RUSSIAN PRESIDENTIAL ACADEMY OF NATIONAL ECONOMY AND PUBLIC ADMINISTRATION
International Laboratory of Political Demography and Social Macro-Dynamics

RUSSIAN ACADEMY OF SCIENCES
Keldysh Institute of Applied Mathematics
Institute of Oriental Studies

INTERNATIONAL CENTER FOR EDUCATION, SOCIAL AND HUMANITARIAN STUDIES

HISTORY & MATHEMATICS

Big History Aspects

Edited by
Leonid Grinin,
Andrey Korotayev

'Uchitel'
Publishing House
Volgograd

ББК 22.318 60.5

History & Mathematics: Big History Aspects. Yearbook / Edited by Leonid Grinin, Andrey Korotayev. – Volgograd: 'Uchitel' Publishing House, 2019. – 344 pp.

Our Yearbook 'History and Mathematics' has already celebrated its 10th anniversary and has confidently entered its second decade. The common feature of all our Yearbooks, including the present volume, is the usage of formal methods and social studies methods in their synthesis to analyze different historical phenomena.

The present Yearbook (which is the seventh in the series) is subtitled 'Big History Aspects'. This issue is devoted to the problems of evolutionary development of the world. In no way will it be a digression from the direction which we have initially defined for our Yearbook, but just an extension of the scope of the research. The matter is that there are two kinds of history: the history of nature (or more exactly the Universe and the Earth) and the history of humans and mankind. It is not surprising that the idea of historicism penetrated almost every scientific field. At the same time the search for common foundations of this endless in its diversity world has intensified. One of the directions of this interdisciplinary search for the unity of the world in its diversity is Universal Evolutionism (Big History). Mathematical and formal methods help to understand much deeply both natural and human history.

This issue of the Yearbook consists of four main sections: (I) Patterns of Big History; (II) Hypotheses of Deep Big History; (III) Biological Aspects; (IV) History and Future of Social Systems.

We hope that this issue will be interesting and useful both for historians and mathematicians, as well as for all those dealing with various social and natural sciences.

'Uchitel' Publishing House
143 Kirova St.,
400079 Volgograd,
Russia

ISBN 978-5-7057-5464-9

Volgograd 2019

Contents

IV. History and Future of Social Systems

Introduction. From the History of Humankind to Big History*

Leonid E. Grinin and Andrey V. Korotayev

Our Yearbook 'History and Mathematics' has already celebrated its 10th anniversary and has confidently entered its second decade. In this regard it will be quite reasonable to try to extend the scope of its research field. Therefore, this issue is designed in an unusual manner while still preserving the relevance for our approach (see about it below). It also makes sense to remind the reader of the various aspects that were touched upon in the previous issues.

The first issue of the Yearbook entitled 'Analyzing and Modeling Global Development' came out in 2006 (see Grinin, de Munck, and Korotayev 2006). This volume initiated a series of edited volumes dedicated to various aspects of the application of mathematical methods to the study of history and society. It comprised articles that apply mathematical methods to the study of various epochs and scales: from deep historical reconstruction to the pressing problems of the modern world. On the other hand, all the articles of this issue were dedicated to the analysis, periodization, or modeling of global development. It was shown that the mathematical modeling of historical macroprocesses suggests a fresh approach to the periodization issue. The authors studied these problems from different perspectives (technological, economic, demographic, sociostructural, cultural-psychological, linguistic). New quantitative insights on the dynamics of contemporary processes were presented. These insights allowed the authors to make a number of important forecasts on this basis.

The second issue was entitled 'Historical Dynamics and Development of Complex Societies'. It demonstrated that the application of mathematical methods not only facilitates the processing of historical information, but can also give to a historian a deeper understanding of historical processes (see Turchin *et al.* 2006).

The third issue of the Yearbook 'Processes and Models of Global Dynamics' presented a qualitative and quantitative analysis of global historical, political, economic and demographic processes, as well as their mathematical models (see Grinin *et al.* 2010).

---

* The research is supported by the Russian Foundation for Basic Research (Project No. 17-06-00464).

The fourth issue 'Trends and Cycles' was devoted to cyclical and trend dynamics in society and nature; special attention was paid to economic and demographic aspects, in particular to the mathematical modeling of the Malthusian and post-Malthusian traps' dynamics (see Grinin and Korotayev 2014).

The fifth issue 'Political Demography & Global Ageing' brought together a number of interesting articles by scholars from Europe, Asia, and America. They examined such an important modern historical macroprocess as global ageing from a variety of perspectives (see Goldstone, Grinin, and Korotayev 2015).

The sixth (Anniversary) issue 'Economy, Demography, Culture, and Cosmic Civilizations' revealed the extraordinary potential of the application of mathematical methods to the study of historical processes (see Grinin and Korotayev 2017).

The common feature of all our Yearbooks, including the present volume, is the usage of formal methods and social studies methods in their synthesis to analyze different historical phenomena.

The present Yearbook (which is the seventh in the series) is subtitled 'Big History Aspects'. This issue is devoted to the problems of evolutionary development of the world. In no way will it be a digression from the direction which we have initially defined for our Yearbook, but just an extension of the scope of the research.

As Karl Marx and Friedrich Engels wrote in their *German Ideology*, 'We know only a single science, the science of history… the history of nature and the history of men' (Marx and Engels 1976 [1846]: 34). Since that time the history of nature went even further back into the past, but the approach has not changed. On the contrary, at present the idea of historicism penetrated almost every scientific field. At the same time the search for common foundations of this endless in its diversity world has intensified. One of the directions of this interdisciplinary search for the unity of the world in its diversity is Universal Evolutionism (Big History).

In the paper 'Big History Trends and Patterns' an American astrophysicist **Eric J. Chaisson** writes, 'In the 20th century, several independent efforts came forth virtually simultaneously, as Sagan (1980), Jantsch (1980), Reeves (1981), and Chaisson (1981) all advanced the idea of complex systems naturally emerging with the pace of natural history'. Due to the efforts of scientists from different countries, there appeared a field of study which literally unified the history of nature and people. This is the Big (or Universal) History which explores the history of the Universe and humankind from the Big Bang to the present day (including our future). In regard to this approach, we wrote,

> One of the clearest manifestations of the evolutionary approach is the form of universal evolutionism that considers the process of evolution as a continuous and integral process – from the Big Bang all the way down to the current state of human affairs and beyond. Universal evolutionism implies that cosmic, chemical, geological, biological, and social types of

> macroevolution exhibit forms of genetic and structural continuity <...> The significance of this approach (which has both the widest possible scope, and a scientific basis) is evident. It strives to encompass within a single theoretical framework all the major phases of the Universe, from the Big Bang to the forecasts for the entire foreseeable future, while showing that the present state of humankind is a result of the self-organization of matter (Grinin, Markov *et al*. 2009: 8–9).

Thus, Evolutionistics that we develop in our works and Yearbooks (‘Эволюция’ and ‘Evolution’) is considered as an interdisciplinary common field (as well as intended combination of history and mathematics), which shows the unity of the world in its diversity. And what is better than mathematics at all times proved this unity of the world? Thus, we think that the integration of Big History's and mathematical dimensions in our Yearbook is fully justified and reasonable.

The present Yearbook consists of four sections.

Section I ‘Patterns of Big History’ includes the article by **Eric J. Chaisson** (‘Big History Trends and Patterns’). According to the author, evolution – ascent with change of Nature's many varied systems – has become a powerful unifying concept throughout the sciences. In its broadest sense, cosmic evolution, which includes the subject of Big History, comprises a holistic explanatory narrative of countless changes within and among organized systems extending from the Big Bang to humankind. This interdisciplinary scenario has the potential to unite physical, biological, and social sciences, thereby creating for people of all cultures at the start of the new millennium a consistent, objective, and comprehensive worldview of material reality.

The second article of this ***Section*** (‘Potential Nested Accelerating Returns Logistic Growth in Big History’) by **David J. LePoire** presents the models for the interpretation of Big History. This interpretation includes the increasing rates of the evolutionary events and phases of life, humans, and civilization. These three phases, previously identified by others, have different information processing mechanisms (genes, brains, and writing). The accelerating returns aspect of the new model replicates the exponential part of the progress as the transitions in these three phases started roughly 5 billion, 5 million, and 5,000 years ago. Each of these three phases might be composed of a further level of about six nested transitions with each transition proceeding faster by a factor of about three with corresponding changes in free energy flow and organization to handle the increased generation rate of entropy from the system.

The article by **Antony Harper** (‘A Toy Model Mechanism for Greater-Than-Exponential Human Population Growth’) proposes an underlying mechanism which generates this greater-than-exponential growth. The mechanism is represented by a toy model of two differential equations of interacting populations, the interactions of which enhance the reproductive abilities of the other population. The end result of this enhancement due to positive human interaction, a quintessential characteristic of our species, is a pattern of growth motivated by a greater-than-exponential rate of growth. It should be noted that the model proposed here is one of many potential models and not the sole, the only, possible model.

Section II is devoted to cosmic evolution. It mainly consists of the contributions connected with the hypotheses about prehistory of our Universe, therefore it is subtitled 'Hypotheses of Deep Big History'. It opens with **David Baker's** paper ('10^{500}. The Darwinian Algorithm and a Possible Candidate for a "Unifying Theme" of Big History'). This article postulates another aspect of the long sought-after 'unifying theme' of Big History, in addition to the rise of complexity and energy flows. It looks briefly at the manifestation of the 'Darwinian algorithm', that is to say an algorithm of random variation and non-random selection, in many physical processes in the Universe: cosmology, geology, biology, culture, and even the occurrence of universes themselves. This algorithm also seems to gradually open more forms of variation and more selection paths over time, leading to a higher level of free energy rate density, or what we know as 'complexity'. In fact the complexity of the object under discussion seems to correspond to the available number of selection paths. The article closes with a bit of philosophical reflection on what the Darwinian algorithm and the rise of complexity could possibly mean for humanity and the future of the cosmos.

The article by **Tom Gehrels** ('The Chandra Multiverse') extends the already colossal time horizons of Big History in a truly fantastic way. While reading this article, it is difficult to avoid exclaiming something like: 'This is a really BIG history!' Of course, this is a hypothesis, with which many might not agree. But this is a very bold hypothesis that extends the Big History horizon by many orders of magnitude. According to Gehrels, equations of Planck and Chandrasekhar lead to the conclusion that our universe is a member of a quantized system of universes, which he calls the 'Chandra Multiverse'. It is a trial-and-error evolutionary system. All universes have the same critical mass and finely tuned physics that our universe has. The origin and demise of our universe is described. In our astronomical environment, everything ages and decays; even the proton may have a limited half-life. The decay products of all the universes expand into the inter-universal medium (IUM), clouds form in the IUM, from which new universes are started. When the density at the center of our proto-universe cloud reached proton density, then photons, protons and neutrons were re-energized. A Photon Burst marks the beginning of our universe at 10^{-6} sec (1037 Planck times) later than a Big Bang, and the evolution of forces, sub-atomic particles and finely tuned physics occurs in the Chandra Multiverse. This theory of the multiverse also makes identification of dark energy and dark matter possible.

The section concludes with the contribution by **Leonid E. Grinin** ('Was There a Big Bang?'). The idea that our Universe emerged as a result of the extraordinary power of the Big Bang from singularity (*i.e.*, a state of an infinitely small quantity and infinitely high concentration of matter) is still very popular today. It was one of the main postulates of the Big Bang theory that completely formed in the 1960s–1970s. However, at present this idea as well as the Big Bang theory is outdated, although it is still shared by many scientists. Being widespread since the end of the 1970s the Inflation theory appears more modern. The main reason for the emergence of the Inflation theory was that the Big Bang theory could not satisfactorily explain a number of the contemporary parameters of the Universe.

The Inflation theory makes still widespread views of the Big Bang theory archaic as regards the following points: 1) the history of the Universe started with the Big Bang; 2) it started with the singularity. According to the Inflation theory, the Big Bang was not the beginning and the moment of the origin of the Universe, but it was preceded by at least two epochs: inflation and post-inflationary heating. That is, the Big Bang or precisely the hot Big Bang is just a phase transition from the state of cold inflation to the hot phase. Since the Inflation theory does not consider the Big Bang as the initial phase there emerges an intricate problem of the role of the Big Bang in the process of the formation the Universe as a whole. The paper considers the confusion with the Big Bang notion, a number and sequence of 'bangs' and why the theory can dispense easily without the notion the Big Bang. The advantages and disadvantages of the Inflation theory are discussed in this contribution.

Section III 'Biological Aspects' opens with the paper by **Edmundas Lekevičius** ('Ecological Darwinism or Preliminary Answers to Some Crucial though Seldom Asked Questions'). The author asserts that evolutionary regularities might be deduced from basic principles describing how life functions, most notably part-whole relationships and control mechanisms. Lekevičius suggests adding the concept of functional hierarchy to the concept of the struggle for existence: no solitary individual or species is functionally autonomous. Life as we know can exist only in the form of a nutrient cycle. Only two purely biotic forces – 'biotic attraction' and 'biotic repulsion' – act in the living world. The first one maintains and increases diversity and organizes solitary parts into systems integrated to a greater or lesser degree. The second one, in the form of competition, lessens biodiversity but at the same time provides life with necessary plasticity. On that basis, tentative answers to the following questions are given: (1) Why does life exhibit such a peculiar organization with strong integration at lower levels of organization and weak integration at higher ones?; (2) Why did particular species and guilds appear on the evolutionary stage at that particular time and not at any other?; (3) Why was the functional structure of ecosystems prone to convergence despite a multitude of stochastic factors?

In their contribution ('Relationship between Genome Size and Organismal Complexity in the Lineage Leading from Prokaryotes to Mammals') **Alexander V. Markov, Valery A. Anisimov, and Andrey V. Korotayev** emphasize that there is a direct relationship between the level of organization and the minimal genome size (MGS) in the lineage leading from prokaryotes to mammals, in which the tendency towards increasing complexity is especially clear. The dynamics of MGS in this lineage can be adequately described by the model of hyperexponential growth. This implies the existence of nonlinear positive feedbacks that account for the acceleration of MGS growth. The nature of these feedbacks is discussed, including the formation of new genes by means of recombination of the fragments of existing genes, formation of 'niches' for new genes in the course of evolution of gene networks, and the expansion of regulatory regions. Hyperexponential growth of different variables related to the level of organization of the biosphere and society (biodiversity, MGS, size and complexity of organisms, world population, technological development, urbanization, *etc.*) suggests that the evolution of the biosphere and humanity in the direction of increasing complexity is a self-accelerating (autocatalytic) process.

Section IV 'History and Future of Social Systems' gives a series of forecasts. It opens with the paper by **Andrey V. Korotayev and Leonid E. Grinin** ('A Mathematical Model of Influence of the Interaction between Civilization Center and Barbarian Periphery on the World-System Development'). This article offers an analysis and mathematical modeling of the influence of one of the major factors of the World System macrodynamics throughout most part of its history (since the 'urban revolution') – the factor of interaction of civilizations with their barbarian periphery. The proposed mathematical model is intended to describe possible influence of interaction between civilizational core of the World System and its barbarian periphery on the formation of the specific curve of the world urbanization dynamics. It imitates completion of the phase transition, behavior of the system in the attraction basin and beginning of the phase transition to the attraction basin of the new attractor and is aimed to identify the role of the factor of interaction between the civilizational core and barbarian periphery in the formation of attractor effect during the completion of phase transition, that is for clarification of the reason why there was observed not only slowdown of growth rates of the main indicators of the World System development after completion of phase transitions during its development, but also their falling with the subsequent temporary stabilization near some equilibrium level. Achievements of modern barbarology, including complexity of barbarian periphery itself and its heterogeneity are considered. The basic principle of the proposed dynamic model is that sizes, power and level of complexity in realization of external policy functions in nomadic unions (empires) closely correspond to sizes, power and level of political culture and activity of the core states with which nomads constantly had to do (this point has been established in works of the known experts in nomadic studies). Various alternatives are shown in the model, when depending on power and size of one of the two components of the system 'civilization – barbarian periphery' studied by the authors, another one also changes significantly as it has to respond to the challenge properly, or can make less efforts without feeling threat or resistance. This principle is observed throughout the long period of the history of the World System. It is shown that interaction between the civilization center and barbarian periphery really can explain some characteristic features of the World System dynamics in the 4th millennium BCE – the 2nd millennium CE. The ways of further development of the model are outlined.

According to **William R. Thompson** ('Energy, Kondratieff Waves, Lead Economies, and Their Evolutionary Implications'), one approach to interpreting Kondratieff waves, associated with the leadership long cycle research program, emphasizes the role of intermittent but clustered technological innovations primarily pioneered by a lead economy, with various significant impacts on world politics. This approach is further distinguished by asserting that the K-wave pattern is discernible back to the 10th century and the economic breakthrough of Sung Dynasty China. While K-wave behavior has numerous and widespread manifestations, the question raised in this essay is whether explanatory power is improved by giving a greater role to energy and energy transitions in the K-wave process(es). Eight specific implications are traced, ranging from the interaction of technological innovations and energy to cosmological interpretations. In general, the answer to the raised question is affirmative, with one caveat on whether emphasizing new fuels and engines is a hallmark of the hydrocarbon era or a new and evolving feature of K-waves.

