y}{\partial P} \tag{8.12}
\]

These are constraint conditions in complex economics: the marginal increase of commodity on price equals that of money on quantity, while the marginal increase of commodity on quantity is equal to the negative marginal increase of money on price.

With complex numbers, the commodity-money relations become clear and simple, so that the complex economics could promote the developments of economic analysis.

### 8.6 Discussion: A Conjecture on Linking Marginal Effect to Multiplier Effect

The marginal effects originate from micro-economics, where marginal principle suggests that firms or individuals will maximize their profits or incomes by counting only the marginal costs and marginal benefits in their decision. There is the law of diminishing marginal utility (Gossen’s first rule) and the profit-maximizing firms and the utility-maximizing consumers mean that firms’ marginal cost and consumers’ marginal utility will equal each other for equilibrium. When the concepts of MPC and MPS are extended to macro-economics, there is a famous equation \(\text{MPC} + \text{MPS} = 1\).

Essentially, as additional amounts of a commodity or service are added to available resources, their marginal utilities are decreasing. The law of diminishing marginal utility is sometimes treated as a tautology, sometimes as something proven by introspection, or sometimes as a mere instrumental assumption, adopted only for its perceived predictive efficacy.

Actually, an individual will typically be able to partially order the potential uses of a commodity or service. If there is scarcity, then a rational agent will satisfy wants of highest possible priority, so that no want is avoidably sacrificed to satisfy a want of lower priority. The absence of complementarity (Samuelson 1974) across the uses will imply that the priority of use of any additional amount will be lower than that of the established uses. However, if there is a complementarity across uses, then an amount added can bring things past a desired tipping point or an amount subtracted cause them to fall short. In such cases, the marginal utility of a commodity or service might actually be increasing.

Without the presumption that utility is quantified, the diminishing of utility should not be taken to be itself as an arithmetic subtraction. It is the movement from use of higher to lower priority and may be no more than a purely ordinal change. When quantification of utility is assumed, diminishing marginal utility corresponds to a utility function whose slope is continually or continuously decreasing. In the latter
case, if the function also performs smoothly, then the law of diminishing marginal utility may be expressed.

Meanwhile, the rate of substitution is the least favorable rate at which an agent is willing to exchange units of one commodity or service for units of another. The marginal rate of substitution (MRS) is the rate of substitution at the margin, in other words, giving some constraints.

The marginal rate of substitution of commodity or service \( A \) for commodity or service \( B \) (\( \text{MRS}_{AB} \)) is also equivalent to the marginal utility of \( A \) over the marginal utility of \( B \), i.e.,

\[
\text{MRS}_{AB} = \frac{\text{MU}_A}{\text{MU}_B} \quad (8.13)
\]

so the marginal rate of substitution is the slope of the indifference curve.

There is another concept concerning marginal cost. At the highest level of generality, a marginal cost is a marginal opportunity cost. In most contexts, however, marginal cost will refer to marginal pecuniary cost, which means that marginal cost can be measured by forgone money.

Besides, originated from macro-economics, multiplier could be used in creating multiplier-accelerator principle (Hansen-Samuelson model). A multiplier is a factor of proportionality that measures how much an endogenous variable changes in response to a change in some exogenous variable in macro-economics. The multipliers include fiscal multiplier, monetary multiplier and complex multiplier (i.e., an externally induced change in consumption, investment, government expenditure or net exports). Each of those operates to increase or reduce the equilibrium level of income in the economy as the following situations: (1) any increase to an injection will be multiplied to result in a higher level of aggregate expenditure; (2) any decrease in an injection will be multiplied to result in a lower level of aggregate expenditure; (3) any increase in a withdrawal will be multiplied to result in a lower level of aggregate expenditure; (4) any decrease in a withdrawal will be multiplied to result in a higher level of aggregate expenditure. Therefore, the complex multiplier can be measured by Eq. (8.14).

\[
k = \frac{1}{(\text{MPS} + \text{MRT} + \text{MPM})} = 1/\text{MPW} \quad (8.14)
\]

where \( \text{MPS} \) equals marginal propensity to save; \( \text{MRT} \) equals marginal rate of taxation; \( \text{MPM} \) equals marginal propensity to import and \( \text{MPW} \) equals marginal propensity to withdraw.

However, complex multiplier is not mathematically complex number. According to analytical logic, we can introduce and define general complex multiplier with above general marginal effects Eqs. (8.7) and (8.8) as follows.

\[
k_M = \frac{1}{(1 - M_e)} \quad (8.15)
\]

\[
k_m = 1 - M_E \quad (8.16)
\]
where $k_M$ is macro-multiplier, which could enlarge micro-effect of marginal measure; $k_m$ is micro-multiplier, which could reduce macro-effect of marginal measure. Besides, micro-multiplier $k_m$ would diminish economic effects at individual level via macro-marginal effect $M_E$, while macro-multiplier $k_M$ would increase economic effects at national level via micro-marginal effect $M_e$. Since $|M| = 1$ marks the equilibrium, it is suggested that $M_0 = |M|X$ and $M_1 = k_M|M|X$ can be treated as the designed start-point of monetary supply.

Since there are macro-multiplier Eq. (8.15) and micro-multiplier Eq. (8.16), it is proposed that all economic measures could be enlarged by macro-multiplier and reduced by micro-multiplier. That means that any macro-economic effects of economic measures should multiply $k_M$ and any micro-economic effects of economic measures should multiply $k_m$, such as the macro-effect of consumption $C$:

$$C_M = Ck_M \quad (8.17)$$

the macro-effect of investment $I$

$$I_M = Ik_M \quad (8.18)$$

the micro-effect of government purchase $G$

$$G_m = Gk_m \quad (8.19)$$

Dynamic multipliers can also be calculated. That is to say, one can ask how a change in some exogenous variable in year $t$ affects endogenous variables in year $t$, in year $t + 1$, in year $t + 2$ and so forth. A graph showing the impact on some endogenous variable, over time (i.e., the multipliers for times $t$, $t + 1$, $t + 2$, etc.), is called an impulse-response function. The general method for calculating impulse response functions has been sometimes called comparative dynamics.

In a word, an economic measure could be enlarged by multiplier effect acting in macro-economy and a macro-economic policy could be reduced by inverse-multiplier effect. Although money raises the complexity of economic system (Friedman 1970; Lucas 1972, 1973), it also increases the variety of economics (Tobin 1982).

### 8.7 More Analysis: Relations Among the Rates of Inflation, Interest and Exchange

Following the approach of macro-economics, MPC and MPS are linked by equation

$$\text{MPC} + \text{MPS} = 1 \quad (8.20)$$
where $1/MPS = 1/(1 - MPC)$ is just the expenditure multiplier or investment multiplier.

Here, MPC plays a crucial role in Keynesian models of AD, which links the rate of interest $\gamma$ by the following equation

$$MPC = \frac{\gamma}{1 + \gamma - a}$$  \hspace{1cm} (8.21)

where $a$ is a parameter and MPC is an increasing function of parameter $a$. When $a = 0$, we obtain

$$MPC = \frac{\gamma}{1 + \gamma}$$  \hspace{1cm} (8.22)

Combining Eq. (8.22) with Eq. (8.20), we see that $MPS = 1/(1 + \gamma)$.

Since a steady-state economy will be characterized by an inflation rate $\lambda$ and the nominal rate of interest $\gamma$, the real rate of interest is, by definition $\gamma^* = \gamma - \lambda$ (Feldstein 1976). Meanwhile, let $\pi$ be the exchange rate and $\pi'$ the real exchange rate, there is Dombusch’s equation (Dombusch 1976)

$$\gamma = \gamma^* + \frac{\pi'}{\pi} - 1$$  \hspace{1cm} (8.23)

and if the elasticity of expectations $\sigma$ is less than unity, then we have

$$\frac{d\gamma}{d\pi} = \sigma - 1 < 0$$  \hspace{1cm} (8.24)

When domestic output is $Y = E(, y) + T(\beta, y)$, the differentiation of income with regard to the exchange rate becomes

$$\frac{dY}{d\pi} = \frac{1}{1 - \partial E/\partial y - \partial T/\partial y} \left( \frac{\partial E}{\partial \gamma} \frac{d\gamma}{d\pi} + \frac{\partial T}{\partial \gamma} \right)$$  \hspace{1cm} (8.25)

Thus, we know that the rates of inflation, interest and exchange are linked together, which partially reveals reasons of economic complexity.

### 8.8 Concluding Remarks

Combination of the commodity-money analytical framework with the demand-supply analytical framework can provide economic analysis a more complete picture of economic phenomenon and rich details as well. With using complex number in economics, economics might be explored from a different perspective and with comprehensive findings. Any economic measurement could be described by complex, where real part denotes economic commodity-entity and imaginary part indicates eco-
nomic money-shadow. Each economic function \( E(P, Q) = X(P, Q) + iY(P, Q) \) contains real part \( X \) and imaginary part \( Y \). If \( X \) describes real commodity and \( Y \) denotes financial money, \(|Y/X|\) constructs a ratio measurement of financial economy. Analytically, micro-multiplier \( k_m \) would diminish economic effects at individual level via macro-marginal effect \( M_E \), while macro-multiplier \( k_M \) would add economic effects at national level via micro-marginal effect \( M_e \). Although there is no fixed relation between micro-economic function \( m = x(p, q) + iy(p, q) \) and macro-economic function \( M = X(P, Q) + iY(P, Q) \), their modules \(|m| = |y/x|\) and \(|M| = |Y/X|\) construct potential unified measures in economic analysis, which could promote and stimulate further studies.

Acknowledgements This chapter is a revision of the original version published at Euro-Asian Journal of Economics and Finance, 2015, 3(1): 44–52.

References


Dornbusch, R.: Exchange rate expectations and monetary policy. J. Int. Econ. 6(3), 231–244 (1976)


Lucas, R.E.: Expectations and the neutrality of money. J. Econ. Theory 4, 103–124 (1972)


Chapter 9
Economic Complex Analysis
for Approaching Economic Equilibrium
and Economic Stability

Under the framework of complex economics, it is found that the Cauchy-Riemann
equation and the Laplace equation construct the core of economic complex analysis,
which describes economic equilibrium and economic stability. Combining funda-
mental solution of Cauchy-Riemann equation and Laplace equation, the stable eco-
nomic functions are revealed, in which commodity function $X$ and money function
$Y$ as complex economic function $Z = X + iY$ link with price level $P$ and output
quantity $Q$. The numerical simulations provide visual results, with multi-solutions,
while the fundamental solutions keep unique.

9.1 Introduction

In economics, the economic equilibrium is a state where economic forces such as
supply and demand are balanced and the equilibrium values of economic variables
will not change in the absence of external influences. The economic stability refers
to an absence of excessive fluctuations. An economy with fairly constant output
growth and low and stable inflation would be considered economically stable and
the equilibrium is just the outcome of the stability.

For approaching economic equilibrium and economic stability, the general equi-
librium theories have been developed in both micro-economics (Arrow and Debreu
1954) and macro-economics (Tobin 1969). Moreover, the general equilibrium and
economic analysis are core topics in economics, (Fischer 1983; Friedman 1957,
1970, 1975; Lucas 1972, 1973; Papademos and Modigliani 1990) particularly in
monetary economics which also concerns economic stability and monetary mech-
anism (Samuelson 1947; Friedman and Hahn 1990; Friedman 1948; Hahn 1960;
Modigliani 1963). Following the framework of complex economics (Ye 2015a, b), I
try to develop economic complex analysis as follows and expect to benefit and enrich
economic analysis as well as economic theories.
9.2 Complex Economics as Theoretical Framework

Let’s consider a general case: total social real warfare as commodity $X$ (including capital $K$ and product $J$) and imaginary warfare as money $Y$ (including existing money $M_e$ and increasing money $M_i$), with introducing commodity function $X = X(P, Q)$ and money function $Y = Y(P, Q)$ and both $X$ and $Y$ are changed by real variables $P$ and $Q$, where $P$ indicates price level and $Q$ means quantity. We define complex economic function as

$$M = M(X, Y) = Z(P, Q) = X(P, Q) + iY(P, Q) \quad (9.1)$$

If the functions $X$ and $Y$ are analytic, they will fit Cauchy-Riemann condition

$$\frac{\partial X}{\partial P} = \frac{\partial Y}{\partial Q} = 0 \quad (9.2)$$

$$\frac{\partial X}{\partial Q} = \frac{\partial Y}{\partial P} = 0 \quad (9.3)$$

If $X$ and $Y$ are harmonic, they will fit Laplace equation

$$\frac{\partial^2 X}{\partial P^2} + \frac{\partial^2 Y}{\partial Q^2} = \frac{\partial^2 X}{\partial P^2} - i^2 \frac{\partial^2 Y}{\partial Q^2} = 0 \quad (9.4)$$

Now let’s replace $X$ and $Y$ with unified $Z = X + iY$, a complex stable economic system should fit equations

$$\frac{\partial Z}{\partial P} + i \frac{\partial Z}{\partial Q} = 0 \quad (9.5)$$

$$\frac{\partial^2 Z}{\partial P^2} + \frac{\partial^2 Z}{\partial Q^2} = 0 \quad (9.6)$$

where Eq. (9.5) is Cauchy-Riemann equation and Eq. (9.6) Laplace equation.

In general, different initial and boundary conditions will determine different solutions of Eqs. (9.5) and (9.6). However, we can discuss their fundamental analytical solutions according to the following equations

$$\frac{\partial Z}{\partial P} + i \frac{\partial Z}{\partial Q} = \delta(P, Q) \quad (9.7)$$

$$\frac{\partial^2 Z}{\partial P^2} + \frac{\partial^2 Z}{\partial Q^2} = \delta(P, Q) \quad (9.8)$$

Those are first-order and second-order partial differential equation respectively.
9.3 Analytical Solutions

Doing Fourier transformation of both sides to $Q$ for Eq. (9.7), we obtain an ordinary differential equation on $P$ with parameter $s$ as follows.

$$\frac{d\hat{Z}(P, s)}{dP} - s\hat{Z}(P, s) = \delta(P) \quad (9.9)$$

The solution of Eq. (9.9) is

$$\hat{Z}(P, s) = (H(P) + C(s))e^{sP} \quad (9.10)$$

where $C(s)$ is any constant depended by $s$. Then the fundamental solution of Eq. (9.7) becomes

$$Z(P, Q) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isP} \hat{Z}(P, s) ds = \frac{1}{2\pi(P + iQ)} \quad (9.11)$$

For Eq. (9.8), it is two-dimensional Laplace equation and its fundamental solution is

$$Z(P, Q) = \frac{1}{2\pi} \ln(P^2 + Q^2)^{1/2} = \frac{1}{4\pi} \ln(P^2 + Q^2) \quad (9.12)$$

In order to match Eqs. (9.1), (9.11) and (9.12), we need

$$Z(P, Q) = X + iY = \frac{1}{2\pi(P + iQ)} = \frac{1}{4\pi} \ln(P^2 + Q^2) \quad (9.13)$$

so that the equilibrium condition can be written as

$$\frac{1}{(P + iQ)} = \frac{1}{2} \ln(P^2 + Q^2) = \frac{1}{2} \ln(P - iQ)(P + iQ) \quad (9.14)$$

It is equivalent to

$$\frac{(P + iQ)}{2} \ln(P - iQ)(P + iQ) - 1 = 0 \quad (9.15)$$

Equations (9.13) or (9.14) and (9.15) construct the mathematical condition of economic equilibrium and economic stability, which means that stable economic system (characterized by commodity function $X$ and money function $Y$) link with special price level $P$ and output quantity $Q$.

As we know that Laplace equation just describes stable phenomena, Eqs. (9.7) and (9.8), as well as Eqs. (9.11) and (9.12) just fit the economic stability. As there are two equations (second-order Laplace equation and first-order Cauchy-Riemann equation), the solutions describe economic equilibrium.
For visually expression, the numerical methods may be applied to study Eqs. (9.11), (9.12) and (9.15).

### 9.4 Numerical Simulation

In order to understand the economic behaviors of solution Eqs. (9.11), (9.12) and (9.15), we try to simulate their numerical results, with MATLAB.

The visually complex solutions of Eq. (9.11) can be expressed by following program in MATLAB, as shown in Fig. 9.1.

```matlab
figure (9-1)
z = 5*cplxgrid(30);
p = real(z);
q = imag(z);
s = p + q*i;
w = s.^(-1)/(2*pi)
%mesh(p, q, z);
surf(p, q, real(w), imag(w));
colorbar
```

With using the same method, the visually complex solutions of Eq. (9.12) can be shown in Fig. 9.2 by the following program.

```matlab
figure (9-2)
z = cplxgrid(30);
p = real(z);
q = imag(z);
w = log(p.^2 + q.^2)/(4*pi);
%mesh(p, q, z);
surf(p, q, real(w), imag(w));
colorbar
```

![Fig. 9.1](image) The simulation of complex solutions of Eq. (9.11)
9.4 Numerical Simulation

Fig. 9.2 The simulation of complex solutions of Eq. (9.12)

In both Figs. 9.1 and 9.2, the central figures show the real part and the color bar does the imaginary part.

Both Figs. 9.1 and 9.2 reveal the solutions of Eqs. (9.11) and (9.12). The basic shape of their united solution is determined by the following equation

$$Z = \frac{(P + iQ)}{2} \ln(P - iQ)(P + iQ) \quad (9.16)$$

That is the solution surface, which can be shown using the following program, as shown in Fig. 9.3.

```matlab
*figure (9-3)
z=cplxgrid(30);
p=real(z);
q=imag(z);
s=p+q*i;
w=s.*log(p.^2 + q.^2)/2
%mesh(p, q, z);
surf(p, q, real(w), imag(w));
colorbar
```

The visualized results verify the existence of complex solutions. However, the solutions belong to multi-solutions.

vspace*1.5pc
9.5 Discussion

There exist various particular solutions (Polyanin and Manzhirov 2007) of Laplace Equation (9.6), which means that there are a number of cases (multi-conditions) of economic equilibrium and economic stability.

For example, Eq. (9.6) has its polar coordinate form as

\[
\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial Z}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 Z}{\partial \phi^2} = 0 \tag{9.17}
\]

where \( r = \sqrt{P^2 + Q^2} \) and \( P = r \cos \phi; \ Q = r \sin \phi \).

Equation (9.17) has following particular solutions

\[
Z(r) = A \ln r + B \tag{9.18}
\]

\[
Z(r, \phi) = (Ar^m + \frac{B}{r^m})(C \cos m\phi + D \sin m\phi) \tag{9.19}
\]

where \( A, B, C, D \) are arbitrary constants and \( m = 1, 2, \ldots \).

On its Cartesian coordinate form, Eq. (9.6) has following particular solutions.

\[
Z(P, Q) = AP + BQ + C \tag{9.20}
\]

\[
Z(P, Q) = A(P^2 - Q^2) + BPQ \tag{9.21}
\]
\[ Z(P, Q) = A(P^3 - 3PQ^2) + B(3P^2Q - Q^3) \] (9.22)

\[ Z(P, Q) = \frac{AP + BQ}{P^2 + Q^2} + C \] (9.23)

\[ Z(P, Q) = \exp(\pm nP)(A \cos nQ + B \sin nQ) \] (9.24)

\[ Z(P, Q) = (A \cos nP + B \sin nP) \exp(\pm nQ) \] (9.25)

\[ Z(P, Q) = (A \cos nP + B \sin nP)(C \sinh nQ + D \cosh nQ) \] (9.26)

\[ Z(P, Q) = (A \sinh nP + B \cosh nP)(C \cos nQ + D \sin nQ) \] (9.27)

where \( A, B, C, D \) and \( n \) are arbitrary constants.

If \( Z(P, Q) \) is a solution for the Laplace equation, then following functions are also solutions everywhere they are defined, in which \( A, B, C, D, \theta \) and \( \lambda \) are arbitrary constants. The signs at \( \lambda \)'s in the equation for \( Z_1 \) are taken arbitrarily.

\[ Z_1(P, Q) = AZ(\pm \lambda P + C, \pm \lambda Q + D) + B \] (9.28)

\[ Z_2(P, Q) = AZ(P \cos \theta + Q \sin \theta, -P \sin \theta + Q \cos \theta) \] (9.29)

\[ Z_3(P, Q) = AZ \left( \frac{P}{P^2 + Q^2}, \frac{Q}{P^2 + Q^2} \right) \] (9.30)

In general, \( Z = X(P, Q) + iY(P, Q) \), where \( X \) and \( Y \) are harmonic functions of the real variables \( P \) and \( Q \). The real and imaginary parts of \( Z \) satisfy the two-dimensional Laplace equation \((\Delta X = 0, \Delta Y = 0)\). Thus, by specifying analytic functions \( Z \) and taking their real and imaginary parts, one obtains various particular solutions of the two-dimensional Laplace equation. However, for economic equilibrium and economic stability, it is not only dominated by the two-dimensional Laplace Equation (9.6), but also dominated by Cauchy-Riemann Equation (9.5). Thus, the unique fundamental solutions become Eqs. (9.14) and (9.15).

It is true that the world is dominated by differential equations, since we see that most natural and social phenomena can be described by differential equation or equations. Here we reveal that economic equilibrium and economic stability are also mastered by Cauchy-Riemann-Laplace equations, which give strong evidence to believe the truth that differential equations describe the world.

Following the economic complex analysis, the economic system is determined by commodity function \( X \) and money function \( Y \), which are also dominated by price level \( P \) and output quantity \( Q \). Thus, the policy suggestions concentrate on paying more attention to price and quantity. As the financial system is so complex, the
analysis and suggestion above are only a highly simplified estimation of financial markets. We hope that this framework can provide a referable foundation for further studies.

9.6 Concluding Remarks

In conclusion, it is found that the Cauchy-Riemann equation and the Laplace equation construct the core of economic complex analysis, which describes economic equilibrium and economic stability. Combining fundamental solutions of Cauchy-Riemann equation and Laplace equation, the stable function surface is revealed, in which commodity function $X$ and money function $Y$ as complex economic function $Z = X + iY$ link with price level $P$ and output quantity $Q$. It is expected to stimulate further investigations.

Acknowledgements Thank Prof. Dr. Yan Ma, in the University of Rode Island, for her English wording and Mr. Simon S. Lee, in Nanjing University, for his MATLAB coding. This chapter is a revision of the original version published at *Euro-Asian Journal of Economics and Finance*, 2015, 3(3): 133–138.

References

Chapter 10
Complex Economic Metrics Linking to Scaling Money Supply

It is concluded that there is a simple solution for scaling money supply, with using complex economic metrics. When $X$ marks commodity function and $Y$ money function and $Y = kB$ where $B$ is high energy money and $k$ money multiplier, it is found that the $(\gamma/k) < |M| = |Y/X| < (\beta + 1)$ is a stable choice to approach optimum quantity of money supply, where $\gamma$ indicates the ratio of average price to money velocity, $k$ denotes money multiplier and $\beta$ is capital/income ratio.

10.1 Introduction

Money demand and money supply are important issues in monetary economics (Friedman and Hahn 1990; Papademos and Modigliani 1990), for which there were some classic frameworks matching empirical evidence (Dornbusch 1976; Friedman 1959, 1970, 1975; Tobin 1965, 1969; Lucas 1972, 1973). Recently, a new study on capital (Piketty 2014) further explores that research subject.

In my published papers (Ye 2015a, b), the commodity function $X$ and the money function $Y$ were introduced to construct a unified micro-macro-economic analytical framework. In general commodity function $X(P, Q)$ and money function $Y(P, Q)$, $P$ and $Q$ can be general economic factors or measures (e.g., capital, consumption, income, investment, labor, land, price, etc.). When we set $P$ as price level and $Q$ as quantity of commodity and money measured by money unit (e.g., dollar), the $X$ curve and $Y$ curve are equivalent to commodity supply curve and money supply curve respectively, where different commodity and money supplies lead to different results. When $X$ and $Y$ are combined as a complex measure $M = Z = X + iY$, the complex economics can be set up.

In this chapter, while research hypothesis corresponds to the issue of the optimum quantity of money supply (Woodford 1990), we try to find a simple static solution in the framework of complex economics.
10.2 Theoretical Framework for the Upper-Limitations and Lower-Limitations of Money Supply

Let’s focus on the issues of upper-Limitations and lower-Limitations of money supply, as well as the optimizing money supply theoretically.

10.2.1 Three Ways

At first, there are three ways for estimating money supply.

10.2.1.1 Simple Multiplier Model for Bank Money Supply

Supposing $B$ to be the monetary base or high-powered money and $k > 1$ as the money-base multiplier, we have

$$M = kB$$  \hspace{1cm} (10.1)

where $k$ is abided by following Equation

$$k = \frac{1 + h}{h + z} \geq 1$$  \hspace{1cm} (10.2)

where $h$ is a constant fraction of bank deposits and $z$ is a fixed proportion of deposits reflecting either regulatory requirements or internally determined liquidity needs.

That is a simple estimation of the money supply, which is called the money multiplier approach. Since $B$ marks $M_0$ and $M$ matches $M_1$ in monetary economics, we set it as lower-limitation of the money supply.

10.2.1.2 Fisher’s Equation of Exchange

Money supply is also linked to money demand, which fits the Fisher’s equation of exchange as follows.

$$MV = PX \rightarrow M = \frac{PX}{V} = \gamma X > 0$$  \hspace{1cm} (10.3)

where $M$ is the total money in the nation’s money supply; $V$ is the number of times per year each money unit is spent (velocity of money); $P$ is the average price of all the goods and services sold during the year, $\gamma = P/V$ is the ratio of $P$ to $V$; $X$ is the quantity of assets, goods and services sold during the year, which should equal to the total quantity of commodities (GDP).
In mathematical terms, that equation is really an identity which is true by definition rather than description of economic behavior. That is, each term is defined by the values of the other three. Unlike the other terms, the velocity of money has no independent measure and can only be estimated via dividing \( PX \) by \( M \). Some adherents of the quantity theory of money assume that the velocity of money is stable and predictable, being determined mostly by financial institutions. If that assumption is valid then changes in \( M \) can be used to predict changes in \( PX \). If not, then a model of \( V \) is required in order to make the equation of exchange be useful as a macro-economics model or as a predictor of prices.

Most macro-economists replace the equation of exchange with equations for the demand for money which describe more regular and predictable economic behavior. However, predictability of the velocity of money is equivalent to predictability of the demand for money (since in equilibrium real money demand is simply \( X/V \)), so that this is a feasible processing.

In practice, macro-economists almost always use real GDP to measure \( X \), omitting the role of all transactions except for those involving newly produced goods and services (i.e., consumption goods, investment goods, government-purchased goods and exports). That is to say, the only assets counted as part of \( X \) are newly produced investment goods, though the original quantity theory of money does not follow this practice: \( PX \) is the monetary value of all new transactions, whether of real goods and services or of paper assets. So, Fisher’s equation of exchange links \( M \) to \( X \). If \( P = V, M = X \).

### 10.2.1.3 Capital Analysis

According to Piketty’s analysis (Piketty 2014), there is the first fundamental law of capitalism, \( \alpha = r \times \beta \), in which the share of capital income in national income, namely \( \alpha \), is equal to the average rate of return on capital, \( r \), times the capital/income ratio, \( \beta \). It is important to realize that the law \( \alpha = r \times \beta \) is actually a pure accounting identity, valid at all times in all places, by construction.

Meanwhile, in the long run, the capital/income ratio \( \beta \) is related in a simple and transparent way to the savings rate \( s \) and the growth rate \( g \), which leads to the second fundamental law of capitalism, \( \beta = s/g \). The second law is the result of a dynamic process: it represents a state of equilibrium toward which an economy will tend if the savings rate is \( s \) and the growth rate \( g \), but that equilibrium state is never perfectly realized in practice.

The law \( \beta = s/g \) is valid only if asset prices evolve on average in the same way as consumer prices. If the price of real estate or stocks rises faster than other prices, then the ratio \( \beta \) between the market value of national capital and the annual flow of national income can again be quite high without the addition of any new savings. In the short run, variations (capital gains or losses) of relative asset prices (i.e., of asset prices relative to consumer prices) are often quite a bit larger than volume effects (i.e., effects linked to new savings). If we assume, however, that price variations balance
out over the long run, then the law $\beta = s/g$ is necessarily valid, regardless of the reasons why the country in question chooses to save a proportion $s$ of its national income.

Piketty revealed an important inequality in economic development, $r > g$, based on his observations and data. Since $\alpha = r \times \beta = r \times (s/g)$, $\alpha/r = s/g$ and $\alpha \times g = r \times s$, we can derive $r = g \times (\alpha/s)$, $r > g$ is equivalent to that $\alpha/s < 1$ or $\alpha < s$, which means that if the share of capital income in national income $\alpha$ is less than the savings rate $s$, the average annual rate of return on capital $r$ will be larger than the rate of growth of the economy (the annual increase in income or output, $g$).

Supposing $C$ be total capital and $I$ total income in the globe or a special nation, we see that $\beta = C/I = s/g$, which leads to $Cg = Is$ or $g = Is/C$.

As capital estimation above included all the social cumulative historical wealth, when we plus present current commodities $X$, it leads to total quantity of money as

$$M = C + X = I\beta + X \quad (10.4)$$

As that is the sum of historical and present wealth, the capital model gives an upper-limitation estimation of money supply, with considering $I \leq X$, leading to $M \leq (\beta + 1)X$.

Combining Eqs. (10.1), (10.3) and (10.4), we obtain a suitable estimation of $M$ (total money as money demand or money supply) as

$$(\gamma/k)X \leq M \leq I\beta + X \leq (\beta + 1)X \quad (10.5)$$

Equation (10.5) indicates the optimum quantity of money supply.

### 10.2.2 Linking to Complex Metrics

Let’s consider a general case: total social real warfare as commodity $X$ (including capital $K$ and product $J$) and imaginary warfare as money $Y$ (including existing money $M_e$ and increasing money $M_i$), with introducing commodity function $X = X(P, Q)$ and money function $Y = Y(P, Q)$ and both $X$ and $Y$ are changed by real variables $P$ and $Q$, where $P$ could indicate price level and $Q$ means quantity. We define complex economic function as

$$M(X, Y) = Z(P, Q) = X(P, Q) + iY(P, Q) \quad (10.6)$$

where $Z$ is measured by both $X$ and $Y$. 
If the complex function $X$ and function $Y$ are analytic, they will fit to Cauchy-Riemann condition.

$$\frac{\partial X}{\partial P} = \frac{\partial Y}{\partial Q}$$ (10.7)

$$\frac{\partial X}{\partial Q} = -\frac{\partial Y}{\partial P}$$ (10.8)

When $X$ is marks total commodity and $Y$ total money, $X$ is measured by GDP and $Y$ is abided by Eq. (10.5), it is found that $|M| = |Y/X|$ should be kept in

$$(\gamma/k) \leq |M| = |Y/X| \leq (\beta + 1)$$ (10.9)

That means that printed money $Y$ should keep a suitable ratio to real GDP $X$, the ratio is abided between $\beta + 1$ and $\gamma/k$. Equation (10.9) fits scaling money supply yearly.

If the money supply is larger than the upper limitation in one country or one money union, the risk would transfer to other countries and cause the world financial chaos. If the money supply is less than the lower limitation, it could cause markets condense, leading to economic depression.

### 10.3 Numerical Estimation

At global level, since global income = global output, according to Piketty’s statistics (Piketty 2014), the average income per capita is almost exactly 10,000 euros yearly and the global population is close to 7 billion. Consequently, the global output is slightly greater than 70 trillion euros in 2012. If the capital stock is 6 times income on average, we can estimate that the global capital stock is round $6 \times 70$ trillion euros, so that the global wealth can be estimated as $6 \times 70 + 70 = 490$ trillion euros. That gives a basic estimation of maximum money supply (upper limitation) as

$$M < 500 \text{ trillion euros}$$

Meanwhile, as the global output is about 70 trillion euros yearly, if the money multiplier is 7 and $P/V = 1$, we have another basic estimation of minimum money supply (lower limitation) as

$$M > 70 \left(\frac{1}{7}\right) = 10 \text{ trillion euros}$$

Synthetically, the constraint of the global money supply becomes

$$10 \text{ trillion euros} < M < 500 \text{ trillion euros}$$
At national level, the issue mainly occurs in one country. Following Piketty’s statistics (Piketty 2014), there are about 540 million people who live in member countries of the European Union (EU), whose per capita output exceeds 27,000 euros per year. Therefore, we can estimate the EU output as $0.54 \times 27 = 14.58$ trillion euros per year and its capital stock is estimated as $14.58 \times 6 = 87.48$ trillion euros, then the upper limitation of money supply is $M < 87.48 + 14.58 = 102.06$ trillion euros, while the lower limitation as $14.58/7 \approx 2.08$ trillion euros. Similarly, as the USA has 330 million people with a per capita output about 40,000 euros, the estimation will be $1.88 < M < 92.4$ trillion euros. In China, there are 1.3 billion people with a per capita output of 8,000 euros per year, so that the estimation becomes $1.48 < M < 72.8$ trillion euros. If Chinese population becomes 1.4 billion with a per capita output of 10,000 euros per year, the estimation will be $2 < M < 98$.