Leonid E. Grinin and Anton L. Grinin in their contribution ('Technological Dimension of Big History and the Cybernetic Revolution') consider the history of technological development in terms of production (or technological) revolutions. The analysis of its current state and forecasts is made with respect to Big History. Technologies have been playing a significant role in the history of humankind from the very origin of *Homo sapiens*. Numerous facts show that already after 50,000 BP technologies were developed in various fields: from hunting and cooking to primitive painting. Agriculture, building, transportation and many other human achievements could not have happened without certain technologies. Thus, one can argue that technologies play a very important role in Big History. Technologies played a special role in the collective learning which is defined as the sixth threshold of increasing complexity. This *Homo Sapiens'* achievement which happened at the beginning of the Upper Paleolithic was probably one of the most important events in the whole human history, and sometimes is termed as the Human revolution (*e.g.*, Shea 2006). Today we are at the threshold of another important transition which is often called 'posthuman revolution', which could bring quite radical changes to society and even transform the human biological nature.

In his paper ('Global Society as Singularity and Point of Transition to the New Phase of Social Evolution') **Sergey V. Dobrolyubov** considers social evolution as a process consisting of three phases: Adaptive, Structural and Cognitive, which are separated by two phase transitions or by two singularities – the Neolithic and the Global. The mechanism of social evolution at these phases is different and is based on different institutional means of cognition and competition. At the current structural phase, competition of individuals leads to inequality, and competition of societies leads to extension of societies. Social inequality and exploitation of the periphery become institutional tools for the development. The expansion of societies and evolutionary limitations of its growth lead to life cycles of societies. The maximum size of society increases in the process of evolution and tends to cover all humankind. The Global Society is a final point of structural evolution, and transition to it is singularity. It will be a metamorphosis of the society's nature. The mechanism of further social evolution at the cognitive phase will rely directly on individual's need for cognition and self-realization, and not on the special social institutions. Mathematical model of the primary transformations dynamics at the structural phase is described by the equation $T(n) = -11214 + 1893n$, where $T(n)$ – is the moment of evolutionary transformation, and n – is the ordinal number of transformation. Global singularity is predicted by this model in AD 3930.

The Section concludes with the article by **Antony Harper** ('The Punctuated Equilibrium Macropattern of World System Urbanization and the Factors that Give Rise to that Macropattern'). This paper is about the World System evolution as it is reflected in the pattern of urbanization over the last 5,000 years. It will be shown that the pattern of urbanization as part of the immensely complex world system exhibits non-linearity in that it is neither smooth nor continuous but rather is punctuated by periods of rapid change interspersed between periods of stasis. This pattern was first described in biological systems by Eldridge and Gould (1972) for speciation, and much of the pattern of urbanization reflects the characteristics of punctuated equilibrium first described by those two authors. Specifically, this paper will investigate the phenomenon of punctuated equilibrium reflected in both the macro-pattern of urbanization over historic time, *i.e.* the evidence for punctuated equilibrium as reflected by data on urbanization and on the level of state development, and possible mechanisms for such punctuated behavior including the general model of self-organized criticality as developed by Per Bak (1996), the role of hypercycle formation in punctuated equilibrium, the role of aromorphic processes, and the interaction between population, carrying capacity, and level of technology as represented by a very general math model.

We hope that the present issue of the Yearbook will be interesting and useful both for historians and mathematicians, as well as for all those dealing with various social and natural sciences.

References

Bak P. 1996. *How Nature Works: The Science of Self-Organized Criticality. Nature* 383(6603): 772–773.

Chaisson E. J. 1981. *Cosmic Dawn: The Origins of Matter and Life*. Boston, MA: Little Brown.

Eldridge N., and Gould S. J. 1972. Punctuated Equilibria: An Alternative to Phyletic Gradualism. *Models in Paleobiology* / Ed. by T. J. M. Schopf, pp. 82–115. San Francisco: W. H. Freeman.

Goldstone J. A., Grinin L. E., and Korotayev A. V. (Eds.) 2015. *History & Mathematics: Political Demography & Global Ageing*. Volgograd: 'Uchitel' Publishing House.

Grinin L. E., and Korotayev A. V. (Eds.) 2014. *History & Mathematics: Trends and Cycles*. Volgograd: 'Uchitel' Publishing House.

Grinin L. E., and Korotayev A. V. (Eds.) 2017. *History & Mathematics: Economy, Demography, Culture, and Cosmic Civilizations*. Volgograd: 'Uchitel' Publishing House.

Grinin L. E., Baker D., Quaedackers E., and Korotayev A. V. 2014. Introduction. Big History's Big Potential. *Teaching & Researching Big History: Exploring a New Scholarly Field* / Ed. by L. E. Grinin, D. Baker, E. Quaedackers, and A. V. Korotayev, pp. 7–18. Volgograd: 'Uchitel' Publishing House.

Grinin L. E., Herrmann P., Korotayev A., and Tausch A. (Eds.) 2010. *History and Mathematics: Processes and Models of Global Dynamics.* Volgograd: 'Uchitel' Publishing House.

Grinin L. E., Markov A. V., Korotayev A. V., and Panov A. D. 2009. Evolutionary Megaparadigm: Potential, Problems, Perspectives. Introduction. *Evolution: Cosmic, Biologicl, Social* / Ed. by L. E. Grinin, A. V. Markov, and A. V. Korotayev, pp. 5–43. Moscow: LIBROCOM. *In Russian* (Гринин Л. Е., Марков А. В., Коротаев А. В., Панов А. Д. Эволюционная мегапарадигма: возможности, проблемы, перспективы. Введение. *Эволюция: Космическая, биологическая, социальная* / Отв. ред. Л. Е. Гринин, А. В. Марков, А. В. Коротаев, с. 5–43. М.: ЛИБРОКОМ).

Grinin L. E., de Munck V. C., and Korotayev A. V. (Eds.) 2006. *History & Mathematics: Analyzing and Modeling Global Development.* Moscow: KomKniga/URSS.

Jantsch E. 1980. *Self-organizing Universe*. Oxford: Pergamon.

Marx K., and Engels F. 1976 [1846]. *The German Ideology.* 3rd rev. ed. Moscow: Progress Publishers.

Reeves H. 1981. *Patience dans l'Azur: l'evolution cosmique*. Paris: Editions du Seuil.

Sagan C. 1980. *Cosmos.* New York, NY: Random House.

Turchin P., Grinin L. E., de Munck V. C., and Korotayev A. V. (Eds.) 2007. *History & Mathematics: Historical Dynamics and Development of Complex Societies*. Moscow: KomKniga/URSS.

I. PATTERNS OF BIG HISTORY

1
Big History Trends and Patterns

Eric J. Chaisson

Abstract

Evolution – ascent with change of Nature's many varied systems – has become a powerful unifying concept throughout the sciences. In its broadest sense, cosmic evolution, which includes the subject of Big History, comprises a holistic explanatory narrative of countless changes within and among organized systems extending from the Big Bang to humankind. This interdisciplinary scenario has the potential to unite physical, biological, and social sciences, thereby creating for people of all cultures at the start of the new millennium a consistent, objective, and comprehensive worldview of material reality.

Keywords: *cosmic evolution, Big History, Universe.*

Historians

A few years ago, while having lunch in the Harvard Faculty Club with a group of science colleagues, I overheard a dispute among scholars at the table next to us. Several famous historians were squabbling about a frivolous territorial issue in their ancient and honorable discipline: Who studies history further back in time? The Greco-Roman expert maintained that the roots of his subject went way back, at least several thousand years. The Egyptian scholar argued that her studies involved events that were surely older, perhaps predating those of ancient Greece by a thousand years or more. And the Sumerian specialist tried to trump them all by claiming that his subject starts even earlier, maybe as long ago as 7,000 years.

As they heatedly volleyed their arguments back and forth with growing indigestion, I could not resist interrupting the historians – an intrusion they did not appreciate, for what right did I, as a scientist, have to say anything of use or interest to them. When I asserted, as an astrophysicist who looks out into space and thus back into time, that I was a 'real historian' whose studies extend into the past to nearly the beginning of time some 14,000,000,000 years ago, they became visibly upset. Their statements had been rendered nonsense, their subject matter reduced to minutia in the larger scheme of all history. At least one of those distinguished historians has not spoken to me since.

Big Historians

The recent onset of the new and exciting subject of Big History has brought forth an outgoing and refreshing breed of historical scholars. Their stories are inspiring, their outlook is uncommonly broad, and their attitudes open to new ideas, big ideas, indeed ideas central to fields well beyond their own. Big historians are helping to show that history writ large comprises many, diverse, yet related events that transpired well before those of recorded history, often extending back virtually to the beginning of time. This is not to say that I concur with all the words and assertions of the Big historians. As a natural scientist, I often experience a mild reaction to their subjective inquiry, indeed I have been trained in quite different methods of scholarship that emphasize objectivity. My scientific work needs to be confirmed with empirical data, or at least be based on statements that are experimentally or observationally testable. Skepticism and validation are my central dogmas.

Nonetheless, it is easy for me to admire the emergence of Big History, whose practitioners are willing and able to cross disciplinary boundaries and whose subject name is simple, clear, and unpretentious. By studying past events that gave rise to humanity on Earth, indeed to Earth itself among the stars and galaxies, Big historians naturally address Nature; to be sure, Big history was once historically called just that – natural history, which is usually defined as 'the study of natural objects and their evolution, origins, description, and interrelationships'. And since I have always regarded natural history expansively as a long and continuous narrative from the early Universe to the present time, not only incorporating the origin and evolution of a wide spectrum of systems and structures but also connecting many of them within an overarching intellectual framework, it is intuitive for me to relate favorably to their important work.

That said, even Big historians' work is limited. Big History, as most often defined – 'human history in its wider context' (Christian 2004) or 'an approach to history that places human history within the context of cosmic history' (Spier 2010) – pertains mostly to the meandering cosmic trek that led specifically to us on Earth. As such, it mainly concerns, in reverse order of appearance, changes that led to humankind, the Earth, the Sun, and the Milky Way Galaxy. Scant treatment is given, or need be given, to other galaxies, stars, or planets throughout the almost unimaginably vast Universe, for the goal of Big History is to place humanity itself into a larger cosmic perspective. Furthermore, Big historians especially need not be burdened with claims of multiple universes on macro-scales unimaginably larger than even those conceived by most physicists today, or of string theory and extra dimensions on micro-scales fully 20 orders of magnitude smaller than anything we can now measure – least of all that we and everything around us are cyberspace avatars in an alien computer simulation running an infinity of parallel worlds and implying that all possible histories conceivable are occurring somewhere, and maybe even everywhere an infinite number of times – none of which mathematical notions currently have any empirically supporting evidence whatsoever (Penrose 2010; Greene 2011).

In declaring these caveats, I wish neither to belittle Big History nor to critique those colleagues who prefer to speculate about the life and times of meta-events beyond the confines of our 14-billion-year-old Universe. Rather, I seek to make clear that most natural scientists still embrace the definition that 'the Universe is all that there is: the totality of all known or supposed objects and phenomena, formerly existing, now present, or to come, taken as a whole', and to suggest that if Big historians are to make headway, indeed to be accepted by traditional historians, they ought to ground their research agenda on empirical facts and tested ideas, where possible, and to focus their subject matter on the role of humanity in the one and only Universe we know.

Cosmic Evolution

Big History is not new, although one might not realize it by reading its current (March 2011) Wikipedia entry; the impression given is that this subject was invented hardly 20 years ago by traditional historians who began realizing that history actually reached well back beyond the onset of civilization. Broad, interdisciplinary explications of natural history have been researched and taught by natural philosophers since Renaissance times, and the specific big-bang-to-humankind story of special interest to Big historians has been championed in recent decades largely by cosmologists, who arguably think more broadly than anyone else on Earth. It is the latter astronomers, who in modern times have christened their subject 'cosmic evolution', but which is alternatively known within various academic disciplines as macroevolution, universal history, and the epic of evolution. (My original, qualitative book exposition Chaisson 1981, was updated in 2006 and made quantitative in 2001; a recent readable summary and a technical review can be found at 2009a and 2009b, respectively – all referenced at the end of this paper.[1])

Cosmic evolution is the study of the sum total of the many varied developmental and generational changes in the assembly and composition of radiation, matter, and life throughout the history of the Universe. These are the physical, biological, and cultural changes that have generally produced galaxies, stars, planets, and life-forms – specifically, regarding Big History and its more limited coverage, the Milky Way, Sun, Earth, and life on our planet, especially human life. The result is an inclusive evolutionary synthesis bridging a wide variety of scientific specialties – physics, astronomy, geology, chemistry, biology, and anthropology – a genuine scientific narrative of epic proportions

[1] Most of my recent journal publications including those in the References can be downloaded from my research page: URL: http://www.tufts.edu/as/wright_center/eric/ericrsrch.html.

extending from the beginning of time to the present, from the Big Bang to humankind.

Nor is the general study of change itself new; its essence extends back at least 25 centuries when the philosopher Heraclitus arguably made the best observation ever while noting that ‘everything flows ... nothing stays’. This remarkably simple idea is now essentially confirmed by modern scientific reasoning and much supporting data – indeed the notion that change is ubiquitous in Nature is at the heart of cosmic evolution. Other researchers have addressed life and complexity in a cosmic setting, among them Chambers (1844), who anonymously wrote a pre-Darwinian tome of wide interdisciplinary insight, and Shapley (1930), who pioneered ‘cosmography’ that classified all known structures according to increasing dimensions. Spencer (1896) also broached the idea of growing complexity in biological and cultural evolution, Henderson (1913) regarded the whole evolutionary process, both physical and biological, as one and the same, Whitehead (1925) sought to broaden scientific thinking with his ‘organic philosophy’, von Bertalanffy (1968) championed a systems-theoretic approach to physical, biological, and social studies, and Shklovskii and Sagan (1966) popularized the idea of intelligent life in the cosmos. Later in the 20th century, several independent efforts came forth virtually simultaneously, as Sagan (1980), Jantsch (1980), Reeves (1981), and Chaisson (1981) all advanced the idea of complex systems naturally emerging with the pace of natural history.

Fig. 1 sketches Nature's different kinds of evolution atop the so-called ‘arrow of time’. These three evolutionary subsets constitute the whole of cosmic evolution: physical evolution → biological evolution → cultural evolution, each describing how, in turn, ‘islands’ of growing complexity emerged to become ordered systems, whether massive stars, colorful flowers, or busy cities. Regardless of its shape or orientation, such an arrow symbolizes the *sequence* of events that have changed systems from simplicity to complexity, from inorganic to organic, from chaos in the early Universe to order more recently. That sequence accords well with a long and impressive chain of knowledge linking seven major epochs in time – particulate, galactic, stellar, planetary, chemical, biological, and cultural – wherein each changed chronologically:

- elementary particles into atoms;
- atoms into galaxies and stars;
- stars into heavy elements;
- elements into organic molecules;
- molecules into life;
- life into intelligence;
- intelligence into cultured and technological civilization.

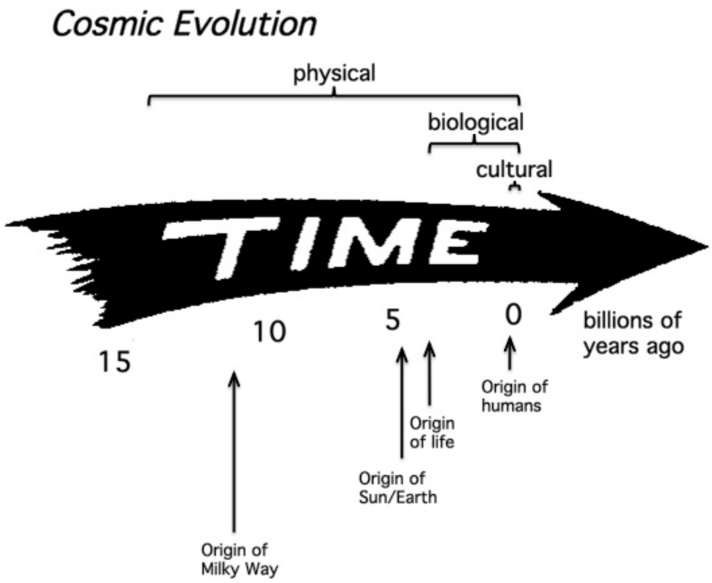


Fig. 1. An arrow of time symbolically chronicles the principal epochs of cosmic history, from the beginning of the Universe ~14 billion years ago (at left) to the present (at right). Labeled across the top are three major types of evolution (physical, biological, and cultural) that have produced, in turn, increasing amounts of order and complexity among material systems observed in the Universe. Cosmic evolution, as a general and inclusive term, comprises all of these subset evolutionary types and temporal phases

Despite the extreme specialization of modern science, evolution marks no disciplinary boundaries; cosmic evolution is a truly interdisciplinary topic. Accordingly, the most familiar kind of evolution – biological evolution, or neo-Darwinism – is just one, albeit important, subset of a broader evolutionary scenario stretching across all of space and all of time. In short, what Darwinian change does for plants and animals, cosmic evolution aspires to do for all things. And if Darwinism created a revolution in understanding by helping to free us from the anthropocentric belief that humans differ from other life-forms on our planet, then cosmic evolution extends that intellectual revolution by treating matter on Earth and in our bodies no differently from that in the stars and galaxies far beyond.

Anthropocentrism is neither intended nor implied by the arrow of time; the arrow is not pointing at humankind. Anthropic principles notwithstanding, no logic supports the idea that the Universe was conceived in order to produce specifically us. Humans are not the pinnacle or culmination of the cosmic-evolutionary scenario, nor are we likely the only technologically competent beings that have emerged in the organically rich Universe. The arrow merely provides an archetypal symbol, artistically conveying the creation of increasingly complex structures, from spiral galaxies to rocky planets to thinking beings.

Note, finally, that time's arrow does not imply that primitive, 'lower' life-forms have biologically changed directly into advanced, 'higher' organisms, any more than galaxies have physically changed into stars, or stars into planets. Rather, with time – much time – the environmental conditions suitable for spawning simple life eventually changed into those favoring the biological origin and evolution of more complex species. Likewise, in the earlier Universe, the physical evolution of environments ripe for galactic formation eventually gave way more recently to conditions conducive to stellar and planetary formation. And now, at least on Earth, cultural evolution dominates, since our local biospheric environment has once more changed to foster robust, societal complexity. Change in surrounding environments usually precedes change in organized systems, and the resulting changes for those systems selected to endure in Nature have *generally* been toward greater amounts of diverse order and inherent complexity.

Energy Flows and Complexity Rises

Of special interest to Big historians are the origin and evolution of the many diverse systems spanning the Universe today, notably those that sequentially and eventually gave rise to humanity on Earth. Particularly intriguing is the increase in complexity of those systems over the course of time, indeed dramatically so (with some exceptions) within the past half-billion years since the Cambrian period on our planet. Both theory and experiment, as well as computer modeling, suggest that islands of increasingly ordered complexity – namely, open, non-equilibrium systems that mainly include galaxies, stars, planets, and life-forms – are numerically more than balanced by great seas of growing disorder elsewhere in the environments beyond those systems. All emergent systems engaged in the cosmic-evolutionary scenario agree quantitatively with the valued principles of thermodynamics, especially its entropy-based 2nd law (Chaisson 2001).[2] Yet what has caused the emergence of systems and their rise in

[2] The 1st law (or principle, or the beginning) of thermodynamics is that the amount of energy in a closed system remains constant, *i.e.* the amount of energy in the Universe is also constant. And the 2nd law of thermodynamics states that spontaneous changes of energy are directed from its uneven to even distribution, *i.e.* the energy moves from the object with large amount of energy to

complexity over time, from the early Universe to the present? Is there an underlying principle, general law, or ongoing process that creates, organizes, and maintains all complex structures in the Universe?

Briefly stated and while keeping technicality minimized, I have suggested for at least a quarter-century that energy flows are at the heart of the cosmic-evolutionary story (Chaisson 1987, 2001, 2004). In particular, specific energy flow (*i.e.* energy rate per unit mass) constitutes a useful complexity metric and potential evolutionary driver for all constructive events throughout universal history. Energy does seem to be a common currency among all such ordered structures; whether living or not, all complex systems acquire, store, and express energy. Energy flow may well be the most unifying process in all of science, helping to provide a cogent explanation for the onset, existence, and complexification of a whole array of systems – notably, how they emerge, mature, and terminate during individual lifetimes as well as across multiple generations.

The chosen metric, however, can be neither energy alone nor even merely energy flow. Life on Earth is surely more complex than any star or galaxy, yet the latter utilize much more total energy than anything now alive on our planet. Accordingly, I have normalized energy flows in complex systems by their inherent mass, thus better enabling more uniform analysis while allowing effective comparison between and among virtually every kind of system encountered in Nature. This, then, has been and continues to be my principal working hypothesis in cosmic evolution: mass-normalized energy flow, termed energy rate density and denoted by Φ_m, is possibly the most universal process capable of building structures, evolving systems, and creating complexity throughout the Universe (Chaisson 2003).

Fig. 2 summarizes much recent research on this subject (Chaisson 2010, 2011), depicting how physical, biological, and cultural evolution over ~14 billion years have changed simple primordial matter into increasingly intricate and complex structures. (For specific power units of W/kg, divide by 10^4.) Values plotted are typical for the general category to which each system belongs, yet as with any eclectic, unifying theme in an imperfect Universe – especially one like cosmic evolution that aspires to address all of Nature – there are variations. And it is likely that from those variations arose the great diversity among complex, evolving systems everywhere.

---

the object with small amount of energy or it means that entropy of the Universe steadily increases (*Editor's note*).

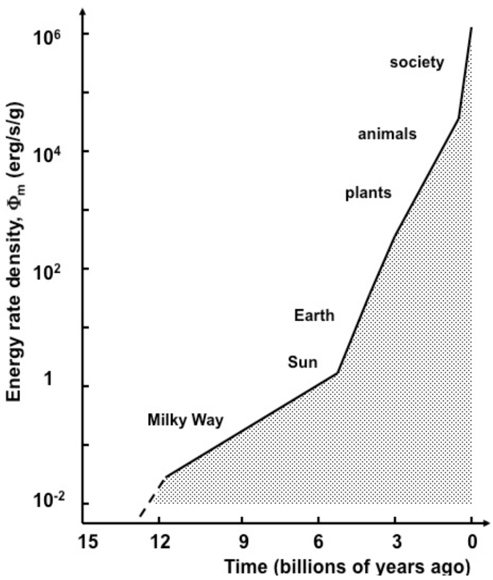


Fig. 2. Energy rate densities, Φ_m, for some complex systems of special interest to Big historians, plotted here semi-logarithmically at the time of their origin, display a clear increase during the ~14 billion-year history of the Universe. The shaded area includes an immense array of changing Φ_m values as myriad systems evolved and complexified (Chaisson 2010, 2011, and 2013)

Following the graphed trend in Fig. 2, which addresses complex systems of greatest interest to Big historians concerned with the specific evolutionary path that likely led to our human society, I have found systematic increases in the energy rate density (expressed here in the metric units of erg/s/g, evaluated against time in billions, millions, and thousands of years ago, Gya, Mya, and kya, respectively):

Within physical evolution:

• The Milky Way Galaxy evolved from protogalactic blobs > 12 Gya ($\Phi_m \approx$ $\approx 10^{-3}$ erg/s/g), which became widespread dwarf galaxies (~10^{-2}), then a mature, normal galaxy ~10 Gya (~0.05), and currently our galaxy's present state (~0.1).

• The Sun evolved from a protostar ~5 Gya ($\Phi_m \approx 1$ erg/s/g) to become a main-sequence star currently (~2), and will continue evolving to subgiant status ~6 Gya in the future (~4), eventually terminating as an aged red-giant star (~10^2).

Within biological evolution:

- Plants evolved from microscopic protists > 470 Mya ($\Phi_m \approx 10^3$ erg/s/g), to seedy gymnosperms ~350 Mya (~5×10^3), to flowering angiosperms ~125 Mya (~7×10^3), and to highly efficient C_4 plants ~30 Mya (~10^4).
- Animals evolved from fish and amphibians 370–500 Mya ($\Phi_m \approx 4\times10^3$), to cold-blooded reptiles ~320 Mya (~$3x10^3$), to warm-blooded mammals ~200 Mya (~4×10^4), and to flying birds ~125 Mya (~9×10^4).

Within cultural evolution:

- Human society evolved from hunter-gatherers ~300 kya ($\Phi_m \approx 4\times10^4$ erg/s/g), to agriculturists ~10 kya (~10^5), to industrialists ~200 ya (~5×10^5), and to technologists of today (~2×10^6).[3]
- Machines evolved from primitive devices ~150 ya[4] ($\Phi_m \approx 10^5$ erg/s/g), to the invention of automobiles ~100 ya (~10^6), to the development of airplanes ~50 ya (~10^7), and to modern jet aircraft and their computers (~5×10^7).

Or, for those readers who prefer words devoid of numbers, a simple 'translation' of the above technical summary suggests a ranked order of increasingly complex systems across the many successive phases of cosmic evolution:

- mature galaxies are more complex than their dwarf predecessors;
- red-giant stars are more complex than their main-sequence counterparts;
- eukaryotes are more complex than prokaryotes;
- plants are more complex than protists;
- animals are more complex than plants;
- mammals are more complex than reptiles;
- brains are more complex than bodies;
- society is more complex than individual humans;
- machines are more complex than societies.

Better metrics than energy rate density may well describe each of the system categories within the more restricted domains of physical, biological, and cultural evolution that combine to create the greater whole of cosmic evolution, but no other single metric seems capable of uniformly describing them all. The significance of plotting on a single graph one quantity for such an enormously wide range of systems observed in Nature should not be overlooked. I am unaware of any other single quantity (Φ_m) that can characterize so exten-

[3] Perhaps, pointing the date 300 thousand years ago, the author means the emergence of Neanderthals. However, the transition of primitive and ancient people to hunting and gathering took place much earlier and the people which are anatomically similar with modern people, appeared not earlier than 200 thousand years ago (see Markov 2012). (*Editor's note*).

[4] Machines appeared much earlier. Some machine-like-mechanisms can be mentioned referring to the period of Ancient times, but at the end of Middle Ages and beginning of Early Modern Period (the 14th – 16th centuries) one can mention about machines and even systems of machines in the proper sense of the word (see Grinin 2012). (*Editor's note*).

sively and uniformly so many varied complex systems spanning ~20 orders of magnitude in spatial dimension and nearly as many in time.

What seems inherently attractive is that energy flow as a universal process helps suppress entropy within increasingly ordered, localized systems evolving amidst increasingly disordered, wider environments, indeed a process that arguably governed the emergence and maturity of our galaxy, our star, our planet, and ourselves. If correct, energy itself is the mechanism of change in the expanding Universe. And energy rate density is an unambiguous, objective measure of energy flow enabling us to gauge all complex systems in like manner, as well as to examine how over the course of time some systems evolved to command energy and survive, while others apparently could not and did not. The optimization of such energy flows might well act as the motor of evolution broadly conceived, thereby affecting each of cosmic evolution's subset domains of physical, biological, and cultural evolution.

Teaching Cosmic Evolution

My philosophy of approach firmly grounds my research in empiricism, mines data from a wealth of observations, and aims to synthesize history in a seamless story that unifies much of what is actually known to exist in Nature. Fig. 2 contains a huge amount of data, computations, and modeling, summarizing many years of effort to interpret, at a quantitative level, my original exposition of the modern cosmic-evolutionary scenario (Chaisson 1981). Cosmic evolution has become a natural way for me to cross stultifying academic boundaries and to understand – at some level, in chronological order, and in a unified way – many of the complex, organized systems in the known Universe. To be honest, it has been a personal intellectual journey to learn about who I am and whence I came.

My interests in interdisciplinary science are deeply rooted in my earlier career, extending back several decades when I first arrived as a student at Harvard. It was then that I aimed to enroll in the course that I had always wanted to take, but found that it did not exist. I was seeking a broad survey course that cut across the boundaries of all the natural sciences, not only because I was unsure which of the sciences I might like later to study in depth but also because I was personally seeking an overarching, integrated worldview. I was eager to make sense of all that I saw around me in the air, land, sea, and sky, and I was especially struggling to place myself into the big picture of Nature writ large.

Sadly, nearly everyone I met 40 years ago – much as still the case today – was into 'their own thing'. Peers studied narrow disciplines, faculty researched specialized domains, and few people showed much interest in others' fields of knowledge. That universities are so lacking in universal learning and teaching was my biggest disappointment at the time, and still is. There had been a few earlier exceptions: Observatory director Harlow Shapley had taught a wide survey on 'cosmography' from the 1920s to the 1950s, and (my predecessor) Carl

Sagan had taught 'life in the universe' to big crowds in the 1960s; but by the time I arrived as a student, Shapley was dead, Sagan banished, and the broad course I sought was nowhere to be found in the Harvard curriculum.

Less than a decade later, when I was appointed to the Harvard faculty in the mid-1970s, I was fortunate to be able to co-(re)create that broad survey course along with a senior professor, George Field, who had also long wanted to teach the sciences in integrated fashion. We called the course 'cosmic evolution' and we resolved to make it intentionally 'a mile wide and an inch deep', regardless of expected criticism. This would be a true survey of the sciences from big bang to humankind – an interdisciplinary sweep across physics, astronomy, geology, chemistry and biology, with social studies included as well. We were unsure if any students would show up.

Within three years, Cosmic Evolution had become the largest science course on the Harvard campus, limited only by the fire codes of the biggest lecture hall. Its immediate acceptance and rapid growth were partly due to our having taken the art of teaching seriously, but mostly because students 'voted with their feet'. When asked, the students were quick to reply that they, too, were seeking the bigger picture – trying to grasp a larger perspective of all else studied at college, and especially trying to create for themselves a grand system of understanding.

I have now taught cosmic evolution at Harvard for 28 of the past 35 years since its creation, almost all of those years (as now) alone. For the first few years, I imported many guest speakers, including Steve Jay Gould, E. O. Wilson, George Wald, and several other experts outside my own expertise of physical science. The guest talks were fine as individual appearances, but together they lacked educational continuity. So, when I received a Sloan Fellowship in the 1980s, I surprised my colleagues by using those funds to take a year's leave to learn for myself all the science needed to teach the epic myself. Solo teaching of the course has led to much greater satisfaction personally as it has forced me to keep abreast of advances in a wide spectrum of subject areas; and it has provided much richer pedagogy and continuity for attending students by having a single person present the bulk of the course content[5]. A few years ago, after many unsuccessful attempts to inaugurate a course in cosmic evolution at Tufts University (owing to the usual turf battles with specialized faculty members), I finally succeeded while co-conspiring with a senior scientist, David Walt, provided that the course was team-taught by representatives from each of the science departments. Today, 'From the Big Bang to Humankind' is a popular offering at Tufts, where I co-teach it with an organic chemist, glacial geologist, developmental biologist, and cultural anthropologist. Such a team effort does lack educational continuity from speaker to speaker, but its decided advantage

[5] This course's syllabus and multi-media web site are freely accessible at URL: http://www.tufts.edu/as/wright_center/cosmic_evolution.

is that students meet a variety of leading researchers, each of whom has substantial expertise in their respective disciplines.

Our principal reason for creating this broad survey course at Tufts is that a distinct minority of students there studies natural science. Although about a third of incoming freshman each year indicates intent to major in math/science, less than 10 % graduate with a degree in it. It is not much different on many college campuses across the nation – Americans are opting out of science in droves. My contention has been – to the distress of many colleagues – that the science faculty is the main problem here. Blame need not be placed on elementary-school teachers or high-school curricula; rather, it is more likely that college professors, having shirked our duties to teach well, to teach broadly, indeed often to teach at all at the introductory level, have abrogated our responsibilities to disseminate the excitement and enthusiasm that we have for our subjects.

Even so, the hope was that such a survey that sweepingly integrates many science disciplines would renew student interest in science – and it most certainly has. The numbers are rising, the students once again voting with their feet. And they are very much inspired by the big picture, as witnessed this past semester when, after one of my lectures, a young woman paid me high tribute while remarking with tears in her eyes, ‘Thank you for helping me remember the love I once had for science’. That is the kind of sentiment that makes teaching this stuff for 35 years worthwhile!

Summary

The subject of cosmic evolution has been at the core of my entire academic career. It is the only thing I know – yet fortunately it includes vast facts, ideas, and implications. As I built the course at Harvard over decades (along with its extensive suite of online-supporting materials), my scholarly research agenda gradually shifted from mainstream astrophysics to fully embrace this interdisciplinary topic, and the science-education program that I direct at Tufts' Wright Center has adopted it as our intellectual theme. What started out as a search for a single course by a wandering student displaying hardly more than persistent curiosity became a life-long pursuit to understand our world, our universe, and ourselves.

Even after decades of researching, teaching, and writing about the epic of evolution, I am still unsure if I know who I am or how I really fit into the larger scheme of things. But I have found a lifetime of satisfaction exploring the general theme of cosmic evolution, publishing quantitative science to bolster the big-bang-to-humankind story, and especially sharing the details, excitement, and significance of that awesome story with countless people eager to discover their own worldviews. It has been, for me, the best of all scholarly endeavors: I have selfishly sought to know myself, yet in the process I have apparently helped myriad others to explore themselves and their sense of place in the amazing cosmos.

References

Chaisson E. J. 1981. *Cosmic Dawn: The Origins of Matter and Life*. Boston, MA: Little-Brown.

Chaisson E. J. 1987. *The Life Era*. New York: Atlantic Monthly Press.

Chaisson E. J. 2001. *Cosmic Evolution: The Rise of Complexity in Nature*. Cambridge – London: Harvard University Press.

Chaisson E. J. 2003. A Unifying Concept for Astrobiology. *International Journal of Astrobiology* 2: 91–101.

Chaisson E. J. 2004. Complexity: An Energetics Agenda. *Complexity* 9: 14–21.