10.4 Discussion: Money Supply and Money Policy

Money is used as a medium of exchange, a unit of account and a ready store of value. Its different functions are associated with different empirical measures of the money supply. There is no single “correct” measure of the money supply. Instead, there are several measures, classified along a spectrum or continuum between narrow and broad monetary aggregates. Narrow measures include only the most liquid assets, the ones most easily used to spend (currency, checkable deposits). Broader measures add less liquid types of assets (certificates of deposit, etc.).

That continuum corresponds to the way that different types of money are more or less controlled by monetary policy. Narrow measures include those more directly affected and controlled by monetary policy, whereas broader measures are less closely related to monetary-policy actions. It is a matter of perennial debate about whether narrower or broader versions of the money supply have a more predictable link to nominal GDP.

The main functions of the central bank are to maintain low inflation and a low level of unemployment, although those goals are sometimes in conflict (according to Phillips curve). A central bank may attempt to do that by artificially influencing the demand for goods by increasing or decreasing the nation’s money supply (relative to trend), which lowers or raises interest rates and stimulates or restrains spending on goods and services.

An important debate among economists in the second half of the 20th century concerned the central bank’s ability to predict how much money should be in circulation, given current employment rates and inflation rates. Economists such as Milton Friedman believed that the central bank would always get it wrong, leading to wider swings in the economy than that if it were just left alone. That is why they advocated a non-interventionist approach—one of targeting a pre-specified path for the money supply independent of current economic conditions—even though in practice that might involve regular intervention with open market operations (or other monetary-policy tools) to keep the money supply on target.
Considering consumption function (Friedman 1957), consumption always links with multi-economic factors, particularly income and savings. In our outputs (GDP), we consume some and save some every year, in which the consumption cause wealth diminishing and savings transfer into capital. That allows us to consider small changes \( \varepsilon \) in scaling money supply as

\[
(\gamma/k + \varepsilon) \leq |M| = |Y/X| \leq (\beta + 1 - \varepsilon)
\] (10.10)

Meanwhile, different countries adopt different types of money, which are typically classified as \( M \)'s. The \( M \)'s usually range from \( M_0 \) (narrowest) to \( M_3 \) (broadest, even \( M_4 \) in some cases) but which \( M \)'s are actually stressed in policy formulation depends on the country’s central bank.\(^1\) The typical layout for each of the \( M \)'s concludes: the different forms of money in government money supply statistics arise from the practice of fractional-reserve banking. Whenever a bank gives out a loan in a fractional-reserve banking system, a new sum of money is created (Fischer 1977, 1983). That new type of money is what makes up the non-\( M_0 \) components in the \( M_1-M_3 \) statistics.

\(^1\)Different countries adopt different types of \( M \)'s. In Eurozone, the European Central Bank’s definition of euro area monetary aggregates: \( M_1 \) as currency in circulation + overnight deposits, \( M_2 = M_1 + \) deposits with an agreed maturity up to 2 years + deposits redeemable at a period of notice up to 3 months. And \( M_3 = M_2 + \) repurchase agreements + money market fund (MMF) shares/units + debt securities up to 2 years. However, there are just two official UK measures. \( M_0 \) is referred to the “wide monetary base” or “narrow money” and \( M_4 \) is referred to “broad money” or simply “the money supply”. \( M_0 = \) cash outside Bank of England + banks’ operational deposits with Bank of England. \( M_4 = \) cash outside banks (i.e., in circulation with the public and non-bank firms) + private-sector retail bank and building society deposits + private-sector wholesale bank and building society deposits and certificates of deposit. In the UK, \( M_0 \) includes bank reserves, so \( M_0 \) is referred to the monetary base, or narrow money. \( M_1 \) is referred to bank reserves. \( M_2 \) is a broader classification of money than \( M_1 \) and is a key economic indicator used to forecast inflation. And \( M_3 \) is \( M_2 \) plus large and long-term deposits. Meanwhile, MB is referred to the monetary base or total currency and MZM is money with zero maturity, which measures the supply of financial assets redeemable at par on demand. The ratio of a pair of those measures, most often \( M_2/M_0 \), is called an actual, empirical money multiplier. In China, \( M_0 \) indicates cash, \( M_1 \) means cash plus active deposits and \( M_2 \) denotes \( M_1 \) plus all savings accounts. In the US, \( M_0 \) is the total of all physical currency including coinage, i.e., \( M_0 = \) federal reserve notes + US notes + coins. It is not relevant whether the currency is held inside or outside of the private banking system as reserves. \( M_1 \) is the total amount of \( M_0 \) (cash/coin) outside of the private banking system plus the amount of demand deposits, travelers checks and other checkable deposits. \( M_2 = M_1 + \) most savings accounts, money market accounts, retail money market mutual funds and small denomination time deposits (certificates of deposit of under 100,000 US dollars). \( M_3 = M_2 + \) all other CDs (large time deposits, institutional money market mutual fund balances), deposits of eurodollar and repurchase agreements. \( M_4 = M_3 + \) all other CDs (large time deposits, institutional money market mutual fund balances), deposits of eurodollar and repurchase agreements. \( M_4 = M_3 + \) commercial paper. \( M_4 = M_3 + \) T-Bills (or \( M_3 + \) commercial paper + T-Bills). MB is the total of all physical currency plus federal reserve deposits (special deposits that only banks can have at the Fed). MB = coins + US notes + federal reserve notes + federal reserve deposits. MZM as “money zero maturity” is one of the most popular aggregates in use by the Fed because its velocity has historically been the most accurate predictor of inflation. It is \( M_2 - \) time deposits + money market funds. Then \( L \) is the broadest measure of liquidity that the federal reserve no longer tracks, which equals to pretty much \( M_4 + \) Bankers’ acceptance. Money multiplier = \( M_1/MB \).
In the money supply statistics, central bank money is MB while the commercial bank money is divided up into $M_1$–$M_3$ components. Generally, the types of commercial bank money that tend to be valued at lower amounts are classified in the narrow category of $M_1$ while the types of commercial bank money that tend to exist in larger amounts are categorized in $M_2$ and $M_3$, with $M_3$ having the largest. In short, there are two types of money in a fractional-reserve banking system: central bank money (obligations of a central bank, including currency and central bank depositary accounts); commercial bank money (obligations of commercial banks, including checking accounts and savings accounts).

The QE (quantitative easing) policy in the USA or EU input about 1 trillion dollars or euros into the USA or EU, which is not excessive compared to limitations above, so that it has not yet caused chaos of the local finance. However, if the QE policy input much more money into markets, beyond the total capital and commodity in the local society even the world, potential crisis could be caused and big issues may be introduced to national economy, even in the global economy. Although one can raise interest rate for condensing or fixing part money, the QE policy also faces big risk. According to theoretical framework and numerical estimation above, it is suggested to restrict QE under the lower-limitation of money supply.

There are several different definitions of money supply to reflect the differing stores of money. Due to the nature of bank deposits, especially time-restricted savings account deposits, $M_4$ represents the most illiquid measure of money. $M_0$, by contrast, is the least liquid measure of the money supply. For the consideration with easy, money policy can focus on $M_1$, so we can have estimation as

$$M_1 = k|M|X \geq \gamma X \quad (10.11)$$

Meanwhile, as mentioned above, when $X$ is marks commodity function and $Y$ money function and $M = kB$ where $B$ is high energy money and $k$ money multiplier, it is found that the $(\gamma/k) < |M| = |Y/X| < (\beta + 1)$ is a stable choice to approach optimum quantity of money. When the $k$ links with marginal effect as

$$k = 1/\left(1 - \frac{\partial M}{\partial Z}\right) \quad (10.12)$$

where $Z$ is complex variable of $X$ and $Y$, the multiplier effect, the marginal effect and the complex metrics are linked together.

10.5 Conclusion

Although complex numbers seems more complicated than real numbers, the complex metrics actually simplifies the measures of money supply. Therefore, we can have a simple solution for scaling money supply with using complex economic metrics.
While the upper-limitations and lower-limitations of money supply are scaled and numerically estimated, the money policies can also been restricted. Since the financial system is a complex system, the analysis and discussion above are only highly simplified estimation of financial markets. It is expected to provide a useful foundation for further studies.

Acknowledgements This chapter is a revision of the original version published at *Euro-Asian Journal of Economics and Finance*, 2015, 3(3): 188–194.

References

Dornbusch, R.: Exchange rate expectations and monetary policy. J. Int. Econ. 6(3), 231–244 (1976)
Lucas, R.E.: Expectations and the neutrality of money. J. Econ. Theory 4, 103–124 (1972)
Chapter 11
A Synthetic Macro-Economic Model Integrating Interest, Exchange and Tax Rates

In economics, interest rate, exchange rate and tax rate are three key ratios for economic analysis. However, they have no unified relations in existing economic models. Combining Mundell-Fleming model, money supply-demand model and Laffer curve, a synthetic macro-economic model is suggested for integrating the three rates. While tax rate regulates wealth distribution policy for social equality, interest policy represents national finance policy for money optimization and exchange rate describes international economic policy for national benefits, the synthetic model reveals that there are only limited tools of policy for dealing with economic issues because the three rates are correlated to each other, leading to equivalent policy effects. As both interest and exchange rates are related to inflation, a possible extended model including inflation rate is introduced. More widely, the social economic implements of the three rates are also discussed in the view of political economics.

11.1 Introduction

The economic models can be classified into quantitative or qualitative models according to whether the variables are quantitative or classified into deterministic (non-stochastic) and stochastic models according to whether all the variables are deterministic. Here, we mainly focus on deterministic quantitative models, though some variables belong to stochastic variables. In the deterministic quantitative economic models, within hyperbolic coordinates, quantitative variables are deterministically related to each other. In some cases, economic predictions in a coincidence of a model merely assert the direction of movement of economic variables. Therefore, the functional relationships are used only stoical in a qualitative sense (for example, if the price of an item increases, then the demand for that item will decrease), where economists often use two-dimensional graphs instead of functions.

Many economic models have been set up for processing and understanding economic issues. Specially, we mention Mundell-Fleming model (Mundell 1963; Fleming 1962) in macro-economics and Arrow-Debreu model (Arrow and Debreu 1954)
in micro-economics, as well as their possible synthesis (Ye 2015a), extension and improvements (Ye 2015b, c, d). Here, along the economic tradition, we pay attention to macro-economic quantitative model and try to combine interest rate, exchange rate and tax rate together, leading to new quantitative modeling for approaching economic analysis.

11.2 Concepts and Relations of Three Rates

At first, let’s clarify interest rate, exchange rate and tax rate (Sánchez 2005; Floyd 2010), including both nominal and real ones.

An interest rate is the rate at which interest is paid by borrowers (debtors) for the use of money that they borrow from lenders (creditors). Specifically, the interest rate is a percentage of principal paid a certain number of times per period for all periods during the total term of the loan or credit. Interest rates are normally expressed as a percentage of the principal for a period, including a year, a month or a day. While there are different interest rates for the same or comparable time periods (depending on the default probability of the borrower, the residual term, the payback currency and many more determinants of a loan or credit), there are nominal and real interest rates. The nominal interest rate is the amount, in percentage terms, of interest payable and the real interest rate, which measures the purchasing power of interest receipts, is calculated by adjusting the nominal rate charged to take inflation into account. The relation between the real interest rate and the nominal interest is given by the Fisher equation as follows.

\[ 1 + \gamma^* = \frac{1 + \gamma}{1 + \lambda} \quad (11.1) \]

where \( \gamma, \gamma^* \) are nominal and real interest rate respectively and \( \lambda \) is inflation rate. Since \( 1 + \gamma = (1 + \gamma^*)(1 + \lambda) = 1 + \gamma^* + \lambda + \gamma^*\lambda \) and \( \gamma^*\lambda \ll \gamma^* + \lambda \), there is linear approximation \( \gamma^* = \gamma - \lambda \).

An exchange rate (also known as a foreign-exchange rate or forex rate) between two currencies is the rate at which one currency will be exchanged for another. Each country, through varying mechanisms, manages the value of its currency. Generally, exchange rates are determined in the foreign exchange market, including both nominal and real exchange rates. A market-based exchange rate will change whenever the values of either of the two component currencies change, while a currency will tend to become more valuable whenever demand for it is greater than the available supply. If a currency is free-floating, its exchange rate is allowed to vary against that of other currencies and is determined by the market forces of supply and demand. The real exchange rate (RER) is the purchasing power of a currency relative to another at current exchange rates and prices. It is also regarded as the value of one country’s currency in terms of another currency. Therefore, the real exchange rate is the exchange rate times the relative prices of a market basket of goods in the two countries. As exchange rate can be defined by national money or foreign money, here
we set exchange as national money vs. foreign money, for matching interest change direction. According to the statement above, a steady-state economy will be characterized by an inflation rate $\lambda$ and the nominal interest rate $\gamma$, with approximately linear relation to the real interest rate $\gamma^* = \gamma - \lambda$. Let $\pi$ be the nominal exchange rate and $\pi^*$ real exchange rate, there is Dornbusch’s approximate equation (Dornbusch 1976) as

$$\gamma - \gamma^* = \frac{\pi^*}{\pi} - 1$$  \hfill (11.2)

Since we have approximate $\gamma - \gamma^* = \lambda$ from Eq. (11.1), we can substitute it into Eq. (11.2), yielding

$$\frac{\pi^*}{\pi} = \lambda + 1$$  \hfill (11.3)

Equation (11.3) links the nominal exchange rate $\pi$ and real exchange rate $\pi^*$ to inflation rate $\lambda$.

Following Dornbusch’s analysis (Dornbusch 1976), if the elasticity of expectations $\sigma$, is less than unity, there exists Eq. (11.4).

$$\frac{d\gamma}{d\pi} = \sigma - 1 < 0$$  \hfill (11.4)

That is another link between interest rate and exchange rate via elasticity $\sigma$.

A tax rate describes the ratio (usually expressed as a percentage) at which a business or person is taxed. There are several methods used to present a tax rate: statutory, average, marginal and effective. Those rates can also be presented using different definitions applied to a tax base: inclusive and exclusive. A statutory tax rate is the legally imposed rate. An income tax could have multiple statutory rates for different income levels, where a sales tax may have a flat statutory rate. An average tax rate is the ratio of the total amount of taxes paid to the total tax base (taxable income or spending), expressed by a percentage. A proportional tax rate is fixed and the average tax rate equals that tax rate. A marginal tax rate refers to the tax rate that an individual would pay on one additional income.

Let $T$ be the total tax liability and $I$ be the total tax base (income). The average tax rate $= T/I$. The nominal marginal tax rate $\tau$ on income can be expressed mathematically as

$$\tau = \frac{\Delta T}{\Delta I}$$  \hfill (11.5)

where $I$ equals total income and $\Delta$ refers to a numerical change. In accounting practice, the tax denominator in Eq. (11.5) usually includes taxes at federal, state, provincial and/or municipal levels.

Meanwhile, the Laffer curve is typically represented as a graph which starts at 0 tax with zero revenue, rises to a maximum rate of revenue at an intermediate rate of taxation and then falls again to zero revenue at a 100% tax rate, though the shape of the curve is uncertain and disputed. One potential result of the Laffer curve is
that increasing tax rates beyond a certain point will be counter-productive for raising further tax revenue. A hypothetical Laffer curve for any given economy can only be estimated and such estimates are controversial and government revenue per top-income household follows Eq. (11.6).

\[ R = (Y_{\text{taxfree}}(1 - \tau)\sigma - Y_{\text{top}})\tau \]  

(11.6)

where \( R \) is government revenue; \( Y_{\text{taxfree}} \) is the income that the average top-income household would earn if the top income-tax rate was zero; \( Y_{\text{top}} \) is the minimum income needed to be part of the top one percent; \( \tau \) is the top income-tax rate; \( \sigma \) is the elasticity that measures the behavioral response. Equation (11.6) links government revenue \( R \) to elasticity \( \sigma \), where the first term in the bracket \( Y_{\text{taxfree}}(1 - \tau)\sigma \) is the income of the average top-income household, while the full term in brackets is the income that is subject to the top income-tax rate.

In economics (Samuelson and Nordhaus 2005; Friedman and Hahn 2000; Jones and Kenen 1985), marginal tax rates are important because they impact the incentive of increased income, which are applied to income in countries with progressive taxation schemes, with incremental increases in income taxed in progressively higher tax brackets. With higher marginal tax rates, individuals have less incentive to earn more. That is the basis of the Laffer curve, which theorizes that population-wide taxable income decreases as a function of the marginal tax rate, making net governmental tax revenues decrease beyond a certain taxation point.

Then, integrating interest, exchange and tax rates, a synthetic model can be introduced, on the basis of existing macro-economic models.

11.3 Synthetic Model

Combining Mundell-Fleming model, money supply-demand model and Laffer curve under the framework of interest rate, exchange rate and tax rate, an analytical model can be suggested as shown in Fig. 11.1. Replacing common price dimension (Taylor 1979; Woodford 2003), those three rates supply themselves as frame dimensions. For avoiding confusing with imaginary unit and time, I apply Greek alphabetic \( \gamma \), \( \pi \) and \( \tau \) which denote interest rate, exchange rate and tax rate respectively.

The macro-economic analytical model consists of three parts: the first part is Mundell-Fleming model, i.e., IS-LM-BP curves, described by the domestic interest rate plotted vertically and real GDP (\( Y \)) plotted horizontally, where IS curve represents the equilibrium of product market, LM curve means the equilibrium of money market and BP addresses the balance of payments or international income-payment. Geometrically, the IS curve is downward sloped and the LM curve is upward sloped, while the BP curve is upward sloped unless there is perfect capital mobility, in which case it is horizontal at the level of the world interest rate. In the IS-LM-BP graph, under less than perfect capital mobility the positions of both the IS curve and the BP curve depend on the exchange rate, since the IS-LM graph is actually a two-
Fig. 11.1  The synthetic model integrating three ratios

![Diagram](https://example.com/diagram.png)

dimensional cross-section of a three-dimensional space involving all of the interest rate, income and the exchange rate. Under perfect capital mobility, the BP curve is simply horizontal at a level of the domestic interest rate equal to the level of the world interest rate. In pure IS-LM model, the domestic interest rate is a key component for keeping both the money market and the commodity market in equilibrium. However, differing from the pure IS-LM model, Mundell-Fleming model added international financial elements, for fitting to the open economy assumption.

Using equational statement, according to the mainstream economics (Samuelson and Nordhaus 2005; Mishkin 2004), the GDP includes four components of expenditure: consumption $C$, investment $I$, government purchases $G$ and net exports $X$. Let $Y$ be GDP and $A = C + I + G$, the IS curve means

$$Y = C + I + G + X = A(\gamma, Y) + X(\pi, Y)$$  \hspace{1cm} (11.7)

while LM curve indicates at price level $P$

$$\frac{M}{P} = L(\gamma, Y)$$  \hspace{1cm} (11.8)

And the BP curve denotes

$$BP = CA + KA$$  \hspace{1cm} (11.9)

where BP is the balance of payments surplus; CA is the current account surplus; KA is the capital account surplus.

The second part is the money demand and supply curve, where the demand curve of money illustrates the quantity of money demanded at a given interest rate. Generally, the demand curve of money is downward sloping, which means that people want
to hold less of their wealth in the form of money the higher that interest rates on bonds and other alternative investments are. The supply curve of money illustrates the quantity of money supplied at a given interest rate. Unlike a typical supply curve in the commodity market, the supply curve of money could be vertical, because it might not depend on interest rates. It depends entirely on decisions made by the central bank, as the central bank controls the supply of money. However, the market mechanism could also introduce the demand-supply interaction in the money market, so we can also illustrate the money demand and supply with using normal demand-supply curves, as shown in Fig. 11.1. Equilibrium in the money market takes place when the quantity of money demanded is equal to the quantity supplied. Since there are complex relations between interest and exchange rates (Sánchez 2005; Floyd 2010), for simplifying discussed issues, it is designated that real exchange rate and nominal interest rate are linked by Eq. (11.3), i.e., $\gamma \uparrow \rightarrow \pi^* \uparrow$, characterized by national money versus foreign money, i.e., interest and exchange rates keep changes in the same direction.

The third part is Laffer curve, which is one possible representation of the relationship between rates of taxation and the hypothetical resulting levels of government revenue. Because the government revenue is proportional to GDP, we translate the Laffer curve into the coordinate system with tax rate plotted down-vertically and real GDP ($Y$) plotted horizontally as shown in Fig. 11.1. Following claims to illustrate the concept of taxable income elasticity, taxable income will change in response to changes in the rate of taxation.

Thus, the synthetic model integrates interest, exchange and tax rates within it, where interest and exchange are plotted up-vertically, with different calibration in same direction, while tax rate is plotted down-vertically. The Mundell-Fleming model (IS-LM-BP curves) is set in the first quadrant, and the money supply-demand model is put into in the second quadrant, and the Laffer curve is arranged in the fourth quadrant. It is valuable to emphasize that interest rate can be changed by artificial operations via bank in discontinuously ways, while the exchange rate randomly changes following market change and tax rate is rigidly determined by laws, generally.

### 11.4 Policy Analysis

Traditionally, macro-economic policy is usually implemented through two sets of tools: monetary policy and fiscal policy. Both forms of the policy are used to stabilize the economy, which usually means boosting the economy to the level consistent with economic resources.

On monetary policy is that central banks implement the policy by controlling the money supply through several mechanisms. Typically, central banks take action by issuing money to buy bonds (or other assets), which boosts the supply of money and lowers interest rates, called expansionary money policy; or in the case of contraction monetary policy, banks sell bonds and take money out of circulation. Also, bank can continuously shift the money supply to maintain a fixed interest rate target. Some
banks allow the interest rate to fluctuate and focus on targeting inflation rates instead. Central banks generally try to achieve high output without letting loose monetary policy to create large amounts of inflation.

Besides, fiscal policy is the use of government revenue and expenditure as instruments to influence the economy, including tools such as expenditure, taxes and debt.

When the economy is producing less than potential output, government spending can be used to employ idle resources and boost output. Although government spending does not have to make up for the entire output gap, the multiplier effect could boost the impact of government spending.

However, the effects of fiscal policy may be limited by crowding out. When government takes on spending projects, it limits the amount of resources available for the private sector to use. Crowding out occurs when government spending simply replaces private sector output instead of adding additional output to the economy. Crowding out also occurs when government spending raises interest rates with restricting investment. Defenders of fiscal stimulus argue that crowding out is not a concern when the economy is depressed, plenty of resources are left idle and interest rates are low.

Fiscal policy is also implemented through automatic stabilizers. Without suffering from the policy lags of discretionary fiscal policy, automatic stabilizers use conventional fiscal mechanisms but take effect as soon as the economy takes a downturn: spending on unemployment benefits automatically increases when unemployment rises and, in a progressive income tax system, the effective tax rate automatically falls when incomes decline.

Economists usually favor monetary over fiscal policy because it has two major advantages. Firstly, monetary policy is generally implemented by independent central banks instead of the political institutions that control fiscal policy, where independent central banks are less likely to make decisions based on political motives. Secondly, monetary policy suffers shorter inside lags and outside lags than fiscal policy does, as central banks can quickly make and implement decisions while discretionary fiscal policy may take time to pass and even longer to carry out.

Yet, actually, the economic effects of monetary or fiscal policy could be restricted, as the changes of three ratios will affect each other. With integrating analytical framework of interest, exchange, tax rates in synthetic model above, the situation is clear as a case of raising interest rate (by bank) as shown in Fig. 11.2.

The policy of raising interest rate could cause feedback chain: \( \gamma \uparrow \rightarrow MS \downarrow \rightarrow Y \uparrow \), which is equivalent to tax cuts as \( \gamma \uparrow \rightarrow \tau \downarrow \), as it is same to \( \tau \downarrow \rightarrow \gamma \uparrow \rightarrow MS \downarrow \), which means that the policy effect of raising interest rate is equivalent to that of tax cut, or, the policy effect of tax cuts is equivalent to the raising interest rate, leading to less money supply. Inversely, increasing money supply is equivalent to lower interest rate or lifting tax, i.e., \( MS \uparrow \rightarrow \gamma \downarrow \rightarrow \tau \uparrow \). However, once money supply increases, the inflation could occur, resulting in offsetting the policy effect. Thus, the policy tools are limited by the correlated links, where any rate impacts each other, leading to restricted economic effects.

Meanwhile, the tax rate, interest rate and exchange rate really represent different views of political economics. The tax rate marks national policy for the redistribution
of social wealth, which expresses social equality. The interest rate describes national financial policy, which is a way to optimize national financial market. The exchange rate characterizes the interactions of international economies, which is determined by international trade and international financial market.

Generally, tax is dominated by government, which is rigidly restricted by laws. For social equity, it is better to set high tax rate for charging rich people and low tax rate to protect poor persons. Interest rate can be changed by central bank, referring to economic level and situation. And exchange rate seems randomly changeable following international money market, though some nations could also do impacts on its changes.

Politically, tax, interest and exchange rates serve for national objects, so that a national government couldn’t abandon to dominate tax rate, interest rate and exchange rate. In a national point of view, interest rate focuses on money efficiency, which could optimizes the money market and the liquidity. Exchange rate expresses national finance policy, which could regulate international trade and international finance.

In the view of complex economics Ye (2015b,c), interest rate will affect both the real commodity and the imaginary money. Generally, if interest rate looks high, people will deposit money in banks, that is to say, levels of personal savings will go up. Consequently, the imaginary money will shrink, with stock prices going down. Meanwhile, high interest rate will reduce liquidity, which may cause the “liquidity shortage”. That implies that a higher interest rate leads to lower money demand. And
11.4 Policy Analysis

The tax rate mainly affects the real output and social wealth redistribution, where a high tax rate could restrict the development of production and consumption and a low tax rate might stimulate economic recovery.

11.5 Possible Extension

As a general rule, a country with a consistently lower inflation rate exhibits a rising currency value, as its purchasing power increases with other currencies. Thus, inflation is also a key factor in macro-economic analysis (Sidrauski 1967; Lucas 1973; Dornbusch and Frenkel 1973), no matter what it is cost-push inflation or demand-pull inflation. Gentle inflation reflects healthy economy. Meanwhile, since interest rate and exchange rate are correlated to inflation as Eqs. (11.1) and (11.3), we may add the factor of inflation into the synthetic model, with different calibration in down-vertically direction, as shown in Fig. 11.3.

While the money supply does not exceed the equilibrium point of MS and MD, the inflation will keep at gentle level. When the money supply exceeds the equilibrium point, inflation will slightly increase. Once the money supply surpluses the lower-limitation for QE (Ye 2015d), the higher inflation may occur, leading to quick increase of inflation rate. The curve in the third quadrant describes the situation, which is completely different from Phillips curve, where inflation rate links to unemployment. Generally, when the unemployment rate increases, the government should decrease tax. If there would be a healthy stock market that could become a regulator for balancing economy, a tax-free policy might be applied when the stock market values were too low.

Fig. 11.3 An extended synthetic model
11.6 Discussion: Quantitative Easing and Three Ratios

Conventional monetary policy can be ineffective in situations such as a liquidity trap. When interest rates and inflation are near zero, the central bank can not loosen monetary policy through conventional means. It is designed that central banks may use unconventional monetary policy such as quantitative easing (QE) to increase output. However, it is difficult to reach synthetic equilibrium in the synthetic model, as shown in Fig. 11.4, where money supply curve shifted \( MS \rightarrow MS' \rightarrow MS'' \) and \( LM \rightarrow LM' \rightarrow LM'' \) while \( \gamma \rightarrow 0 \).

Thus, the QE policy could input much more money into local markets, beyond the total capital and commodity, leading to potential crisis, so that it is suggested to restrict QE under the lower-limitation of money supply (Ye 2015d) theoretically.

Moreover, aside from factors such as interest rates and inflation, the exchange rate is one of the most important determinants of a national relative level of economic health. Exchange rates play a vital role in a national level of trade, which is critical to most every free market economy in the world. There are numerous factors that determine exchange rates, in which most ones are related to the trading relationship between two countries. Then interest rates, inflation and exchange rates are highly correlated. By manipulating interest rates, central banks exert influence over both inflation and exchange rates and changing interest rates impact inflation and currency values. Higher interest rates offer lenders in an economy a higher return relative to other countries. Therefore, higher interest rates attract foreign capital and cause the exchange rate to rise. The impact of higher interest rates is mitigated, however, if inflation in the country is much higher than in others, or if additional factors serve

![Fig. 11.4 The case of quantitative easing](image-url)
11.6 Discussion: Quantitative Easing and Three Ratios

to drive the currency down. The opposite relationship exists for decreasing interest rates, i.e., lower interest rates tend to decrease exchange rates.

The QE policy might disturb exchange rate via interest rate, so that it could cause the chaos in economic system. That is also a potential crisis.

Meanwhile, it is valuable to point out that tax rates are rigidly regulated by law, with having long-term stability; interest rates are controlled by the central bank, with possessing short-term stability; the exchange rate will randomly fluctuate by the market changes. The studies and discussions above are based on the standard of a nation/country. As tax rates and interest rates are formulated in a nation/country and exchange rates happen among nations/countries, the economic factors constitute complex system. Our analysis was highly simplified, so that there are limitations for practical applications.

11.7 More Discussion: Social Equality and Economic Efficiency

In political economics, wealth distribution and capital profit are always important issues for discussion, in which the former concerns social equality and the latter represents economic efficiency. In some cases, if we pay attention to social equality of wealth redistribution, we might lose partly economic efficiency. For maintaining economic efficiency, we might lose partly social equality. Economic policy actually looks like an art for balancing equality and efficiency, where the social wealth redistribution (regulation via tax rate) and economic efficiency (realization via interest rate) become important factors for applications.

In any societies, there are always rich people and poor people. The governments get tax revenues to support public affairs and to regulate rich-poor difference, for realizing social equality via wealth redistribution. Both social and personal surplus money would cumulatively transfer to capital and the capital demands effective reciprocation via interest or investment. Following Piketty’s analysis (Piketty 2014), more issues on the capital and wealth distribution can be discussed.

At global level, income received from abroad and paid abroad must balance, so that income is by definition equal to output, i.e., global income equals global output. That equality between two annual flows, income and output, is an accounting identity, yet it reflects an important reality. In any given year, it is impossible for total income to exceed the amount of new wealth that is produced (at globally level, a single country may of course borrow from abroad).

At country level, national income may be greater or smaller than its domestic product, depending on whether net income from abroad is positive or negative, with a basic formula as national income = domestic output + net income from abroad = capital income + labor income. With excluding human capital, Piketty concluded that capital reflects the state of development and prevailing social relations of each society, which is just the capital becomes wealth. So, he used the words “capital” and
“wealth” interchangeably. However, the human capital cannot be excluded, actually, as it is a very important economic element.

In the view of wealth, we have national wealth = public wealth + private wealth. There is a typical difference between developed countries and developing countries, where public wealth occupies main portion in developing countries, whereas public wealth in most developed countries is currently insignificant (or even negative, where the public debt exceeds public assets) and private wealth accounts for nearly all of national wealth. But developed countries would engage in large-scale deficit financing to pay for public sector projects and governmental funding. While such activity stimulates the domestic economy, national large public deficits and debts are introduced. However, a large debt encourages inflation and a government may print money to pay part of a large debt, increasing the money supply inevitably causes inflation further. Therefore, the country’s debt rating becomes a crucial determinant of its inflation, as well as exchange rate. Moreover, if a government is not able to service its deficit through domestic means (selling domestic bonds, increasing the money supply), then it must increase the supply of securities for sale to foreigners, thereby the risks are transferred to other countries.

According to Piketty’s statistics (Piketty 2014), the wealth distribution was inequality. The regularity we observe when we try to measure income inequality in practice is that inequality with respect to capital is always greater than inequality with respect to labor. The distribution of capital ownership (and of income from capital) is always more concentrated than the distribution of income from labor. Piketty suggested that government should collect taxes not only on income, but also with more importance, on capital. Because capital produce much more benefit than income.