Chaisson E. J. 2006. *Epic of Evolution: Seven Ages of the Cosmos*. New York: Columbia University Press.

Chaisson E. J. 2009a. Cosmic Evolution – State of the Science. *Cosmos and Culture* / Ed. by S. Dick, and M. Lupisella, pp. 3–23. Washington: NASA Press.

Chaisson E. J. 2009b. Exobiology and Complexity. *Encyclopedia of Complexity and Systems Science* / Ed. by R. Meyers, pp. 3267–3284. Berlin: Springer.

Chaisson E. J. 2010. Energy Rate Density as a Complexity Metric and Evolutionary Driver. *Complexity* 16: 27–40. DOI: 10.1002/cplx.20323.

Chaisson E. J. 2011. Energy Rate Density II: Probing Further a New Complexity Metric. *Complexity* 1 March. DOI: 10.1002/cplx.20373.

Chaisson E. J. 2013. Using Complexity Science to Search for Unity in the Natural Sciences. *Complexity and the Arrow of Time* / Ed. by Ch. H. Lineweaver, P. C. W. Davies, and M. Ruse, pp. 68–79. New York: Cambridge University Press.

Chambers R. 1844. *Vestiges of the Natural History of Creation*. London: J. Churchill.

Christian D. 2004. *Maps of Time.* Berkeley, CA: University of California Press.

Greene B. 2011. *The Hidden Reality*. New York: Knopf.

Grinin L. E. 2012. Global Technological Transformations. *Universal and Global History. Evolution of the Universe, Earth, Life and Society* / Ed. by L. E. Grinin, I. V. Ilyin, and A. V. Korotayev, pp. 383–403. Volgograd: Uchitel. *In Russian* (Гринин Л. Е. Глобальные технологические трансформации. *Универсальная и Глобальная история. Эволюция Вселенной, Земли, жизни и общества* / Ред. Л. Е. Гринин, И. В. Ильин, А. В. Коротаев, с. 383–403. Волгоград: Учитель).

Henderson L. 1913. *Fitness of the Environment*. New York: Macmillan.

Jantsch E. 1980. *Self-Organizing Universe*. Oxford: Pergamon.

Markov A. V. 2012. Anthropogenesis – A Special Chapter of Global History. *Universal and Global History. Evolution of the Universe, Earth, Life and Society* / Ed. by L. E. Grinin, I. V. Ilyin, and A. V. Korotayev, pp. 295–314. Volgograd: Uchitel. *In Russian* (Марков А. В. Антропогенез – особая глава глобальной истории. Универсальная и Глобальная история. Эволюция Вселенной, Земли, жизни и общества / Ред. Л. Е. Гринин, И. В. Ильин и А. В. Коротаев, с. 295–314. Волгоград: Учитель).

Penrose R. 2010. *Cycles of Time*. London: Bodley Head.

Reeves H. 1981. *Patience dans l'Azur: l'evolution cosmique*. Paris: Editions du Seuil.

Sagan C. 1980. *Cosmos*. New York: Random House.

Shapley H. 1930. *Flights from Chaos*. New York: McGraw-Hill.

Shklovskii I. S., and Sagan C. 1966. *Intelligent Life in the Universe*. San Francisco, CA: Holden-Day.

Spencer H. 1896. *A System of Synthetic Philosophy*. London: Williams and Norgate.

Spier F. 2010. *Big History and the Future of Humanity*. Chichester: Wiley-Blackwell.

Von Bertalanffy L. 1968. *General System Theory*. New York: Braziller.

Whitehead A. N. 1925. *Science and the Modern World*. New York: Macmillan.

2
Potential Nested Accelerating Returns Logistic Growth in Big History

David J. LePoire

Abstract

The discussions about the trends in rates of change, especially in technology, have led to a range of interpretative models including accelerating rates of change and logistic progress. These models are reviewed and a new model is constructed that can be used to interpret Big History. This interpretation includes the increasing rates of the evolutionary events and phases of life, humans, and civilization. These three phases, previously identified by others, have different information processing mechanisms (genes, brains, and writing). The accelerating returns aspect of the new model replicates the exponential part of the progress as the transitions in these three phases started roughly 5 billion, 5 million, and 5 thousand years ago. Each of these three phases might be composed of a further level of about six nested transitions with each transition proceeding faster by a factor of about three with corresponding changes in free energy flow and organization to handle the increased generation rate of entropy from the system. Nested logistic transitions have been observed before, for example in the ongoing exploration of fundamental physics, where the progress so far suggests that the complete transition will include about 7 nested transitions (sets of subfields). The reason for this number of nested transitions within a larger transition is not known, although it may be related to the initial step of understanding a fraction of the full problem. Too small of an initial fraction would lead to incomplete problem scope and definition. Too large of an initial step would lead to complications between the development of basic understanding and higher level derivations. An original step of one-seventh of the problem ends up within one standard deviation from the inflection point (mid-way through the transition).

Keywords: *Big History, logistical growth, complex adaptive systems.*

Current Technological Trends

The forecasts of the near future vary widely in scope and outlook, predicting from near utopia to near dystopia. The issues of great concern during this period include: (1) the energy transition problem of moving from an unsustainable fossil fuel-based economy to something else; (2) the widespread nature of the problems currently being discussed in terms of global warming, global trade, global terrorism, and global knowledge transfer; and (3) the possible opportunities and risks of new technologies such as genetics, nanotechnology, and artificially intelligent computers and robots (Bainbridge and Roco 2005). To gain a wider perspective on this transition, this paper further explores the transitions involving energy, environment, leadership and new technologies (Tainter 1988; Diamond 1997; Ponting 2007; LePoire 2010a) in time scales from the current era, modern history, extended past, and potential future.

Recently, various interpretations of trends in technological progress have led to widely differing predictions. Specifically, Ray Kurzweil (2005) hypothesized an ever-increasing rate of technological change, based on his analysis of over a century of progress in computation technologies. Theodore Modis (2002) hypothesized a very different future, one having a decreasing rate of technological change, based on the analyses of events from the 'Big Bang' to the present. Kurzweil investigated the more recent technological acceleration of computing performance. The inclusion of early electronic technologies, such as relays and vacuum tubes, led Kurzweil to propose that the rate of technological change is increasing with time, that is, Moore's Law of the doubling of electronic device densities every 18 months will be surmounted by new technologies that double in performance in less time. An ever increasing rate of technological change could soon lead to a technological 'singularity'. One attempt at a definition of the technological singularity is a 'future time when societal, scientific, and economic change is so fast we cannot even imagine what will happen from our present perspective, and when humanity will become posthumanity' (Vinge 1993).

Another model of technology progression and diffusion that has been studied is based on the logistic equation. This progression assumes that the rate of progress is proportional to both the current level of complexity and the fraction of complexity yet to be discovered. Logistic analysis has been found not only in market adoption and substitution of new products, but also in technology development and ideas (Marchetti 1986, 1980) such as democracy and energy. Theodore Modis (2002) suggests that the history of the Universe might also be considered as a logistic development of complexity. He arranged important events in the history of the Universe from a variety of sources, assumed that each event was equally important, and then made the assumption that the complexity of an event is its importance divided by the transition time to the next event. The dependence of the cumulative fraction of complexity on milestone number (not the event's time) could be interpreted either as (1) the first half of a logistic curve, or (2) a sequence of events that will culminate in a singularity. Modis favored the logistic development interpretation.

These two scenarios can be related to different simple models: Kurzweil's singularity scenario, with continual increasing exponential progress, might derive from a simple complex model, whereas Modis's long-term logistic growth with a tipping point determined by limitations in the learning rate and energy extraction rate, might be related to the more complex but realistic model. If this latter transition is accurate, the rate of technological progress might peak and eventually slow with impacts for economics and leadership (LePoire 2008, 2014).

History may well form a large complex adaptive system (Jantsch 1980; Marchetti 1980; Perry 1995; Spier 1996, 2010). As systems progress, new options that arise for the systems may spontaneously bifurcate into two potential discrete states (see Fig. 1). While the simplest model of complex systems can be driven into chaos, more realistic models with limitations suggest a possible reversal of increasing complexity (Stone 1993). The emergent properties of an evolving complex system might display simple patterns despite the complicated underlying processes (Cohen and Stewart 1994). Another approach is to take a longer view of historical trends and phases. Carl Sagan (1977) presented the stages of information processing, progressing exponentially from the early Universe to the present day. These stages were the development of life, brains, and technology, starting with life origins about 5 billion years ago. A geometrical progression rate would suggest transitions from life evolution to brain evolution around 5 million years ago and further transition to civilization and technological development about 5,000 years ago. The characteristic properties of complex adaptive systems include: (1) a resource which drives the level of complexity, such as energy use (Chaisson 2004); (2) new options at critical stages along development paths (Jantsch 1980); and (3) competition and learning as the options are explored (Dyke 1987).

Complex adaptive systems are found in a variety of fields and display a range of common emergent phenomena, such as bifurcations or transitions (Kauffman 1995). These transitions occur when an input to the system, such as energy flow, increases beyond critical levels. Studies of physical systems far from thermal equilibrium suggest that the energy flow is important in the development of more complex organization. Chaos and adaptation have been investigated in many natural developing systems such as biological evolution, ecosystems, and social systems. In such dissipative systems order can spontaneously form only when the system is maintained far from equilibrium, as measured by the usable energy that flows through a system.

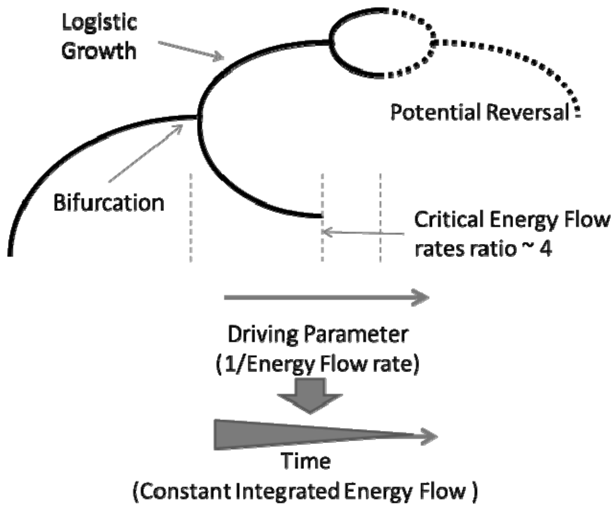


Fig. 1. Characteristics of an evolving Complex System. As the driving parameter (related to energy) increases over time, the organization of the system reaches a point where it is unstable but instead can grow in one of two ways. The growth between bifurcations is logistic

An approach to explore whether these processes are occurring (LePoire 2010b) include: (1) investigating the intensity and timing of energy flow organization; (2) investigating simple systems dynamics models; (3) rate of critical events in human historical transitions; (4) exploring the nested geometrical transition periods and energy uses in biological, mind and technology evolution; (5) exploring possible indications of a reversal in rate of change in fundamental physics and environmental issues; (6) examining the length scales of interacting systems and fundamental agents from the Big Bang to present.

A time contraction factor of about 3 is similar to time and energy contraction factors found by Snooks (2005) and Bejan and Zane (2012). This time contraction factor was used in describing the changes in energy intensity (Fox 1988; Morowitz 2002; Niele 2005; Chaisson 2004; Smil 1994; Bernstein 2004), as summarized in Table 1. Note that just one time contraction factor was realized from the Big Bang to the beginning of life on Earth. The remaining three large phases of life, human, and civilization evolution are separated by bold lines with different shadings. Each of these major phases has five or six subphases. Note that six subphases with a contraction factor of the square root of 10 (about 3) give an overall contraction factor of 1,000 within each major phase.

Table 1. Possible way to organize changes in energy flows through extended evolution covering life, human, and civilization development

Transition Start (Years ago)	Description	Energy Change
15 Billion	Gravitational	Gravitational energy causes clumping and nuclear energy causes energy to be released and element formation
5 Billion	Planet/Life	Life first gathers energy through chemicals or thermal gradients. Later the light from the Sun is captured and turned into chemical energy
1.5 Billion	Complex Cells	Simple prokaryotes form symbiotic relationships to form a larger and more organized eukaryote cell
500 Million	Cambrian	Oxygen levels reach a concentration so that multicellular organisms can be supported. The many body types and survival strategies lead to rapid evolution
150 Million	Mammals	Animals move to land after plants. The larger temperature variations lead to a way to regulate temperature to ensure ability to be active throughout the day and seasons
50 Million	Primates	A generalist strategy using various food sources including fruits leads to greater energy to the brain
15 Million	Hominids	Further generalist strategies and social organization again leads to greater energy use by the brain
5 Million	Humans	Humans adapt to a changing climate by leaving the forest for the savannah along with the capability for walking to expand the range of natural resources
1.5 Million	Speech	Further social organization leads to an expanded food sources including scavenging
500,000	Fire	Fire improves the energy availability from food
150,000	Ecoadaptation	Humans move out into other ecosystems expanding the range of energy resources
50,000	Modern humans	The benefits of specialization and social organization are realized during the ice age

Transition Start (Years ago)	Description	Energy Change
15,000	Agriculture	Domestication of plants and animals leads to a more intense and reliable use of the land
5,000	Civilization	Organization at a city level allows risk reduction and order with increasing population
1,500	Commercial Revolution	Financial and mechanical technological techniques are applied and improved in a sustaining growth organization
500	Scientific/ Exploration	Exploration of lands and ideas leads to expanded energy resources
150	Industrial	Fossil fuel allows large amounts of resources to be used along with increasing specialization
50	Information	Control through systems and computers allows greater efficiency in the use of energy and handling of pollution

There are similarities and differences between this interpretation and previous papers in this series. Specific issues include: 1) the nature of the current inflection point (Panov 2011); 2) the emphasis on non-equilibrium dynamics including bifurcation or energy, technology and also social organization (Nazaretyan 2011); and 3) the organization of evolutionary trends into two or three phases (Grinin, Korotayev, and Markov 2011).

Alexander Panov (2011) also organized evolutionary history with 19 evolutionary crisis transitions with decreasing duration (by about a factor of 3). This is called the scaling law of evolution. If the trend continues, evolution would come to a very rapid rate of evolution at some point in time, the Singularity, which was predicted to occur somewhere within the past two decades. However, as he notes, the rate of evolution cannot approach this infinite rate but instead would be constrained by resources and the ability for evolutionary processes to work by testing various environmental fitness of technologies and cultures. This is similar to the law of Accelerating Returns by Ray Kurzweil. However, just as normal logistic transitions first start with exponential growth and later slow due to limitations to form the S-curve transition, the hyperbolic growth of evolution might also begin to slow. The combined law of accelerating returns and the logistic developed here is one way to model this important inflection point in evolution, the 'crisis of crises' as stated by Panov. Such an inflection would indicate the conflict between conventional economic and resource growth and constraints of global resources, pollution, population, and conflicts.

Panov (2011) continues to discuss the possible 'end' of science. The complexity of the increasingly intertwined processes of science, development, and production might be reflected in the limits of scientific growth within the constraints of society. An outgrowth of this entanglement might be seen in the recent organization of technological hubs which combine research into the scientific basis, product development, and production of new technologies. For example, the U.S. Department of Energy has formed many exploratory technological innovation hubs to pursue new energy technologies such as energy storage (batteries) for both transportation and electrical grid buffering.

This nested logistical pattern, interpreted as alternating evolution of smooth exploration followed by intense reorganizing transitions is compatible with the view of Akop Nazaretyan (2011). He emphasized the non-equilibrium aspects of the evolution and the role of mental capacities to form new organizations that would lead through the crises to a new sustainable growth. It is not just the energy flow that increases complexity, but the ability to control the environmental impacts of that flow. Control was important in many transitions including the use of fire, the use of engineered water projects, and the current issues with potential climate change. He identified the 21st 'bifurcation' century as a test where a major inflection in the complexity, suggested by Big History, might raise many issues concerning conflict, energy, and the environment.

There is also the question of whether the evolutionary process occurred in three major phases discussed here – biological, human, and civilization, or two – biological and social phases as discussed by Grinin, Korotayev, and Markov (2011). The reason the phases were split into three in this paper is because of the three information storage and passing mechanisms identified (Sagan 1977), *i.e.* DNA, the human mind, and writing/artifacts. It is true that genetic changes were key in the development of humans but much of the learning was passed through social signals leading to the new tool of spoken language, while also beginning to control aspects of nature such as fire, dogs, and plants. Grinin *et al.* (2011) mention this aspect in the afterward of their paper by discussing intermediate subphases such as the biological-social type which would be similar to the human mind phase discussed here. Then the civilization, or pure social phase, is distinct because of the hierarchy of the social groups in combination with written records supporting the new organization.

Growth Pattern Characteristics

The first topic concerns the combination of accelerating returns model and the logistic model. The accelerating rate of return can be written $y' = ky^2$. This tends to grow faster than an exponential with a singularity at some point in time. The logistic equation can be written $y' = ky(1 - y)$, which has an inflection point but more linear progression of time scales. The super-exponential early growth can be combined with the logistic equation by combining the features of each $y' = k[y(1 - y)]^2$. This is shown in Fig. 2 in linear and log form. This is appropriate for a long-term development discovery system, whereas the simple logistic equation is more appropriate for diffusion of information. More insight can be gained by noting the exponential growth rate, *y'/y*, which is proportional to $1 - y$ for a traditional logistic growth, y for an accelerating returns growth, and $y(1 - y)^2$ for the combination.