Following the first fundamental law of capitalism, \( \alpha = r \times \beta \), where \( \alpha \) is the share of income from capital in national income; \( r \) the rate of return on capital; \( \beta \) the capital/income ratio, and the second fundamental law of capitalism, \( \beta = s/g \), where \( s \) is the savings rate; \( g \) the growth rate, Piketty revealed that the fundamental force of inequality focuses on \( r > g \) (where \( r \) stands for the average annual rate of return on capital, including profits, dividends, interest, rents and other income from capital, expressed as a percentage of its total value and \( g \) stands for the rate of growth of the economy, i.e., the annual increase in income or output). Since \( \alpha = r \times \beta = r \times (s/g) \), \( \alpha/r = s/g \), \( \alpha \times g = r \times s \), leading to \( r = g(\alpha/s) \), so \( r > g \) is equivalent to that \( \alpha/s < 1 \) or \( r < s \), which means that if the share of capital income in national income \( \alpha \) is less than the savings rate \( s \), the average annual rate of return on capital \( r \) will larger than the rate of growth of the economy \( g \).

Taxation is not only a technical matter, but also preeminently a political and philosophical issue, perhaps the most important political economic issue. Without taxes, society has no common destiny and collective action is impossible. It is certainly correct that progressive taxation plays a key role in wealth redistribution. A progressive tax is a crucial component for social equality, which plays a central role in its development and in the transformation of the structure of inequality in the world and remains important for ensuring the viability of social equality in the future.
11.8 Conclusion

Conclusively, those three rates (interest, exchange and tax) are important factors in economic analysis and they possess special political and economic meanings.

Tax, interest and exchanges are three important control parameters in an economic system, where tax, interest and exchange rates interact with each other. Tax expresses the economic policy of governance, which reflects social equality. Interest rate is a key parameter for regulating money supply, which determines the national finance. Exchange rate is a vital parameter for international business, which determines the international economy.

Tax rates are regulated by law, with having long-term stability. Interest rates are controlled by the bank, with possessing short-term stability. The exchange rate will randomly fluctuate with the market changes. While tax rate regulates wealth redistribution for approaching social equality, interest policy represents national finance policy for optimizing money market and exchange rate describes international economic policy for realizing national benefits.

If we choose to use zero interest rate and one kind of currency around the globe, for example, the world dollar, the economy will be simple, with production-consumption and goods-goods exchange only. Thus as long as interest and exchange rates exist, the economic and financial system will be a huge and complex system, on which explorations keeps extending.

Acknowledgements This chapter is a revision of the original version published at Euro-Asian Journal of Economics and Finance, 2015, 3(4): 216–217.

References


Dornbusch, R.: Exchange rate expectations and monetary policy. J. Int. Econ. 6(3), 231–244 (1976)


Ye, F.Y.: Complex economic metrics linking to scaling money supply. Euro Asian J. Econ. Finan. 3(3), 188–194 (2015d)
Chapter 12
A Cubic Integrated Economic Model for Macro-Economic Analysis

A novel cubic integrated economic model is developed by extending synthetic economic model, with arranging total economic quantities at horizontal coordinate and relative economic rates at other coordinates. That model is suggested as a new framework for macro-economic analysis. While the cubic model integrates the existing models of Mundell-Fleming model, money supply-demand curve, Laffer curve and Phillips curve, it is also foreseen that unknown interrelationships might be probed in future research.

12.1 Introduction

There are many economic models for describing economic phenomena and processes, particularly for economic analysis. An economic model is a theoretical construct representing economic processes by a set of variables and logical or quantitative relationships between them. It could provide a simplified framework to illustrate complicated economic phenomena. As powerful tools, economic models may assist to understand some economic relationships, thus, we can have some renowned models in economics, such as IS-LM model, supply-demand model and so on.

Under the standard model of Arrow-Debreu-Mundell-Fleming (Ye 2015a), economics could achieve coherence at both micro-level and macro-level. In the synthetic macro-economic model (Ye 2015e), the variables of interest, exchange and tax rates are synthesized into one model, which potentially promotes economic modeling. However, some important economic factors have not yet been included, such as labor or employment and so on. Especially, in macro-economics, the interrelations among the labor market, the commodity market and the money market should be unified, for probing how those interactions influence macro-level variables such as employment levels, aggregate income and GDP. As a complex system, there are many factors interacting with each other in economic system, leading to complicated phenomena and results.
In present economic models, there are some important models that reveal key economic relationships. Two of those models, namely, Mundell-Fleming model (Mundell 1963; Fleming 1962) in macro-economics and Arrow-Debreu model (Arrow and Debreu 1954) in micro-economics, as well as their possible synthesis, extension and improvements Ye (2015a,b,c,d), are given special attentions of theoretical interest. If we extend the synthetic macro-economic model by including more economic elements, we can introduce the cubic integrated economic model for macro-economic analysis, as follows.

### 12.2 A Brief Review of the Analytical Foundations

Following the idea that the Arrow-Debreu model and the Mundell-Fleming model are acknowledged as standard economic models (Ye 2015a), more non-standard models can be considered to merge into the unified analytical framework. Particularly, Tobin’s improvements (Tobin 1969), Lucas model (Lucas 1972, 1973, 1975), Fischer model (Fischer 1977, 1983) and Taylor model (Taylor 1979) become important in the imperfect information and imperfect competition conditions, by including classical Keynes’ AD-AS model. Those models specially concern labor or employment in economic system. As labor is an important factor in production-consumption-labor-wage relations, it provides necessary analytical issues in both macro-economics and micro-economics, for revealing unemployment mechanism linking with government policy.

Following Ye (2015b), micro-economic measures are denoted by using small Latin letters and macro-economic measures by capital Latin letters, as well as rates and ratios as small Greek ones. With using symbols of $Y$ to stand for the real output, $M$ for nominal money, $P$ for nominal prices and $W$ for nominal wages, the classical Keynes’ AD-AS model can be expressed as

\[ Y = a(M - P), \quad a > 0 \quad (12.1) \]

\[ Y = b(M - W), \quad b < 0 \quad (12.2) \]

\[ W = W^* \quad (12.3) \]

where $W^*$ marks real wage at macro-level.

In the Keynesian model (12.1) to Keynesian model (12.3), increasing $M$ will increase both output $Y$ and the price level $P$. Tobin summarized wage-price mechanism as follows (Tobin 1969).

\[ Y = a(M - P), \quad a > 0 \quad (12.4) \]

\[ P = W \quad (12.5) \]

\[ W - W_{-1} = b(P_{-1} - P_{-2}) + cY, \quad 0 < b < 1, \quad c > 0 \quad (12.6) \]
where $W$ denotes the nominal wage for the current period and $W_{-1}$ indicates the nominal wage for the previous period, while $P$ and $P_{-1}$ mark the price level at current and previous period respectively. Equation (12.4) represents the AD, Eq. (12.5) is the price equation and embodies the assumption of quick pass-through of wage costs and no effect of demand and Eq. (12.6) is the wage equation, giving wage inflation as a function of lagged price inflation and output, used as a proxy for unemployment. That system reduces to a second-order difference equation in $P$, which is stable when $a > 0$, possibly with complex roots.

By merging micro-economic variables, Lucas (1972) then constructed his macroeconomic model, in which there was a key link between money supply and price and output changes in a simplified economy using rational expectations. It also delivered a new classical explanation of the Phillips curve relationship between unemployment and inflation. Lucas presented his streamlined version as follows.

$$Y = a(M - P) \tag{12.7}$$

$$p_j = p + e_j, \quad j = 1, 2, \ldots, n \tag{12.8}$$

$$y_j = b(p_j - Ep_j) \tag{12.9}$$

Equation (12.7) is the AD at macro-level. Equation (12.8) gives the price facing each firm $p_j$ at micro-level, where $p_j$ differs from the price level $p$ by a random variable $e_j$, which reflects movements in relative demands across markets. Equation (12.9) provides the supply of each firm, in which $Ep_j$ is the expectation of the price level by firm $j$, based on its observation of $p_j$.

Lucas model combines macro-economic consideration with micro-economic observation and is proved to be suitable under imperfect information. Each individual produces some quantity $y_j$, which cumulate to produce total output $Y$ and match the amount of money $M$. Individuals use money for a given number of times to buy a certain quantity of goods which cost a certain price. In the quantity theory of money, that is expressed as famous Fisher’s exchange equation $MV = PY$ at macro-level, where money supply times velocity equals price times output.

Under imperfect competition, Fischer (1977) introduced another model as follows.

$$Y = a(M - P) + V \tag{12.10}$$

$$Y = -(W - P) \tag{12.11}$$

$$W = E(P_{-1}) \tag{12.12}$$

In Fischer’s system, Eq. (12.10) is the AD, with unit elasticity of output with respect to money balances, in which $V$ is a non-policy demand disturbance. Equation (12.11) is the output supply, obtained from profit maximization, with the assumption of unit elasticity of output supply with respect to wage. Equation (12.12) says that
nominal wages preset a constant real wage and constant employment at the beginning of the period, in expectation.

Similarly to the two-period staggered wage-setting model of Fischer, but with one important difference (namely, wages are not only pre-determined, but are fixed for two periods), Taylor (1980) suggested his model as follows.

\[ Y = (M - P) \] (12.13)

\[ P = \frac{1}{2}(W + W_{-1}) \] (12.14)

\[ W = \frac{1}{2}(P + EP_{+1})_{-1} + \frac{1}{2}a(EY_{-1} + EY_{+1})_{-1} \] (12.15)

In that model, Eq. (12.13) gives the AD and Eq. (12.14) gives the price level as a weighted average of two wages, where half of the labor force is paid \( W \) in the current period and half is paid \( W_{-1} \) in the previous period. Equation (12.15) says that the wages \( W \) is determined by price level and output with elasticity \( a \), in which both concern the expectations in previous period and next period.

Fischer-Gray model (Fischer 1977; Gray 1976) was developed to analyze the wage indexation under perfect competition, as follows.

\[ Y = (M - P) \] (12.16)

\[ Y = -(W - P) - U \] (12.17)

\[ W = kP + (1 - k)EP \] (12.18)

where Eq. (12.16) is still AD. Equation (12.17) gives labor demand under perfect competition and allows for a supply or productivity shock, \( U \), with expected value \( EU = 0 \). Equation (12.18) gives the wage-setting rule, with the degree of indexation being equal to \( k \), where \( 0 \leq k \leq 1 \), with \( k = 0 \) indicates no indexation and \( k = 1 \) full indexation.

All the models above enrich economic modeling for macro-economic analysis, and reveal, particularly, labor-wage mechanism. Actually, if there are total labor as \( L \) at macro-level and each individual with wage \( w \) at micro-level, \( W = Lw \) will give a simple macro-micro-relation of labor-wage. Meanwhile, it is shown that the total economic quantity \( Y \) and price level \( P \) are necessary economic measures in all models, so that we should keep them in a new model.

As Ye (2015e) suggested, it is expected to a cubic integrated model by merging all the information into one economic model and combining with the synthetic macro-economic model.
12.3 The Cubic Integrated Model

Ye (2015e) applied three relative ratios (interest rate, exchange rate and tax rate) and two total absolute quantities (GDP $Y$ and money supply $M$) in the synthetic economic model. In this chapter, the researcher clarifies the meanings of variables by defining that capital Latin letters denote total economic quantities and small Greek letters do relative ratios, while Greek alphabetic $\gamma$, $\pi$, $\tau$ and $\lambda$ respectively indicate interest rate, exchange rate, tax rate and inflation rate.

While GDP as $Y$ and money as $M$ are set into the horizontal coordinate, all other coordinates can be arranged by different rates. Following $\pi$ links $\gamma$ positively, $\gamma$ and $\pi$ are arranged at the same vertical coordinate, it can be safely drawn that $\gamma$ and $\pi$ are positively related. Moreover, $\tau$ and $\lambda$ are not simply in a positive relation and many other related factors need to be considered as well. Thus, it should be arranged into another coordinate. As for $\eta$ denoting unemployment rate, it should also be arranged but in the different direction.

Finally, integrating Mundell-Fleming model, money supply-demand model and Laffer curve, a cubic integrated macro-economic model is introduced as Fig. 12.1, where the Mundell-Fleming model (IS-LM-BP curves), money supply-demand model and Laffer curve (Laffer 2004) have been integrated, for macro-economic analysis. The Laffer curve is typically represented as a graph which starts at 0 tax with zero revenue, rises to a maximum rate of revenue at an intermediate rate of taxation and then falls again to zero revenue at a 100% tax rate, though the shape of the curve is uncertain and disputed.

In the cubic framework of integrated economic model, the IS-LM-BP curves described by the nominal interest rate plotted vertically and real GDP ($Y$) plotted horizontally, while the supply-demand curves of money illustrate the quantity of money supplied and demanded at a given interest rate and exchange rate. According to the mainstream economic theories (Samuelson and Nordhaus 2005; Mishkin 2004), the GDP includes four components of expenditure: consumption $C$, investment $I$, government purchases $G$ and net exports $X$. When $Y$ stand for GDP and $A = C + I + G$, the IS curve can be presented as

\[ Y = C + I + G + X = A(\gamma, Y) + X(\pi, Y) \tag{12.19} \]

while LM curve is denoted at price level $P$ as

\[ \frac{M}{P} = L(\gamma, Y) \tag{12.20} \]

And the BP curve denotes

\[ \text{BP} = CA + KA \tag{12.21} \]

where BP is the balance of payments surplus; CA refers to the current account surplus; KA the capital account surplus.
The labor force is defined as the working-age population, namely the number of people who are either employed or actively looking for work. The unemployment rate is defined as the level of unemployment divided by the labor force, while the employment rate is defined as the number of people currently employed divided by the adult population (or by the working-age population). In those statistics, self-employed people are counted as employed and the unemployment level is defined as the labor force minus the number of currently employed population. Therefore, it can be suggested that wage-labor (W-L) relation share similarities with Fischer-Gray model

\[ W = L(1 - \eta)P + L\lambda P = LP(1 - \eta + \lambda) \]  

(12.22)

where \( \eta \) is unemployment rate and \( \lambda \) inflation rate. Equation (12.22) means that wage links with employment rate, inflation rate and price level. The higher the employment rate and inflation rate are, the more wages will be paid.

The Phillips curve (Phillips 1958; Phelps 1967) just reveals the relation of unemployment rate and inflation rate. The curve can be also arranged into the cubic coordinate framework. When rotating the coordinate system in Fig. 12.2, where the Phillips curve links the unemployment rate to the inflation rate, the decreased unemployment (i.e., the increased levels of employment) in an economy will correlate with higher inflation rate.

The cubic integrated economic model contains more information than the plane synthetic economic model, in which the Phillips curve reveals the inverse relationship between unemployment rate and corresponding inflation rate and the Laffer curve describes the relation between tax rate and total output.
12.4 Discussion I: Total Economic Quantities and Relative Economic Rates

In the cubic integrated economic model, there are two total economic quantities of \( Y \) and \( M \), while other factors belong to relative economic rates. Since \( Y \) is the total economic output and \( M \) should scale with \( Y \) (Ye 2015d), the determined factors in economy become those relative economic rates.

Among the relative economic rates, interest rate, exchange rate and tax rate are important factors (Sánchez 2005; Floyd 2010) in determining economic running (Ye 2015e). The inflation rate and unemployment rate are also key factors to reflect economic quality.

However, not every relative economic rate is independent. The relation between the real interest rate, the nominal interest rate, as well as inflation rate is described by famous Fisher equation as

\[
1 + \gamma^* = \frac{1 + \gamma}{1 + \lambda}
\]

In Fisher equation, \( \gamma \), \( \gamma^* \) respectively refer to the nominal and real interest rate and \( \lambda \) means the inflation rate. Since 

\[
1 + \gamma = (1 + \gamma^*)(1 + \lambda) = 1 + \gamma^* + \lambda + \gamma^*\lambda
\]

and \( \gamma^*\lambda \ll \gamma^* + \lambda \), there is linear approximation \( \gamma^* \sim \gamma - \lambda \) (Feldstein 1976). And the nominal exchange rate and real exchange rate is revealed by Dornbusch’s (1976) approximate formula as

\[
\gamma - \gamma^* = \frac{\pi^*}{\pi} - 1
\]
In Eq. (12.24), \( \pi \) denotes nominal exchange rate and \( \pi^* \) real exchange rate. Moreover, the inflation rate and unemployment rate are linked by Phillips curve.

Meanwhile, different relative economic rates have unique dynamic properties. In developing process, the exchange rate expresses random change and the interest rate can be lifted or lowered in an artificial way, while tax rate is rigidly regulated and unemployment rate can be greatly modified by government investment or purchase. That situation is shown in Fig. 12.3.

In Fig. 12.3, \( t \) denotes time and “\( I + G \)” means government investment and purchase. Thus, if we integrate interest, exchange, tax rates, inflation and unemployment rates into a unified cubic economic model, the interactions among various rates can be jointly revealed for macro-economic analysis.

In the economic theories (Friedman and Hahn 2000; Jones and Kenen 1985), marginal tax rate is important because it impacts the incentive of increased income, which is applied to reflect the income level in countries with progressive taxation schemes, with incremental increases in income taxed in progressively higher tax brackets. With higher marginal tax rates, individuals have less incentive to earn more. That is the basis of the Laffer curve, which theorizes that population-wide taxable income decreases as a function of the marginal tax rate, making net governmental tax revenues decrease beyond a certain taxation point.

Similar to marginal tax rates, marginal interest rate, marginal exchange rate and marginal inflation rate might also be important, which could be further explored.

---

**Fig. 12.3** Different dynamic characteristics of \( \gamma, \pi, \tau \) and \( \eta \)
12.5 Discussion II: Possible Applications and Expected Extensions

When many factors are integrated into one three-dimensional coordinate system, the linkages become complicated and all factors are interrelated. Tax rates and interest rates are formulated within a nation and exchange rates are among nations. Thus the cubic integrated economic model, which is featured by combining national economics and international economics, also can be applied to international economic analysis.

Different relative economic rates follow different changeable rules. Tax rates are rigidly regulated by law to achieve long-term stability; interest rates are controlled by the central bank, with the attribute of short-term stability; the exchange rate, inflation rate and unemployment rate randomly fluctuate with changes of the market. The relative economic rates become tightly related to each other when the cubic integrated economic model links them together. Therefore, that provides a highly simplified model.

By applying that model, more unknown interrelationships are suggested to be explored in the future studies, such as:

1. The relations between inflation, \( Y \) and \( M \). If that relation exists, there can be curves in \( \lambda \sim Y(X) \) plane as well as \( \lambda \sim M(X) \) plane. That interrelationship could provide interactive mechanism of inflation and economic output as well as money supply.

2. The relation between unemployment and money supply. If there is such a relation, there should be a curve in \( \eta \sim M(X) \) plane. That interrelationship could create a mechanism of how money supply affects employment or unemployment.

To conclude, the discussions above might promote more considerations for future development of economic modeling and economic analysis.

12.6 Conclusion

Distributing total economic quantities and relative economic rates in a cubic coordinate system, the cubic integrated economic model is set up to be a new analytical framework of complicated economic phenomena. In that model, the key total economic quantities (\( Y \) and \( M \)) and important relative economic rates (\( \gamma \), \( \pi \), \( \tau \) and \( \lambda \)) are integrated in a unified way. Including established curves of IS-LM-BP curves, money supply-demand curves, Laffer curve and Phillips curve, the cubic integrated multiple economic model provides a possible solution for analyzing complex economic phenomena. This chapter presents the novel model and also proposes potential but unknown interrelationships, thus, more theoretical contributions can be made. Its potential applications could be found and its further extensions could be expected in future.

References


Ye, F.Y.: Complex economic metrics linking to scaling money supply. Euro-Asian J. Econ. Financ. 3(3), 188–194 (2015d)
Part III
Scientometrics: Knowledge Metrics
and $h$-type Metrics
By transferring the DIKW hierarchy to the concept of chain, namely data-information-knowledge-wisdom, the knowledge measure is set up as the logarithm of information, while the information is the logarithm of data, so that knowledge metrics are naturally introduced and the mechanism of Brookes’ basic equation of information science is revealed. When knowledge is classified as explicit knowledge and tacit knowledge, qualitative SECI model is changed to quantitative triangle functions on explicit knowledge and tacit knowledge, where the former is measured by the logarithm of data and the latter is measured by the negative entropy of language. The author suggests to treat the unit of knowledge as kit, correspondingly, data as bit and information as byte.

### 13.1 Introduction

While knowledge management has become a hot area, knowledge becomes a scientific key. Although knowledge is well known as a popular concept, it is never defined scientifically and strictly. In Brookes’ theory (Brookes 1980–1981), Popper’s scientific philosophy was applied and a basic equation linking knowledge to information was set up. However, there was no theoretical interpretation of interactive mechanism between information and knowledge, though Shannon information theory was well established (Shannon 1948).

Meanwhile, as a qualitative model, SECI model (i.e., socialization-externalization-combination-internalization) of Nonaka et al. (2000) was an influential model for describing organizational knowledge creation. It contained four modes of knowledge conversion between explicit knowledge and tacit knowledge, where $S$ denoted socialization that described the dimension of tacit to tacit transfer, $E$ meant externalization that described the dimension of tacit to explicit transfer, $C$ indicated combination that described the dimension of explicit to explicit transfer and $I$ was internalization that described the dimension of explicit to tacit transfer. However, there were no quantitative relations in the SECI model.
Scientifically, in order to quantify studies of knowledge, it is necessary to define knowledge and measure knowledge quantitatively. Referring the DIKW hierarchy model (Rowley 2007; originated by Cleveland via Zeleny), I try to develop a mathematical method for measuring knowledge, along with my earlier studies (Ye 1999, 2011).

### 13.2 DIKW Chain and Knowledge Metrics

Following the DIKW hierarchy model (Rowley 2007), data ($D$), information ($I$), knowledge ($K$), and wisdom ($W$) together construct a pyramid structure, as shown in Fig. 13.1.

Although there is argument on DIKW hierarchy (Frické 2009), DIKW hierarchy can be transformed to a logic chain, as shown in Fig. 13.2.

Now the DIKW chain can be quantitatively processed. Let us introduce a median variable $i$ as physical information and $J$ as subjective information, according to Ye (1999), the logic of quantitative DIKW chain becomes: the objective data transfer into physical information ($i$) via natural transmission, so physical information ($i$) is checkable by physical instruments; the physical information ($i$) transfers into objective information ($I$) via social transmission, so that the objective information ($I$) is acceptable by subject; the subjective information ($J$) transforms into subjective information ($J$) via subject absorption, so that the subjective information ($J$) bears subject value judgement. Then the subjective information ($J$) transforms into knowledge ($K$) via structuralization and systemization. The overall applications of knowledge become wisdom ($W$). That is a transmission chain from objective side to subjective side. Data ($D$) and information ($I$) belong to physical objective field, while knowledge ($K$) and wisdom ($W$) belong to cognitive subjective field. It is an objective process from data ($D$) to information ($I$), while it is a subjective process.
13.2 DIKW Chain and Knowledge Metrics

The DIKW chain from knowledge ($K$) to wisdom ($W$) shows a key transformation between information ($I$) and knowledge ($K$), where the subjective value judgment is generated.

Referring to Shannon information theory (Shannon 1948), the physical information ($i$) is defined as

$$i = d \log D$$

where $d$ is the transformation coefficient from data ($D$) to information ($I$).

From physical information ($i$) to objective information ($I$), a transmission channel is needed, where the information motion obeys the wave-heat equations (Ye 1999, 2011). If there are information compression ratio $p$ and loss ratio $q$ in the transmission channel, there is $I = pqi$. Let $pq = b$, thus we have

$$I = bi \quad (13.2)$$

Transforming the objective information ($I$) to the subjective information ($J$) is a key. Following Brookes’ information theory (Brookes 1980–1981) and the principle of logarithmic perspective, with introducing value coefficient $v \in [0, 1]$ (matching Rescher’s model), it is derived as

$$J = \log I^v = v \log I \quad (13.3)$$

Since valuable information increases knowledge (Ye 1999), the unit knowledge increment should be proportional to the subjective information ($J$), yielding

$$\frac{dK}{dI} = kJ = k \ln I^v \quad (13.4)$$

where $k$ is the knowledge transformation coefficient of information ($I$). Therefore, knowledge ($K$) should be the integral

$$K = k \int JdI = k \int v \ln I dI = kvI (\ln I - 1) + K_0 \quad (13.5)$$

where $K_0$ is the integral constant, representing original knowledge ($K$).
This is just the Brookes’ basic equation of information science,

\[ K = K_0 + \Delta K; \ \Delta K = kvI(\ln I - 1) \]  \hspace{1cm} (13.6)

where \( \Delta K = kv(\ln I - 1) \) is the increment of knowledge. The process gives the mechanism of Brookes’ equation.

As for the measurement units, since data \((D)\) uses bit and information \((I)\) does byte, it is suggested to apply “kit” to be the unit of knowledge \((K)\). When \( k = 1 \), it means that 1 byte information \((I)\) can be transferred into 1 kit knowledge \((K)\).

When an intelligent agent \((j)\) has its knowledge elements as \(a_{ik} \in [0, 1]\), there are knowledge vectors

\[ A_j = (a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5}, a_{i6}, a_{i7}, a_{i8}, \ldots) \]  \hspace{1cm} (13.7)

Defining its knowledge matrix as

\[ K_j = A_j^T A_j = (r_{ik}) \]  \hspace{1cm} (13.8)

where \( r_{ik} = \sum_j a_{ij}a_{jk} \) are intelligent correlatives, the wisdom \((W)\) can be measured by intelligent grade as the trace of knowledge matrix

\[ W_j = \text{tr}(K_j) \]  \hspace{1cm} (13.9)

Therefore, it can be seen that this is a feasible foundation for a unified quantitative model linking information to knowledge based on the logic chain data-information-knowledge-wisdom.

### 13.3 Measuring Explicit Knowledge and Tacit Knowledge

Both explicit knowledge and tacit knowledge can be measured in the quantitative way above. However, there are differences between explicit knowledge and tacit knowledge.

In general, while explicit knowledge is codified, tacit knowledge is unmodified. Using the iceberg metaphor, explicit knowledge looks like the peak and tacit knowledge sinks under the sea, by 2/8 rule. So, it is estimated that explicit knowledge can occupy about 20% and tacit knowledge 80% of the total knowledge volume.

In the SECI model, \( S \) (socialization) is realized by sharing tacit knowledge face-to-face or sharing experiences in a community, which typically occurs in a traditional apprenticeship. \( E \) (externalization) is realized by reading documents, where tacit knowledge is made explicit (knowledge is codified) to enable sharing. \( C \) (combination) is realized by combining different types of explicit knowledge, which is
collected and integrated from inside or outside of organization and then combined, edited or processed to form new knowledge. \( I \) (internalization) is realized by an individual, who learns explicit knowledge and will asset to organization then store as tacit knowledge. Those four modes of knowledge dimensions with conversion repeats themselves in a spiral way, as shown in Fig. 13.3, where the sides of the square represent explicit knowledge and tacit knowledge while the inside cross lines mark individual and organization, respectively.

There are two advantages of the SECI model: (1) it appreciates the dynamic nature of knowledge and knowledge creation; (2) it provides a simple framework for management of the relevant processes. However, there are also disadvantages, in which the key issue concentrates on lacks of quantitative analysis. Another issue is based on the study of Japanese organizations that heavily rely on tacit knowledge (Nonaka and von Krogh 2009). To overcome the disadvantages, it is suggested to introduce a quantitative mathematical pattern as shown in Fig. 13.4.

In Fig. 13.4, \( EK_1 \) and \( EK_2 \) denote different explicit knowledge while \( TK_1 \) and \( TK_2 \) do tacit knowledge. \( K \) means knowledge and \( t \) indicates time. A triangle function can be applied to simulate the process as follows:

\[
K = a(t) \sin(EK_i) + (1 - a(t)) \cos(TK_i) \tag{13.10}
\]

where \( a(t) \) is explicit ratio changed by time, \( 1 - a \) is tacit ratio and \( i \) is sum of all items. If \( a = 0.2 \), \( 1 - a = 0.8 \).

When all knowledge becomes explicit knowledge, \( a = 1 \), so that the explicit knowledge is measured by

\[
K = \sin(EK_i) \tag{13.11}
\]
When all knowledge changes to tacit knowledge, $a = 0$, so that the tacit knowledge is measured by

$$K = \cos(TK_i)$$

(13.12)

13.4 Analysis and Discussion

The explicit knowledge may express order while the tacit knowledge shows chaos, so that the knowledge from information and data generated explicit knowledge and the entropy caused tacit knowledge. Thus, the explicit knowledge is proportional to the logarithmic logarithm of data by combining Eqs. (13.1) and (13.5).

$$EK = c \log \log D$$

(13.13)

where $c$ denotes a proportional constant.

Then the tacit knowledge is proportional to the negative entropy of language ($S$), which equals the logarithm of words ($W$).

$$TK = -S = -e \log W$$

(13.14)

where $e$ denotes a constant.
Equation (13.14) looks like Boltzmann formula. That is another metaphor that physics merges into knowledge system. However, the knowledge system has its unique characteristics when the Eqs. (13.13) and (13.14) are inserted into Eq. (13.10).

\[ K = ac \sin(\log \log D_j) + (1 - a)e \cos(\log W_j) \]  

(13.15)

where \( j \) denotes sum of all items.

Another important issue concerns information absorption while information transforms into knowledge. Although we could set \( k \) as knowledge transformation coefficient of information in Eq. (13.4) and calculate the absorption rate into \( k \) together, we can also set a specialized absorption rate separated from transformation coefficient if we need to discuss information absorption in the transformation process from information to knowledge. However, as the information absorption highly relies on existing subjective knowledge, absorption rate is a personally parameter and may not be an objective measurement.

Since knowledge measurement is a comprehensive and difficult issue in scientific metrics and knowledge management, few quantitative models are developed based on strict mathematical methodology. Here we try to set up a mathematical model.

Another related issue concerns innovation measurement in knowledge management, on which we mention triple helix (TH) proposed by Etzkowitz and Leydesdorff (1995). After it came, triple helix has quickly and widely affected the academic world, particularly activated innovation measurement. Triple helix model has also set up quantitative mechanism of interaction among innovative entities and the dynamics with quantitative measures of innovation, based on information theory. Therefore, it will link to knowledge via information. After the emergence of a triple helix of university-industry-government relations (Leydesdorff and Etzkowitz 1996), the triple helix model guides us to study university-industry-government relations as well as their actions and functions in innovation. Later, the triple helix as a model for innovation studies is emphasized that academic, industrial, and governmental institutions interact at both national and international levels (Leydesdorff and Etzkowitz 1998). The triple helix model will also enlighten to study knowledge measurement.

13.5 Conclusion

In this chapter, a mathematical theory of knowledge is suggested to include two parts: the first part links the chain of data-information-knowledge-wisdom and the second part links explicit knowledge and tacit knowledge. The knowledge is estimated by the logarithm of information and information by the logarithm of data, while the conversion between the explicit knowledge and tacit knowledge are simulated by triangle functions.
It is necessary to develop quantitative metrics in knowledge theory and the combination of qualitative concepts and quantitative measurement is important for knowledge research. The preliminary exploration here introduces a potential development of knowledge metrics in the future.

Acknowledgements This chapter is a revision of the original version published at *International Journal of Data Science and Analysis, 2016, 2(2): 32–35.*

References


A simple distribution function \( f(x, t) = c(x + d)^{-\alpha} e^{\lambda t} \) obeys wave and heat equations, which just constructs a theoretical approach to the unification of informetric models, with which we can unify all informetric laws. While the shifted power function with time-type exponential cutoff supplies a unifying informetric framework, the shifted power function with space-type exponential cutoff can be linked to the Pareto distribution and Weibull distribution. In the shifted power function with time-type exponential cutoff, its space-type distributions derive naturally Lotka-type laws in size approaches and Zipf-type laws in rank approaches and its time-type distributions introduce the mechanism of price-type laws and brookes-type laws. The exponent is the crucial parameter in a power function, determining the most important characteristics of the power-law distribution.

14.1 Introduction

It is well known that many distributions in informetrics and information networks (Albert and Barabási 2002) follow a power-law (some with an exponential cutoff) (Egghe 2005a; Newman 2003, 2010). While the shifted Lotka theory (Egghe and Rousseau 2012; Burrell 2008) was discussed recently, a shifted power function can be considered as a general unifying function for informetric distributions. Using the concepts of size-frequency and rank-frequency (Egghe and Rousseau 1990, 2003), we can classify all informetric laws into Lotka-type laws (size approaches) and Zipf-type laws (rank approaches). For their unification of the continuous versions (Bookstein 1990; Rousseau 1990; Egghe 2005a), we find a simple unified model, which can synthesize most classic informetric models (Leimkuhler 1967; Merton 1968; Fairthorne 1969; Price 1976; Brookes 1977; Garfield 1980; Egghe and Rousseau 1995; Egghe 2005b; Newman 2005) and Mandelbrot distribution (Mandelbrot 1982), which is the shifted power function with an exponential cutoff.
14.2 Theoretical Methodology

The logic process to set up theoretical methodology is presented as follows.

14.2.1 Shifted Power Function Fits Wave-Heat Equations

A simple unified distribution function is chosen as

\[ f(x, t) = c(x + d)^{-\alpha} e^{\lambda t} \]  

(14.1)

where \( c \) and \( d \) are parameters, \( \alpha \) is Lotkanian exponent and \( \lambda \) real exponent.

Solving its partial differentials, we obtain

\[ \frac{\partial f}{\partial x} = -\alpha c(x + d)^{-\alpha - 1} e^{\lambda t} = -\frac{\alpha}{x + d} f(x, t) \]  

(14.2)

\[ \frac{\partial^2 f}{\partial x^2} = \alpha (\alpha + 1) c(x + d)^{-\alpha - 2} e^{\lambda t} = \frac{\alpha (\alpha + 1)}{(x + d)^2} f(x, t) \]  

(14.3)

\[ \frac{\partial f}{\partial t} = \alpha c(x + d)^{-\beta} e^{\alpha t} = \alpha f(x, t) \]  

(14.4)

\[ \frac{\partial^2 f}{\partial t^2} = \alpha^2 c(x + d)^{-\beta} e^{\alpha t} = \alpha^2 f(x, t) \]  

(14.5)

So we have

\[ \frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = \left[ \frac{\alpha (\alpha + 1)}{(x + d)^2} - \frac{\lambda^2}{c^2} \right] f(x, t) = a(x) f(x, t) \]  

(14.6)

\[ \frac{\partial^2 f}{\partial x^2} - \frac{c}{\partial t} \frac{\partial f}{\partial t} = \left[ \frac{\alpha (\alpha + 1)}{(x + d)^2} - \frac{\lambda}{c} \right] f(x, t) = b(x) f(x, t) \]  

(14.7)

Equations (14.6) and (14.7) mean that Eq. (14.1) will fit both wave and heat (or scatter) equations when

\[ a(x) = \frac{\alpha (\alpha + 1)}{(x + d)^2} - \frac{\lambda^2}{c^2} \]  

(14.8)

\[ b(x) = \frac{\alpha (\alpha + 1)}{(x + d)^2} - \frac{\lambda}{c} \]  

(14.9)
Thus, wave and heat equations could construct a simple unified informetric model as

\[
\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = a(x) f(x, t) \tag{14.10}
\]

\[
\frac{\partial^2 f}{\partial x^2} - \frac{1}{c} \frac{\partial f}{\partial t} = b(x) f(x, t) \tag{14.11}
\]

Equation (14.10) is the wave equation and Eq. (14.11) the heat one, where \(a(x)\) and \(b(x)\) are spatial functions, constant \(c\) will be both wave velocity and scatter coefficient. And Eq. (14.1) is just a common fundamental solution of wave-heat equations.

### 14.2.2 Fundamental Solutions and the Function \(F(x, T)\) as a Fundamental Solution

Let \(L(D)\) be a differential operator with constant coefficients. Then a fundamental solution (Renardy and Rogers 2004) for \(L\) is a function \(G\) (more generally a distribution or generalized function) satisfying the equation \(L(D)G = \delta\), where \(\delta\) is the Dirac \(\delta\) function. Fundamental solutions are unique only up to a solution of the homogeneous equation \(L(D)u = 0\). The significance of the fundamental solution lies in the fact that \(L(D)(G \ast f) = (L(D)G) \ast f = \delta \ast f = f\), provided that the convolution \(G \ast f\) is defined. Recall (Rousseau 1998) (where the importance of convolutions in information science is illustrated) that a convolution of two real functions \(g\) and \(h\) are defined as

\[
(f \ast h)(t) = \int_{-\infty}^{+\infty} f(t - u)h(u)du \tag{14.12}
\]

For functions of two variables, that becomes

\[
(f \ast s)(x, t) = \int_{\mathbb{R}^2} f(x - u, t - v)s(u, v)dudv \tag{14.13}
\]

We have shown that \(f(x, t)\) is a solution of a wave-type and of a heat-type partial differential equation. In order to be a fundamental solution we need to have constants coefficients \(a\) and \(b\). That leads to the additional requirements

\[
\frac{\alpha(\alpha + 1)}{(x + d)^2} = \frac{\lambda^2}{c^2} = \frac{\lambda}{c} \tag{14.14}
\]
14.2.3 Space-Type Distributions and Their Size-Rank Transformation

Let us focus on the general shifted power function \( f(x) \) changed by variable \( x \)

\[
f(x) = c(x + d)^{-\alpha}
\]

(14.15)

where \( \alpha > 0 \) is power exponent and \( c \) constant, which is just Lotka’s law when \( d = 0 \).

At first, we recall three theorems for size-rank transformation proofed by Egghe and Rousseau, with different forms by emphasizing the roles of different integral intervals.

14.2.3.1 Transformation Theorem for Linking Lotka Size Distribution to Zipf-Mandelbrot Rank Function

**Theorem 1** (Egghe 2005a, b) When \( \alpha > 0 \) and \( \alpha \neq 1 \), size variable \( x \) changes to rank variable \( r \) with interval transformation \( x \in [1, N] \rightarrow r \in [1, T] \), which will result in size-rank transformation from the standard Lotka size distribution

\[
f(x) = cx^{-\alpha}
\]

(14.16)

to Zipf-Mandelbrot rank function

\[
g(r) = p(r + q)^{-\beta}
\]

(14.17)

where \( N \) is the maximal item per source density; \( T \) is the total number of sources; \( c, p, q \) are constants.

**Proof** In Eq. (14.16), \( f(x) \) is a size-frequency distribution. Suppose its corresponding rank-frequency distribution function be \( g(r) \), following Egghe (2005a, b)

\[
r = g^{-1}(x) = \int_x^N f(t)dt \tag{14.18}
\]

(14.18)

\[
f(x) = -\frac{1}{g'(g^{-1}(x))} \tag{14.19}
\]

(14.19)

While \( x \rightarrow r \), interval transformation is \( x \in [1, N] \rightarrow r \in [1, T] \). Put Eq. (14.16) into Eq. (14.18), we derive

\[
r = \int_x^N ct^{-\alpha}dt = \frac{c}{1 - \alpha} (N^{1-\alpha} - x^{1-\alpha}) \tag{14.20}
\]

(14.20)
Then we solve

\[ x = g(r) = \left( \frac{\alpha - 1}{c} r + \frac{1}{N^{\alpha - 1}} \right)^{\frac{1}{\alpha - 1}} = \left[ \frac{cN^{\alpha - 1}}{(\alpha - 1)N^{\alpha - 1}r + c} \right]^{\frac{1}{\alpha - 1}} \]

(14.21)

Let

\[ p = \left( \frac{c}{\alpha - 1} \right)^{\frac{1}{\alpha - 1}} \] (14.22)

\[ q = \frac{c}{\alpha - 1} N^{1 - \alpha} \] (14.23)

\[ \beta = \frac{1}{\alpha - 1} \] (14.24)

We obtain Zipf-Mandelbrot distribution Equation (14.17) from Eq. (14.21). Inversely, from Eq. (14.17), following Eq. (14.19), we have

\[ f(x) = -\frac{1}{g'(g^{-1}(x))} = -\frac{1}{-\beta p \left( \left( \frac{p}{x} \right)^{1/\beta} - q + q \right)^{-(\beta + 1)}} = \frac{1}{\beta p} \frac{\beta + 1}{x^{\beta + 1}} \] (14.25)

Let

\[ c = \frac{1}{\beta p} \beta + 1 \] (14.26)

\[ \alpha = \frac{\beta + 1}{\beta} \] (14.27)

We reduce to Eq. (14.16).

**Corollary 1** Interval transformation \( x \in [1, \infty] \rightarrow r \in [0, T] \) will result in size-rank transformation from standard Lotka size distribution Equation (14.16) to standard Zipf rank function

\[ g(r) = br^{-\beta} \] (14.28)
Proof If the interval transformation is \( x \in [1, \infty] \rightarrow r \in [0, T] \) while \( x \rightarrow r \), the result becomes
\[
r = \int_x^\infty c t^{-\alpha} \, dt = \frac{c}{\alpha - 1} x^{1-\alpha}
\] (14.29)
then
\[
x = g(r) = \left( \frac{\alpha - 1}{c} r \right)^{\frac{1}{\alpha - 1}} = \frac{c}{\alpha - 1} \left( \frac{r}{c} \right)^{\frac{1}{\alpha - 1}}
\] (14.30)
Introducing
\[
b = \left( \frac{c}{\alpha - 1} \right)^{\beta}
\] (14.31)
and keeping \( \beta = (\alpha - 1)^{-1} \), we just obtain standard Zipf function, Eq. (14.28). With similar process by using Eq. (14.17), we can reduce to Eq. (14.16) from Eq. (14.28).

14.2.3.2 Transformation Theorem for Linking Shifted Lotka Distribution to Shifted Mandelbrot Rank Function

Theorem 2 When \( \alpha > 0 \) and \( \alpha \neq 1 \), size variable \( x \) changes to rank variable \( r \) with interval transformation \( x \in [1, N] \rightarrow r \in [1, T] \), which will result in size-rank transformation from shifted Lotka size distribution
\[
f(x) = c(x + d)^{-\alpha}
\] (14.32)
to shifted Mandelbrot rank function
\[
g(r) = p(r + q)^{-\beta} - d
\] (14.33)
where \( N \) is the maximal item per source density; \( T \) is the total number of sources; \( c, d, p, q \) are constants.

Proof Following similar process of shifted Lotka theory (Egghe and Rousseau 2012), by using same Eqs. (14.18) and (14.19), we derive results as follows.

While \( x \rightarrow r \), interval transformation is \( x \in [1, N] \rightarrow r \in [1, T] \). We derive
\[
r = \int_x^N c(t + d)^{-\alpha} \, dt = \frac{c}{1 - \alpha} [(N + d)^{1-\alpha} - (x + d)^{1-\alpha}]
\] (14.34)
Then we solve

\[ x = g(r) = \left( \frac{\alpha - 1}{c} \frac{1}{(N + d)^{\alpha - 1}} \right)^{\frac{1}{\alpha - 1}} - d = \left[ \frac{c(N + d)^{\alpha - 1}}{(\alpha - 1)(N + d)^{\alpha - 1}r + c} \right]^{\frac{1}{\alpha - 1}} - d \]

\[ = \frac{\left( \frac{c}{\alpha - 1} \right)^{\frac{1}{\alpha - 1}}}{\left[ r + \frac{c(N + d)^{1-\alpha}}{\alpha - 1} \right]^{\frac{1}{\alpha - 1}}} - d \quad (14.35) \]

Let

\[ p = \left( \frac{c}{\alpha - 1} \right)^{\frac{1}{\alpha - 1}} \quad (14.36) \]

\[ q = \frac{c}{\alpha - 1} (N + d)^{1-\alpha} \quad (14.37) \]

\[ \beta = \frac{1}{\alpha - 1} \quad (14.38) \]

So we obtain shifted rank distribution Equation (14.33) according to Eq. (14.35). Inversely, from Eq. (14.33), following Eq. (14.19), we have

\[ f(x) = -\frac{1}{g'(g^{-1}(x))} = -\frac{1}{-\beta p \left[ \left( \frac{p}{x + d} \right)^{1/\beta} - q + q \right]^{-(\beta + 1)}} = \frac{1}{\beta p} \frac{p^{\beta+1}}{(x + d)^{\beta+1}} \]

\[ = \frac{1}{\beta p} \frac{p^{\beta+1}}{(x + d)^{\beta+1}} \quad (14.39) \]

Let

\[ c = \frac{1}{\beta p} p^{\beta+1} \quad (14.40) \]

\[ \alpha = \frac{\beta + 1}{\beta} \quad (14.41) \]

We reduce to Eq. (14.32).

**Corollary 2** Interval transformation \( x \in [1, \infty] \to r \in [0, T] \) will result in size-rank transformation from shifted Lotka size distribution Equation (14.32) to the special shifted rank function

\[ g(r) = br^{-\beta} - d \quad (14.42) \]

**Proof** If the interval transformation is \( x \in [1, \infty] \to r \in [0, T] \) while \( x \to r \), the results become
\[
\int_{x}^{\infty} c(t + d)^{-\alpha} \mathrm{d}t = \frac{c}{\alpha - 1} (x + d)^{1-\alpha}
\] (14.43)

then

\[
x = g(r) = \left( \frac{\alpha - 1}{c} r \right)^{\frac{1}{\alpha}} - d = \left( \frac{c}{\alpha - 1} \right)^{\frac{1}{\alpha - 1}} - d
\] (14.44)

Introducing

\[
b = \left( \frac{c}{\alpha - 1} \right)^{\beta}
\] (14.45)

and keeping \( \beta = (\alpha - 1)^{-1} \), we obtain the special shifted rank function Eq. (14.42). With similar process using Eq. (14.19), we can reduce to Eq. (14.32) from Eq. (14.42).

When \( d = 1 \), it is just the case of shifted Lotkaian function (Egghe and Rousseau 2012).

14.2.3.3 Transformation Theorem for Linking Size-Frequency Power Function and Rank-Frequency Exponential Function

**Theorem 3** When \( \alpha=1 \), size variable \( x \) changes to rank variable \( r \) with interval transformation \( x \in [1, N] \rightarrow r \in [1, T] \), which will result in frequency power function

\[
f(x) = c(x + d)^{-1}
\] (14.46)

to rank-frequency exponential function

\[
g(r) = pe^{-qr} - d
\] (14.47)

where \( c > 0 \), \( \infty > d \geq 0 \), \( p > 1 \), \( \infty > q > 0 \) are constants.

**Proof** Following similar process of proof (Egghe and Rousseau 2003), using Eqs. (14.18) and (14.19), we derive similar proof as follows.

While \( x \rightarrow r \), interval transformation is \( x \in [1, N] \rightarrow r \in [1, T] \). We derive

\[
r = \int_{x}^{N} c(t + d)^{-1} \mathrm{d}t = c \ln \frac{N + d}{x + d}
\] (14.48)

Then we solve

\[
x = g(r) = (N + d)e^{-\frac{r}{c}} - d
\] (14.49)
Let
\[ p = N + d \] (14.50)
\[ q = \frac{1}{c} \] (14.51)

So we obtain shifted rank distribution Equation (14.47) according to Eq. (14.19).

Inversely, from Eq. (14.47), following Eq. (14.19), we have
\[ f(x) = -\frac{1}{g'(g^{-1}(x))} = -\frac{1}{-pqe^{-\frac{1}{q} \ln \frac{x+d}{p}}} = \frac{p}{q(x+d)} \] (14.52)

Let
\[ c = \frac{p}{q} \] (14.53)

We reduce to Eq. (14.46).

When \( d = 0 \), it is just the case proved by Egghe and Rousseau (2003), which shows that a size-frequency power function with \( \alpha = 1 \) is equivalent with an exponentially decreasing rank-frequency function.

Meanwhile, when we solve its derivation of Eq. (14.32), we have
\[ \frac{df(x)}{dx} = -\frac{c}{\alpha + 1} (x + d)^{-(\alpha + 1)} = -\frac{f(x)}{\alpha + 1} (x + d)^{-1} \] (14.54)

which means
\[ \frac{df(x)}{f(x)} = -\left( \frac{1}{\alpha + 1} \right) \frac{dx}{x + d} \] (14.55)

Let \( 1/(\alpha + 1) = k \), the rank-frequency distribution can be found via its integral from \( x = 1 \) to \( r \) as
\[ \int_1^r \frac{df(x)}{f(x)} = -k \int_1^r \frac{dx}{x + d} \] (14.56)

That process also transforms size distribution into rank one so that we get
\[ \frac{f(r)}{f(1)} = \frac{(1+d)^k}{(r+d)^k} \] (14.57)

Let \( p = f(1)(1+q)^k \) and \( q = d \), we obtain Mandelbrot distribution as
\[ f(r) = p(r+q)^{-k} \] (14.58)
If \( k = 1 \), the integral of \( f(x) \) leads cumulative \( F(x) \) as cumulative rank-frequency distribution from \( x = 0 \) to \( r \).

\[
F(x) = \int_0^r f(x) \, dx = \int_0^r c(x + d)^{-1} \, dx = c \ln(r/d + 1) \quad (14.59)
\]

And the distribution from \( x = 1 \) to \( r \) becomes

\[
F(x) = \int_1^r f(x) \, dx = \int_1^r c(x + d)^{-1} \, dx = c \ln[(r + d)/(1 + d)] \quad (14.60)
\]

Those belong to Bradford-Brookes type distributions.

If \( k \neq 1 \), the integral of \( f(x) \) also leads cumulative Mandelbrot-type distributions from \( x = 0 \) to \( r \) and \( x = 1 \) to \( r \) respectively as

\[
F(x) = c[(r + d)^{-k+1} - d^{-k+1}] \quad (14.61)
\]

\[
F(x) = c[(r + d)^{-k+1} - (1 + d)^{-k+1}] \quad (14.62)
\]

In Eq. (14.15), when \( d = 0 \) and \( \alpha = 1 \), it is just Bradford’s law.

Generally, we can consider that Eqs. (14.15) and (14.17) are basic spatial functions for size-frequency distribution and rank-frequency distribution respectively. When we ignore time factor (exponential item), cumulative rank distributions \( G(r) \) can be derived with integral on interval \( r \in [1, r] \) or \( r \in [0, r] \) (there is no \( \infty \) for rank) when \( \beta \neq 1 \) as

\[
G(r) = \int_1^r p(r + q)^{-\beta} \, dr = \frac{p}{1 - \beta}[(r + q)^{1-\beta} - (1 + q)^{1-\beta}] \quad (14.63)
\]

\[
G(r) = \int_0^r p(r + q)^{-\beta} \, dr = \frac{p}{1 - \beta}[(r + q)^{1-\beta} - q^{1-\beta}] \quad (14.64)
\]

Those are cumulative Zipf-Mandelbrot-type distributions, which also belong to generalized Leimkuhler functions (Rousseau 1988, 2005).

When \( \beta = 1 \), the integral produce regular Leimkuhler-type distributions

\[
G(r) = \int_1^r p(r + q)^{-1} \, dr = p \ln[(r + q)/(1 + q)] \quad (14.65)
\]

\[
G(r) = \int_0^r p(r + q)^{-1} \, dr = p \ln(1 + r/q) \quad (14.66)
\]

In Eq. (14.17), when \( \beta = 1 \) and \( q = 0 \), it is just Zipf’s law.

So, the space-type distributions from the unified informetric functions Eqs. (14.15) and (14.17) cover all forms of Bradford, Lotka, Zipf, Leimkuhler and Mandelbrot laws as shown in Table 14.1.
Table 14.1  Typical space-type distribution as informetric laws

<table>
<thead>
<tr>
<th>$x$ or $r$</th>
<th>$\alpha$ or $\beta$</th>
<th>$d$ or $q$</th>
<th>Informetric law</th>
<th>Mathematical form</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$ : size approach</td>
<td>$\alpha \neq 1$</td>
<td>$d \neq 0$</td>
<td>Shifted Lotka-type</td>
<td>$f(x) = c(x + d)^{-\alpha}$</td>
</tr>
<tr>
<td></td>
<td>$\alpha = 1$</td>
<td>$d \neq 0$</td>
<td>Bradford-Brookes type</td>
<td>$F(x) = c \ln[(x + d)/(1 + d)]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$F(x) = c \ln(1 + x/d)$</td>
</tr>
<tr>
<td></td>
<td>$\alpha \sim 2$</td>
<td>$d = 0$</td>
<td>Lotka</td>
<td>$f(x) = cx^{-\alpha}$</td>
</tr>
<tr>
<td>$r$ : rank approach</td>
<td>$\beta \neq 1$</td>
<td>$q \neq 0$</td>
<td>Zipf-Mandelbrot; Generalized Leimkuhler</td>
<td>$g(r) = p(r + q)^{-\beta}$; $G(r) = \frac{p}{1 - \beta}[(r + q)^{1-\beta} - (1 + q)^{1-\beta}]$; $G(r) = \frac{p}{1 - \beta}[(r + q)^{1-\beta} - q^{1-\beta}]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$G(r) = p \ln[(r + q)/(1 + q)]$</td>
</tr>
<tr>
<td></td>
<td>$\beta = 1$</td>
<td>$q \neq 0$</td>
<td>Leimkuhler-type</td>
<td>$G(r) = p \ln[(r + q)/(1 + q)]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$G(r) = p \ln(1 + r/q)$</td>
</tr>
<tr>
<td></td>
<td>$\beta = 1$</td>
<td>$q = 0$</td>
<td>Zipf</td>
<td>$g(r) = pr^{-1}$</td>
</tr>
</tbody>
</table>

Since the space-type distribution of the unified informetric model covers all forms of Bradford, Lotka, Zipf, Leimkuhler and Mandelbrot distribution (Fairthorne 1969; Bookstein 1990; Rousseau 1990), we can say that it reveals a unified mechanism.

14.2.4 Time-Type Distributions

When we only consider time distribution, the distribution is independence of space so that we have (Ye 1998, 2011)

$$\frac{d^2 f}{dr^2} - \frac{c^2}{d^2} \frac{df}{dr} = c^2(b - a) f(t)$$  \hspace{1cm} (14.67)

where $a$, $b$, $c$ and $d$ are real constants.

The characteristic equation of Eq. (14.45) is $r^2 - (c^2/d^2)r - c^2(b - a) = 0$. Its two solutions are $r_{1,2} = \left[(c^2/d^2) \pm \sqrt{(c^2/d^4) + 4c^2(b - a)}\right]/2$ so that the general solution of Eq. (14.67) becomes

$$f(t) = c_1e^{r_1t} + c_2e^{-r_2t}$$  \hspace{1cm} (14.68)

In the special case that $c^2 = 4d^4(a - b)$, the characteristic equation has a double root equal to $c^2/d^2$ and

$$f(t) = c_3e^{(c^2/d^2)t} + c_4te^{(c^2/d^2)t}$$  \hspace{1cm} (14.69)
Table 14.2 Some special time-type distributions of informetrics

<table>
<thead>
<tr>
<th>$r_1, r_2$</th>
<th>$c_1, c_2$</th>
<th>Informetric type</th>
<th>Mathematical form</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_1, r_2 &gt; 0$</td>
<td>$c_1, c_2 \neq 0$</td>
<td>Avramescu-type</td>
<td>$f(t) = c_1 e^{r_1 t} + c_2 e^{-r_2 t}$</td>
</tr>
<tr>
<td>$r_1 &gt; 0, r_2 = 0$</td>
<td>$c_2 \neq 0, c_2 = 0$</td>
<td>Exponential-type</td>
<td>$f(t) = c_1 e^{r_1 t} + c_2$</td>
</tr>
<tr>
<td>$r_1 = 0, r_2 &gt; 0$</td>
<td>$c_1 \neq 0, c_1 = 0$</td>
<td>Price-type (growth)</td>
<td>$f(t) = c_1 e^{r_1 t}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative exponentail-type</td>
<td>$f(t) = c_1 + c_2 e^{-r_2 t}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bernal-Brookes-type (ageing)</td>
<td>$f(t) = c_2 e^{-r_2 t}$</td>
</tr>
</tbody>
</table>

In the solutions of Eq. (14.68), it covers some special time-type distributions of informetrics, as shown in Table 14.2.

Combining Eq. (14.68) with Table 14.2, we see that most known time-type informetric distributions are included.

Another possible or potential study concerns whether integrated transformation can link space-type distributions with time-type ones.

Clearly the function $f(x, t) = c(x + d)^{-\alpha} e^{kt}$ or $g(r, t) = p(r + q)^{-\beta} e^{\lambda t}$ is a simple and interesting common solution of the partial differential equations (Ye and Rousseau 2010; Ye 2011), belonging to wave-heat equations. The unified informetric model avoids some artificial suppositions such as another unified scientometric model (Bailón-Moreno et al. 2005). Meanwhile, both space-type distribution and time-type distribution may introduce new types beyond the present informetric laws, which provide a theoretical framework and could stimulate further studies.

14.3 Theoretical Extension

Now, we give some interpretations on that unified informetric model. In the following sections, $f(x)$ always denotes various size-frequency distributions and $g(r)$ various rank-frequency distributions, while $F(x)$ and $G(r)$ denote the corresponding cumulative distributions of $f(x)$ and $g(r)$ respectively. While $f(x)$ means space-type distribution, $f(t)$ does time-type distribution.

When $x \to r$ and interval transformation as $x \in [1, N] \to r \in [1, T]$ or $r \in [0, T]$, based on Eq. (14.17), we suppose similar result according to Theorem 2 above.

$$g(r, t) = p(r + q)^{-\beta} e^{\lambda t} - d$$

(14.70)

where a constant difference $(-d)$ may happen at most. Because the exponential item can be accounted into constant $p$ (there is no changes about $t$), the time-type exponential cutoff never affect the size-rank transformation.

As size function Eq. (14.15) satisfied wave-heat equations (Ye 1998, 2011) and the differential of a constant is zero, rank function Eq. (14.17) will also satisfies the
wave-heat equations. That is why we have similar mathematical structure for both size variable \( x \) and rank variable \( r \) (Ye and Rousseau 2010).

When we consider the shifted power function with exponential spatial cutoff

\[
f(x) = c(x + d)^{-\alpha}e^{-\kappa x}
\]

it provides a theoretical path to link with other known distributions.

As a probability density function (PDF) \( f(x) \), Eq. (14.71) provides the way to approach other probability distributions.

There are one constant \( c \) and three independent parameters as \( \alpha \), \( \kappa \) and \( d \) in Eq. (14.71). If \( \kappa = 0 \) and \( d = 0 \), it becomes standard power function as Pareto distribution (Type I).

When \( \alpha = 0 \), the Eq. (14.71) becomes a pure exponential distribution and the constant \( c \) is simple according to normalization condition for \( x \geq x_m \)

\[
c = \kappa e^{\kappa x_m}
\]

where \( x_m \) is the minimum \( x \).

When \( \kappa = 0 \), the Eq. (14.71) becomes shifted power function Equation (14.15) and the normalized constant \( c \) is

\[
c = (\alpha - 1)(x_m + d)^{\alpha - 1}
\]

So the shifted power law density distribution Equation (14.71) becomes

\[
f(x) = (\alpha - 1)(x_m + d)^{\alpha - 1}(x + d)^{-\alpha}
\]

with normalized condition

\[
\int_{x_m}^{\infty} f(x)dx = \int_{x_m}^{\infty} c(x + d)^{-\alpha}e^{-\kappa x}dx = 1
\]

Corresponding cumulative distribution function (CDF) is

\[
F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} (\alpha - 1)(x_m + d)^{\alpha - 1}(t + d)^{-\alpha}dt
\]

If \( x \) belongs to discrete variables, as \( c \sum_{\alpha=1}^{\infty} x^{-\alpha} = c\zeta(\alpha) = 1 \), we have

\[
c = 1/\zeta(\alpha)
\]

where \( \zeta(\alpha) \) is Riemann \( \zeta \)-function.
In very special case, when we set $x \rightarrow x/\lambda$, $\alpha = 1 - k$, $\kappa x = (x/\lambda)^k$, $d = 0$ and $c = k/\lambda$, the Eq. (14.71) becomes Weibull distribution ((Liang et al. 1996)

$$f_{k,\lambda}(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}$$

(14.78)

where shape parameter $k > 0$ and scale parameter $\lambda > 0$ and $x \geq 0$ (when $x < 0$, $f(x) = 0$).

If $\kappa \neq 0$, generally, according to normalized condition, the constant $c$ in Eq. (14.51) can be determined by integral

$$\int_{x_m}^{\infty} f(x)dx = \int_{x_m}^{\infty} c(x + d)^{-\alpha} e^{-\kappa x} = 1$$

(14.79)

For $\lambda > 0$, integral equation is

$$\int_{0}^{\infty} x^{n-1} e^{-\kappa x} dx = \frac{1}{\kappa^n} \Gamma(n)$$

(14.80)

So we have (Clauset et al. 2009)

$$c = \frac{\kappa^{1-\alpha}}{\Gamma(1-\alpha, \kappa x_m)}$$

(14.81)

where $\Gamma$ denotes $\Gamma$-function. According to Stirling’s approximation for the gamma function, it satisfies

$$\Gamma(x) \approx \sqrt{2 \pi} e^{-x} x^{x-1/2}$$

(14.82)

Obviously, all power functions or Pareto distribution is crucially determined by exponent $\alpha$, so that the exponent $\alpha$ is the most important measure in power-law distributions. Statistically, its moments are given by (for $k < \alpha - 1$).

$$\langle x^k \rangle = \int_{x_m}^{\infty} x^k f(x)dx = \frac{\alpha - 1}{\alpha - 1 - m} x_m^k$$

(14.83)

which is strongly linked with power exponent $\alpha$. When $1 < \alpha < 2$, the first moment, i.e., the mean, is infinite, along with all higher moments. When $2 < \alpha < 3$, the first moment is finite, but the second one (the variance) and higher moments are infinite. That is, all moments $k \geq \alpha - 1$ diverge: when $\alpha < 2$, the average and all higher-order moments are infinite; when $2 < \alpha < 3$, the mean exists, but the variance and higher-order moments are infinite. For finite-size samples drawn from such distribution, that behavior implies that the central moment estimators (like the mean and the variance) for diverging moments will never converge.

Meanwhile, the Gini coefficient is another measure of the deviation of the Lorenz curve, linking Pareto distribution, from the equidistribution line which is a line
connecting [0, 0] and [1, 1]. It is also determined by $\alpha$. The Gini coefficient for the Pareto distribution is calculated according to Eq. (14.84).

$$G = 1 - 2 \int_0^1 L(F) dF = \frac{1}{2\alpha - 1} \quad (14.84)$$

Generally, two cases are included, i.e., the convex form as Lorenz curve ranking from smallest to largest and concave form as Leimkuhler curve ranking from the largest to the smallest (Burrell 2005; 2007).

The Lorenz curve is often used to characterize income and wealth distributions. For any distribution, when the Pareto index is $\alpha = \log_4 5 = \log 5/\log 4$, approximately 1.161, 80/20 rule can be derived from the Lorenz curve formula given above. That excludes Pareto distribution in which $0 < \alpha \leq 1$, which has infinite expected value, so it can not reasonably model income distribution.

To find the estimator for $\alpha$, we can compute the corresponding partial derivative and determine where it is zero.

$$\frac{dl}{d\alpha} = \frac{n}{\alpha} + n \ln x_m - \sum_{i=1}^n \ln x_i = 0 \quad (14.85)$$

Thus Eq. (14.86) becomes the maximum likelihood estimator for $\alpha$, with expected statistical error $\sigma = \alpha_{\text{max}}/\sqrt{n}$

$$\alpha_{\text{max}} = n \sum_{i=1}^n (\ln x_i - \ln x_m) \quad (14.86)$$

The power exponent also links with fractal dimension $D$ as a self-similar fractal as $D = \alpha - 1$ (Egghe 2009). Therefore, we can say, in most general senses, that the power exponent $\alpha$ is so important that it determines main properties of power-law distributions.

### 14.4 Discussion

Besides, supposing the power function $g(r)$ distribute from high to low according to the rank variable $r$, referring to Hirsch (2005), the intersection of the 45° line with the curves will give a series of $h$-index, as shown in Fig. 14.1.

In Fig. 14.1, $h$-points are produced by the crosses of power function $g(r) \sim r^{-\beta}$ and the 45° line $g(r) = (\tan \theta)r$, where $\theta = 45^\circ$ and $\tan 45^\circ = 1$.

Statistically, $h$-index is a characteristic indicator for differentiating the curvatures of different power functions, while $\alpha$ marks the shape of a power function.

In addition, the limitations of processing above should be mentioned. As shifted power function was set to an ideal-like or simple model, it might differentiate from
real situations, so that there could be differences between the ideal model and real states. Nevertheless, as an ideal or standard model is benefit for theoretical analysis, the simple approach is useful for further studies.

Clearly the function $f(x, t) = c(x + d)^{-\alpha} e^{\lambda t}$ is a simple and interesting common solution of the partial differential Equations (14.10) and (14.11). The unified informetric model avoids some artificial suppositions such as another unified scientometric model (Bailón-Moreno et al. 2005). We propose a generalized informetric theory based on the wave and heat equations and raise open question about which $a(x)$ and $b(x)$ lead to meaningful solutions and which are the domains on which those functions must be studied. Are there other solutions? How can they be described?

We expect the combined wave and heat equations to become a unified research framework, which may introduce some new phenomena. Certainly, that is only one theoretical approach to fit informetric distributions following mathematical analysis. And other approaches are following dynamical IPPs (information production processes) (Egghe and Rousseau 1995; Egghe 2007) and thermodynamic way. Because IPPs are complex random processes, different approaches provide different views, which construct different reference systems. And theoretical approach above provides a simple unified framework for interpretation of complex informetric distributions.