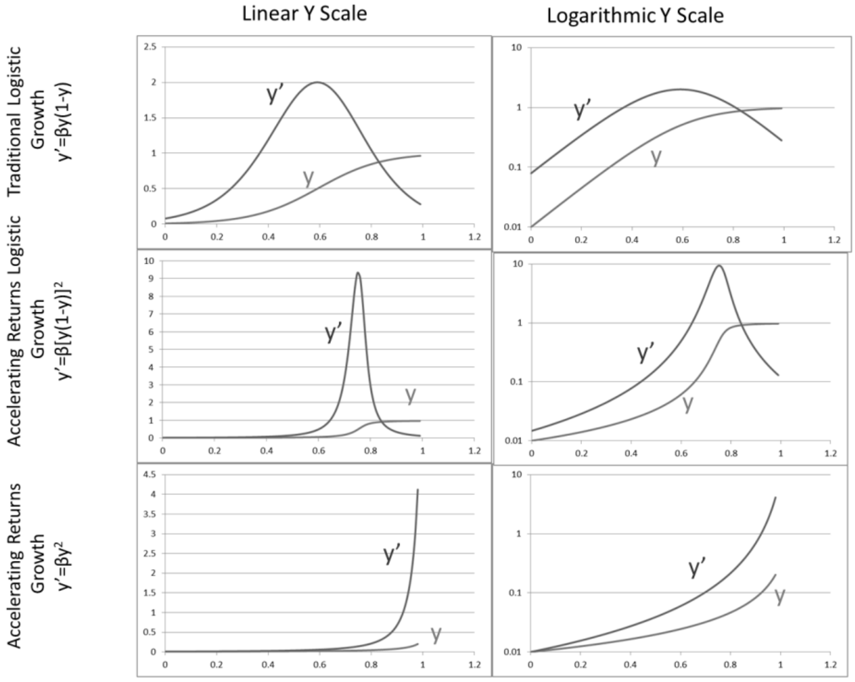


Fig. 2. Patterns of growth – Logistic (top), accelerating returns logistic growth (middle), and accelerating returns (bottom)

The second characteristic is the nested logistic growth pattern in the both overall information mechanism transition, and in energy flow. The historical complexity and energy use growth seems to have about 3 development phases with different information systems until the inflection point leading to a rate of 6 per full transition. Each of the three phases seems to be formed by 6 or 7 subphases where the information mechanism is the same (*i.e.*, DNA, brain, writing) but the energy flow is continually increasing.

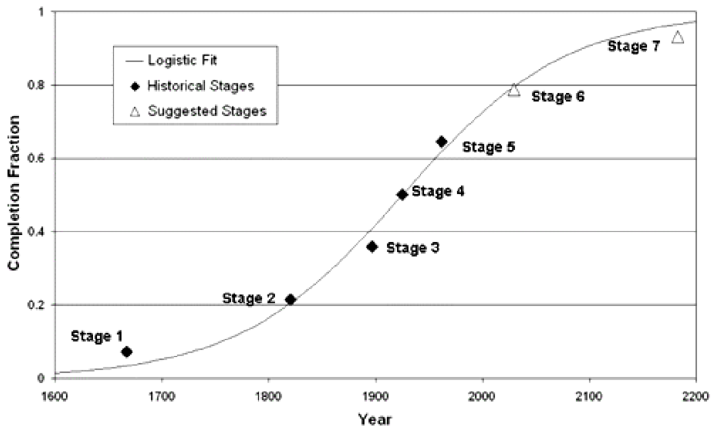


Fig. 3. Nested logistic growth in the area of fundamental physics. Each stage represents a logistic growth in at least one field of fundamental physics starting with gravitation and classical mechanics. Stage 4 occurred with the fastest pace as general relativity, quantum mechanics, and nuclear physics were being developed

This nested pattern of logistic growth with about 7 subphases was observed in the development of fundamental physics (see Fig. 3) (LePoire 2005). This analysis considered the rate of discoveries in various physics subfields. The subfields showed logistic growth patterns with initial slow progress followed by a period of steady progress, ending with another slow rate of progress as the subfield was fully integrated in a consistent manner. Within this one field of science, the nature of logistic development is seen in both the subfields and the complete field. If the logistic interpretation is correct and is followed, the data suggest that string physics is likely to be 50 % complete in 2030 and 80 % complete in 2090. However, if the development curve is logistic, then the development curve would be symmetric around the midpoint, identified as the 'Golden Age' of physics in the 1920s with the simultaneous developments in general relativity, quantum mechanics, and nuclear physics. This would imply that there should be symmetric stages that correspond to each other, that is, if three stages are identified before the midpoint, then there should be three after the midpoint. String physics is only the second identified stage, leading to the suggestion that another stage in the development of fundamental physics might come after string physics. If symmetry holds, the last stage's 20, 50, and 80 % completion times would be 2100, 2180, and 2260.

Why would a logistic transition be broken into 7 substeps? In the process of exploring a new field, one of the more difficult steps is the first in defining the subject, the scope, and process. While there were many discoveries in physics before Galileo, the sustained nature of scientific progress afterwards points to this difficulty. Some difficult concepts to frame were inertia – bodies tend to keep on moving, the relationship between force and acceleration (versus the force in the static mechanics). This led Galileo to the experiment with simple toy-like apparatus' like rolling balls down inclined planes and measuring objects at the end of strings. If the first step was in the wrong direction, too small or too large, the progress could be halted. Some of the Greek philosophers tried to solve everything in one hypothesis but were not able to defend it with measurements. If Galileo had observed the moons of Jupiter and then asserted the laws of the heavens had inertia but that did not apply on Earth, the rolling balls down incline planes would not have generalized the concept, *i.e.* it would have been too small. Instead the first phase of fundamental physics was the idea of laws of motion that applied on Earth and in the sky was set out by Galileo and took about 150 years through the great mathematics of the 18th century to develop the tools such as calculus, variational calculus, wave theory, and generalized laws of motion such as the Lagrangian. These tools that were developed in the first phase would become instrumental in further developments, for example, the Lagrangian formulation was instrumental in quantum mechanics whereas Newton's laws were found only applicable in classical mechanics.

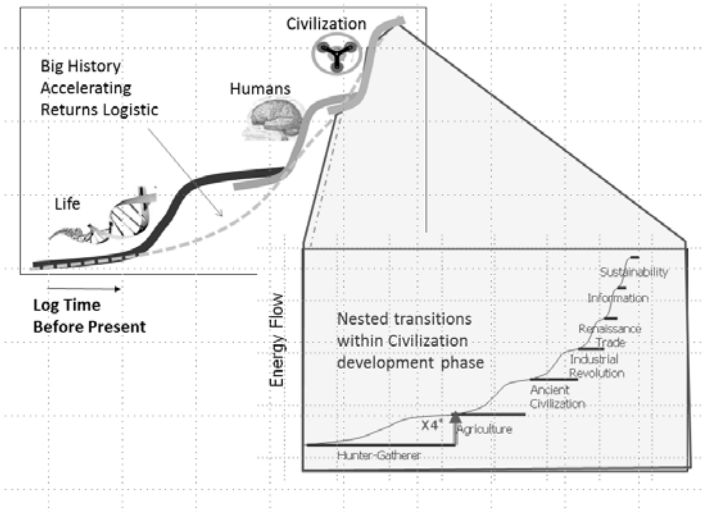


Fig. 4. Example how the major three phases of extended evolution might be considered to be each formed by a nested set of subphases (civilization subphases are shown here)

An appropriate first step in a nested transition might be big enough to tackle fundamental issues. For example, in the physics case above, the fundamental laws of physics were developed and the common force of gravity was explained.

However, the first step cannot be too big, due to the dependency of the steps. For example, again in the physics case above, the theory and experimental techniques developed in the first phase were a necessary prerequisite for the second step (electromagnetism) to be explored. An appropriate step size might be less than half of the transition. A simple measure of the width of the transition is its standard deviation, *i.e.* a characteristic duration of the full transition. So the first step size of 16 % of the full transition means that the second step would start one standard deviation from the inflection (mid-point) of the transition (see Fig. 5). If equal fractions of the problem (*e.g.*, 16 % as in the first step) are later tackled, it would take about 6 steps to complete the logistic transition.

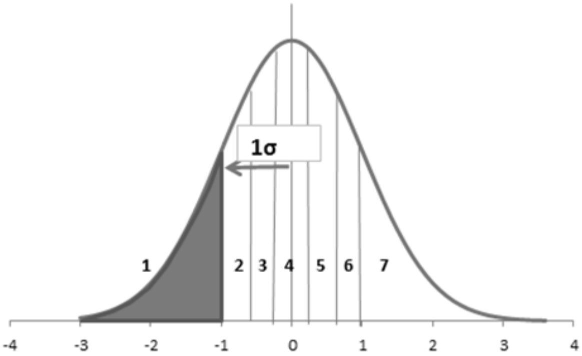


Fig. 5. Nested steps in approaching a normal distribution. If the full exploration of a phase is the whole normal distribution, one possible strategy to initiate the process is to break the problem into subphases, starting with a fraction that is neither too small nor too large. Here it shows that if the whole is divided into seven (or six) equal area sections, then the first section would represent exploring up to one standard deviation from the inflection point (middle of the normal distribution)

Possible Pattern Extensions

One way to project what a logistic world would look like after the inflection point would be to mirror past rates of change. For example, if 2000 was the inflection point, meaning that technological change continues for a while at the same rapid progress but does not accelerate, then the 20th century might be a good indication of some of the changes we might see in the 21st. In the 20th century there was such growth in population and resource use to move the world beyond sustainability. The major energy supply was from non-renewable fossil fuels. The expansion was driven by relatively inexpensive energy, new insights from physics leading to electrical, aerodynamic and material technology. These innovations led to creative destruction in capital formation and development as exemplified by the semiconductor industry which followed

Moore's Law. The technology also caused problems such as arms races, environmental impacts, and expensive medical options. The 21st century might adjust by slowing to a more sustainable society with more efficient energy use, a stable or declining population, reduction in the gap between incomes, and exploration of ways to mitigate global environmental impacts while maintaining robust fair trade.

While there are possible indications that a large transition such as a general technological slowdown is underway (LePoire 2014) there are reasons why we might be experiencing a transition inflection point in general evolution from a bootstrap natural evolution type system to a technologically design driven evolution which might show initial speed in development and then slow as leadership is assumed.

The problems with a bootstrap natural selection is that it has a difficult beginning and is stuck in historically determined structures, for example, the long period (billions of years) between the development of simple life and earth and multicellular life during the Cambrian explosion. The energy mechanisms, information storage and expression of biological systems were determined and mostly stable throughout further development. The systems also have difficulty scaling due to individual perspectives and limited information resources for collective action but instead are good at individual exploration, competition, and growth. The techniques for dealing with uncertainty are intuitive and based on evolutionary trajectories, for example, a simple intuitive fight or flight mechanism.

The assumption in the technological singularity is that technology will be able to quickly resolve human and biological constraints and continue an exponential growth path. However, this is unlikely as many constraints are tied to human systems and are also filled with uncertainty which might be better handled but will not be eliminated.

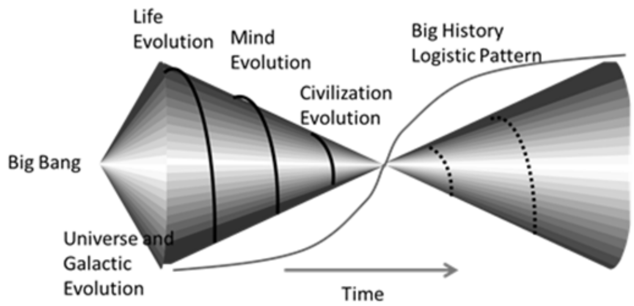


Fig. 6. Hypothesis of the Big History logistic growth pattern. The left hand side is what has been discussed in this paper. The trend towards shorter duration phases cannot continue. One way to extend the pattern is using an all-encompassing Big History logistic pattern which shows an inflection point where the rate of change is the largest

If the natural evolution is considered as an ever quickening set of phases from molecular evolutions, cellular evolution, multicellular, primate, human civilization and technological stages then it might be expected that the technological evolution might show similar phases. Since the design of the technology is driven by humans, there is a period where the two are tightly coupled. For example, the development of fundamental components similar to the molecular evolution for energy, information, and movement structures in organelles would be compared to the development of the understanding of theoretical logic, mechanical relays, electromechanical relays, transistor tubes, and isolated semiconductor transistors along with the resistors, capacitors, and inductors necessary for electronic designs. The next phase of integration, which might be compared with cellular evolution, was done at first in large computer rooms such as ENIAC, then IBMs with semiconductor transistors, but a major breakthrough in integration came with Texas Instrument's integrated chip technology which formed the ability for scalable integration and functionality. The network enabled quick growth of shared information and applications through the Internet, wireless and cellular networks might be compared to evolution of multicellular organisms. However, the Internet resulted in unintended consequences of malware, high dependence requiring high reliability, and ID theft. Concern over integrating with grids, decision support, and automated robotic support are leading to cautious rates of applications. It will be interesting if the artificial intelligence evolves at an accelerating rate or moves more cautiously as this artificial intelligence and robotics start complementing the intelligence that took billions of years to develop on Earth.

On a much longer time-scale the reflection of the rate of change would look something similar to Fig. 6. This paper focused on the left cone which includes the three phases (shown on a log time scale) of life evolution, human evolution, and civilization. If this follows the large logistic trends identified, the pattern in the future would look something like the cone on the right, again three phases with the characteristic slowing-down. However, it is not clear how and whether the logistic development pattern will continue.

Conclusion

The three major phases after cosmological development (*i.e.*, life, human, and civilization) had durations of about 1,000th that of the previous. Each might have six subtransitions with durations being reduced by a factor of 3. Energy and organization might qualitatively change between the transitions to handle the additional entropy flow through the systems. These characteristics seem to be consistent with the interpretation of a complex adaptive system with evolution through a sequence of bifurcations and logistic learning. A logistic accelerating rate of return logistic pattern, formed by combining the accelerating rate of return growth and the traditional logistic growth pattern shows similar exponentially shorter transitions. The pattern eventually reverses as has been observed in ecological systems. The decomposition of each phase into six transitions is motivated by the reorganization to maintain increasing energy flows. This nested logistic transition has been observed in the history of the discoveries in fundamental physics. While it is not known why a nested logistic transition would decompose into this number of subtransitions, it was observed that with seven subtransitions, the first step explores up to one standard deviation of the inflection point. Future inflection, symmetrical from boot-strapped to engineered system may gain more control over energy and physical systems.

References

Bainbridge W. S., and Roco M. 2005. *Managing Nano-Bio-Info-Cogno Innovations: Converging Technologies in Society.* Springer.

Bejan A., and Zane J. P. 2012. *Design in Nature: How the Constructal Law Governs Evolution in Biology, Physics, Technology, and Social Organization*. New York: Doubleday.

Bernstein W. J. 2004. *The Birth of Plenty: How the Prosperity of the Modern World Was Created.* New York: McGraw-Hill.

Chaisson E. J. 2004. Complexity: An Energetics Agenda; Energy as the Motor of Evolution. *Complexity* 9(3): 14–21.

Cohen J., and Stewart I. 1994. *The Collapse of Chaos: Discovering Simplicity in a Complex World*. New York: Viking.

Diamond J. M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. New York: W.W. Norton & Co.

Dyke C. 1987. *The Evolutionary Dynamics of Complex Systems: A Study in Biosocial Complexity*. New York: Oxford University Press.

Fox R. F. 1988. *Energy and the Evolution of Life.* New York: W. H. Freeman.

Grinin L. E., Korotayev A. V., and Markov A. V. 2011. Biological and Social Phases of Big History: Similarities and Differences of Evolutionary Principles and Mechanisms. *Evolution: A Big History Perspective* / Ed. by L. E. Grinin, A. V. Korotayev, and B. H. Rodrigue, pp. 158–198. Volgograd: Uchitel.

Jantsch E. 1980. *The Self-Organizing Universe: Scientific and Human Implications of the Emerging Paradigm of Evolution*. New York: Pergamon Press.

Kauffman S. A. 1995. *At Home in the Universe: The Search for Laws of Self-organization and Complexity*. New York: Oxford University Press.

Kurzweil R. 2005. *The Singularity is Near: When Humans Transcend Biology*. New York: Viking.

LePoire D. J. 2005. Application of Logistic Analysis to the History of Physics. *Technological Forecasting and Social Change* 72(4): 471–479.

LePoire D. J. 2008. Exploration of Connections between Energy Use and Leadership Transitions. *Systemic Transitions: Past, Present, and Future* / Ed. by W. Thompson, pp. 205–222. New York: Palgrave Macmillan.

LePoire D. J. 2010a. Long-Term Population, Productivity, and Energy Use Trends in the Sequence of Leading Capitalist Nations. *Technological Forecasting and Social Change* 77(8): 1303–1310.