14.5 Conclusion

We conclude that in view of the results of that contribution, the wave-heat equations lead to a theoretical approach of unified framework to study informetric laws in terms of partial differential equations, formally similar to that of well-known physical laws. Moreover, we would like to point out that the wave and the heat equation describe
physical phenomena, while our informetric framework uses a generalized wave-type and heat-type set of equations as a theoretical interpretation of informetric laws.

The general shifted power function leads to a unified core and provides a unified and rich mechanism of informetric distributions. Based on the size-to-rank transformation, most theoretical issues can be linked together. While the shifted power function with exponential time cutoff supplies a unified informetric function, the shifted power function with exponential spatial cutoff can—under special situations—be linked to the Pareto distribution and Weibull distribution. The power exponent and $h$-index are two crucial parameters of the power functions, which determine the most important characteristics of the power law distributions. If the general shifted power function becomes the core of standard model for informetrics, more developments can be expected.

Acknowledgements I acknowledge the financial support of the National Natural Science Foundation of China (NSFC grant numbers 70773101, 7101017006 and 71173187) and thank helpful comments from Prof. Dr. Leo Egghe and Prof. Dr. Ronald Rousseau. This paper is integrated and revised by merging two published papers: (1) Ye 2011. A theoretical approach to the unification of informetric models by wave-heat equations. *Journal of the American Society for Information Science and Technology*, 62(6): 1208–1211; (2) Ye, F.Y. (2014). A progress on the shifted power function for modeling informetric laws. *Malaysian Journal of Library and Information Science*, 19(1): 1–15.

References

Chapter 15
The Empirical Investigation and Theoretical Unification of Mathematical Models for the $h$-Index

Among existing theoretical models for the $h$-index, Hirsch’s original formula, the Egghe-Rousseau model and the Glänzel-Schubert model are the three main representatives. Supposing $h_c$, $h_p$ and $h_{pc}$ stand for respectively the Hirsch estimation, Egghe-Rousseau estimation and Glänzel-Schubert estimation empirical data show an inequality $h_p < h \sim h_{pc} < h_c$ at meso-level of institutions and journals. Assuming a power-law relation or Heaps’ law between publications and citations, a unified theoretical explanation for these three models is provided. The results support Glänzel-Schubert’s model as a better estimation of the $h$-index.

15.1 Introduction

Since Hirsch (2005) introduced the idea of the $h$-index, mathematical models of the $h$-index have been discussed by Egghe and Rousseau (2006), Egghe (2007), Glänzel and Schubert (2007, Glänzel (2006), Burrell (2007), and others (Csajbok et al. 2007; Ye and Rousseau 2008). Its application has been extended to journals, countries, patents and other cases (Braun et al. 2006; Schubert and Glänzel 2007; Guan and Gao 2008; van Raan 2006). For further information the reader is referred to the following reviews (Alonso et al. 2009; Egghe 2010).

In all mathematical estimations of the $h$-index, Hirsch’s original formula, Egghe-Rousseau’s model and Glänzel-Schubert’s model are three representative ones. However, which the best fitting is still needs to be examined.
15.2 Methodology

15.2.1 Method

When we define the general $h$-index as “an information source has its index $h$, if $h$ of its $N$ produced outputs have at least $h$ citations each and others ($N - h$) outputs have no more than $h$ citations each”, then all assignees, authors, research groups, institutions, journals and countries have their $h$-indices.

In Hirsch’s original paper (Hirsch 2005), Hirsch proposed his mathematical model for the $h$-index, in which the $h$-index only links total citations $C$ and measures with

$$ h = \sqrt{\frac{C}{a}} $$

(15.1)

where $a$ is a constant ranging between 3 and 5.

Egghe and Rousseau developed a model that we suggest to call Egghe-Rousseau model (Egghe and Rousseau 2006), with the $h$-index linking only total source publications $P$, using

$$ h = P^{1/a} $$

(15.2)

where $\alpha > 1$ is Lotka’s exponent.

Glänzel and Schubert set up the Glänzel-Schubert model (Glänzel 2006; Schubert and Glänzel 2007; Csajbok et al. 2007) with Eq. (15.3)

$$ h = c P^{1/3} (CPP)^{2/3} $$

(15.3)

where $CPP = C/P$ is citations per publication (for journals, $C/P$ is associated with the impact factor, IF); $c$ is a constant.

Using each of those models in turn, we can estimate the $h$-index with Eqs. (15.4)–(15.6), when $\alpha = 2$, $a = 5$ and $c = 1$.

$$ h_c \sim \sqrt[5]{C} $$

(15.4)

$$ h_p \sim \sqrt{P} $$

(15.5)

$$ h_{pc} \sim P^{1/3} (C/P)^{2/3} $$

(15.6)

Let us call $h_p$, $h_c$ and $h_{pc}$, respectively, the Hirsch estimation, Egghe-Rousseau estimation and Glänzel-Schubert estimation of the $h$-index. They can be examined with empirical data.
15.2 Methodology

15.2.2 Data

In the database ISI web of Science (WoS), we can search the $h$-index of a journal via publication name and of an institution via its address. Meanwhile, in the database ISI Essential Science Indicators (ESI), we have selected total publications $P$, citations $C$ and citations per publication (CPP) in a time window of 10–11 years. When we choose the time span 1997–2007 in WoS to correspond to ESI 1997–2007, we can obtain comparable data. For more comparison, we also choose the data ESI 1999–2009 for getting ESI $h$-index directly, with universities only. The three data sets, coming from same ISI source, enable us to construct comparable data sets for the investigation. While searching $h$-indices, WoS data are applied to be real $h$-indices and the ESI data are applied to be corresponding computing ones and 1999–2009 ESI data are kept as a special data set.

The data of journal titles and institutions constitutes two large samples, in which the top 200 journals and institutions are chosen according to their citation ranking in ESI. Using the same journal titles and institutions, we search WoS and obtain the $h$-indices by sorting “times cited”. For simplifying data collection, different spellings of institutions are ignored, as their main spellings contribute their main $h$-indices, thus we can still get the credible results on a statistical level.

15.3 Results

We can compare the searching $h$-indices with the computing $h$-indices via Eqs. (15.4)–(15.6) using data from ESI. In Table 15.1, as below, the example data and results of the top 10 journals, institutions (1997–2007) and universities (1999–2009) are listed according to their $h$-index ranking (for titles with identical $h$-indices, we rank according to $h_{pc}$. In computing $h_{pc}$, we set $c = 0.9$ for journals and $c = 1$ for institutions and record the corresponding $h_{pc}$ as $0.9h_{pc}$ for journals and $h_{pc}$ for institutions).

For observation more data samples, we can broaden the top 10 to the top 100 journals, institutions and universities for their $h$-index fittings as shown in Figs. 15.1, 15.2 and 15.3 respectively.

It is visually clear that the Glänzel-Schubert estimation corresponds best to real data. In Figs. 15.1 and 15.2, we see that all $h_{pc}$ correlate with the searching $h$-indices better, while almost all $h_p$ form the bottom and all $h_c$ leap to the top. In Fig. 15.3, most computing $h$-indices of $h_p$ and $h_c$ distribute lower or higher than real $h$-indices, while $h_{pc}$ does little higher than real ones.
### Table 15.1  Searching $h$-indices and computing $h$-indices: the top 10 journals, institutions and universities

<table>
<thead>
<tr>
<th>Journal</th>
<th>$P$</th>
<th>$C$</th>
<th>CPP</th>
<th>$h$</th>
<th>$h_p$</th>
<th>$h_c$</th>
<th>$0.9h_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature</td>
<td>11,274</td>
<td>1,337,209</td>
<td>118.61</td>
<td>487</td>
<td>106.179</td>
<td>517.147</td>
<td>487.172</td>
</tr>
<tr>
<td>Science</td>
<td>10,404</td>
<td>1,263,175</td>
<td>121.41</td>
<td>476</td>
<td>102</td>
<td>502.628</td>
<td>481.739</td>
</tr>
<tr>
<td>N Engl J Med</td>
<td>3,879</td>
<td>569,640</td>
<td>146.85</td>
<td>374</td>
<td>62.281</td>
<td>337.532</td>
<td>393.609</td>
</tr>
<tr>
<td>Cell</td>
<td>3,824</td>
<td>552,923</td>
<td>144.59</td>
<td>350</td>
<td>61.838</td>
<td>322.542</td>
<td>387.103</td>
</tr>
<tr>
<td>Proc Nat Acad Sci USA</td>
<td>31,437</td>
<td>1,485,447</td>
<td>47.25</td>
<td>315</td>
<td>177.304</td>
<td>545.059</td>
<td>371.242</td>
</tr>
<tr>
<td>Lancet</td>
<td>7,320</td>
<td>438,190</td>
<td>59.86</td>
<td>286</td>
<td>85.557</td>
<td>296.037</td>
<td>267.405</td>
</tr>
<tr>
<td>J Biol Chem</td>
<td>59,611</td>
<td>1,864,004</td>
<td>31.27</td>
<td>264</td>
<td>244.153</td>
<td>545.059</td>
<td>371.242</td>
</tr>
<tr>
<td>JAMA-J Am Med Assn</td>
<td>4,076</td>
<td>340,127</td>
<td>83.45</td>
<td>264</td>
<td>63.843</td>
<td>260.816</td>
<td>274.539</td>
</tr>
<tr>
<td>Nat Genet</td>
<td>780</td>
<td>232,109</td>
<td>29.76</td>
<td>323</td>
<td>88.317</td>
<td>215.457</td>
<td>190.452</td>
</tr>
<tr>
<td>Circulation</td>
<td>10,271</td>
<td>495,513</td>
<td>48.24</td>
<td>250</td>
<td>101.345</td>
<td>314.805</td>
<td>259.249</td>
</tr>
<tr>
<td>Institution</td>
<td>$P$</td>
<td>$C$</td>
<td>CPP</td>
<td>$h$</td>
<td>$h_p$</td>
<td>$h_c$</td>
<td>$h_{pc}$</td>
</tr>
<tr>
<td>Harvard University</td>
<td>95,457</td>
<td>2,651,015</td>
<td>27.77</td>
<td>330</td>
<td>308.961</td>
<td>728.150</td>
<td>419.102</td>
</tr>
<tr>
<td>NIH</td>
<td>7,800</td>
<td>232,109</td>
<td>29.76</td>
<td>323</td>
<td>88.317</td>
<td>215.457</td>
<td>190.452</td>
</tr>
<tr>
<td>Stanford University</td>
<td>49,363</td>
<td>1,131,732</td>
<td>22.93</td>
<td>307</td>
<td>222.177</td>
<td>475.758</td>
<td>296.076</td>
</tr>
<tr>
<td>University Calif San Francisco</td>
<td>36,621</td>
<td>981,823</td>
<td>26.81</td>
<td>299</td>
<td>191.366</td>
<td>443.130</td>
<td>297.468</td>
</tr>
<tr>
<td>Johns Hopkins University</td>
<td>53,594</td>
<td>1,211,258</td>
<td>22.6</td>
<td>297</td>
<td>231.503</td>
<td>492.190</td>
<td>301.377</td>
</tr>
<tr>
<td>University Washington</td>
<td>55,003</td>
<td>1,131,765</td>
<td>20.58</td>
<td>297</td>
<td>234.527</td>
<td>475.765</td>
<td>285.600</td>
</tr>
<tr>
<td>MIT</td>
<td>36,315</td>
<td>814,312</td>
<td>22.42</td>
<td>291</td>
<td>190.564</td>
<td>403.562</td>
<td>263.300</td>
</tr>
<tr>
<td>University Calif San Diego</td>
<td>41,318</td>
<td>920,778</td>
<td>22.29</td>
<td>290</td>
<td>203.268</td>
<td>429.133</td>
<td>273.812</td>
</tr>
<tr>
<td>Brigham &amp; Women’s Hospital</td>
<td>14,940</td>
<td>482,231</td>
<td>32.28</td>
<td>289</td>
<td>122.229</td>
<td>310.557</td>
<td>249.692</td>
</tr>
<tr>
<td>Max Planck Society</td>
<td>72,087</td>
<td>1,346,597</td>
<td>18.68</td>
<td>284</td>
<td>268.490</td>
<td>518.959</td>
<td>293.001</td>
</tr>
<tr>
<td>University</td>
<td>$P$</td>
<td>$C$</td>
<td>CPP</td>
<td>$h$</td>
<td>$h_p$</td>
<td>$h_c$</td>
<td>$h_{pc}$</td>
</tr>
<tr>
<td>Johns Hopkins University</td>
<td>57,208</td>
<td>1,410,157</td>
<td>24.65</td>
<td>322</td>
<td>239.181</td>
<td>288.208</td>
<td>326.356</td>
</tr>
<tr>
<td>Stanford University</td>
<td>51,462</td>
<td>1,307,580</td>
<td>25.41</td>
<td>320</td>
<td>226.852</td>
<td>279.632</td>
<td>321.477</td>
</tr>
<tr>
<td>University California-San Diego</td>
<td>43,037</td>
<td>1,031,652</td>
<td>23.97</td>
<td>300</td>
<td>207.453</td>
<td>254.339</td>
<td>291.345</td>
</tr>
<tr>
<td>University Washington</td>
<td>57,128</td>
<td>1,318,456</td>
<td>23.08</td>
<td>300</td>
<td>239.014</td>
<td>280.560</td>
<td>312.196</td>
</tr>
<tr>
<td>University California-San Diego</td>
<td>38,246</td>
<td>1,075,166</td>
<td>28.11</td>
<td>298</td>
<td>195.565</td>
<td>258.572</td>
<td>311.497</td>
</tr>
<tr>
<td>University California-Berkeley</td>
<td>36,977</td>
<td>961,455</td>
<td>26.06</td>
<td>295</td>
<td>192.085</td>
<td>247.270</td>
<td>292.609</td>
</tr>
<tr>
<td>University California-Los Angeles</td>
<td>49,349</td>
<td>1,089,988</td>
<td>22.09</td>
<td>284</td>
<td>221.146</td>
<td>259.997</td>
<td>288.749</td>
</tr>
<tr>
<td>University California-Berkeley</td>
<td>45,900</td>
<td>995,116</td>
<td>21.68</td>
<td>269</td>
<td>214.242</td>
<td>250.697</td>
<td>278.384</td>
</tr>
</tbody>
</table>
15.3 Results

Fig. 15.1 $h$-index fitting of the top 100 journals

Fig. 15.2 $h$-index fitting of top 100 institutions

Fig. 15.3 ESI $h$-index fittings of top 100 universities
15.4 Analysis and Discussion

Among the models for the $h$-index, Hirsch’s original formula, the Egghe-Rousseau model and the Glänzel-Schubert model are the three main representatives. Empirically, we found that the Glänzel-Schubert model matched practical data best, particularly at higher levels of aggregation (Ye 2009).

Under the assumptions of the power-law model (Lotkaian informetrics), Egghe-Rousseau’s model is written as

$$ h = P^{1/\alpha} \tag{15.7} $$

where $\alpha$ is the Lotka coefficient for the citation distribution and $P$ is the number of publications (or, in general, the number of sources).

Assuming a power law relation between $P$ and $C$, i.e., Heaps’ law or Herdan’s law exists between $P$ and $C$ (Egghe 2007), we obtain

$$ C = a P^\beta \tag{15.8} $$

Substituting Eq. (15.8) into Eq. (15.7), leads to

$$ h = \left( \frac{C}{a} \right)^{1/\alpha \beta} \tag{15.9} $$

That is Hirsch’s original formula of the $h$-index (Hirsch 2005) when $\alpha \beta = 2$.

$$ h = \sqrt{\frac{C}{a}} \tag{15.10} $$

Combining Eqs. (15.7) and (15.8), we obtain

$$ h = P^{1/\alpha} C^{\alpha/(\alpha+1)} C^{-\alpha/(\alpha+1)} = a^{-\alpha/(\alpha+1)} P^{1/\alpha} C^{-\alpha/(\alpha+1)} (P^{-\alpha/(\alpha+1)})^\beta \tag{15.11} $$

If now $\beta$ is related to $\alpha$ as follows.

$$ \beta = 1 + \frac{1}{\alpha^2} = \frac{\alpha^2 + 1}{\alpha^2} \tag{15.12} $$

and letting $c = a^{-\alpha/(\alpha+1)}$, we get

$$ h = a^{-\alpha/(\alpha+1)} P^{1/\alpha} (C/P)^{\alpha/(\alpha+1)} (P^{-1/\alpha/(\alpha+1)}) = c P^{1/(\alpha+1)} (C/P)^{\alpha/(\alpha+1)} \tag{15.13} $$

That is the Glänzel-Schubert model for the $h$-index, where $c > 0$ is a constant.

Thus, we see that the $h$-index formulae of Hirsch, Egghe-Rousseau, Glänzel-Schubert are linked based on the power-law relation or Heaps’ law Equation (15.8).
In general, the $h$-index is a function of $P$ and $C$. We know that $P$ and $\alpha$ are independent variables, leading to that the Egghe-Rousseau formula is the tightest link between $P$ and $h$. That is why that pure power-law model seems difficult to fit practical data (Ye and Rousseau 2008). With more free parameters, it becomes easier to fit real data, so that Glänzel-Schubert model looks better, where there are two free parameters, $c$ and $\alpha$.

When $\alpha = 2$ and $c = 1$, we have $\beta = 1.25$ and $a = 1$. That is a special situation, on which we obtain unified simple estimations for the $h$-index, including the Hirsch-type estimation

$$h_H = h_c \sim C^{2/5} = C^{0.4} \quad (15.14)$$

the Egghe-Rousseau estimation

$$h_{E-R} = h_p \sim P^{1/2} = P^{0.5} \quad (15.15)$$

and the Glänzel-Schubert estimation

$$h_{G-S} = h_{pc} \sim P^{1/3}(C/P)^{2/3} \quad (15.16)$$

Consequently, the Egghe-Rousseau model and the Glänzel-Schubert model for the $h$-index focus on the existence of a power law relation between $P$ and $C$ as shown by Eq. (15.8) and a relation between $\alpha$ and $\beta$ as given by Eq. (15.12).

The empirical results above support the Glänzel-Schubert model as a better estimation of the $h$-index at both journal and institution level. Because we know that the $h$-index relates both publications and citations, it is realistic and correct to support the Glänzel-Schubert model. The fitting results also show that there is an inequality for most cases, as follows, at meso-level of both journals and institutions.

$$h_p < h < h_{pc} < h_c \quad (15.17)$$

Meanwhile, we can compute the Pearson correlation coefficients of the top 100 data columns, using SPSS, shown in Table 15.2 (for journals), Table 15.3 (for institutions, 1997–2007) and Table 15.4 (for universities, 1999–2009). That is confirmed by calculating both Pearson and Spearman correlation coefficients (as the Pearson coefficient measures a linear relation and the Spearman correlation measures if ranks corresponds, we use both coefficients for synthetic measures) between the estimates and the real $h$-indices (Rousseau and Ye 2008), integrated as a correlation matrix.

Thus, all data suggest that the Glänzel-Schubert model is a better estimation of the $h$-index, so that it is feasible to apply the Glänzel-Schubert model to estimate the $h$-indices of countries and other information sources (Csajbok et al. 2007).
### Table 15.2  The correlation matrix of both Spearman and Pearson (at the 0.01 level) for journals

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Spearman (Sig.(2-tailed))</th>
<th>Pearson (Sig.(2-tailed))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$h$</td>
<td>$0.9h_{pc}$</td>
</tr>
<tr>
<td>$h$</td>
<td>1</td>
<td>0.942(0.000)</td>
</tr>
<tr>
<td>$0.9h_{pc}$</td>
<td>0.975(0.000)</td>
<td>1</td>
</tr>
<tr>
<td>$h_p$</td>
<td>0.124(0.220)$^a$</td>
<td>0.169(0.093)$^a$</td>
</tr>
<tr>
<td>$h_c$</td>
<td>0.740(0.000)</td>
<td>0.723(0.000)</td>
</tr>
</tbody>
</table>

$^a$indicates no correlation. The others show significant correlations.

### Table 15.3  The correlation matrix of both Spearman and Pearson (at the 0.01 level) for institutions

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Spearman (Sig.(2-tailed))</th>
<th>Pearson (Sig.(2-tailed))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$h$</td>
<td>$h_{pc}$</td>
</tr>
<tr>
<td>$h$</td>
<td>1</td>
<td>0.907(0.000)</td>
</tr>
<tr>
<td>$h_{pc}$</td>
<td>0.885(0.000)</td>
<td>1</td>
</tr>
<tr>
<td>$h_p$</td>
<td>0.449(0.000)</td>
<td>0.576(0.000)</td>
</tr>
<tr>
<td>$h_c$</td>
<td>0.756(0.000)</td>
<td>0.899(0.000)</td>
</tr>
</tbody>
</table>

### Table 15.4  The correlation matrix of both Spearman and Pearson (at the 0.01 level) for universities

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Spearman (Sig.(2-tailed))</th>
<th>Pearson (Sig.(2-tailed))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ESI $h$</td>
<td>$h_{pc}$</td>
</tr>
<tr>
<td>ESI $h$</td>
<td>1</td>
<td>0.976(0.000)</td>
</tr>
<tr>
<td>$h_{pc}$</td>
<td>0.986(0.000)</td>
<td>1</td>
</tr>
<tr>
<td>$h_p$</td>
<td>0.745(0.000)</td>
<td>0.749(0.000)</td>
</tr>
<tr>
<td>$h_c$</td>
<td>0.947(0.000)</td>
<td>0.957(0.000)</td>
</tr>
</tbody>
</table>

### 15.5 Conclusion

Employing the Hirsch formula, the Egghe-Rousseau model and the Glänzel-Schubert model respectively, we estimate the $h$-index as $h_c \sim \sqrt{C/5}$, $h_p \sim \sqrt{P}$ and $h_{pc} \sim c(P)^{1/2}(C/P)^{2/3}$, based on three data sets from ISI databases. The results support the Glänzel-Schubert model as a better estimation of the $h$-index at both journal and institution levels, so that we can also apply Glänzel-Schubert model to estimate the $h$-indices of countries and other information sources. And an inequality for most cases is suggested by $h_p < h \sim h_{pc} < h_c$.

We hope that this unification of models provides a new and useful perspective on the study of $h$-indices.
Acknowledgements  I acknowledge NSFC Grants (number 70773101 and number 7101017006) and am grateful to Dr. Ronald Rousseau for his comments. This chapter began in China and was completed in Germany, I thank financial support from Humboldt University and conditions provided by Prof. Stefan Hornbostel, Dr. Sybille Hinze and iFQ colleagues. This chapter is integrated and revised by merging two published papers: (1) Ye, F. Y. 2009. An investigation on mathematical models of the \textit{h}-index. \textit{Scientometrics}, 81 (2): 493–498; (2) Ye, F. Y. 2011. A unification of three models for the \textit{h}-index. \textit{Journal of the American Society for Information Science and Technology}, 62(1): 205–207.

References

Ye, F.Y.: An investigation on mathematical models of the \textit{h}-index. Scientometrics \textbf{81}(2), 493–498 (2009)
Chapter 16
The \( h \)-Type Core Structure in Single-Layer and Multi-layer Weighted Information Networks

Applying \( h \)-type indicators into weighted networks, a series of \( h \)-type measures, including \( h \)-degree and \( h \)-strength, are introduced, for characterizing core structure in weighted networks. While the algorithms of \( h \)-degree and \( h \)-strength act on single-layer weighted networks for identifying network \( h \)-core, \( h \)-subnet and \( h \)-bone, the algorithms of \( h \)-degree and \( h \)-strength can also act on multi-layer weighted networks for extracting \( h \)-crystal, by constructing layer-bridges between the layers’ network \( h \)-cores and \( h \)-subnets. It is empirically found that all the \( h \)-type core structures exist, while their features and properties are revealed.

16.1 Introduction

Stimulated by Watts and Strogatz (1998) and Barabási and Albert (1999), the complex networks became a heated research field, leading to new developments of network science, covering graph theory, social networks, ecological networks, molecular networks, cognitive networks as well as information networks (Albert and Barabási 2002; Barrat et al. 2004; Börner et al. 2007; Newman, 2001, 2003; Clauset et al. 2008; Serrano et al. 2009; Ahn et al. 2010). Moreover, the studies of complex networks have led from single-layer and homogeneous networks to multi-layer and heterogeneous networks (Sun and Han 2012), including hierarchical, multi-layer and multiplex networks (Bianconi 2013; Menichetti et al. 2014; Boccaletti et al. 2014). Many kinds of typical information networks, such as citation and co-citation networks (Egghe and Rousseau 2002; Ding et al. 2009; Chen and Redner 2010), coauthor networks (Liu et al. 2005; Rodriguez and Pepe 2008), keywords networks (Su and Lee 2010) and patent networks (Liu and Shih 2011) are being investigated. With providing quantitative methods to express relational data and to resolve the structure of relations, network analysis is a useful tool for investigating complex phenomena, including those with informetric content. Those methods can be also
used to study different forms of single-layer and multi-layer networks, where a central issue of network analysis is to identify a minimum set of core nodes and links in a large and complex network and \( h \)-type indicators provide a kind of highly efficient method to extracting the core from ranked data.

Meanwhile, since \( h \)-index (Hirsch 2005) was introduced, its theoretical aspects have been thoroughly studied (Egghe and Rousseau 2006; Glänzel 2006; Schubert and Glänzel 2007; Ye, 2009, 2011) and its applications have been widely extended to various other source-item relations (Egghe 2005; Braun et al. 2006; Alonso et al. 2009; Norris and Oppenheim 2010). Combining network analysis with the \( h \)-index became an interesting idea. In 2009, Korn et al. (2009) and Schubert et al. (2009) introduced the lobby index as a centrality parameter for nodes and the \( h \)-index of a network as an indicator for complete networks (Schubert and Soos 2010). The lobby index of node \( n \) is defined as the largest integer \( k \) such that \( n \) has at least \( k \) neighbors with a degree, \( d(n) \), of at least \( k \). The \( h \)-index of a network is defined as “a network’s \( h \)-index is \( h \), if not more than \( h \) of its nodes have a degree not less than \( h \)” (Schubert 2010). Those indicators can be further developed leading to enriched network analysis methods in numerous applications (Rousseau and Ye 2011).

Along that direction, we also extended the \( h \)-type terminology: \( h \)-degree, producing network \( h \)-core (Zhao et al. 2011); \( h \)-strength, producing \( h \)-subnet (Zhao et al. 2014). Combining \( h \)-degree and \( h \)-strength, we can derive more \( h \)-type core structures, leading to \( h \)-bone for single-layer networks and \( h \)-crystal for multi-layer networks. In this chapter, we give a logic infrastructure and strengthen the \( h \)-strength leading to \( h \)-subnet in single-layer weighted networks and \( h \)-crystal in multi-layer weighted networks, for revealing the core structures in information networks. That could be beneficial to much wider subjects and applications.

### 16.2 Methodology

A network is constructed by nodes (vertices) and links (edges) (Otte and Rousseau 2002; Boccaletti et al. 2006; Newman 2010). When nodes and links represent information-related objects, we refer to such networks as information networks. Figure 16.1 shows that some objects used to build information networks of scientific literature and different objects may belong to different types.

Co-authorship networks, co-citation networks, bibliographic coupling networks and similar networks are examples of single-layer networks, which can only host the same kind of information. Following the concepts of Boccaletti et al. (2014), we consider that several aspects simultaneously lead to a multi-layer ed structure. For example, we may consider the three layers: a co-citation layer, a bibliographic coupling layer and a co-keyword layer. That construction is an example of a multi-layer network in which constituents, i.e., nodes and links, are of a different nature. In those examples, nodes are papers (Layer 1 and Layer 2) and terms (Layer 3), and links have different meanings as well, namely referring to co-citation, bibliographic coupling and being co-keywords.
In the following sections, we characterize the core structures for the single layer networks as well as for the multi-layer networks, with defining the extend $h$-type measures in weighted networks.

### 16.2.1 $h$-Degree, Leading to Network $h$-Core

In many real networks, the strength of links is an important parameter, leading to the notion of weighted (or valued) networks. We define the node strength of a node in a weighted network as the sum of the strengths (or weights) of its all links (Barrat et al. 2004). Often the term node degree is used in the case of unweighted networks as well as in weighted ones (sometimes for making a clear distinction, we use the term node degree only for unweighted networks and use the term node strength in weighted networks.).

The lobby index mentioned in the introduction can be adapted to a weighted network as follows: the $w$-lobby index (weighted network lobby index) of node $n$, denoted as $l(n)$ is defined as the largest integer $k$ such that node $n$ has at least $k$ neighbors with node strength at least $k$.

Figure 16.2 shows that node strength as defined above does not distinguish between obviously different situations. Node A in Fig. 16.2 only has one link with one other node; node B has 5 links with 5 other nodes. The node strengths of A and B are the same and equal to 11, although their position and role in the network are significantly different.

That example illustrates that more suitable basic parameters are needed to describe structural properties in a weighted network. Definition 1 introduces a new type of
degree centrality, which takes a step in that direction. It is called the \( h \)-degree, denoted as \( d_h \) and is defined as follows.

**Definition 1** (Zhao et al. 2011): the \( h \)-degree \( (d_h) \) of node \( n \) in a weighted network is equal to \( d_h(n) \) if \( n \) has at least \( d_h(n) \) links each with strength at least equal to \( d_h(n) \).

The following properties establish a relation between the \( h \)-degree and the \( w \)-lobby index.

**Proposition 1** a node’s \( h \)-degree is always smaller than or equal to its \( w \)-lobby index: \( d_h(n) \leq l(n) \)

*Proof* if node \( n \)’s \( h \)-degree is \( k \) then node \( n \) has at least \( k \) neighbors (as it has \( k \) links). Those neighbors’ node strength is at least \( k \) (the link with \( n \) alone yields at least a value equal to \( k \)). Hence \( l(n) \) is at least \( k \).

**Proposition 2** if node \( n \) has exactly \( d_h(n) \) links then \( d_h(n) = l(n) \).

*Proof* if \( n \) has exactly \( d_h(n) \) links then its \( w \)-lobby index can at most be equal to \( d_h(n) \). However, by proposition 1 \( d_h(n) \leq l(n) \). It proves the equality.

Let \( N \) denote the total number of nodes in a network and let \( N_a \) be the number of adjacent nodes of a given node. Clearly, \( N_a \leq N - 1 < N \). Then, the following basic inequalities hold.

**Proposition 3** for non-isolated nodes in a weighted network, the following inequality involving the \( h \)-degree \( (d_h) \) always holds

\[
1 \leq d_h \leq N_a < N
\] (16.1)

*Proof* for non-isolated nodes, there is at least one adjacent node linked to it and the strength of that link is larger than or equal to 1, hence \( 1 \leq d_h \). \( N_a \) is the total number of links of that node, leading to: \( d_h \leq N_a < N \).

**Proposition 4** if node \( n \) has node strength \( d(n) \) and \( h \)-degree \( d_h(n) \) then

\[
d(n) \geq (d_h(n))^2
\] (16.2)

*Proof* when a node has \( h \)-degree \( d_h \), then the minimum number of links is \( d_h \), and each of these links has at least a strength equal to \( d_h \). Hence \( d(n) \geq (d_h(n))^2 \).
Similar to node degree in unweighted networks and node strength in weighted networks, the $h$-degree is a basic parameter for network analysis. Using the notion of an $h$-degree leads to a network’s $h$-core, a substructure of the complete network.

**Definition 2** a network’s $h$-core is set of nodes and their links, that all have an $h$-degree at least $h$.

It is important to point out that the $h$-degree ($d_h$) is a node-based measure, while a network’s $h$-core is a set of nodes and their links.

### 16.2.2 $h$-Strength, Leading to $h$-Subnet and $h$-Bone

Following network significant ideas, we also introduce a measure to characterize the major part of a network by the link strengths, called $h$-strength. The definition is given below.

**Definition 3** (Zhao et al. 2014): the $h$-strength of a network is equal to $h_s$, if $h_s$ is the largest natural number such that there are $h_s$ links each with strength at least equal to $h_s$ in the network.

The $h$-strength can be computed by descending the strength of links, as shown in Fig. 16.3.

There are some theoretic or practical features of $h$-strength. Firstly, it aims to abstract the links’ strength in weighted networks and its computation directly focuses on the links (not nodes). As links and their strengths characterize the structure of a network and its weights, the measure could be specific statistics to represent high strength links and their distribution in the network. A high $h$-strength means that in the network some pairs of actors have strong paths, connections or interactions. Those links should be the interesting focus when exploring a network and its main structure. Secondly, $h$-strength characterizes the set of links with high strength (usually implying the importance or high relevance) and also fits the common phenomenon of power law (Barabási and Albert 1999; Barabási 2009; Clauset et al. 2009; Zhao and Ye 2013). Thirdly, inheriting from the characteristic of $h$-index, $h$-strength naturally
balances the number of links and the links’ strength. Thus its numerical value is determined by the quantities in the network itself, rather than other artificial threshold values (usually are 5, 10 or 100, etc.) set by researchers. Another advantage of \( h \)-strength is the simplicity of calculating. In network analysis, a network can be mathematically represented in two main ways, adjacency matrix and link list. It is believed that adjacency matrix is more appropriate in application, while link list is only suitable for storing the structure of a network on computer, but cumbersome for computing (Newman 2010). However, such a situation could be improved with the introduction of \( h \)-strength.

\( h \)-strength is totally designed for weighted networks and it always equals to 1 in unweighted networks since there is at least one link. In brief, as a basic measure, the motivation of applying \( h \)-strength in weighted networks is to provide a specific perspective to study the distribution of links’ structure and strength, which may lead to some interesting extensions. In present work, we suggest two applications of \( h \)-strength in detecting the main structure of networks by \( h \)-subnet.

According to the interesting findings of \( h \)-core (Rousseau 2006; Kuan et al. 2011; Glänzel 2012) and the concept of \( h \)-strength, the core set of links including the top part of links with high strength can be identified. For example, as listed in Fig. 16.3, the set of high strength links, \((C, E), (F, G), (C, F)\) and \((E, F)\), can be extracted when computing \( h \)-strength. Thus, there is a way to abstract the core sub-network of a network, called \( h \)-subnet, which can be defined as follows.

**Definition 4** the \( h \)-subnet of a network is a subnetwork which includes the links whose strengths are larger than or equal to \( h \)-strength of the network and the nodes adjacent to these links.

Note that \( h \)-strength \((h_*)\) is a measure defined by links and \( h \)-subnet includes all the links and their nodes. The \( h \)-strength is a link-based measure, while \( h \)-subnet is a set of the links and their nodes. All those informetric indicators characterize the importance of the nodes within the single-layer network (within one layer). The \( h \)-subnet contains the high strength links and the nodes linked by them. If a network has \( h \)-strength \( k \), sometimes there is more than one link with strength \( k \). We suggest that \( h \)-subnet should count all the links with strength \( k \) and thus the number of links in \( h \)-subnet would possibly be larger than \( h \)-strength in actual cases.

For keeping the connection in \( h \)-subnet, to form \( h \)-bone, we need another concept called key-bridge(s).

**Definition 5** key-bridge(s) of a whole network is the link(s) that link through network \( h \)-core and \( h \)-subnet by connecting the nodes with the highest \( h \)-degree in the network, according to the shortest paths.

As both network \( h \)-core and \( h \)-subnet supply two sub-structures of weighted networks, the key-bridge(s) provide the linkage(s) of those two sub-structures. Based on \( h \)-degree and \( h \)-strength, linked by key bridge(s), with combining network \( h \)-core and \( h \)-subnet, a connected core structure of whole weighted network, called \( h \)-bone, can be defined as follows.
Definition 6 the $h$-bone of a single-layer network is a core structure, which includes all the network $h$-core and the $h$-subnet in the weighted network, with connected by key-bridge(s).

In some cases, $h$-subnet itself is connected, so it is just $h$-bone. If the core structure of $h$-bone is broken or unconnected, we say that $h$-bone does not exist.

Concepts above, from $h$-degree to $h$-bone, provide a methodology for finding the core structure or sub-structure in a complex single-layer weighted network.

16.2.3 $h$-Crystal

It is also necessary and important to find the core structure or sub-structure of a complex multi-layer network (Serrano et al. 2009; Battiston et al. 2014), for judging the relatively important nodes and links as well as their functions in the complex network. For multi-layer networks, we firstly introduce the notion of a layer bridge to connect the core structures of all single layer networks so that we can approach the multi-layer core structure.

Definition 7 if two nodes in two layers of a multi-layer network represent the same object (either the same document or documents with the same indexing terms), these nodes are artificially linked. These nodes are called layer-bridges of the multi-layer network.

When two layers in a multi-layer network have layer-bridges they are connected, otherwise, they are not. The network $h$-core and the $h$-subnet are sub-structures of weighted networks, which might not be connected. When all layers are connected through layer-bridges, that leads to the complete connected graph of the multi-layer network.

Combing $h$-degree, $h$-strength with layer-bridges, linking through multi-layer network $h$-cores and $h$-subnets, a core structure of whole multi-layer weighted network, called $h$-crystal, can be defined as follows.

Definition 8 the $h$-crystal is a core structure existing in a multi-layer weighted network, which is consisted of all network $h$-cores and $h$-subnets in each layer of the network, connected by layer-bridges between two layers.

By definition, $h$-crystal is a core structure, linking through network $h$-cores and $h$-subnets in all layers via layer-bridges, where $h$-crystal must be connected. If there is only unconnected or broken structure, we say that $h$-crystal does not exist in the multi-layer weighted network. Here, the layer-bridges between two layers link the really same nodes (same documents). As the $h$-core and $h$-subnet among single-layer networks are connected by the bridges, $h$-crystal will be connected and lead unique core structure in the multi-layer (sometimes with heterogeneous) weighted network so long as $h$-crystal exists. For more general cases, we can introduce extended $h$-crystal.
Definition 9 the extended $h$-crystal is a connected graph as core structure of a multi-layer weighted multi-layer network, consisting of all $h$-bones in each layer of the network, linked by layer-bridges between two layers.

Since $h$-bone is connected, extended $h$-crystal will be connected if it exists. Here, the layer-bridges between two layers link the really same nodes (same documents). As the $h$-core and $h$-subnet among single-layer networks are connected by the bridges, $h$-crystals may be connected and lead minimum core structure in the multi-layer network and extended $h$-crystal becomes a super-set of $h$-crystal.

16.2.4 Algorithm and Procedure

On the basis of concepts above, we can develop an operational method or program for finding $h$-bone and $h$-crystal.

The procedure of finding $h$-bone consists of three steps, as shown in Fig. 16.4, where the step 1 (left side) realizes to extract network $h$-core via the algorithm of $h$-degree and the step 2 (right side) abstracts $h$-subnet as another sub-structure of network via the algorithm of $h$-strength. Mostly, $h$-subnet and network $h$-core are not same sub-structure in complex networks. Then, an important operation aims at checking network connection. With using software UCINET (https://sites.google.com/site/ucinetsoftware/home) via its function multi-layer network in which constituents, i.e., nodes and links, are of a different nature. In those examples, nodes are papers (Layer 1 and Layer 2) and terms (Layer 3), and links have different meanings as well, namely referring to co-citation, bibliographic coupling and being co-keywords. “network-cohesion-distance” (or “network-path”), we have the matrix of geodesic distances. If there is element 0 in the matrix, its corresponding distance is infinite, which means that the network is broken. So, we need re-construct core sub-network by key-bridge(s). In the step 3 (in the middle), we re-construct the connected sub-network by the shortest path(s) among the nodes with the highest $h$-degree, leading to form key-bridge(s). After completing the key bridge(s), linking through network $h$-core and $h$-subnet, the core structure of connected network, $h$-bone, is setup.

In step 2, if we find that networking and $h$-subnet are connected as a unified sub-network without broken structure, the $h$-subnet is just the $h$-bone, which is a single-core network, while others belong to multi-core networks.

Meanwhile, the $h$-crystal of a multi-layer network can be also identified in three steps as illustrated in Fig. 16.5. Step 1: extracting in each layer the network $h$-core using the extracting in each layer the network $h$-core using the algorithm of $h$-degree; step 2: finding the $h$-subnet using the algorithm of $h$-strength in each layer of multi-layer weighted information network; step 3: constructing layer-bridges through linking the same nodes in two layers of the multi-layer weighted information network.

The dark black nodes and edges in Fig. 16.5 mark the network’s $h$-core; the white nodes with edges form the $h$-subnet. The same paper nodes in Layer 1 and Layer 2 construct layer-bridges between Layer 1 and Layer 2, and the same keywords in
16.2 Methodology

Fig. 16.4 Finding the $h$-bone in a single-layer weighted network

Fig. 16.5 Identifying the $h$-crystal in a multi-layer network (Li et al. 2016)

different papers in Layer 2 and Layer 3 produce layer-bridges between Layer 2 and Layer 3. As the co-citation layer, bibliographic coupling layer and co-keyword layer are linked by layer-bridges, we got connected $h$-crystal.

Pseudo-codes for identifying $h$-crystal in a multi-layer network are provided in Appendix, in which $h$-subnet algorithm has been included. The two-dimensional diagram and three-dimensional diagram shown in this chapter are, respectively, created using the open-source software packages NetDraw (https://sites.google.com/site/netdrawsoftware/home) and Mage (http://kinemage.biochem.duke.edu/software/mage.php).
16.2.5 Data Sets and Experiments

Our empirical data originate from related two data sets, for single-layer and multi-layer networks respectively.

The first data set referred to the “$h$-index”, which is the paper co-citation network with the research topic of “$h$-index”. The nodes here correspond to the research papers about “$h$-index” in Web of Science (WoS) 2005–2012, which have more than 9 citations and their references, and the links are the co-citation relationships among papers (nodes). That network has a considerable size with 2,868 nodes and 162,947 links, but low average link strength (1.095).

The second data set referred to the “$h$-set”, which is retrieved from the WoS by the following search strategy: TS= ($h$-index or $h$-type ind* or $h$-like ind* or Hirsch index) or TI=“an index to quantify*” in the WoS for the publication period 2005–2012. Results were restricted to the two fields of information science & library science and multi-disciplinary sciences.

The first data set will be applied to finding $h$-subnet in single-layer single layer information network and the second data set will be applied to identifying $h$-crystal in multi-layer information networks by combining a co-citation network, a bibliographic coupling network and a co-keyword network (including the keywords in the ID and DS records). Table 16.1 shows the main features of these multi-layer graphs in the data set “$h$-set”.

<table>
<thead>
<tr>
<th>Table 16.1 Multi-layer information network parameters of the “$h$-set”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Number of nodes</td>
</tr>
<tr>
<td>Number of edges</td>
</tr>
<tr>
<td>$h$-degree</td>
</tr>
<tr>
<td>Number of nodes in the network’s $h$-core</td>
</tr>
<tr>
<td>$h$-strength</td>
</tr>
<tr>
<td>Number of nodes in the network’s $h$-core and $h$-subnet</td>
</tr>
<tr>
<td>Number of edges in the network’s $h$-core and $h$-subnet</td>
</tr>
</tbody>
</table>
16.3 Results

16.3.1 The $h$-Subnet of Co-citation Network on “$h$-Index”

With the application of $h$-strength and $h$-subnet, Fig. 16.6 unveils the major structure of the network more clearly. The $h$-subnet contains 24 links and 17 nodes and can be regarded as a representation of the center of network. In that $h$-subnet, the maximum link strength is 55 between the nodes “Hirsch J, 2005, PNAS” and “Egghe L, 2006, Scientometrics, 69: 131”, and the minimum strength is 24 among 4 pairs of nodes. The original publication of $h$-index, “Hirsch J, 2005, PNAS”, naturally becomes the center of the network. Overall, $h$-strength and $h$-subnet significantly simply a complex network and readily present the core structure in an effective way.

In addition, that case study shows that $h$-subnet not only abstracts the main high strength links, but also extracts a large share of important nodes at the same time, which constructs the $h$-bone. We found that all the 10 papers with the highest $h$-degree are included in the $h$-subnet of Fig. 16.6, and 13 out of the 17 important nodes whose $h$-degree are not less than 5 are included in the $h$-subnet. $h$-subnet in Fig. 16.6 covers the majority of primary works of $h$-index: the most important original research paper of $h$-index (Hirsch 2005), the earliest test for the validity of $h$-index (Bornmann and Daniel 2005; van Raan 2006), the famous modifications ($g$-index and $R$-index) (Egghe 2006; Jin et al. 2007), the earliest theoretical models of $h$-index (Egghe and Rousseau 2006; Glänzel 2006), the first extended application of $h$-index (Braun et al. 2006), etc. In that case, the method of $h$-subnet uses only 0.014% of the links but extracts most of the important nodes.

![Fig. 16.6 $h$-subnet in “$h$-index” co-citation network ($h$-strength=24) (Zhao et al. 2014)](image-url)
16.3.2 The h-Crystal of the “h-Set”

The h-crystal identified by the h-crystal procedure for the “h-set” is shown in Fig. 16.7, where a three-dimensional representation is drawn. In Fig. 16.7, A refers to the co-citation layer, B to the bibliographic coupling layer and C to the co-keyword layer. Linked nodes between layers A and B represent the same papers, i.e., A270 equals B195 and so on.

In Fig. 16.7, the A-type nodes express the core nodes of the co-citation network (A-core); B-type nodes mark the core nodes of the bibliographic coupling network (B-core), while C-type nodes refer to the core nodes of the co-keyword network (C-core). To illustrate the network structural information, we select the top five core nodes according to their betweenness centrality, shown in Table 16.2, as core sample in the multi-layer network.

![Fig. 16.7 The h-crystal of the data set “h-set” (Li et al. 2016)](image)

<table>
<thead>
<tr>
<th>Layer of co-citation network</th>
<th>Layer of bibliographic coupling network</th>
<th>Layer of co-keyword network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core node</td>
<td>bc</td>
<td>Core node</td>
</tr>
<tr>
<td>A4063</td>
<td>0.46</td>
<td>B296</td>
</tr>
<tr>
<td>A2778</td>
<td>0.356,9</td>
<td>B116</td>
</tr>
<tr>
<td>A2387</td>
<td>0.298,5</td>
<td>B53</td>
</tr>
<tr>
<td>A1802</td>
<td>0.055,4</td>
<td>B120</td>
</tr>
<tr>
<td>A1446</td>
<td>0.055,4</td>
<td>B391</td>
</tr>
</tbody>
</table>
The empirical case above demonstrates that an “$h$-crystal” can be identified in real-world multi-layer information networks. It forms the core structure of a multi-layer network in scholarly literature. It is clear that those nodes provide the key information for the constructed multi-layer graphs. For example, the first ranked items at the A-layers are Hirsch’s original paper (A2778) introducing the $h$-index in “$h$-set”.

### 16.4 Analysis and Discussion

Some analytic properties and theoretical derives can be discussed as follows.

#### 16.4.1 $h$-Centrality Measures for Nodes

The use of centrality measures, originating from social network analysis (Wasserman and Faust 1994; Scott 2000) has led to valuable methods in all types of networks (Otte and Rousseau 2002; Bollen et al. 2009; Borgatti et al. 2009). Although those methods are mainly applied to identity and characterize key nodes in a network, various centrality measures focus on different roles played by nodes in a network. The best known centrality measures are degree centrality, based on node degree; closeness centrality, based on distance to other nodes; betweenness centrality, based on the ability of nodes to control flows in a network (Freeman, 1977, 1979; Everett et al. 2004). Those measures were originally designed for unweighted networks. For example, when calculating the closeness centrality of nodes in a weighted network, the strength of links is often ignored and the weighted network is actually converted to an unweighted network.

Degree centrality of node $n$ in an unweighted network, denoted as $C_d(n)$, is defined as (Freeman 1979)

$$C_d(n) = \frac{d(n)}{N - 1} \quad (16.3)$$

Adapting that definition to the context of the $h$-degree we propose the following definition.

**$h$-centrality**: in a weighted network with $N$ nodes, the $h$-centrality, $C_h$, of node $n$ is defined as

$$C_h(n) = \frac{d_h(n)}{N - 1} \quad (16.4)$$

where $d_h(n)$ is the $h$-degree of node $n$. $h$-centrality is just a normalized form of the $h$-degree, which measures the importance of a node in a weighted network. Moreover, because it takes the size of the network into account, that measure makes nodes from different networks comparable. When considering a node, one difference between
the w-lobby index and \( h \)-centrality is that the w-lobby index of a node is based on its neighboring nodes’ degree while \( h \)-centrality is based on the number of links and their strength. Some basic properties of \( h \)-centrality are given below.

**Proposition 5** for non-isolated nodes in weighted networks, \( h \)-centrality \((C_h)\) always satisfies the following inequality

\[
0 < \frac{1}{N - 1} \leq C_h \leq \frac{N_a}{N - 1} \leq 1 \quad (16.5)
\]

**Proof** by Proposition 3, we have

\[
\frac{1}{N - 1} \leq \frac{d_h}{N - 1} \leq \frac{N_a}{N - 1} \leq \frac{N - 1}{N - 1}.
\]

As \( \frac{d_h}{N - 1} = C_h \), Equation (16.5) follows immediately.

**Proposition 6** if a node has node strength \( d \) and \( h \)-degree \( d_h \) then the following inequality holds

\[
C_h \leq \frac{\sqrt{d}}{N - 1} \quad (16.6)
\]

**Proof** by Proposition 4 we know that \( d_h \leq \sqrt{d} \). Dividing by \( N - 1 \) yields Eq. (16.6).

### 16.4.2 \( h \)-Centralization for Whole Networks

Freeman (1977) observed that there are two meanings for the notion of centrality of a network: it could mean the extent to which all nodes are central, or it could refer to the dominance of a single point. Applying Freeman’s centralization procedure (Freeman 1979; Everett et al. 2004) we take the second approach. Given a vertex centrality index \( F \) a centralization index \( F_1 \) for the whole graph \( G \) with \( N \) nodes is defined as

\[
F_1(G) = \frac{\sum_{i=1}^{N} [\max(G) - F(n_i)]}{\max(N)} \quad (16.7)
\]

where \( \max(G) \) is the maximum value attained by \( F \) in the graph \( G \) and \( \max(N) \) is the maximum value attained by \( F \) in all possible graphs with \( N \) nodes.

Based on this principle we define the \( h \)-centralization of a network based on the \( h \)-degree.
**h-centralization**: in a weighted network $G$ with $N$ nodes, the $h$-degree centralization, $C_h(G)$ of that network is

$$C_h(G) = \frac{\sum_{i=1}^{N} \left[ \max(G) - d_h(n_i) \right]}{(N - 1)(N - 2)}$$ \hspace{1em} (16.8)

The denominator of Equation (16.8) is obtained as follows. The largest possible value for $\max(G)$ is $N - 1$ (which can be reached in a star network). Then there are $N - 1$ nodes with $d_h = 1$, and hence a difference with the largest value of $N - 2$, leading to a denominator equal to $(N - 1)(N - 2)$.

$h$-centralization describes the distribution of weights in a network. A network has a high $h$-centralization if, compared to the node with largest $h$-centrality, the $h$-centrality of the other nodes is low. In those circumstances the links or weights in this high $h$-centralization network are concentrated in the central node. The distribution of weights in that network is unbalanced.

### 16.4.3 Some Theoretical Properties of $h$-Strength

In a weighted network with $N$ nodes and $L$ links, $h_s$ denotes the $h$-strength of network, $d(n)$ and $s(m)$ are the degree (strength) of node $n$ and strength of link $m$, respectively. Then we have the following inequalities

$$0 \leq h_s \leq L \leq \sum_{m=1}^{L} s(m)$$ \hspace{1em} (16.9)

$$h_s \leq \sqrt{\sum_{m=1}^{L} s(m)}$$ \hspace{1em} (16.10)

where $\sum_{m=1}^{L} s(m)$ represents the total strength of links in network. Those inequalities suggest that $h$-strength is limited by the total number of links and the square root of total link strength in a network, i.e., the size and strength of the links.

Because in a weighted network one link strength produces two node degrees, we always have

$$\sum_{m=1}^{L} s(m) = 2 \sum_{n=1}^{N} d(n)$$ \hspace{1em} (16.11)

Thus, from Eqs. (16.9) to (16.11), (16.12) follows immediately
Equation (16.12) indicates that, although there is no direct relationship between \( h \)-strength and the node degree, the total degree still affects the possible maximum value of \( h \)-strength. Therefore, in a network the size and strength of links and nodes might indirectly influence the value of \( h \)-strength.

For the \( h \)-subnet of a weighted network, \( N_b \) and \( L_b \) denote the nodes and links in \( h \)-subnet, respectively; \( e \) is the number of links whose strengths equal to \( h_s \). It is easy to check Eqs. (16.13)–(16.15).

\[
hs + e - 1 = L_b \quad (16.13)
\]

\[
2(hs + e - 1) \geq N_b \quad (16.14)
\]

\[
hs \leq \sqrt{\frac{L_b}{\sum_{m=1}^{N_b} s(m)}} = \sqrt{\frac{2\sum_{n=1}^{N_b} d(n)}} \quad (16.15)
\]

Those equations show that the size of \( h \)-subnet heavily depends on the value of \( h \)-strength. Therefore, \( I_h \)-index is computed for the changes of \( h \)-strength and provides a perspective to represent the dynamic evolution of the network mainstay.

### 16.4.4 Some Theoretical Properties of \( h \)-Crystal

As the \( h \)-crystal consists of core nodes (nodes in the network’s \( h \)-core), core edges (edges in the \( h \)-subnet) in each single layer weighted networks and layer-bridges between two layers, we can derive some of their theoretical properties following the ideas of an \( h \)-degree and \( h \)-strength (Zhao et al. 2011, 2014). Actually, the analytical properties of \( h \)-crystal can be obtained by merging all the layers’ network’s \( h \)-cores and \( h \)-subnets, where network’s \( h \)-cores generated by nodes following \( h \)-degree algorithm and \( h \)-subnets by edges following \( h \)-strength algorithm. However, as there are complex mathematical structure in multi-layer networks (de Domenico, 2013), we have not yet reached general results.

#### 16.4.4.1 The nodes of the \( h \)-crystal

Consider a multi-layer weighted network consisting of three layers A, B and C, with \( N_A \), \( N_B \) and \( N_C \) nodes, and the three layers have the \( h \)-degree \( N_{Ah} \), \( N_{Bh} \) and \( N_{Ch} \), respectively. Suppose that the numbers of nodes that the node’s \( h \)-degree equals \( N_{Ah} \), \( N_{Bh} \) and \( N_{Ch} \) are \( N_{Ad} \), \( N_{Bd} \) and \( N_{Cd} \) respectively. Let \( N_{Ah-cs} \), \( N_{Bh-cs} \) and \( N_{Ch-cs} \)
represent number of nodes in the network’s $h$-core and $h$-subnet, $N_{Ah-s}$, $N_{Bh-s}$ and $N_{Ch-s}$ denote nodes in the $h$-subnet only (except nodes in the network’s $h$-core). If the number of same nodes of $N_{Ah}$ and $N_{Bh}$ is $N_{AB}$ and that of $N_{Bh}$ and $N_{Ch}$ is $N_{BC}$, the total number of nodes in the $h$-crystal ($N_{h-crystal}$) can be calculated by Eqs. (16.16)–(16.21).

\[
N_{h-crystal} = N_{Ah-cs} + N_{Bh-cs} + N_{Ch-cs} - N_{AB} - N_{BC} \tag{16.16}
\]

\[
N_{Ah-cs} = N_{Ah-c} + N_{Ah-s} = N_{Ah} + N_{Ad} - 1 + N_{Ah-s} \tag{16.17}
\]

\[
N_{Bh-cs} = N_{Bh-c} + N_{Bh-s} = N_{Bh} + N_{Bd} - 1 + N_{Bh-s} \tag{16.18}
\]

\[
N_{Ch-cs} = N_{Ch-c} + N_{Ch-s} = N_{Ch} + N_{Cd} - 1 + N_{Ch-s} \tag{16.19}
\]

\[
N_{Ad} \geq 1, N_{Bd} \geq 1, N_{Cd} \geq 1 \tag{16.20}
\]

\[
N_{Ah-s} \geq 0, N_{Bh-s} \geq 0, N_{Ch-s} \geq 0 \tag{16.21}
\]

In our examples, we find that the $N_{h-crystal} = 27 + 38 + 20 - 4 - 0 = 81$ for the “$h$-set” theoretically.

If the total number of nodes $N = N_A + N_B + N_C$ and $N_h = N_{Ah} + N_{Bh} + N_{Ch}$, in which node $i$ has degree $d_i$, with $h$-degree $h_i$, we get

\[
N_{h-crystal} \geq N_h \tag{16.22}
\]

\[
0 \leq N_h \leq N \leq \sum_{i=1}^{N} h_i \leq \sum_{i=1}^{N} d_i \tag{16.23}
\]

\[
N_h \leq \sqrt{\sum_{i=1}^{N} h_i} \leq \sqrt{\sum_{i=1}^{N} d_i} \tag{16.24}
\]

That means that $N_{h-crystal}$ is restricted by $N$ and $N_h$.

16.4.4.2 The Edges of the $h$-Crystal

Suppose that a multi-layer weighted network consists of three layers A, B and C, with number of edges $L_A$, $L_B$ and $L_C$ respectively, each with $h$-strength $L_{Ah}$, $L_{Bh}$ and $L_{Ch}$ and the numbers of edges for which the edge’s $h$-strength equals $L_{Ah}$, $L_{Bh}$ and $L_{Ch}$ are $L_{Ah-c}$, $L_{Bh-c}$ and $L_{Ch-c}$ respectively. Let $L_{Ah-cs}$, $L_{Bh-cs}$ and $L_{Ch-cs}$ represent number of edges in the network’s $h$-core and $h$-subnet. If the number of links of two layers A, B is $L_{AB}$ and that of two layers B, C is $L_{BC}$, the total number of edges of the $h$-crystal ($L_{h-crystal}$) is
\[ L_{h\text{-crystal}} = L_{Ah\text{-cs}} + L_{Bh\text{-cs}} + L_{Ch\text{-cs}} + L_{AB} + L_{BC} \] (16.25)

In our cases, for “h-set”, the \( L_{h\text{-crystal}} = 199 + 263 + 103 + 0 + 207 = 772 \).

If the total number of edges \( L = L_A + L_B + L_C \) and \( L_h = L_{Ah} + L_{Bh} + L_{Ch} \), with weights \( s_j \) in edge \( j \) and \( L \leq N \cdot (N - 1) \), we obtain

\[ 0 \leq L_h \leq L \leq \sum_{j=1}^{L} s_j \] (16.26)

\[ L_h \leq \sqrt{\frac{\sum_{j=1}^{L} s_j}{\sum_{i=1}^{N} d_i}} \] (16.27)

where \( \sum_{j=1}^{L} s_j \) is the sum of the total weight of all edges in the network and \( L \) indicates total edges.

It is valuable to point out that our results show general ways to reach the core structure of both multi-layer and heterogeneous networks, as the \( h \)-crystal covers heterogeneous structure (where co-keyword is different from co-citation and bibliographic coupling), though each of the three layers belong to homogeneous layer.

### 16.4.5 Limitations

In this chapter, we defined \( h \)-type measures and constructed core structures in weighted networks, including both single-layer and multi-layer networks. However, the methodology is not suitable for unweighted networks. In some network studies, the numerical scale of links’ weight is not suitable for ranking, for instance when weights belong to the interval \([0, 1]\) and certainly when weights can be negative. At that point, a numerical transformation is needed (but then scientists should agree on the exact formulation of this transformation). Although there are some limitations, more extensions could be developed in future.

### 16.5 Conclusion

In summary, we introduced \( h \)-type measures for extracting core structures of weighted information networks. For single-layer weighted networks, we suggest \( h \)-degree, \( h \)-strength or \( h \)-bone to realize the objects, where \( h \)-subnet or \( h \)-bone provides a promising method for abstracting and characterizing core sub-networks from the whole complex networks. For multi-layer weighted networks, we propose
to identify the core structure in the complex multi-layer networks, where the $h$-crystal integrates top nodes and links of the networks.

We conclude that the $h$-type measures, including $h$-degree, $h$-strength, $h$-bone, and $h$-crystal as well as $h$-centrality and $h$-centralization, could provide a set of useful indicators for probing into weighted networks.

Other core structures may exist in weighted networks, depending on the operationalization of the notion of a core. Yet, $h$-type cores can be considered basic ones and it will likely be minimum cores and have practical implications. Like all the other $h$-type measures, $h$-subnet, $h$-bone or $h$-crystal will be efficient methods for identifying and selecting objects from large information networks. Although weighted information networks are only one kind of weighted networks, the $h$-type measure methodology can be generalized to other types of networks. Currently, our case study addresses only undirected weighted information networks, leaving directed weighted networks with three-dimensional visualization displays for future investigations.


### 16.6 Appendix Pseudo-codes of the $h$-Type Core Algorithm

**Algorithm 1 generate bibliographic coupling network**

**Input:** the record data of WoS  
**Output:** bibliographic coupling network and the node information  
**Initialize:** the empty bibliographic coupling network  
**Initialize:** the sparse matrix $M$

for each record $r$ in the record data of WoS do
  for each reference $c$ in the record $r$ do
    $M$ ← add the ID of the record $r$, the ID of the reference $c$ and 1  
    the node information ← add the node information of the record
  end for
end for  
$M$ ← $M * M^T$

for each nonzero element $a_{ij}$ in $M$ do
  bibliographic coupling network ← add $i$, $j$, $a_{ij}$ into the Graph
end for

Return bibliographic coupling network, the node information
Algorithm 2 generate co-citation network

Input: the record data of WoS
Output: co-citation network and the node information
Initialize the empty co-citation network
for each record in the record data of WoS do
    for the reference $c_i$ in the references of the record do
        for the reference $c_j$ in the references of the record do
            if $i < j$ then
                if the reference $c_i$ and the reference $c_j$ in the co-citation network then
                    the weight of edge between the reference $c_i$ and the reference $c_j$< -1 else if then
                        co-citation network ← add the ID of reference $c_i$, the ID of the reference $c_j$ and 1
                        the node information ← add the node information of the reference
                end if
            end if
        end for
    end for
end for
Return co-citation network, the node information

Algorithm 3 generate co-term network

Input: the content of ID and DE field in the record data of WoS
Output: co-term network and the node information
Initialize the empty co-term network
for each record in the record data of WoS do
    for the term $t_i$ in the content of ID and DE field of the record do
        for the term $t_j$ in the content of ID and DE field of the record do
            if $i < j$ then
                if the term $t_i$ and the term $t_j$ in the co-term network then
                    then the weight of edge about the term $t_i$ and the term $t_j$< -1 else if then
                        co-term network ← add the ID of term $t_i$, the ID of the term $t_j$ and 1
                        the node information ← add the node information of the term
                end if
            end if
        end for
    end for
end for
Return co-term network, the node information

Algorithm 4 compute h-core and h-subnet of each layer network

Input: G(bibliographic coupling network/co-citation network/co-term network)
Output: h-core and h-subnet of the corresponding network
Initialize the empty h-core and h-subnet of G
for each node $n$ in G do
    h-degree of the node ← $n$ has at least $h(n)$ links each with strength at least equal to $h(n)$
end for
h degree of the G ← h-index of h-degree of every node in G
for each node $n$ in G do
    if h-degree of node $n$ >= h-degree of the G then
        h core of G ← add node $n$ into h-core of G, meanwhile add the link
if node n have a link with other nodes in h-core of G
    end if
end for
h strength of the G ← h-index of the strength of every edge in G
for each edge e in G do
    if the strength of edge e ≥ h-strength of the G then
        h subnet of G ← add edge e into h-subnet of G
    end if
end for
h-core and h-subnet of G ← combine h-core of G with h-subnet of G
Return h-core and h-subnet of G

Algorithm 5 construct bridge relations between bibliographic coupling network and co-citation network

Input: h-core and h-subnet of bibliographic coupling network, h-core and h-subnet of co-citation network
Output: bridge relations between bibliographic coupling network and co-citation network
for each node n_i in h-core and h-subnet of bibliographic coupling network do
    for each node n_j in h-core and h-subnet of co-citation network do
        if n_i equal n_j then
            bridge relations ← construct a link between node n_i and node n_j
        end if
    end for
end for
Return bridge relations between bibliographic coupling network and co-citation network

Algorithm 6 construct bridge relations between bibliographic coupling network and co-term network

Input: h-core and h-subnet of bibliographic coupling network, h-core and h-subnet of co-term network
Output: bridge relations between bibliographic coupling network and co-term network
for each node n_i in h-core and h-subnet of bibliographic coupling network do
    for each node n_j in h-core and h-subnet of co-term network do
        if n_i associated with n_j then
            bridge relations ← construct a link between node n_i and node n_j
        end if
    end for
end for
Return bridge relations between bibliographic coupling network and co-term network

Algorithm 7 construct h-crystal of multi-layer weighted networks

Input: h-core and h-subnet of co-citation network, h-core and h-subnet of bibliographic coupling network, h-core and h-subnet of co-term network, bridge relations between bibliographic coupling network and co-citation network, bridge relations between bibliographic coupling network and co-term network
Output: h-crystal of multi-layer weighted networks
Initialize the empty graph(h-crystal of multi-layer weighted networks)
h-crystal of multi-layer weighted networks ← combine(h-core and h-subnet of co-citation network, h-core and h-subnet of bibliographic coupling network, h-core and h-subnet of co-term network, bridge relations between bibliographic coupling network and co-citation network, bridge relations between bibliographic coupling network and co-term network)
if h-crystal of multi-layer weighted networks is connected then
   h-crystal of multi-layer weighted networks is existing
end if
Return h-crystal of multi-layer weighted networks

References


Rousseau, R., Ye, F.Y.: Subgraphs derived from the h-core in undirected, unweighted networks. ISSI Newslett. 7(1), 5–9 (2011)


Ye, F.Y.: An investigation on mathematical models of the \( h \)-index. Scientometrics 81(2), 493–498 (2009)
Chapter 17
The $h$-Core and $h$-Tail Distribution with Dynamic Metrics

In the ranked publication and citation distribution, $h$-core and $h$-tail are naturally generated via $h$-index, on which related dynamic core-tail measures are introduced, including shape descriptors and shape centroids, $k$-index and $k'$-index. With practical data in the fields of physics and sociology, the dynamic core-tail measures are discussed, which reveals that there are obvious differences between sciences (physics) and social sciences (sociology) when $c$-descriptor, $h$-core centroid and $k$-index are applied as dynamic core-tail measures, while a few differences exist by using $t$-descriptor, $h$-tail centroid and $k'$-index. Moreover, the academic matrices, academic vectors and academic tensor are suggested to be overall measures for comparative academic evaluation.