LePoire D. J. 2010b. Exploration of Energy and Environment Roles in Driving Large-Scale Historical System Transitions. *Integrating Complexity: Environment and History.* London – Ontario: University of Western Ontario.

LePoire D. J. 2014. Potential Economic and Energy Indicators of Inflection in Complexity. *Evolution: Development within Big History, Evolutionary and World-System Paradigms* / Ed. by L. E. Grinin, and A. V. Korotayev, pp. 108–118. Volgograd: Uchitel.

Marchetti C. 1980. Society as a Learning System: Discovery, Invention, and Innovation Cycles Revisited. *Technological Forecasting and Social Change* 18: 267–282.

Marchetti C. 1986. Fifty-Year Pulsation in Human Affairs: Analysis of Some Physical Indicators. *Futures* 18(3): 376–388.

Modis T. 2002. Forecasting the Growth of Complexity and Change. *Technological Forecasting and Social Change* 69: 377–404.

Morowitz H. J. 2002. *The Emergence of Everything: How the World Became Complex*. New York: Oxford University Press.

Nazaretyan A. P. 2011. Mega-Evolution and Big History. *Evolution: A Big History Perspective* / Ed. by L. E. Grinin, A. V. Korotayev, and B. H. Rodrigue, pp. 82–100. Volgograd: Uchitel.

Niele F. 2005. *Energy: Engine of Evolution*. Amsterdam – Boston: Elsevier.

Panov A. D. 2011. Post-Singular Evolutions and Post-Singular Civilizations. *Evolution: A Big History Perspective* / Ed. by L. E. Grinin, A. V. Korotayev, and B. H. Rodrigue, pp. 212–231. Volgograd: Uchitel.

Perry D. A. 1995. Self-Organizing Systems Across Scales. *Trees-Structure and Function* 10(6): 241–244.

Ponting C. 2007. *A New Green History of the World: The Environment and the Collapse of Great Civilizations*. New York: Penguin Books.

Sagan C. 1977. *The Dragons of Eden: Speculations on the Evolution of Human Intelligence.* New York: Random House.

Smil V. 1994. *Energy in World History*. Boulder: Westview Press.

Snooks G. D. 2005. Big History or Big Theory? Uncovering the Laws of Life. *Social Evolution and History* 4(1): 160–188.

Spier F. 1996. *The Structure of Big History from the Big Bang until Today*. Amsterdam: Amsterdam University Press.

Spier F. 2010. *Big History and the Future of Humanity*. Chichester – Malden, MA: Wiley-Blackwell.

Stone L. 1993. Period-Doubling Reversals and Chaos in Simple Ecological Models. *Nature Cell Biology* 365: 617–620.

Tainter J. A. 1988. *The Collapse of Complex Societies*. Cambridge – New York: Cambridge University Press.

Vinge V. **1993.** The Coming Technological Singularity: How to Survive in the Post-Human Era. *Proc. Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*: 11–22. NASA: Lewis Research Center.

3
A Toy Model Mechanism for Greater-Than-Exponential Human Population Growth

Antony Harper

Abstract

The term, exponential, has long been associated with the growth of organismal populations from microbial populations to the populations of complex multicellular eukaryotes. It can be shown, however, that human population growth occurs at greater-than-exponential rates. Von Foerster et al. (1960) but followed more elegantly by Korotayev et al. (2006a), have proposed models to more accurately represent this characteristic mode of human population growth. In this paper an underlying mechanism is proposed which generates this greater-than-exponential growth. The mechanism is represented by a toy model of two differential equations of interacting populations, the interactions of which enhance the reproductive abilities of the other population. The end result of this enhancement due to positive human interaction, a quintessential characteristic of our species, is a pattern of growth motivated by a greater-than-exponential rate of growth. It should be noted that the model being proposed is one of many potential models and not the sole, the only, possible model.

Keywords: *population, exponential, greater-than-exponential growth (GTEG), hyperbolic.*

Introduction

In 1960 von Foerster *et al.* proposed a model of global population growth in which the form of the growth was greater than exponential. Korotayev, Malkov, and Khaltourina (2006a, 2006b) expanded on this model and showed clearly that the form of what they named hyperbolic growth fit the data of human population growth over the past ten thousand years quite well. Specifically, the integral form of the equation, $dN/dt = aN^2$, gives $N_t = a/(t_0 - t_n)$, where a is a fitted constant, t_n represents some time before t_0, and t_0 represents what von Foerster *et al.* called Doom's Day, the time at which, to use a very appropriate Russian phrase, the population enters its 'blow-up' phase, *i.e.* hyperbolic growth reaches a critical point in finite time. Note also that a hyperbolic growth model is also recognized as a valid fit for human population data by Joel Cohen in his book *How Many People Can the Earth Support?* (Cohen 1995).

In light of the fact that Korotayev *et al*. (2006a, 2006b) have shown how important this form of growth, labeled hyperbolic growth, is to understanding human demography over time, it will be important to begin to understand the mechanism behind such growth. It should be noted here that Korotayev *et al.* (*Ibid.*) have shown that this form of growth applies over both short- and long-term views of human population change ranging from a period of ten thousand years to much shorter periods of a few hundred years. Interestingly, if growth form does not change with time scale, an implication of such growth is that it is (probably) scale-free in context. This form of growth also manifests itself above the population level of biological organization (Markov and Korotayev 2007). However, that will not be the focus of this short paper.

The focus of this paper is to present a mechanism for the type of population behavior giving rise to the blow-up phase or regime, or stated another way, to give rise to greater-than-exponential growth (GTEG). (The acronym, GTEG, will be used throughout this paper to represent all patterns of population growth that are greater than exponential growth including but not limited to bi-exponential and hyperbolic growth.) Exponential growth will be compared to GTEG, and it will be shown that while the growth of most animal populations can be represented by exponential growth, human population growth cannot and cannot because of a quintessentially human characteristic, that of human interaction *which occurs* at a much higher level than the interaction between members of other animal species. Specifically, the toy model will be used to show that total population growth within a set of interacting sub-populations is greater than total growth that is the sum of non-interacting sub-populations.

Alternative Models of Population Growth

Exponential population growth is recognized as a basic form of growth exhibited by a variety of organisms (*e.g.*, Hutchinson 1978; Gotelli 2001). As noted previously, exponential population growth is given by the integral solution to the differential equation, $dN/dt = rN$, which yields, $N_t = N_0e^{rt}$, with N_0 = the initial size of the population, N_t = the size of the population at some future point in time, t, and r = a fitted constant which is the growth rate of the population. The log-transform of this equation is: $ln\ N_t = lnN_0 + rt$. It should be noted that the form of this equation is linear, and therefore raw data plotted either on semi-log graph paper or log-transformed population data plotted against time must give a straight line plot.

Both exponential growth and the log-transform of exponential growth are represented in the following graphs. In Fig. 1 one can see that the graph turns sharply upward and would continue to grow in that fashion if time units greater than 64 were used. In Fig. 2 the log-transformed population data are plotted against the same set of time values. Here the plot is linear, and this form is representative of all exponential growth models that are log-transformed. So, any log-transform of population data yielding a linear plot represents a log-transformation of data of an exponentially growing population. This would not be the case of population data of GTEG populations. In this case the log-transformation of the data would yield a curve similar in shape to the un-transformed exponential data. In Fig. 3 the raw data of a population growing at GTEG are represented, and in Fig. 4 the natural log-transformed data are plotted over the same time period.

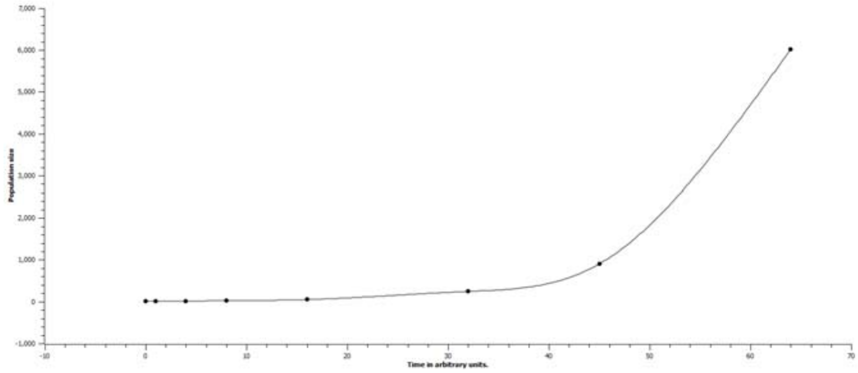


Fig. 1. This graph represents the growth of a population over 64 time units based on the equation, $N = 10e^{.1t}$, where the initial population size is 10, e is the base of the natural logarithms, and t represents the arbitrary units of time

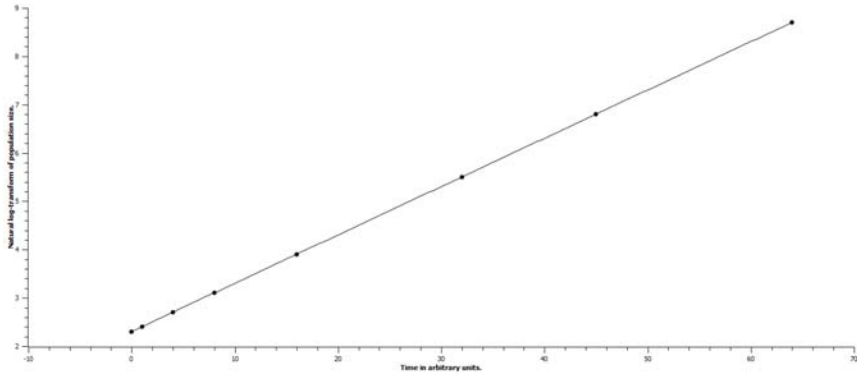


Fig. 2. This graph represents the natural log-transformed data of population size used in Fig. 1 plotted against time in arbitrary units. The equation used is: $\ln N = \ln 10 + .1t$. Variables are defined as in Fig. 1

In Fig. 3 the form of the curve is not unlike that of an exponential curve in that it begins to rise slowly and then toward the end of the time period accelerates in growth, however, when the natural log-transformed data are plotted, a curve rather than a straight line is produced (see Fig. 4.). In other words, mathematical words, the relationship can be represented by a power function, *i.e.* $lnN = at^b$, in which, when $b > 0$, the relationship is non-linear and when $b > 1$, the relationship will produce a curve which is concave up as in Fig. 4. This implies that the antilog-transformed equation, $N = Ae^{t^\wedge b}$, will grow at greater-than-an exponential rate as the exponent is not constant but is itself a power function.

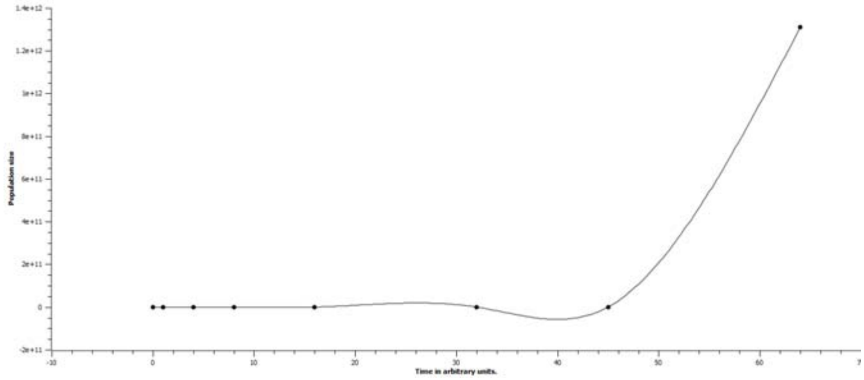


Fig. 3. Population growth in which the growth rate is GTE is represented in this graph. Note that the time period is the same as in Figs 1, 2, and 4. The equation used to generate this data is: $N_t = 10e^{.1t^\wedge 1.5}$

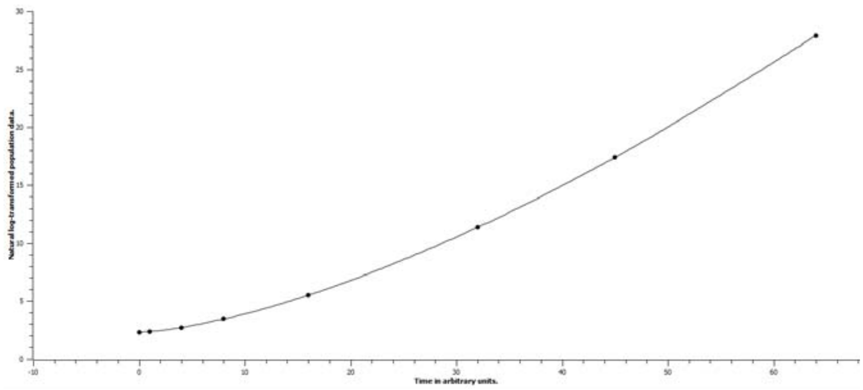


Fig. 4. This graph represents the natural log-transformed data of Fig. 3 and can be produced by the equation, $lnN = ln10 + .05t^{1.5}$. The plot is unquestionably non-linear

Real World-System Data

Using the previously defined types of growth, either exponential growth or GTEG, the data on population size over time for the world-system *in toto* can be evaluated to determine which of these two models more realistically represents the pattern of growth of an actual human population. In Figs 5 and 6 data for the world-system population from 800 CE through 2000 CE are plotted either as unaltered population size (see Fig. 5), or as natural log-transformed population size (see Fig. 6), in which both sets of data are plotted against time. Fig. 5 superficially resembles the curves represented in both Fig. 1 and Fig. 3. So, simply on visual inspection, it would be difficult to determine which of the two curves more appropriately matched that of Fig. 5. However, even casual inspection of the graph in Fig. 6 unquestionably shows that the log-transformed population data are not linear. In Fig. 6 both a linear and an exponential fit are represented, and even without the aid of formal statistical analysis, the exponential curve can be seen to be a much better fit. This implies that the rate of growth of the world-system population is GTEG. Interestingly, if the populations of other organisms are assessed, they are found to be exponential, so, what is there about the mechanism of human population growth that produces a GTEG pattern? This question will be addressed in the following section.

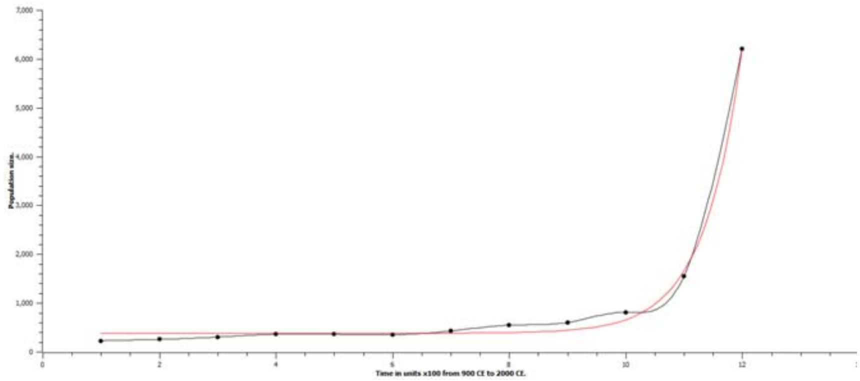


Fig. 5. Population size for the world system is graphed against time for the period, 900 CE to 2000 CE

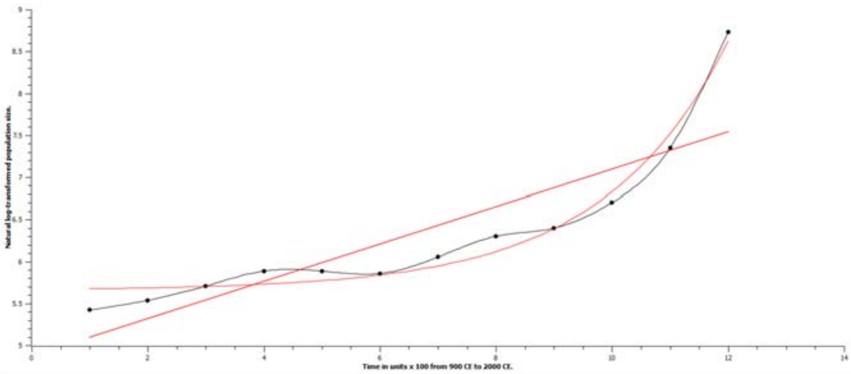


Fig. 6. The population data in Fig. 5 are natural log-transformed and plotted over the same period of time. Note that both a linear and an exponential regression are fitted to this data with the exponential regression the better fit of the two

A Toy Model Mechanism for GTEG

Based on the information of the previous section a generalized equation representing GTEG has the form, $N_t = Ae^{t^\wedge b}$, and the question then becomes: What is the reality of the exponent, b, of the exponent, t or time? Why is it that animal populations other than human can be represented more simply by the equation, $N_t = Ae^t$? Clearly, the log-transform of both equations gives, $ln\ N_t = lnA + t$ and $ln\ N_t = lnA + t^b$ respectively. The first transformed equation is linear, while the second one is exponential, and we need to consider what it is about humanity that gives the exponent of the exponent, b, its reality.