17.1 Introduction

Since $h$-index was introduced in 2005 (Hirsch 2005), it had quickly been applied as an academic measure (Alonso et al. 2009; Egghe 2010) and had led to a simple and meaningful unification of publications and citations (Ye 2009; Ye 2011). Then, some $h$-type indices have been introduced for improving $h$-index (Egghe 2006; Glänzel 2006; Jin et al. 2007).

At the beginning, when Hirsch (2005, 2007) proposed $h$-index, he had found that there were differences among different fields, with ten cases in physics and biology respectively. Then, many scholars revealed the informetric differences in various fields (Batista et al. 2006; Schubert and Glänzel 2007; Iglesias and Pecharroman 2007; Lillquist and Green 2010). Among them, Batista et al. (2006) gave $h_1$-index for correcting $h$-index in physics, chemistry, biology/biomedical and mathematics and Iglesias and Pecharroman (2007) compared different fields with a simple method for scaling $h$-index so that we could compare $h$-indices across fields. When Liang (2006) introduced $h$-index sequence and $h$-index matrix for overcoming the faults of $h$-index in special time span, Rousseau and Ye (2008) also proposed dynamic $h$-type index for...
measuring the dynamic nonlinear properties. All the studies show that scholars have paid attention to dynamic process of $h$-index.

As single $h$-index ignores citation distribution (Vanclay 2007) and there is $h$-inconsistency at some cases (Waltman and van Eck 2012), so that it is necessary to notice the distribution of $h$-core and $h$-tail (Rousseau 2006; Ye and Rousseau 2010), for academic comparison. Then, Kuan et al. (2011a) suggested two indicators, $c$-descriptor and $t$-descriptor, for analyzing patent performance of assignees, according to their rank-citation curves based on practical data. As $c$-descriptor and $t$-descriptor can not instantly show relative patent performance of all assignees, Kuan et al. (2011b) introduced $h$-core centroid and $h$-tail centroid, which were respectively located at the geometric centers of $h$-core area and $h$-tail area following the rank-citation curves, as two indicators.

Therefore, based on studies above, we try to probe into the dynamic core-tail measures for $h$-core and $h$-tail, with indicators such as $c$-descriptor, $t$-descriptor, $h$-core centroid ($c_x, c_y$) and $h$-tail centroid ($t_x, t_y$), as well as $k$ and $k'$, on the basis of data in time span from 1 year to 10 years during 2001–2010, for revealing the dynamic changes of publications, citations and rank-citation curves in different fields (e.g., physics and sociology). Meanwhile, combining core-tail distribution with I3 (integrated impact indicators) (Leydesdorff and Bornman 2011; Bornman et al. 2013), academic matrices and academic tensor are suggested as synthetic and overall measures for academic comparison.

### 17.2 Methodology

The method is based on ranked publication-citation distribution, which means that publications $P$ and citations $C$ could be re-arranged into a diagram when they were ranked according to total citations of each publication from high to low, where $h$-index always locates on $C$-$P$ curve and forms the $h$-core and $h$-tail distribute in the $C$-$P$ plane, as shown in Fig. 17.1.

Two kinds of informetric measures can be introduced as follows.
17.2 Methodology

17.2.1 Dynamic Core-Tail Measures

We are interested in the portion of $h$-core and $h$-tail and the difference between science and social science, for which $e$-index (Zhang 2009) and $k$-index (Ye and Rousseau 2010) could be mentioned as

$$C_H = h^2 + e^2$$  \hspace{1cm} (17.1)

$$k = \frac{C/P}{C_T/C_H}$$  \hspace{1cm} (17.2)

where $C_H$ and $C_T$ denotes respectively citations in $h$-core and in $h$-tail. Since $C/P$ as average impact is also problematic (Rousseau and Leydesdorff 2011) and $C-P$ is consistent logically (Rousseau and Ye 2011), we can modify $k$ as $k'$ for measuring core-tail ratio with same decreasing tendency as follows.

$$k' = \frac{C - P}{C_T - C_H}$$  \hspace{1cm} (17.3)

Using shape descriptors (both $c$-descriptor and $t$-descriptor) proposed by Kuan et al. (2011a), we have

$$c\text{-descriptor} = \frac{\sum_{i=1}^{h} C(P_i)^2}{\sum_{i=1}^{h} C(P_i)}$$  \hspace{1cm} (17.4)

$$t\text{-descriptor} = \frac{\sum_{i=h+1}^{Nc} iC(P_i)^2}{\sum_{i=h+1}^{Nc} C(P_i)}$$  \hspace{1cm} (17.5)

where $Nc$ stands for the number of cited publications (with at least one citation). Meanwhile, shape centroids, including $h$-core centroid ($c_x, c_y$) and $h$-tail centroid ($t_x, t_y$), could be obtained (Kuan et al. 2011b) as follows.

$$c_x = \frac{\sum_{i=1}^{h} (i - 0.5)C(P_i)}{\sum_{i=1}^{h} C(P_i)}$$  \hspace{1cm} (17.6)
\[
\begin{align*}
\text{c}_y &= \frac{1}{2} \sum_{i=1}^{h} C(P_i)^2 = \frac{1}{2} \text{c-descriptor} \\
\text{t}_x &= \frac{\sum_{i=h+1}^{Nc} (i - 0.5) C(P_i)}{\sum_{i=h+1}^{Nc} C(P_i)} = \text{t-descriptor} - 0.5 \\
\text{t}_y &= \frac{1}{2} \sum_{i=h+1}^{Nc} C(P_i)^2 \\
\end{align*}
\]

Since the \( c_y \) and \( t_y \) are same mathematical form, we can only apply \( c_x \) and \( t_x \) as the independent measures of \( h \)-core centroid and \( h \)-tail centroid respectively.

### 17.2.2 Distributed Core-Tail Measures

Following the definition of I3 (Leydesdorff and Bornmann 2011; Leydesdorff et al. 2011; Rousseau and Ye 2012), one can formalize I3 as

\[
\text{I3} = \sum_{i=1}^{C} f(X_i) \cdot X_i \tag{17.10}
\]

where \( X_i \) is the percentile ranks and \( f(X_i) \) is the frequencies of the ranks with \( i = [1, C] \) as the percentile rank classes. In other words, the measures \( X_i \) are divided into \( C \) classes each with a scoring function \( f(X_i) \) or weight \( w_i \) so that one can aggregate as follows (Rousseau 2012, 2013).

\[
\text{I3} = \sum_{i} w_i X_i \tag{17.11}
\]

More generally, when one ranks publications according to their citations from high to low, one obtains a \( C-P \) rank distribution—the citation curve—as shown in Fig. 17.1. We added the three sections that are relevant for the \( h \)-index: the \( h \)-core, \( h \)-tail and the uncited (zero citations) publications \( P_z \) respectively (Ye and Rousseau...
2010; Chen et al. 2013; Liu et al. 2013). Furthermore, Zhang (2009) proposed to call the area above the h-core a representation of “excess citations”, that is, citations which are gathered, but do not further contribute to the h-value.

Let \( P = P_c + P_t + P_z \) be the total number of publications and \( C = C_c + C_t + C_e \) be the total number of citations, while \( P_c \) denotes the number of publications in the h-core, \( P_t \) the number of publications in the h-tail, and \( P_z \) the number of uncited (zero citation) publications, and \( C_h = C_c + C_e \) indicates the total number of citations in the h-core, \( C_t \) the number of citations in the h-tail, and \( C_e \) the number of citations in the excess area (Zhang 2009, 2013). In order to account for the full set of three classes of publications and three classes of citations, Ye and Leydesdorff (2014) proposed to fill the performance matrices (main matrix \( M_1 \) and associate matrix \( M_2 \)):

\[
M_1 = \begin{pmatrix}
Y_1 & Y_2 & Y_3 \\
X_1 & X_2 & X_3 \\
Z_1 & Z_2 & Z_3
\end{pmatrix} = \begin{pmatrix}
Y \\
X \\
Z
\end{pmatrix} = (Y X Z)^T \tag{17.12}
\]

\[
M_2 = \begin{pmatrix}
X_1 & X_2 & X_3 \\
Y_1 & Y_2 & Y_3 \\
Z_1 & Z_2 & Z_3
\end{pmatrix} = \begin{pmatrix}
X \\
Y \\
Z
\end{pmatrix} = (X Y Z)^T \tag{17.13}
\]

where vectors \( X = (X_1, X_2, X_3) \) and \( Y = (Y_1, Y_2, Y_3) \) are applied as normalized proportions of publications and citations, in which each element is I3-type measure:

\[
X_1 = \frac{P_c}{P_c + P_t + P_z} \cdot P_c \tag{17.14}
\]

\[
X_2 = \frac{P_t}{P_c + P_t + P_z} \cdot P_t \tag{17.15}
\]

\[
X_3 = \frac{P_z}{P_c + P_t + P_z} \cdot P_z \tag{17.16}
\]

\[
Y_1 = \frac{C_c}{C_c + C_t + C_e} \cdot C_c \tag{17.17}
\]

\[
Y_2 = \frac{C_t}{C_c + C_t + C_e} \cdot C_t \tag{17.18}
\]

\[
Y_3 = \frac{C_e}{C_c + C_t + C_e} \cdot C_e \tag{17.19}
\]

Then an additional vector \( Z \) can be defined in terms of the difference between the total number of citations and the total number of patents for each area:

\[
Z = (Z_1, Z_2, Z_3) = (Y_1 - X_1, Y_2 - X_2, Y_3 - X_3) \tag{17.20}
\]
The terms of $Z$ can be appreciated as the fraction of citations (with the dimensionality of citation) minus the fractions of publications (with the dimensionality of publication), so that $Z$ is a set of meaningful indicators, where $C_i$ can have the values of $C_c$, $C_t$ and $C_e$, and $P_i$ the values $P_c$, $P_t$ and $P_z$. $Z_3$ is a complex indicator because one considers the excess citations as a possible compensation for the uncited publications. The fraction of uncited publications contributes negatively to $Z_3$, but this can be compensated by the fraction excess citations in a set. $Z_3 = (C_e^2 / C) - (P_z^2 / P)$ can also be negative when the number of uncited publications is larger than the sum of the excess citations. In order to avoid the problems of understanding a negative value for the performance (in the case of many uncited publications), one can define a sub-matrix $SM$ without the third vector $Z_3$, as follows:

$$SM = \begin{pmatrix} C_c^2 / C & C_t^2 / C \\ P_c^2 / P & P_t^2 / P \end{pmatrix}$$ (17.21)

The main matrix $M_1$, associate matrix $M_2$ and sub-matrix $SM$ contribute as academic matrices for measuring distributed data in $C$-$P$ plane, leading to synthetic overall measures.

In Matrix (17.12), Matrix (17.13) and Matrix (17.21), the elements of main diagonals are typical indicators for measuring academic performance, which characterized $h$-core and $h$-tail distribution of an academic subjects or objects, so that we can define following academic vectors for applications (Prathap 2014).

Primary vector (or main vector): the diagonals vector of main matrix $V_1$ is called as primary vector (or main vector) as

$$V_1 = (Y_1, X_2, Z_3) = \left( \frac{C_c^2}{C}, \frac{P_t^2}{P}, \left( \frac{C_e^2}{C} - \frac{P_z^2}{P} \right) \right)$$ (17.22)

Secondary vector (or associate vector): the diagonals vector of associate matrix $V_2$ is called as secondary vector (or associate vector) as

$$V_2 = (X_1, Y_2, Z_3) = \left( \frac{P_c^2}{P}, \frac{C_t^2}{C}, \left( \frac{C_e^2}{C} - \frac{P_z^2}{P} \right) \right)$$ (17.23)

Sub-vector: the diagonals vector of sub-matrix $SV$ is called as sub-vector as

$$SV = \left( \frac{C_h^2}{C}, \frac{P_t^2}{P} \right)$$ (17.24)

If we need one indicator to measure the academic performance, traces of the academic matrices could be applied as follows.
Primary trace (or main trace): the trace of main matrix $T_1$ is called as primary trace (or main trace) as

$$T_1 = \text{tr}(M_1) = \frac{C^2_c}{C} + \frac{P^2_t}{P} + \left( \frac{C^2_c}{C} - \frac{P^2_c}{P} \right)$$  \hspace{1cm} (17.25)

Secondary trace (or associate trace): the trace of associate matrix $T_2$ is called as secondary trace (or associate trace) as

$$T_2 = \text{tr}(M_2) = \frac{P^2_c}{P} + \frac{C^2_c}{C} + \left( \frac{C^2_c}{C} - \frac{P^2_c}{P} \right)$$  \hspace{1cm} (17.26)

Sub-trace: the trace of sub-matrix $ST$ is called as sub-trace as

$$ST = \text{tr}(SM) = \frac{C^2_h}{C} + \frac{P^2_t}{P}$$  \hspace{1cm} (17.27)

Both $T_1$ and $T_2$ summarize the representative information distributed over the $e$-area, the $h$-area, the $t$-area and the uncited area. $ST$ ignores uncited publications and focuses on the cited ones in terms of the core-tail distribution, so that the sub-trace will always be positive since $C > C_h > 0$ and $P > P_t > 0$.

More generally, academic tensor $T_{\mu\nu}$ can be defined as an extension of academic matrices as

$$T_{\mu\nu} = \{ P_{\mu\nu}, C_{\mu\nu} \}$$  \hspace{1cm} (17.28)

where $\{ P_{\mu\nu}, C_{\mu\nu} \}$ represent all possible combinations of publications and citations. When we compare two academic subjects or objects A and B, if all elements in the $T_{\mu\nu}$ of A are preceded more than all elements in the $T_{\mu\nu}$ of B ($T_{A\mu\nu} \succ T_{B\mu\nu}$), we can say that A is superior to B.

### 17.3 Empirical Study

For numerical comparison, we searched sample data from *Journal Citation Indicator* (JCR) 2010 in the fields of physical and sociological journals.

In physics, data are collected from the JCR category “Physics, particles and fields” with 27 journals, while sociology collected from the category “Sociology” with 132 journals. All changes of journal names are considered and changed journals are also searched.

Then we use the journal titles to search the data from Science Citation Index (SCI) and Social Science Citation Index (SSCI) respectively, including all document types from 2001–2010.

Check Table 17.1 for total data, in which Jn means number of journals.
Table 17.1 Data sets

<table>
<thead>
<tr>
<th>Field</th>
<th>Jn</th>
<th>P</th>
<th>C</th>
<th>h</th>
<th>CH</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physics</td>
<td>27</td>
<td>91,252</td>
<td>1,146,184</td>
<td>219</td>
<td>84,775</td>
<td>1,061,409</td>
</tr>
<tr>
<td>Sociology</td>
<td>132</td>
<td>57,718</td>
<td>209,443</td>
<td>103</td>
<td>17,953</td>
<td>191,490</td>
</tr>
</tbody>
</table>

Table 17.2 Data on time stages in physics

<table>
<thead>
<tr>
<th>Physics</th>
<th>Jn</th>
<th>P</th>
<th>C</th>
<th>h</th>
<th>CH</th>
<th>CT</th>
<th>c-max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001–2001</td>
<td>19</td>
<td>7,754</td>
<td>146,312</td>
<td>132</td>
<td>30,220</td>
<td>116,092</td>
<td>575</td>
</tr>
<tr>
<td>2001–2002</td>
<td>20</td>
<td>15,182</td>
<td>296,897</td>
<td>163</td>
<td>49,109</td>
<td>247,788</td>
<td>2,673</td>
</tr>
<tr>
<td>2001–2003</td>
<td>21</td>
<td>23,189</td>
<td>442,056</td>
<td>185</td>
<td>63,140</td>
<td>378,916</td>
<td>2,673</td>
</tr>
<tr>
<td>2001–2004</td>
<td>22</td>
<td>32,554</td>
<td>589,053</td>
<td>201</td>
<td>71,378</td>
<td>517,675</td>
<td>2,673</td>
</tr>
<tr>
<td>2001–2007</td>
<td>26</td>
<td>60,389</td>
<td>942,261</td>
<td>216</td>
<td>81,807</td>
<td>860,454</td>
<td>3,545</td>
</tr>
<tr>
<td>2001–2008</td>
<td>27</td>
<td>70,314</td>
<td>1,036,031</td>
<td>218</td>
<td>82,941</td>
<td>953,090</td>
<td>3,545</td>
</tr>
<tr>
<td>2001–2009</td>
<td>27</td>
<td>80,808</td>
<td>1,106,150</td>
<td>219</td>
<td>83,674</td>
<td>1,022,476</td>
<td>3,545</td>
</tr>
<tr>
<td>2001–2010</td>
<td>27</td>
<td>91,252</td>
<td>1,146,184</td>
<td>219</td>
<td>84,775</td>
<td>1,061,409</td>
<td>3,545</td>
</tr>
</tbody>
</table>

Table 17.3 Data on time stages in sociology

<table>
<thead>
<tr>
<th>Sociology</th>
<th>Jn</th>
<th>P</th>
<th>C</th>
<th>h</th>
<th>CH</th>
<th>CT</th>
<th>c-max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001–2001</td>
<td>87</td>
<td>5,043</td>
<td>35,150</td>
<td>70</td>
<td>9,799</td>
<td>25,351</td>
<td>921</td>
</tr>
<tr>
<td>2001–2002</td>
<td>87</td>
<td>9,965</td>
<td>66,254</td>
<td>82</td>
<td>13,314</td>
<td>52,940</td>
<td>921</td>
</tr>
<tr>
<td>2001–2003</td>
<td>92</td>
<td>15,040</td>
<td>95,922</td>
<td>93</td>
<td>15,053</td>
<td>80,869</td>
<td>921</td>
</tr>
<tr>
<td>2001–2005</td>
<td>94</td>
<td>25,399</td>
<td>149,378</td>
<td>101</td>
<td>17,800</td>
<td>131,578</td>
<td>921</td>
</tr>
<tr>
<td>2001–2006</td>
<td>97</td>
<td>30,918</td>
<td>169,568</td>
<td>102</td>
<td>18,030</td>
<td>151,538</td>
<td>921</td>
</tr>
<tr>
<td>2001–2007</td>
<td>110</td>
<td>37,018</td>
<td>186,374</td>
<td>103</td>
<td>17,953</td>
<td>168,421</td>
<td>921</td>
</tr>
<tr>
<td>2001–2008</td>
<td>129</td>
<td>43,895</td>
<td>198,995</td>
<td>103</td>
<td>17,953</td>
<td>181,042</td>
<td>921</td>
</tr>
<tr>
<td>2001–2009</td>
<td>129</td>
<td>50,734</td>
<td>206,361</td>
<td>103</td>
<td>17,953</td>
<td>188,408</td>
<td>921</td>
</tr>
<tr>
<td>2001–2010</td>
<td>132</td>
<td>57,718</td>
<td>209,443</td>
<td>103</td>
<td>17,953</td>
<td>191,490</td>
<td>921</td>
</tr>
</tbody>
</table>

We see that there are 27 journals in the field of physics (particles & fields) with 91,252 papers, 1,146,184 citations and 219 $h$, while there are 132 journals in the field of sociology with 57,718 papers, 209,443 citations and 103 $h$, during 2001–2010.

For comparing the changes of data, we also collected data according to time stages on 2001–2001 to 2001–2010, as shown in Tables 17.2 and 17.3.

In Tables 17.2 and 17.3, c-max denotes maximum citations of the most cited paper.

For comparing academic matrices, we also collect data at various levels, in which we only provide few examples for comparison: JASIS (Journal of the American
17.3 Empirical Study

Table 17.4: Results with core-tail measures on total data

<table>
<thead>
<tr>
<th>Field</th>
<th>Shape centroids</th>
<th>Shape descriptors</th>
<th>$k$ and $k'$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$c_x$</td>
<td>$c_y$</td>
<td>$t_x$</td>
</tr>
<tr>
<td>Physics</td>
<td>79.35</td>
<td>349.45</td>
<td>14,963.60</td>
</tr>
<tr>
<td>Sociology</td>
<td>37.85</td>
<td>129.08</td>
<td>4,873.73</td>
</tr>
</tbody>
</table>

Table 17.5: Results with core-tail measures in physics on time stages

<table>
<thead>
<tr>
<th>Physics</th>
<th>Shape centroids</th>
<th>Shape descriptors</th>
<th>$k$ and $k'$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$c_x$</td>
<td>$c_y$</td>
<td>$t_x$</td>
</tr>
<tr>
<td>2001–2001</td>
<td>51.08</td>
<td>134.91</td>
<td>1,535.43</td>
</tr>
<tr>
<td>2001–2002</td>
<td>59.20</td>
<td>240.20</td>
<td>2,921.75</td>
</tr>
<tr>
<td>2001–2003</td>
<td>66.94</td>
<td>287.60</td>
<td>4,315.29</td>
</tr>
<tr>
<td>2001–2004</td>
<td>74.11</td>
<td>286.99</td>
<td>5,905.98</td>
</tr>
<tr>
<td>2001–2005</td>
<td>78.07</td>
<td>283.87</td>
<td>7,470.65</td>
</tr>
<tr>
<td>2001–2006</td>
<td>76.97</td>
<td>351.96</td>
<td>9,039.15</td>
</tr>
<tr>
<td>2001–2007</td>
<td>78.12</td>
<td>349.45</td>
<td>10,665.88</td>
</tr>
<tr>
<td>2001–2008</td>
<td>79.21</td>
<td>347.73</td>
<td>12,258.53</td>
</tr>
<tr>
<td>2001–2009</td>
<td>79.82</td>
<td>346.80</td>
<td>13,787.20</td>
</tr>
<tr>
<td>2001–2010</td>
<td>79.35</td>
<td>347.45</td>
<td>14,963.60</td>
</tr>
</tbody>
</table>

According to data in Table 17.1, there are 91,252 papers published on 27 journals in the field of physics (particles and fields), while there are 57,718 papers published on 132 journals in the field of sociology, during 2001–2010. Totally, number of citations in physics is about 5.5 times as many as that in sociology, with number of papers as about 1.6 times and journals only about 1/5. If each journal has its special focus, we see that sociology has more focuses than physics.

According to data in Tables 17.2 and 17.3, we see that the papers in both physics and sociology increase by increasing of time span. However, when the time span is larger than 5 years, the increment becomes fewer and fewer. $h$-index also becomes stable when the time span is larger than 7 years in the field of sociology or 9 years in the field of physics, which fits the dynamic $h$-index model proposed by Egghe (2007).

On the basis of data above, we compute shape centroids, shape descriptors and the core-tail radio $k$-indice and $k'$-indice, and the results are merged into Tables 17.4, 17.5 and 17.6.

From Tables 17.4 and 17.5, when we synthesize the information of $h$-index, $c$-descriptor and $h$-core centroid $(c_x, c_y)$, it is shown that there are differences between
Table 17.6 Results with core-tail measures in sociology on time stages

<table>
<thead>
<tr>
<th>Sociology</th>
<th>Shape centroids</th>
<th>Shape descriptors</th>
<th>( k ) and ( k' )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( c_x )</td>
<td>( c_y )</td>
<td>( t_x )</td>
</tr>
<tr>
<td>2001–2001</td>
<td>23.42</td>
<td>127.450</td>
<td>554.05</td>
</tr>
<tr>
<td>2001–2002</td>
<td>28.52</td>
<td>130.515</td>
<td>1,037.20</td>
</tr>
<tr>
<td>2001–2003</td>
<td>33.98</td>
<td>125.735</td>
<td>1,529.22</td>
</tr>
<tr>
<td>2001–2004</td>
<td>36.00</td>
<td>129.025</td>
<td>2,031.28</td>
</tr>
<tr>
<td>2001–2005</td>
<td>36.93</td>
<td>129.040</td>
<td>2,560.26</td>
</tr>
<tr>
<td>2001–2006</td>
<td>37.43</td>
<td>128.655</td>
<td>3,115.09</td>
</tr>
<tr>
<td>2001–2007</td>
<td>37.85</td>
<td>129.080</td>
<td>3,679.75</td>
</tr>
<tr>
<td>2001–2008</td>
<td>37.85</td>
<td>129.080</td>
<td>4,244.13</td>
</tr>
<tr>
<td>2001–2009</td>
<td>37.85</td>
<td>129.080</td>
<td>4,658.91</td>
</tr>
<tr>
<td>2001–2010</td>
<td>37.85</td>
<td>129.080</td>
<td>4,873.73</td>
</tr>
</tbody>
</table>

Fig. 17.2 The changes of \( k \)-indice and \( k' \)-indice on time stages in physics and sociology. In physics, its \( h \)-index becomes stable when time span is larger than 9 years, but its \( c \)-descriptor and \( h \)-core centroid \( (c_x, c_y) \) continue to change a few. In sociology, its \( h \)-index becomes stable when time span is larger than 7 years, while its \( c \)-descriptor and \( h \)-core centroid \( (c_x, c_y) \) are also stable, without change. That is a sign of the differences between sciences and social sciences.

Results above also show the decreasing tendency of \( k \)-indice and \( k' \)-indice, which is the same to revealed cases (Ye and Rousseau 2010), as shown in Fig. 17.2.

In both physics and sociology, decreasing tendency of \( k \)-indice and \( k' \)-indice keeps almost the same pattern, in which the ratio of \( k \)-indice (physics)/\( k \)-indice (sociology) increases when time span increases, from 1.8 to 2.9, while their \( k' \)-indice keeps almost the same, with few changes.

For probing into the evolutional reasons of different \( k \)-indice and \( k' \)-indice, we try to draw the evolutional figures of \( C/P \) and \( C_T/C_H \), \( C-P \) and \( C_T-C_H \), as shown in Figs. 17.3 and 17.4, in which we see that \( C/P \) in physics is as large as 3 times in sociology while their \( C_T/C_H \) increases from 3 to 10 times when time span increases from 1 to 10 years, so that their \( k \)-indices become changeable. Meanwhile, \( C-P \) and \( C_T-C_H \) keep few changes in every time span, so that \( k' \)-indices also keep few changes.
Fig. 17.3  The changes of $C/P$ and $CT/C_H$ on time stages in physics and sociology

Fig. 17.4  The changes of $C-P$ and $CT-C_H$ on time stages in physics and sociology

For the $h$-core—$h$-tail measures of distributed data, we show following examples of academic matrices:

$$M_{1JASIST} = \begin{pmatrix} 66.56 & 1,190.9 & 40.49 \\ 0.82 & 222.3 & 39.1 \\ 65.73 & 968.61 & 1,388 \end{pmatrix} \quad (17.29)$$

$$M_{1JOI} = \begin{pmatrix} 92.73 & 275.06 & 55.21 \\ 3.09 & 64.04 & 0.24 \\ 89.65 & 211.02 & 54.97 \end{pmatrix} \quad (17.30)$$

$$M_{1Stanford} = \begin{pmatrix} 3,274.72 & 373,369.8 & 2,358.78 \\ 0.922 & 6 & 205,995 & 4,332.12 \\ 3,241.8 & 352,770.3 & -1,973.34 \end{pmatrix} \quad (17.31)$$

$$M_{1Zhejiang} = \begin{pmatrix} 428.32 & 160,340.9 & 223.62 \\ 0.265 & 2 & 16,452.8 & 3,102.16 \\ 428.06 & 143,888.1 & -2,878.54 \end{pmatrix} \quad (17.32)$$
In academic matrices, the academic information of publications and citations for various academic subjects and objects can be concluded.

### 17.4 Analysis and Discussion

Suppose the $C-P$ curve is a continuous function $C(x)$, where $x$ denotes the publications ranked by citations. We can write the analytical definitions of shape descriptors and shape centroids as follows.

\[
c - \text{descriptor} = \frac{\int_1^h C^2(x) \, dx}{\int_1^h C(x) \, dx}
\]

\[
t - \text{descriptor} = \frac{\int_{h+1}^{N_c} x C^2(x) \, dx}{\int_{h+1}^{N_c} C(x) \, dx}
\]

\[
c_x = \frac{\int_1^h (x - 0.5) C(x) \, dx}{\int_1^h C(x) \, dx}
\]

\[
c_y = \frac{1}{2} \frac{\int_1^h C^2(x) \, dx}{\int_1^h C(x) \, dx}
\]

\[
t_x = \frac{\int_{h+1}^{N_c} (x - 0.5) C(x) \, dx}{\int_{h+1}^{N_c} C(x) \, dx}
\]
The moving changes of $C$-curve and shape centroids looks like Fig. 17.5. Although definitions above never contain time as a factor, the changes of shape centroids reflected the evolution. Following time stages, $h$-index will increase and $h$-core and $h$-tail will be larger and larger so that shape centroids will move along up-right direction.

However, the moving situation may become complex. Because another coordinate in $h$-core centroid ($c_x$, $c_y$) or $h$-tail centroid ($t_x$, $t_y$) could also be changeable. Let us divide two parts, $h$-core and $h$-tail, for analysis and discussion of dynamic evolution.

### 17.4.1 The Evolution of $h$-Core

For characterizing $h$-core, we apply its $c$-max, $c$-descriptor and $h$-core centroid ($c_x$, $c_y$) as dynamic core-tail measures. For example, in the field of physics, $c$-max underwent 3 changes, from 575 at time span as 1 year, via 2,673 during time span as 2 to 5 years, to 3,545 when time span became 6 to 10 years, while $c$-descriptor and $h$-core centroid ($c_x$, $c_y$) underwent 2 stage changes, quick increase and almost stop, as shown in Fig. 17.6. It is valuable to mention the phenomenon that $h$-core centroid ($c_x$, $c_y$) moves to right direction when $c$-descriptor stops. In sociology, $c$-max always keeps at 921 and $c$-descriptor maintains no changes, while $h$-core centroid ($c_x$, $c_y$) also moves to right direction.
17.4.2 *The Evolution of h-Tail*

Although the fields of physics and sociology are so different from each other, their evolitional $h$-tails show same pattern. Using the $t$-descriptor and $h$-tail centroid ($t_x$, $t_y$) as dynamic core-tail measures, their $t$-descriptors increase stably while their $h$-tail centroid ($t_x$, $t_y$) move to up-right direction during time span from 1 to 4 years and move to down-right direction during time span from 5 to 10 years, as shown in Fig. 17.7.

Evolutional measures of $h$-core and $h$-tail above reveal part of interesting characteristics in both sciences (physics) and social sciences (sociology).

Meanwhile, for academic matrices with their traces, we suppose that the $C$-$P$ rank distribution is a continuous function $C(x)$, where $x$ denotes the publications ranked by citations. We can then generalize $C_e$, $C_c$ and $C_t$ as follows ($P_c = h$).

$$C_h = C_e + C_c = \int_1^h C(x)\,dx$$ \hspace{1cm} (17.41)

Fig. 17.7 The changes of $h$-tail centroid ($t_x$, $t_y$) on time stages
\[ C_t = \int_{h+1}^{P} C(x) \, dx \]  

(17.42)

\( C_h \) and \( C_t \) determine the shape of \( C(x) \).

If Heaps’ law or Herdan’s law exists between \( P \) and \( C \) (Egghe 2007), we may suppose

\[
C = aP^\beta \text{ or } P = bC^{-\beta}; \quad b = 1/a
\]

(17.43)

where \( a, b \) and \( \beta \) are constants. As \( e, h \) and \( t \) are certainly determined by the same citation distribution (Rousseau 2013), the citations \( C \) can be ordered following a power-law as follows:

\[
C(r) = cr^{-\alpha}
\]

(17.44)

where \( r \) is the rank of the publications, \( \alpha (\neq 1) \) is the Lotka exponent in Lotkaian informetrics (Egghe 2005; Egghe and Rousseau 2006) and \( c \) is a constant. We can derive

\[
C_h = e^2 + h^2 = \int_{h}^{P} C(r) \, dr = \frac{c}{1-\alpha} h^{1-\alpha}
\]

(17.45)

\[
C_t = t^2 = \int_{h}^{P} C(r) \, dr = \frac{c}{1-\alpha} (P^{1-\alpha} - h^{1-\alpha})
\]

(17.46)

By inserting Eqs. (17.42) and (17.43) into Eq. (17.24), we obtain

\[
ST = \left( \frac{c}{1-\alpha} \right)^2 \left[ \frac{h^{2(1-\alpha)}}{C} + \frac{(P^{1-\alpha} - h^{1-\alpha})^2}{P} \right]
\]

(17.47)

Thus, in power-law system, \( ST \) is determined by two parameters \( c \) and \( \alpha \) and three measures: total citations \( C \), total publications \( P \) (except uncited ones) and \( h \)-index. We keep the analytical issue as further research questions.

### 17.5 Conclusion

Empirical data show that the applied indicators related to \( h \)-index such as \( c \)-descriptor, \( t \)-descriptor, \( h \)-core centroid \((c_x, c_y)\), \( h \)-tail centroid \((t_x, t_y)\), \( k \)-index and \( k' \)-index as well as \( c \)-max as dynamic core-tail measures for \( h \)-core and \( h \)-tail really change in time spans. The research reveals that there are obvious differences between sciences (physics) and social sciences (sociology) when we apply \( c \)-descriptor, \( h \)-core centroid \((c_x, c_y)\) and \( k \)-index as dynamic core-tail measures, and that there are few differences between sciences (physics) and social sciences (sociology) when we use \( t \)-descriptor, \( h \)-tail centroid \((t_x, t_y)\) and \( k' \)-index as dynamic core-tail measures, following time span from 1 to 10 years.
Meanwhile, the academic matrices (extending to academic tensor) can be applied to compare core-tail distributed data as synthetic overall measures in analysis for academic performances, which could stimulate further studies.

Acknowledgements I acknowledge the National Natural Science Foundation of China Grants (number 70773101, number 7101017006 and number 71173187) for the financial support. Thanks for the collaborations of Prof. Dr. Dar-zen Chen, Prof. Dr. Muhuan Huang, Prof. Dr. Loet Leydesdorff, Prof. Dr. Ronald Rousseau et al. in the studies, and Dr. Max C H Kuan, Ms. Huei-Ru Dong et al. in data assistances. This chapter is extracted and integrated by three published papers: (1) Ye and Rousseau (2010). Probing the h-core: an investigation of the tail-core ratio for rank distributions. *Scientometrics*, 84(2): 431–439; (2) Chen, D. Z., Huang, M. H. and Ye, F. Y. 2013. A probe into dynamic core-tail measures for h-core and h-tail. *Journal of Informetrics*, 7(1): 129–137; (3) Ye and Leydesdorff (2014). The “academic trace” of the performance matrix: a mathematical synthesis of the h-index and the integrated impact indicator (13). *Journal of the Association for Information Science and Technology*, 65(4): 742–750.

References


Rousseau, R., Ye, F.Y.: A formal relation between the $h$-index of a set of articles and their I3 score. J. Inf. 6(1), 34–35 (2012)


Chapter 18
A Quantitative Relationship Between Per Capita GDP and Scientometric Criteria

It is discovered that there exists a quantitative relationship, which can be expressed as $G = kF(\lg P)N$, where $G$ is per capita GDP, $F$ gross expenditure on R & D as percentage of GDP, $P$ patent applications, $N$ Internet users per 10,000 inhabitants, and $k$ a constant ranging from 0.4 to 1.2 in most countries. The mechanism of the relationship is explained in this chapter.

18.1 Introduction

It is well known that S & T (scientific and technical) and R & D (research and development) levels relate to the economic level in the information age. But we have not yet developed quantitative evaluation measures to assess the correlation between economic development levels and S & T or R & D. For finding the quantitative relationships between economic indicators and certain scientometric criteria, let us consider that S & T contribute to boom every economy and that R & D promotes economic development in every country. There should be relationships between economic indicators such as GDP and some scientometric criteria such as gross expenditure on R & D as percentage of GDP. Therefore, it is very important whether a quantitative relationship exists between per capita GDP and scientometric criteria.

In informetrics (Egghe and Rousseau 1990), some bibliometric laws have been integrated. However, there are not enough empirical laws among scientometric criteria and economic indicators. Thus, scientometrics seems to be only a name without rich contents. For constructing the systematical theory of scientometrics, I attempt to find quantitative patterns between economic indicators and scientometric criteria, referring to past informetric empirical laws and current theoretical research (Burrell 2005; Egghe 2005; Egghe and Rao 2002; Egghe and Rousseau 2005) as well as econometric analysis (Greene 2003). In this chapter, a quantitative relationship between per capita GDP and scientometric criteria is set up, based on statistical
data published by IMF (International Monetary Fund), WIPO (World Intellectual Property Organization) and UNESCO (United Nations Education, Science and Culture Organization).

18.2 Methodology

18.2.1 Subject

IMF, WIPO and UNESCO have collected a lot of data from all countries around the world. Their statistical data are the foundation of this chapter. While systematically checking the statistical data from IMF, WIPO and UNESCO through the viewpoint of informetrics and scientometrics, I find that there exists a quantitative relationship between per capita GDP and some scientometric criteria such as gross expenditure on R & D as percentage of GDP, patent applications and Internet users per 10,000 inhabitants.

During the research, I chose per capita GDP as an economic indicator and gross expenditure on R & D as percentage of GDP, patent applications, Internet users per 10,000 inhabitants and others as scientometric criteria, then tried to find a quantitative pattern between the economic indicators and these scientometric criteria. At last, a quantitative relationship is revealed between per capita GDP and gross expenditure on R & D as percentage of GDP with patent applications and Internet users per 10,000 inhabitants. Rather than being the result of a systematic research, it is much more like an accidental discovery.

18.2.2 Methods

The basic method is data analysis with test calculation. Through calculation and comparison, quantitative relationships can be observed and established. Software MS Excel was used as a simple analysis tool.

18.2.3 Procedures

(1) Collection of data. Firstly, I collected statistical data from IMF, WIPO and UNESCO web sites.
(2) Comparison of data. Secondly, I compared all data from various countries, using the same economic indicators and scientometric criteria.
(3) Analysis of data. Thirdly, I tried to find the quantitative relationships among the data.
18.2 Methodology

18.2.4 Data

Typical data which are collected for this chapter are shown in Table 18.1.

Table 18.1  Statistical data (2001)

<table>
<thead>
<tr>
<th>Country</th>
<th>Per capita GDP(^a)</th>
<th>Gross expenditure on R &amp; D as percentage of GDP(^b)</th>
<th>Patent applications(^c)</th>
<th>Grants of patents(^d)</th>
<th>Internet users per 10,000 inhabitants(^e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>22,730.2</td>
<td>2.11</td>
<td>98,489</td>
<td>12,019</td>
<td>4,514</td>
</tr>
<tr>
<td>the USA</td>
<td>35,366.6</td>
<td>2.74</td>
<td>375,657</td>
<td>166,038</td>
<td>4,995</td>
</tr>
<tr>
<td>Mexico</td>
<td>6,030.9</td>
<td>0.39</td>
<td>82,470</td>
<td>5,476</td>
<td>362</td>
</tr>
<tr>
<td>Brazil</td>
<td>2,984.1</td>
<td>1.04(2000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>18,458.7</td>
<td>1.55(2000)</td>
<td>94,893</td>
<td>13,983</td>
<td>3,723</td>
</tr>
<tr>
<td>Austria</td>
<td>23,217.6</td>
<td>2.07</td>
<td>233,181</td>
<td>14,328</td>
<td>3,901</td>
</tr>
<tr>
<td>Denmark</td>
<td>30,148</td>
<td>2.4</td>
<td>232,921</td>
<td>10,618</td>
<td>4,313</td>
</tr>
<tr>
<td>Finland</td>
<td>23,718.9</td>
<td>3.4</td>
<td>230,441</td>
<td>5,362</td>
<td>4,317</td>
</tr>
<tr>
<td>France</td>
<td>21,535</td>
<td>2.23</td>
<td>175,122</td>
<td>42,963</td>
<td>2,633</td>
</tr>
<tr>
<td>Germany</td>
<td>22,529.9</td>
<td>2.65</td>
<td>292,398</td>
<td>48,207</td>
<td>3,756</td>
</tr>
<tr>
<td>Greece</td>
<td>10,679.4</td>
<td>0.65</td>
<td>155,346</td>
<td>7,870</td>
<td>1,318</td>
</tr>
<tr>
<td>Italy</td>
<td>18,910.3</td>
<td>1.11</td>
<td>156,858</td>
<td>25,130</td>
<td>2,713</td>
</tr>
<tr>
<td>Netherlands</td>
<td>23,957.4</td>
<td>1.89</td>
<td>158,932</td>
<td>20,624</td>
<td>4,959</td>
</tr>
<tr>
<td>Norway</td>
<td>37,232.7</td>
<td>1.6</td>
<td>84,373</td>
<td>2,448</td>
<td>4,679</td>
</tr>
<tr>
<td>Poland</td>
<td>4,561.3</td>
<td>0.65</td>
<td>81,074</td>
<td>2,022</td>
<td>985</td>
</tr>
<tr>
<td>Russia</td>
<td>2,134.7</td>
<td>1.16</td>
<td>107,678</td>
<td>16,292</td>
<td>297</td>
</tr>
<tr>
<td>Spain</td>
<td>14,979.1</td>
<td>0.96</td>
<td>234,543</td>
<td>19,709</td>
<td>1,851</td>
</tr>
<tr>
<td>Sweden</td>
<td>23,546.7</td>
<td>4.27</td>
<td>231,483</td>
<td>14,873</td>
<td>5,208</td>
</tr>
<tr>
<td>Switzerland</td>
<td>33,951.4</td>
<td>2.63(2000)</td>
<td>233,652</td>
<td>15,639</td>
<td>3,102</td>
</tr>
<tr>
<td>the UK</td>
<td>23,735.7</td>
<td>1.86</td>
<td>264,706</td>
<td>39,649</td>
<td>3,325</td>
</tr>
<tr>
<td>China</td>
<td>910.8</td>
<td>1.07</td>
<td>149,294</td>
<td>16,296</td>
<td>262</td>
</tr>
<tr>
<td>Japan</td>
<td>32,765.8</td>
<td>3.06</td>
<td>496,621</td>
<td>121,742</td>
<td>3,840</td>
</tr>
<tr>
<td>Turkey</td>
<td>2,283.5</td>
<td>0.72</td>
<td>229,339</td>
<td>2,137</td>
<td>591</td>
</tr>
<tr>
<td>Egypt</td>
<td>1,490</td>
<td>0.19(2000)</td>
<td>1,387</td>
<td>430</td>
<td>87</td>
</tr>
</tbody>
</table>

Data sources:
\(^a\)http://www.imf.org/external/pubs/ft/weo/2002/02/data/index.htm;
\(^b\)http://www.uis.unesco.org/ev\_en.php?ID = 2867_201&ID2=DO_TOPIC;
\(^c\)\(^d\)http://www.wipo.int/ipstats/en/statistics/patents/index.html;
\(^e\)http://www.uis.unesco.org/ev.php?URL_ID = 5495&URL_DO=DO_TOPIC&URL_SECTION=201. The data from IMF, WIPO and UNESCO seem unique. Accidentally, when I integrate the data with simple calculations, a quantitative relationship is revealed naturally.
18.3 Results

With the data in Table 18.1, I made some calculations. Typical results are shown in Table 18.2.

The data in Table 18.2 show that there exists a quantitative relationship between per capita GDP and gross expenditure on R & D as percentage of GDP with patent applications and Internet users per 10,000 inhabitant in most countries, which can be expressed as \( G = kF(\log P)N \), where \( G \) is Per Capita GDP, \( F \) is Gross Expenditure on R & D as percentage of GDP (GERD%), \( P \) is patent applications, \( N \) is Internet users per 10,000 inhabitants, and \( k \) is a constant among 0.5–1. That result can be shown in Fig. 18.1.

Although that is only an approximate relationship, it is valuable in scientometrics and the social sciences as an initial finding.

### Table 18.2

<table>
<thead>
<tr>
<th>Country</th>
<th>Per capita GDP</th>
<th>( F(P/Q)N )</th>
<th>( F(\log P)N )</th>
<th>( 0.5F(\log P)N )</th>
<th>( 2F(\log P/\log Q)N )</th>
<th>( k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>22,730.2</td>
<td>78,048.29</td>
<td>47,559.72</td>
<td>23,779.86</td>
<td>23,314.34</td>
<td>0.47793</td>
</tr>
<tr>
<td>the USA</td>
<td>35,366.6</td>
<td>30,964.93</td>
<td>76,298.27</td>
<td>38,149.13</td>
<td>29,231.89</td>
<td>0.463531</td>
</tr>
<tr>
<td>Mexico</td>
<td>6,030.9</td>
<td>2,126.208</td>
<td>694.082.7</td>
<td>347.041.3</td>
<td>371.319.8</td>
<td>8.689022</td>
</tr>
<tr>
<td>Brazil</td>
<td>2,984.1</td>
<td>13,401.48</td>
<td>2,399.848</td>
<td>1,199.924</td>
<td>1,359.857</td>
<td>1.243454</td>
</tr>
<tr>
<td>Australia</td>
<td>18,458.7</td>
<td>39,161.43</td>
<td>28,721.88</td>
<td>14,360.94</td>
<td>13,856.56</td>
<td>0.64267</td>
</tr>
<tr>
<td>Austria</td>
<td>23,217.6</td>
<td>131,417.7</td>
<td>43,344.5</td>
<td>21,672.25</td>
<td>20,857.83</td>
<td>0.535653</td>
</tr>
<tr>
<td>Denmark</td>
<td>30,148</td>
<td>227,068.4</td>
<td>55,557.05</td>
<td>27,778.53</td>
<td>27,598.84</td>
<td>0.542649</td>
</tr>
<tr>
<td>Finland</td>
<td>23,718.9</td>
<td>630,803.2</td>
<td>78,710.58</td>
<td>39,355.29</td>
<td>42,211.68</td>
<td>0.301343</td>
</tr>
<tr>
<td>France</td>
<td>21,535</td>
<td>23,933.26</td>
<td>30,786.75</td>
<td>15,393.37</td>
<td>13,289.93</td>
<td>0.699489</td>
</tr>
<tr>
<td>Germany</td>
<td>22,529.9</td>
<td>60,372.03</td>
<td>54,405.03</td>
<td>27,202.51</td>
<td>23,234.57</td>
<td>0.414114</td>
</tr>
<tr>
<td>Greece</td>
<td>10,679.4</td>
<td>16,910.41</td>
<td>4,447.387</td>
<td>2,223.693</td>
<td>2,283.068</td>
<td>2.401275</td>
</tr>
<tr>
<td>Italy</td>
<td>18,910.3</td>
<td>18,796.93</td>
<td>15,645.9</td>
<td>7,822.952</td>
<td>7,111.464</td>
<td>1.208643</td>
</tr>
<tr>
<td>Netherlands</td>
<td>23,957.4</td>
<td>72,226.13</td>
<td>48,748.41</td>
<td>24,374.2</td>
<td>22,598.14</td>
<td>0.49145</td>
</tr>
<tr>
<td>Norway</td>
<td>37,232.7</td>
<td>258,027</td>
<td>36,879.53</td>
<td>18,439.76</td>
<td>21,765.47</td>
<td>1.009576</td>
</tr>
<tr>
<td>Poland</td>
<td>4,561.3</td>
<td>25,671.43</td>
<td>3,142.911</td>
<td>1,571.456</td>
<td>1,901.464</td>
<td>1.451298</td>
</tr>
<tr>
<td>Russia</td>
<td>2,134.7</td>
<td>2,277.021</td>
<td>1,733.668</td>
<td>866.834.2</td>
<td>823.209.4</td>
<td>1.23132</td>
</tr>
<tr>
<td>Spain</td>
<td>14,979.1</td>
<td>21,146.36</td>
<td>9,542.671</td>
<td>4,771.335</td>
<td>4,443.965</td>
<td>1.569697</td>
</tr>
<tr>
<td>Sweden</td>
<td>23,546.7</td>
<td>346,114.2</td>
<td>119,297</td>
<td>59,648.52</td>
<td>57,183.91</td>
<td>0.197379</td>
</tr>
<tr>
<td>Switzerland</td>
<td>33,951.4</td>
<td>121,887.2</td>
<td>43,798.19</td>
<td>21,899.09</td>
<td>20,885.08</td>
<td>0.775178</td>
</tr>
<tr>
<td>the UK</td>
<td>23,735.7</td>
<td>41,289.17</td>
<td>33,537.08</td>
<td>16,768.54</td>
<td>14,586.95</td>
<td>0.707745</td>
</tr>
<tr>
<td>China</td>
<td>910.8</td>
<td>2,568.304</td>
<td>1,450.491</td>
<td>725.245.5</td>
<td>688.728.9</td>
<td>0.627925</td>
</tr>
<tr>
<td>Japan</td>
<td>32,765.8</td>
<td>47,933.3</td>
<td>66,930.57</td>
<td>33,465.29</td>
<td>26,322.43</td>
<td>0.489549</td>
</tr>
<tr>
<td>Turkey</td>
<td>2,283.5</td>
<td>45,666.04</td>
<td>2,280.991</td>
<td>1,140.495</td>
<td>1,370.045</td>
<td>1.0011</td>
</tr>
<tr>
<td>Egypt</td>
<td>1,490</td>
<td>53,318.86</td>
<td>51,938.52</td>
<td>25,969.26</td>
<td>39,444.96</td>
<td>28.68776</td>
</tr>
</tbody>
</table>
18.4 Analysis and Discussion

On the basis of the data above and following calculations, it is shown that per capita GDP is among 0.5$F(lgP)N$ and $F(lgP)N$. As the GERD% expresses a level for investing in R & D, $P$ is directly the measurement of R & D and IUPI (Internet users per 10,000 inhabitants) indicates the informational level of the economy. Thus, we can conclude that the economic level, as expressed by per capita GDP, is proportional to GERD%, $lgP$ and IUPI.

In scientometric criteria, GERD% is obviously an important factor for R & D. Patent applications are an important factor for both S & T and R & D, since patent information is one of the most crucial elements for R & D. In the information era, the Internet is a necessary tool for economic development, thus IUPI (Internet users per 10,000 inhabitants) is also an important indicator for the economy and R & D.

The quantitative relationship shows that economic level is proportional to investment level in R & D, information level of society and the logarithm of patent activities.

In other words, the factor $F$ is similar to an accelerator for R & D as a percent of investment to R & D. The factor $P$ is an important index for R & D. And the factor $N$ is a technical indicator, or information indicator, of a society.

The significance of the constant $k$ can be also clarified. For most cases, we see the constant $k$ is between 0.4 to 1.2 in most countries. Normally, $k \sim 0.5$ seems better, and 0.4 to 1.6 seems acceptable.

If $k > 1$, it shows that GERD% is weaker and fewer than that in normal cases. When $k > 1.6$, there are few patents, so that innovative activities seem weak in the country.

If $k < 0.4$, it shows that there are many more patent applications than grants of patents, so that the creative quality seems dropping in the country.
Generally, constant $k$ is a measure for innovative level in a country. Thus, we can call it the innovative coefficient.

Meanwhile, there is no theoretical explanation on the relationship between per capita GDP and scientometric criteria, for which we can only link the science level with economic level qualitatively. The co-integration (Engle and Granger 1987) looks like a good idea for linking scientometric criteria to economic indicators and more quantitative relations could be stimulated, with linking to interdisciplinary studies such as combining informetrics with econometrics.

### 18.5 Conclusion

In this chapter, I recommend $G$ to be per capita GDP, $F$ to be GERD% (gross expenditure on R & D as percentage of GDP), $P$ to be patent applications, $N$ to be Internet users per 10,000 inhabitants and $k$ to be a constant. The discovery above means that there exists an approximate relationship in most countries, which can be expressed as $G = kF(lgP)N$. Moreover, the mechanism of this relationship is discussed in this chapter, which shows that GERD% means the investment for R & D, patent applications concern the scientific and technical research level and Internet users per 10,000 inhabitants correspond to the level of the information economy. That is a result based on the data in year 2001 only. More years of data are needed for to confirm the results. In the next step, I will look for more data to check the results above, if the IMF and UNESCO statistics becomes richer.

The advantage of scientometric methodology as a means of research is that it provides a global view of science and economy and applicable concepts or viewpoints for an informetric study. The methodology will be effective for analysis of other economic indicators and scientometric criteria.

**Acknowledgements** Thank sincerely Dr. Yan Ma, professor at University of Rhode Island in the USA and Ms. Regina P. Entorf, reference librarian at Wittenberg University in the USA, for their English wording. This chapter was originally published as follows and this version had been modified. Ye, F. Y. 2007. A quantitative relationship between per capita GDP and scientometric criteria. *Scientometrics*, 71(3): 407–413.

**References**


Appendix A
The Outline of Triad Philosophy: A Philosophical Idea for Merging Western and Eastern Thoughts

The triad philosophy is introduced based on triad logic. The systematical core ideas include triad ontology, triad epistemology and triad axiology. The triad ontology is characterized by objective knowledge (Li), objective matter (Ch’i) and subjective spirit (Hs’in). Triad epistemology is characterized by experiencing, studying and thinking. Triad axiology consists of honesty, charity and forgiveness.

A.1 Introduction

Currently people need a contemporary philosophy in which Western and Eastern philosophical ideas of excellence can be melted together for processing all things in the world. Synthesizing Western, Chinese and Eastern thought, I introduce the triad philosophy based on triad logic.

In triad logic, there exist triad operators (+, *, R) and three values(1, 0, −1). Their value table fits an extended Lukasiewicz L3 (Lukasiewicz, 1930) system shown as follows.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>0</th>
<th>−1</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>1</td>
<td>0</td>
<td>−1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>−1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>−1</td>
</tr>
<tr>
<td>−1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

That is a non-symmetric ternary logic (Haaparanta 2000).

On the basis of the value table above, the following logic laws can be set up (Eldon 2000).

1Here, Kant’s philosophy, Popper’s philosophy, philosophy of I Ching and Buddhism are particularly considered.

2Lukasiewicz logic used values 0, 1, 2.
A.1.1 The exchange law

\[ A + B = B + A \quad A \times B = B \times A \quad ARB = BRA \]

A.1.2 The combination law

\[ (A + B) + C = A + (B + C) \quad (A \times B) \times C = A \times (B \times C) \]

A.1.3 The distribution law

\[ A \times (B + C) = (A \times B) + (A \times C) \quad AR(B + C) = (ARB) \times (ARC) \]

\[ AR(B \times C) = (ARB) + (ARC) \]

Based on the logic system, sentences will be classified as positive sentences (true value, 1), indefinite sentences (no value, 0) and negative sentences (false value, −1) in language. Thus, triad philosophy can be developed based on triad language logic.

A.2 Triad Ontology

A.2.1 Category Origin

The concept Onto(s), which comes from Western philosophy, means the essence(s) of existing things. In Western philosophy, there are objective knowledge, which means the idea of Li (rightness) in Chinese philosophy or the idea of Brahman in Indian philosophy; objective matter, which means the idea of Ch’i (vapor) in Chinese philosophy or the idea of Maya in Indian philosophy; subjective spirit, which means the idea of Hs’in (mind) in Chinese philosophy or the idea of Atman in Indian philosophy. Referring to Popper’s idea (Popper 1972) of objective matter (World 1), subjective spirit (World 2) and objective knowledge (World 3), I can construct triad ontology as follows.

A.2.2 Structure Construction

Synthesizing objective knowledge (rightness, Brahman), objective matter (vapor, Maya) and subjective spirit (mind, Atman) in Western, Chinese and Indian philosophy, triad ontology is constructed as in Fig. A.1.
Appendix A: The Outline of Triad Philosophy: A Philosophical Idea for Merging … 243

Fig. A.1 Triad ontology

Referring to onto-hermeneutics (Cheng 1996), Li in Chinese philosophy and Brahman in Indian philosophy can be interpreted as objective knowledge in Western philosophy, so are Ch’i and Maya as objective matter and Hs’in and Atman as subjective spirit. Figure A-1 means that objective knowledge (namely rightness or Brahman), objective matter (namely vapor or Maya) and subjective spirit (namely mind or Atman) exist at the same time.³

A.2.3 Function Development

Considering the relations in triad ontology, there is Fig. A.2 for reference.

That is an idea from the relations of five elements in Chinese philosophy.

Triad methodology, or the triad thinking pattern, can be introduced by triad ontology. The triad method is a system method in which there are three main groups or elements. The three groups or elements construct a triad relation as in Fig. A.3.

In that system, Group A includes foundation/basis element(s), Group B includes motivation/power element(s) and Group C contains regulation/adjustment element(s).

Every triad method system or subsystem may be dominated by three main elements. When a system is controlled by three elements A, B, C, the system may be called a one-order triad system. When A is divided as A1, A2, A3, B as B1, B2, B3 and C as C1, C2, C3, the system that contains nine elements may be called a two-order triad system (Fig. A.4).

Generally, a system which consists of $m$ elements may be called $(n - 1) + \frac{m}{3^n}$

³Differing from Popper’s philosophy, here objective knowledge, objective matter and subjective spirit constructed a triad structure.
order triad system. When $m \in [1, 3]$, $n = 1$; when $m \in [4, 9]$, $n = 2$; when $m \in [10, 27]$, $n = 3$.

In Wittgenstein’s Tractatus Logic-philosophy (Wittgenstein 1974), the basic philosophic problem is the relation of said and shown. In triad philosophy, it is a triad relation.

A.3 Triad Epistemology

A.3.1 Category Origin

The generation and development of knowledge must be a process. At the beginning, knowledge came from experiencing. After people concluded some principles or rules from experiences, they could think based on the principles or rules and got knowledge. When knowledge was accumulated and organized into a system, people could get knowledge through learning or studying.

A.3.2 Structure Construction

Thus, there are three sources of knowledge, i.e., experiencing (acting), thinking (pondering) and studying (learning), which together construct triad epistemology as Fig. A.5.
Considering Kant’s philosophy (Kant 1929), experiencing, thinking and studying may undergo the process of “perception $\rightarrow$ understanding $\rightarrow$ reason”. Combining Kant’s epistemological line with Chinese “nine squares” as a “cognition frame” for memory, triad epistemology can be shown in Fig. A.6.

The numbers express the importance order. The larger the number is, the more important is, for the kind of knowledge that contributes to modern knowledge.

There are some examples for explaining the order above.

**A.3.2.1 Examples of reason-thinking (order 9)**

1. Euler’s formula: $\exp(i\pi) = \cos \pi + is\sin \pi$.
2. Newton’s gravity law: $F = G \frac{m_1 m_2}{r^2}$.
3. Knowledge is power.

**A.3.2.2 Examples of perception-experiencing (order 8)**

1. All that glitters are not gold.
2. A rolling stone gathers no moss.
3. Rome was not built in a day.

**A.3.2.3 Examples of understanding-studying (order 7)**

1. $0 + 1 = 1, 1 + 1 = 2, \cdots, 9 + 1 = 10, 10 + 1 = 11, \cdots, 99 + 1 = 100 \cdots$
2. $0 \times 1 = 0, 1 \times 1 = 1, \cdots, 2 \times 2 = 4, 2 \times 3 = 6, \cdots, 9 \times 9 = 81 \cdots$
3. One day equals 24 h and one hour equals 60 min.
A.3.3 **Function Development**

Human knowledge can be induced by perception, understanding and reason on the basis of nature, society and mankind. At the beginning, knowledge came from experiencing with perception and gradually literature, arts and natural philosophy were produced. Then natural, social sciences and humanities were introduced based on thinking and studying with understanding and reason. Step by step, physics became the core of natural sciences, economics of social sciences and philosophy of humanities.

Now, I can conclude three principles and three laws as the core of human knowledge.

**Principle I (cycle principle):** there are a lot of cycles in the world. That is the basis of existing things. The cycle principle is expressed as the chemical periodic table, TCA cycle, economic cycle and many other periodic phenomena.

**Principle II (harmony principle):** all parts of the world relate to each other. That is the adjustment for existing things. The harmony principle is expressed as symmetry, poetic rhyme, music harmonics and other harmonic phenomena.

**Principle III (optimization principle):** optimization is the direction for development in the world. That is the power of existing things. The optimization principle can be expressed as the least action law, maximum profit and minimum cost rules and other optimum phenomena.

**Law 1:** substance never vanishes. That is the foundation of objective matter.
**Law 2:** spirit never dies away. That is the foundation of subjective spirit.
**Law 3:** logic never confuses. That is the foundation of objective knowledge.

The three principles and three laws set up the basis of the world and construct the framework of contemporary knowledge.  

A.4 **Triad Axiology**

A.4.1 **Category Origin**

In Western philosophy, truth, goodness and beauty are keywords for axiology. And in Chinese and Eastern philosophy, honesty, charity and forgiveness are strengthened. Honesty reflects the true; charity means the good; and forgiveness shows the beauty. So, honesty (the true), charity (the good) and forgiveness (the beauty) construct the triad axiology. 

---

4 If a contemporary philosophy can not explain contemporary knowledge, it is untenable.
5 In my point of view, mass axiology means ethics and personal axiology expresses the outlook of life.
**A.4.2 Structure Construction**

In triad axiology, the triad relation is shown in Fig. A.7.

Triad axiology means that we should generally be charitable to everyone as a condition of maintaining basic ecological balance and people should forgive each other and should be sincere in society.

**A.4.3 Function Development**

An ideal humanity or personality is the unity of charity, forgiveness and honesty. I propose that humanity has a complex multi-level inner structure, which contains physiological, psychological and philosophical levels. It is a two-order triad system (Fig. A.8).

In that system, charity (goodness) is the dynamic source of kindness, volition and imagination, which construct the psychological level. So are forgiveness (beauty) of consciousness, love, emotion for the physiological level and honesty (truth) of perception, reason, understanding for the philosophical level. The triad outlook of life derives from the structure of humanity.

Logic, science and law result from reason. Courage, willpower and war are caused by volition. And joy, anger, anxiety, fear (terror) and the arts come from emotion.

At the physiological level, volition is motivation; at the psychological level, love is primary; at the philosophical level, reason is power. Humanity may mainly be a
trinity of volition, love and reason. Certainly, we can not ignore other natures, for example, there will be no hope without imagination.

Paying attention to kindness and developing reason are necessary paths for training an ideal personality.

The moral levels will adapt to the social civilization and economic levels. In the future, people can choose suitable life standards based on their conditions. There are three standards for choice:

(1) High standard (the sage view of life): charity for everything and forgiveness for everyone. Someone who is glad to help others and to act for society will say: “I like to be charitable to the world, to forgive all people and to be honest in society.”

(2) Middle standard (the gentleman’s view of life): keep goodness in mind and be upright in action. Someone who is peaceful and modest will say: “I will keep charity and forgiveness for good people and honesty in society.”

(3) Low standard (the masses view of life): do not destroy social safety. Someone who dislikes helping others, but does not hurt others either will say: “I am charitable only to myself, forgiveness is only for my friends; but I am honest in society.”

If most people can get to the middle standard in a society, the social morality will be good. Even the low standard will lead to peace and quietness. The base is to keep “honesty in society.”

I think that agricultural civilization, industrial civilization and knowledge (information) civilization will co-exist in the future. The triad civilization (society) will need triad axiology.

A.5 Conclusion

Triad ontology, triad epistemology and triad axiology construct the main framework of triad philosophy. In triad ontology, the main elements are objective knowledge (rightness, Brahman), objective matter (vapor, Maya) and subjective spirit (mind, Atman). In triad epistemology, the main elements are experiencing (acting), thinking (pondering) and studying (learning). And in triad axiology, the main elements are honesty (truth), charity (goodness) and forgiveness (beauty).

That is an outline of triad philosophy. Deeper research into triad ontology, triad epistemology and triad axiology will be separately studied. I hope the triad philosophy will provide a reference system for contemporary philosophy.

Acknowledgements: During Prof. Michael Cohen’s visit in Zhejiang University from University of Michigan in 2006, he mentioned Peirce’s triadic philosophy to me, which enriched the reference system.

References


Note: this chapter is revised and simplified by the following paper: Eldon, F. Y. 2000. Triad philosophy and triad science based on triad logic. Journal of Zhejiang University (Science), 1(4): 467–475
Appendix B
Main Publications (1998–2016)
of Prof. Dr. Fred Y. Ye

Peer Reviewed Articles in Physics

Peer Reviewed Articles in Economics
Peer Reviewed Articles in Scientometrics


Main Books