Humans are more closely connected with each other both locally and at distance, and I wish to propose that it is this higher level of interaction that is ultimately responsible for GTEG. Consider this simple model of interaction between two rural communities. One community is primarily devoted to farming, actually producing food for human consumption, while the second community is devoted to producing farming equipment. If both communities interact then the farming community with the aid of farming equipment, for example tractors, reapers, *etc.*, will produce food for both communities, while the second community will, as noted, supply the first with farm equipment. This synergism will aid both communities, and without it, both communities will have to both produce their own farm equipment and raise their own food.

Mathematically with respect to population growth, the following set of differential equations is analogous to the synergism described in the previous paragraph:

$$dN_1/dt = (r_1 + aN_2)N_1, \qquad \text{(Eq. 1)}$$

$$\text{and } dN_2/dt = (r_2 + bN_1)N_2, \qquad \text{(Eq. 2)}$$

where r_1 and r_2 are the growth rates of the respective populations, N_1 and N_2, and a and b are constants representing the degree of synergism between N_1 and N_2. As can be seen, the growth factor for each population includes their own rate of growth and a positive contribution from the other population, either aN_2 or bN_1. The contributions of each population to the other's growth is represented as a linear contribution only because linearity represents the simplest case. With further research into real cases, the component that each population gives to the other may in fact be non-linear, however, the focus of this paper is to provide a possible simple mechanism by which GTEG can be produced by synergism between populations and do so as simply as possible. Without this synergism, the above equations would simply represent exponential growth and would have the form: $dN_1/dt = r_1N_1$, and $dN_2/dt = r_2N_2$. But, what is the actual reality that the above coupled equations will produce GTEG?

Inspection of the set of differential equations shows that the growth component of each equation, $r + xN$, contains a constant component, r, and a component that is not constant but increases. More explicitly, this component is xN, where x is a constant and N grows at least at a rate, $dN/dt = rN$, which on solution gives, $N_t = N_0e^{rt}$. In other words, this component grows at least exponentially, and therefore the growth of the rate at which this population grows is at least exponential. However, since the contributing population is also growing at GTEG due to the contributions of the first population, then the first population must also grow at GTEG and vice versa. This can be shown numerically using the simulated data in Table 1. By graphing the data of the summed populations from the table, a graph of population over time is produced (see Fig. 7). If these same data are natural log-transformed, then if the growth is exponential a linear plot should be expected, while if the growth is GTEG, then a curve representing exponential growth of the growth rate should be expected. It is the latter type of graph that is produced, so the growth is GTEG and is due to the interaction components of the equations, *i.e.* the xN component of the growth component, $r + xN$. It should be noted here that without the xN components in each of the coupled differential equations, these differential equations would have the form, $dN/dt = rN$, which would, of course, yield exponential growth and not GTEG.

Table 1.

TIME	1	2	3	4	5	6	7	8	9	10	11	12
N_1	1	1.15	1.33	1.55	1.83	2.19	2.66	3.23	4.10	5.36	7.29	16.34
N_2	1	1.17	1.38	1.63	1.93	2.30	2.77	3.36	4.14	5.20	6.74	13.57
N_1 & N_2	2	2.32	2.71	3.18	3.76	4.49	5.43	6.59	8.24	10.56	14.03	29.91
Ln N_1 & N_2	.69	.84	1.00	1.16	1.32	1.50	1.69	1.89	2.11	2.36	2.64	3.40

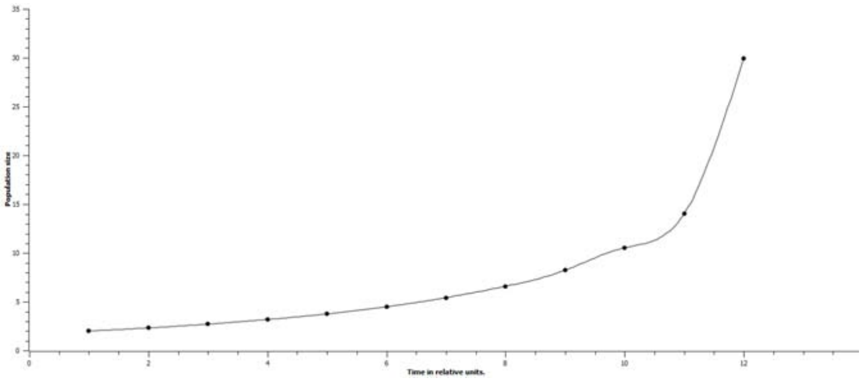


Fig. 7. This graph represents the combined population data, N_1 and N_2, given in Table 1 plotted against time. While this plot is unquestionably greater than linear, it is impossible to determine by inspection whether or not the graph is exponential or GTEG

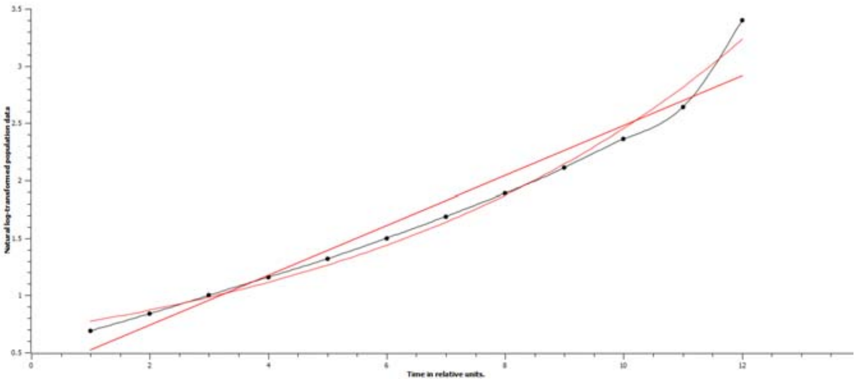


Fig. 8. The population data in Fig. 7 are natural log-transformed and plotted over the same period of time. Both a linear and an exponential regression are fitted to this data, and it can be seen that the exponential regression is the better fit implying that the actual mode of population growth is not exponential but GTEG

Since the growth of the simulation, as represented in Figs 7–8, and that of the world system, as represented in Figs 5–6, are both greater-than-exponential, it is proposed that the mechanism by which the simulation data are produced could also be the mechanism by which the actual world-system population grows at a greater-than-exponential rate. In other words, human interaction at the inter-group level can be modeled to produce GTEG by coupled differential equations in which the growth component contains within it a component of exponential growth due to the interaction with another population. Clearly, the world-system consists of many interacting populations not just two. However, what has been presented here is a toy model of a proposed mechanism by which the world-system population pattern of growth is GTEG and not exponential.

Summary

1. Von Foerster *et al.* (1960) and Korotayev *et al.* (2006a) have shown that human populations grow at a rate greater than exponential.

2. Exponential growth increases at a constant rate, and as a consequence the log-transformed population data give a linear plot against time.

3. Greater-than-exponential growth (GTEG) yields a curved plot in which the growth rate increases with increasing size of the population.

4. Real world-system data when log-transformed yield as expected an exponential curve.

5. A mechanism is proposed by which human interaction between groups yields GTEG.

6. This mechanism in its simplest form is represented by the following two differential equations: $dN_1/dt = (r_1 + aN_2)N_1$, and $dN_2/dt = (r_2 + bN_1)N_2$.

7. It is shown by numerical simulation that the combined growth of the populations represented by these two equations gives GTEG, suggesting that these coupled equations represent a model for human population growth.

References

Cohen J. 1995. *How Many People can the Earth Support?* New York: Norton.

Foerster H. von, Mora P., and Amiot L. M. 1960. Doomsday, Friday 13, November, 2026. *Science* 132: 1291–1295.

Gotelli N. J. 2001. *A Primer of Ecology.* Sunderland, MA: Sinauer Press.

Hutchinson G. E. 1978. *An Introduction to Population Ecology*. New Haven: Yale University Press.

Korotayev A., Malkov A., and Khaltourina D. 2006a. *Introduction to Social Macrodynamics: Compact Macromodels of the World System Growth.* Moscow: Uchitel.

Korotayev A., Malkov A., and Khaltourina D. 2006b. *Introduction to Social Macrodynamics: Secular Cycles and Millennial Trends.* Moscow: Uchitel.

Markov A. V., and Korotayev A. V. 2007. Phanerozoic Marine Biodiversity Follows a Hyperbolic Trend. *Paleoworld* 16(4): 311–318.

II. HYPOTHESES OF DEEP BIG HISTORY

4
10^{500}. The Darwinian Algorithm and a Possible Candidate for a 'Unifying Theme' of Big History

David C. Baker

Abstract

This article postulates another aspect of the long sought-after 'unifying theme' of Big History, in addition to the rise of complexity and energy flows. It looks briefly at the manifestation of the 'Darwinian algorithm', that is to say an algorithm of random variation and non-random selection, in many physical processes in the Universe: cosmology, geology, biology, culture, and even the occurrence of universes themselves. This algorithm also seems to gradually open more forms of variation and more selection paths over time, leading to a higher level of free energy rate density, or what we know as 'complexity'. In fact the complexity of the object under discussion seems to correspond to the available number of selection paths. The article closes with a bit of philosophical reflection on what the Darwinian algorithm and the rise of complexity could possibly mean for humanity and the future of the cosmos.

Keywords: *Universal Darwinism, random variation, non-random selection, complexity.*

One thing that the inaugural International Big History Conference in August, 2012 made clear was that one of the major tasks of Big History in the coming years is to prove it can sustain research projects, just like any other genre of historical scholarship. As someone who entered the field to do precisely that, I know that such research is not only possible, it is essential – both to bridging the gap between the sciences and humanities and to our understanding of the history of life and the cosmos. The unique approach of Big History has suddenly opened up a vast horizon of research agendas, or, to put it another way, triggered a speciation event where hundreds of new niches have opened up, waiting to be filled. The ecological terrain is vast and the numbers that currently populate it are few. I urge anyone interested in researching in Big History to do so. The research comes in a variety of forms. There are, of course, Esther Quaedackers's *Little Big Histories* that cover the full 13.8 billion years of any subject – extending the Big History perspective to any line of inquiry. There are also research agendas that pursue debates and questions about a certain chunk of the grand narrative, but nevertheless hearken back to broad trends. Many of these are highlighted in the course we teach in Sydney, and many of these would make excellent fodder for graduate research projects that can be realistically achieved within a set timeframe. There are also more ambitious ideas that deal with the unifying themes of Big History, themes which encompass the full trajectory of the universe and underscore the full chronology of 13.8 billion years. In this short article, I have no intention of asserting that this is true, but I do wish to illuminate a research agenda to figure out if it is.

To explain what we are dealing with, let us go back. Our story begins with a bang. And there is no point asking what happened 'before' the Big Bang. That is the wrong sort of question. Thanks to space-time relativity, there was no 'before' the Big Bang. Time as we know did not exist before the universe did. What is more, at the moment of the Big Bang, we are talking about a singularity of such intense heat and such intense pressure that the laws of physics would have broken down. Trying to describe what happened 'before' the Big Bang using the rules with which humans are familiar is rather like trying to describe colour to a dog.[1] Accordingly, the Big Bang is the earliest start date on any historical timeline a human being may care to construct. A tiny fragment of a second later, or 0.0000000000000000000000000000000000000000001 seconds to be more precise, we already have the first major tick on our timeline. I insert the decimal to give the reader a full idea of the infinitesimal scale, something that the exponent leaves somewhat understated, but this is Planck time (10^{-43}) the smallest length of time that has any physical meaning. Gravity had come into being. Then we have the next major event on our timeline between 0.000000000000000000000000000000000001 seconds (10^{-36}) (and 0.00000000000000000000000000000001 seconds (10^{-32}) after the Big Bang. The universe cooled ever so slightly from Absolute Hot by a few degrees Kelvin, allowing strong nuclear and electroweak forces to become more distinct, completing the collection of fundamental forces that control the physical processes of our universe. Around the same time the universe inflated due to the creation of a false vacuum and the gravitational repulsion and negative pressure of scalar fields, and grew faster than the speed of light (which is around 300,000 km per second) to an enormous size while continuing to cool, and then the scalar fields decayed into energy reheating the universe to its ultra-hot state. During the period of inflation, quantum fluctuations shaped the future growth of our universe, by creating minute variations in density, which were then in-

[1] And this absence of a conventional line of causality is what makes a twentieth century pseudo-scientific rehash of a medieval argument involving a supernatural First Cause so absurd (Craig 1979).

flated to such a large scale that they created the clumps of hydrogen and helium which in turn created our galaxies. These slight variations are mirrored in the temperature of Cosmic Background Radiation (CBR).[2] During the period of 10^{-36} and 10^{-32} seconds, most of the heavy lifting that set the physical processes of the universe in motion was accomplished. The clock was wound, the rules of the game were set, and the rest of the tale can be told with staggering accuracy using the familiar laws of physics.

What of other regions beyond the cosmic horizon of the visible universe? Our region endured a brief surge of inflation that explains small irregularities, the expansion rate, and the nature of further development. It would appear that there are other regions, each undergoing a different amount of inflation and developing physical properties vastly different to our own. Inflationary cosmology predicts that once inflation takes hold in one region, it causes accelerated expansion and inflation in other regions, producing a ripple effect (Guth 2007). Inflation is still underway in regions beyond our cosmic horizon. We are just one bubble where inflation has slowed down, like a hole in a block of Swiss cheese. Other regions in the 'multiverse', totally inaccessible to us, will also slow down in a runaway reproduction of universes. Until recently, we thought that one set of physical laws governed by a Grand Unified Theory was possible, and then we thought string theory illuminated a small number of possible sets of physical laws, now M-Theory shows that a vast number of functioning sets of physical laws with different properties, dimensions, and fundamental forces can exist and function (Duff 1998). The estimated number of sets of physical laws that is favoured by physicists at the present time is 10^{500} (Hawking and Mlodinow 2010: 118). That number embraces all the possible sets of physical laws that could form the basis for a universe when it cools down enough for those physical laws to become distinct, and it pops into existence, like a bubble in boiling water. These sets of laws fall together in the inflationary stage as a universe cools just a split second after the Big Bang (see, *e.g.*, Hawking 2001; Emiliani 1995: 82; Christian 2004: 27; Chaisson 2001: 126; Davies 1995: 28–35; Greene 2004: 312–313; Guth 1997: 20). Their formation appears to be the outcome of a random process (Barrow 2011: 214). Each set of laws determines density, temperature, fundamental forces, constants, dimensions, and whether or not things like matter exist. The actual number of universes based on those sets of physical laws is probably much higher, with many variations, but they all fall within the selection constraints of 10^{500}. Those universes that do not fall within those constraints do not get to exist.

Here is the fundamental basis for the Darwinian algorithm, a major research area in Big History. 10^{500} is therefore a very important number. It is

[2] The account of these events is given a decent treatment in many works, *e.g.*, David Christian (2004: 24–27) and John Barrow (2011).

the number of working sets of physical laws, the number of parameters in which a universe can occur. It is the primordial niche of all evolution, the foundation for an algorithm of random variation and non-random selection, a process that seems to arise time and time again alongside the rise of complexity in the universe (Dennett 1996: 48–61). The algorithm even seems to govern the formation of universes themselves. A Darwinian algorithm is anything that obeys a process of random variation and non-random selection. The game of 'universal natural selection' appears to be the first instance in the cosmic story where such an algorithm happens. The selection constraints appear to be the number of sets of physical laws in which a universe can start to exist. Universes appear randomly in inflationary space and only those universes that fall within the constraints of 10^{500} are non-randomly selected to form stable functioning universes. In such a scenario, universes are not constrained by any form of direct competition, but a form of 'niche selection' where the physical attributes of a universe that are capable of dwelling within a cosmic set of constraints make a form of non-random selection possible.

Nevertheless, that primordial niche is extremely wide, as you might expect from a form of selection that goes back to the birth of universes. To give you an idea of how many variations of sets of physical laws could exist, take a trillion of them, and then multiply that trillion by a trillion. Then another trillion × × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × trillion × × trillion × trillion. By comparison, 10^{14} is the number of years before the end of star formation, when every single last star will flicker out and the universe will wander in a cosmic graveyard of pitch black.[3] 10^{40} is roughly the number of years before the death of matter (Adams and Laughlin 1997). And 10^{100} is roughly the number of years before the total heat death of our universe.[4] The number of different sets of physical laws that form the game of cosmic selection is greater still. That is the magnitude of 100^{500}. Some of those universes that arise would operate without electromagnetism and they would never form clumps of hydrogen and helium, and by extension stars and galaxies. Some of those universes would never form atoms at all. And some of those universes would be based on properties, fundamental forces, and dimensions that, once again, in trying to understand how they operated using the physical concepts with which we are familiar in this universe would be like trying to explain colour to a dog.

---

3 100 trillion years (Adams and Laughlin 1999: 35–39).

4 If total heat death is indeed what awaits it. There are a number of possible scenarios and perhaps others undiscovered (*Idem.* 1997).

From this primordial niche comes a vast array of scholarly works that recognise the Darwinian algorithm in a variety of universal processes. At the cosmic level, Lee Smolin and E. R. Harrison have both proposed models for universes themselves, with those more likely to produce black holes or intelligent life, respectively, being favoured (Smolin 1997; Harrison 1995: 193). Both remain highly speculative and favour a hereditary connection between universes. At the end of the day such selection criteria and heritability may not even be required since the number 10^{500} is so large that it covers every variation to make inheritance between universes unnecessary and yet still mathematically finite, making non-random selection possible. Wojciech Zurek has created a model whereby the predictable physics of the Newtonian realm emerge from the chaos of the quantum world – a model that recently gained some new evidence (Zurek 2003; Burke *et al.* 2010: 1–4). If this is correct, then it provides an explanation for the uncertainty in quantum physics. The chaos at the quantum level does not abrogate the idea that the universe functions in a certain way, as would-be scientific determinists have lamented, because the very randomness at the quantum level is fundamental to the prevailing system. In the geological realm, Robert Hazen *et al.* (2008) have proposed an evolutionary model for the generation of new minerals. While making sure to clarify that the model differs from biology, the authors highlight several places where selection, punctuation, and gradients for change are present, exponentially increasing the number of mineral types throughout geological history, from stellar nebulae, through planetary accretion, and all the changes thereafter. But the most thorough examination of the Darwinian algorithm in areas beyond the realm of biology has been within cultural evolution. The idea was first pioneered by Donald Campbell (1960) and later revived by Richard Dawkins (1976), and then most effectively, in my opinion, developed by Peter Richerson and Robert Boyd (1985). In cultural evolution, any ideas, knowledge, beliefs, values, skills, and attitudes that are more practical or more appealing, are easier to learn or are better geared toward survival, are more likely to lead to social prominence than others, spread more easily from person to person. Those cultural practices that lead to early death or social stigma are less frequent or simply disappear (Richerson and Boyd 2005: 5–12). From so simple a beginning, came a flood of works on cultural evolution in recent years.[5] It also provoked

[5] Many works have been written on the subject, though I believe Richerson and Boyd remain the most successful at explaining it. Richard Dawkins, by contrast, as recently as *The God Delusion* (Dawkins 2006: 228) claimed Susan Blackmore (1999) held that honour. A number of other works have also been written on the subject: Stephen Shennan (2002), Ruth Mace, Clare J. Holden, and Stephen Shennan (2005) – particularly David Bryant, Flavia Filimon, and Russell Gray (2005) who advocate the NeighborNet program to plot trees for both vertical and horizontal transmission for all the Indo-European languages, John Ziman (2000), and a close runner up to Richerson, Boyd, and Black more is Stephen Shennan (2009), notable for its many in-depth investigations.

a great deal of debate.[6] Two of the most rigorous bits of research, in my opinion, have been Lake and Venti's work on nineteenth century bicycle technology and Ritt's work on the formation of dialects and new languages (Lake and Venti 2009; Ritt 2004). Finally, at the recent conference, my colleague and fellow Big historian, Christian Jennings, and I have discussed how Darwinian algorithms are used to fill a range of useful functions in the computer realm. Not all mechanisms of information in a computer are processed in a Darwinian algorithm. But since the 1970s, numerous programs have employed a 'genetic algorithm' which is a search heuristic that mimics the process of natural evolution. The computer automatically finds better ways to run programming through a game of variation and selection. It is currently employed in bioinformatics, engineering, economics, chemistry, mathematics, and more. Various entities of the universe are simply different forms of information – whether energy flows, DNA, or cultural ideas – and they seem to be processed by the same algorithm just as information in a computer. It is not the place of this article to confirm or deny the accuracy of the assertions cited above, but rather to exhort Big historians to future research, especially on any project that ties these various manifestations of the Darwinian algorithm together into one theory. The spectre of the algorithm has already been spotted by a number of scholars working in a number of disciplines. This could be what unites them all – an elegantly simple process, a form of variation, selection, and preservation that underwrites all things.

It may also have a trajectory. If the Darwinian algorithm is present, if not instrumental, at every stage in the rise of complexity in the universe, it may be that this pattern tends ever more to greater forms of complexity. And it would appear that the number of possible outcomes is relative to the complexity of the process under discussion, hence why relatively few outcomes make it from the quantum to the Newtonian level, why only a few thousand variations emerge from the geological level, whereas in biological evolution the number of possible selection paths is increased manifold, and the number of cultural variations is exponentially greater still. When we arrive at something as complex as culture and modern human society, with a free energy rate density 25 times higher than the average product of genetic evolution and 500,000 times higher than the Milky Way, there are a mind-boggling number of combinations of ideas and innovations. The rate of complexity seems to increase with the number of viable selection paths.

---

[6] For instance, see Joseph Fracchia and R. Lewontin (2005: 1–13, 14–29, and 30–41), in several back and forth exchanges. Fracchia and Lewontin's misunderstanding of what cultural evolution actually led both sides to more or less repeat the same arguments at each other.

Table 1.

Generic Structure	Free Energy Rate Density (erg s$^{\uparrow}$(–1) g$^{\uparrow}$ (–1))
Galaxies (*e.g.*, Milky Way)	1
Stars (*e.g.*, Sun)	2
Planets (*e.g.*, Earth)	75
Plants (biosphere)	900
Animals (*e.g.*, human body)	20,000
Brains (*e.g.*, human cranium)	150,000
Society (*e.g.*, modern human culture)	500,000

Source: Chaisson 2001: 139.

At the recent conference, I asked futurist and IBHA board member, Joseph Voros what sort of complexity we might expect to see from an intelligent species capable of harnessing increasing levels of energy, that is, the power of stars and the galaxy, known as Type II and Type III civilisations. He said that another exponential increase in free energy density was less likely than an increase in the complexity of networks. It would appear, for the time being, cultural evolution and the complexity it bestows is the highest point in this process of which we are yet aware. Others may open up that we cannot predict, but it is worthwhile to understand exactly what cultural evolution involves. There are two tiers of human evolution. The first is genetics, which operates in the same way as for other organisms. Those genes gave humans a large capacity for imitation and communication. Those two things enabled the second tier. Culture operates under similar laws, but on a much faster scale. Cultural variations are subject to selection and the most beneficial variations are chosen. Unlike genes these variations can be transmitted between populations of the same generation and can be modified numerous times *within* that generation. Like a highway overpass looming over older roads, cultural evolution can blaze along at a much faster rate of speed. Ultimately, culture accumulates. Population pressure compels some of this accumulation to be geared toward increasing the human ability to extract resources from the environment. This process raises the carrying capacity, which produces more people, more accumulation, which in turn raises the carrying capacity. The cycle continues and grows in complexity. If expressed as a general principle, it may be said that the rate of growth of the carrying capacity of a human population is relative to the number and connectivity of variant innovations.

The second evolutionary tier of culture, a swifter form of evolution, should not come as a surprise in a Darwinian algorithm. Gradually through natural selection, not only do species become better at surviving, they become better at evolving. This follows the logic that improving the *rate* of your improvement of your survival chances is just as naturally selected for, since it also improves the rate of your survival. Logically, a third tier is likely to emerge where our growing knowledge allows us to directly guide the evolution of our genes. If we discard the manmade concept of tiers, in a sense through a relatively short evolutionary process of 200,000 years, our genes have evolved the ability to develop more rapidly and efficiently. The universe is composed of webs of energy of varying complexity. Life-forms are entities that harvest energy to perpetuate their complexity, to spread it, and even to increase it. Human history has been dominated by this hunt for resources. Our evolution, both genetic and cultural, has ultimately been geared toward aiding this hunt. Standard evolution can be defined as the change in the traits of a population of organisms through successive generations to sustain or increase their complexity. Human evolution can be described as the change in traits and behaviours between populations of the same generation and through successive generations to sustain or increase their complexity.

We now know that there is no hard-and-fast division between the organic and inorganic world. As such, life can be (somewhat coldly) defined as a series of physical processes that contain a hereditary program for defining and directing molecular mechanisms that actively extract matter and energy from the environment that are converted into building blocks for the perpetuation and reproduction of those physical processes (Spier 2010: 77). Life is the only thing in the Universe that does this. Stars, minerals, and the rest of the inorganic world do not actively seek out matter and energy from the environment. Even objects as gigantic as stars burn their fuel like lamps and candles and eventually flicker out. This has been proceeding since the beginning of the universe. Eventually every single last tiny slow burning star will be extinguished. Only life has the agency to go out and extract energy from the environment to keep itself going. We do not just sit still and wait for death to take us. We fight – for a time. If we want to preserve our vast complexity, we have to continue harvesting matter and energy to keep ourselves going. All other considerations are secondary. It is the bottom line of human history. During most, if not all, of our history, the quest to extract matter and energy to perpetuate our existence has been the overriding theme (Spier 2010: 116). It is the battle with disorder, chaos, entropy, and the second law of thermodynamics which we have carried on since the very beginning of our existence, and it is a battle that physicists believe we must eventually and inevitably lose.

This brings me to the topic of how the Darwinian algorithm relates to how we perceive ourselves in the grand narrative and how Akop Nazaretyan at the recent conference exhorted Big historians to provide the world with non-exclusive 'meanings of life' – beyond religion and ideology that inevitably vilify the infidel and the 'other' – to ideas of meaning that bind the entire human race together in common cause (Nazaretyan 2010). If the Darwinian algorithm prevails at many stages in the rise of complexity in the universe, then it is possible that the evolution of life and species capable of cultural evolution is just another stage in this trajectory, just like star formation or planetary accretion. At risk of sounding sensationalist and glib, two things that I abhor, I must state that research in this direction may *possibly* provide us with something approaching a secular and objective 'meaning of life' that unites us all.

There are as many as 10^{500} possible sets of physical laws for universes. Each of these sets of physical laws governs the evolution of a universe in various ways. Our universe is 13.7 billion years old, very complex life on Earth about 550 million, and the human race as we know it only about 200,000. Our local star is middle-aged and will last only another 5 billion years and will boil the Earth's surface dry in well under 3 billion. If the human race does not destroy itself in the meantime, it has hundreds of millions of years to exist and evolve on Earth, after which time we could venture out into other solar systems and long outlast the death of our own. We could huddle around the fires of hundreds of thousands of stars in the habitable section of the Milky Way for nearly a trillion years and more stars would be produced in the centre of the galaxy and eventually spread out and be used too. But unless we somehow learn to create stars ourselves, in 100 trillion years every single last dim little star will have flickered out and the universe will become a cosmic graveyard, where bodies of dead stars and planets will wander in pitch black. Until, of course, the energy that creates matter itself (which, remember, is really just a congealed form of energy) in 10^{40} years will grow feeble and matter will cease to exist, and then after a period of 10^{100} years, even black holes will cease to exist, and the universe will be an empty orb of weak cosmic radiation – a victim of Heath Death.

Here is the grim fate to which we must resign ourselves that also seems to indicate that our story and the story of the universe itself is ultimately and objectively pointless. Yet, the notion of the Darwinian algorithm of random variation and non-random selection governing processes in the universe as disparate as geology, biology, and culture, indicates another interesting possibility. Life is the only entity in the universe that actively harvests energy rather than just burning down and in only the last 250 years human beings have mastered the atom and figured out how to harness energy in impressive magnitudes. The next 'spontaneous' rise of complexity in the universe will be down to intelligent life. Current physical processes in the universe indicate a future of heat death. But those calculations do not take the evolution of intelligent life into account. That grim fate for the universe may be avoided. It is very difficult to see why the wheels are churning when we ourselves are inside the machine. We have millions, billions, if not trillions of years before us, to devise a way to keep the lamps of the galaxy lit, energy flowing, and the universe itself from 'dying'. And, perhaps most profoundly of all, life itself may have been another one of those ever-present Goldilocks conditions: an entity that keeps harvesting and creating energy to perpetuate the complexity of the universe. Like tiny white cells in the human body, our small and seemingly insignificant species may nevertheless have an extremely important role in the universe. Our fates might be bound together. It may be why we are here. In that sense, the 'meaning of life' is a fairly easy question. The question of the 'meaning of the universe', on the other hand, is a much more difficult proposition.

At any rate it remains an open possibility – and it has significance for us today, not just trillions of years from now. Albert Camus (1913–1960), a French writer and philosopher, once said that in all philosophy there is only one problem, and that is suicide – judging whether life is or is not worth living amounts to the most fundamental question of philosophy (Camus 1942: 15). In the secular scientific narrative of Big History, we are robbed of traditional answers to that question. In a cold, often cruel, empirical universe based on fact and not on fantasy, we do not have access to the pre-packaged meaning, morality, and life purpose that animates religious culture. What we are left with is a universe that evolved from impersonal physical laws and is so vast as to reduce all the trials of daily life and indeed all human history to a state of woeful insignificance. The universe does not owe you a sense of purpose. It does not owe you a sense of comfort. Lacking an objective scientifically reinforced meaning of life and purpose to existence, where the universe has no higher role for living things, there ultimately is no point. In such a state of affairs that is the hard, grim, inevitable fact. You are an accident of physics, kept alive by an evolutionarily instilled fear of death that translates into a multitude of subjective, often paltry, excuses for why you have not yet opened your throat. Even good answers to that question, like the noble scientific curiosity to explore the universe, or love, or duty, or stubbornness (KBO, the motto from British trenches in the First World War, keep buggering on), just sound like provisional reasons so we can move on and stop thinking about it. Even now the reader's mind may be racing, reminding themselves of their own reasons for living. And perhaps these subjective excuses are all we can ever hope to achieve. The Darwinian algorithm, however, returns to the question of an objective secular scientific meaning of life and whether life is or is not worth living – the fundamental question of philosophy.

Research on the Darwinian algorithm may be crucial in a variety of ways. From it we might attain a greater sense of where humanity fits in the history of the universe. We might identify some of the processes that govern human development and also identify the universal context in which humanity faces the distant future. From here it may be possible to establish an objective sense of human purpose in the universe, though the validity of this last step is far from certain. And when I use words like 'meaning' and 'purpose' I do not engage with the idea of strong emergentism and those scientists who are using concepts of strong emergence to revive 'religiosity' and the feelings of 'awe', 'creation', 'enchantment', 'transcendence', 'reverence', 'gratitude', and 'objective and universal morality', normally associated with the traditional religions.[7] I have no desire to replace religion with science or anything else for that matter in an increasingly secular age.[8] I am content to let such feelings of 'religiosity' go. Empirical work on the Darwinian algorithm should not be optimistic or indulge in mysticism. To mature intellectually in the 21st century, one must stop being such a child and admit that the questions of existence and morality are not as clear-cut as old religions had led us to believe, the answers are not often uplifting, and it is harder to take refuge in feelings of reverent religiosity today than it was in the time of your ancestors. Science will not revive those feelings. Being perpetually confused and scared is part of being an adult in the 21st century. We are in the process of casting aside old fairy tales. Now is not the time to be inventing new ones. But it is my fear that research on the Darwinian algorithm will be just another desperate grasp at a comforting myth. I remain characteristically pessimistic about its prospects, but it is too interesting a possibility to pass up. But the possibility may fail and join the ranks of other pathetic exploits in pseudo-science, in which case there is very little besides subjective reasoning between you and the stark contemplation of the grand unfolding tale of 13.8 billion years.

Paul Dirac (1902–1984), English theoretical physicist who predicted the existence of antimatter, and whose brother, Felix, committed suicide in 1925, wrote his entire philosophy of life on three pages of a notebook in 1933, in which he said:

> My article of faith is that the human race will continue to live forever and will develop and progress without limit. This is an assumption that I must make for my peace of mind. Living is worthwhile if one can contribute in some small way to this endless chain of progress (quoted in Farmelo 2009: 221).

There is, of course, absolutely no guarantee that humanity or our descendant species will not go extinct, and much to indicate the contrary as we enter the bottleneck of the 21st century. But perhaps Dirac is right, despite this assumption. Perhaps, within the Darwinian algorithm, life is worthwhile if we can contribute in some small way to the rise of complexity in the universe – a strange, blind, but inexorable process that has been proceeding for 13.8 billion years.

---

[7] The most explicit statement to this effect is Ursula Goodenough and Terrence Deacon, *The Sacred Emergence of Nature* (Goodenough and Deacon 2006).

[8] See also Stuart Kauffman (2006), and also the seminal (and less proselytising) works on emergence (*Idem.* 1993, 1995), and also Terrence Deacon's recent ode to emergence, *Incomplete Nature: How Mind Emerged from Matter* (Deacon 2011), which was not well received by experts, *e.g.*, Jerry Fodor (2012), in addition to lingering allegations of plagiarism.

References

Adams F., and Laughlin G. 1997. A Dying Universe: The Long-Term Fate and Evolution of Astrophysical Objects. *Reviews of Modern Physics* 69: 337–372.

Adams F., and Laughlin G. 1999. *The Five Ages of the Universe*. New York: The Free Press.

Barrow J. 2011. *The Book of Universes: Exploring the Limits of the Cosmos*. London: W.W. Norton and Company.

Blackmore S. 1999. *The Meme Machine*. Oxford: Oxford University Press.

Boyd R., and Richerson P. 1985. *Culture and the Evolutionary Process*. Chicago, IL: Chicago University Press.

Bryant D., Filimon F., and Gray R. 2005. Untangling Our Past: Languages, Trees, Splits and Networks. *The Evolution of Cultural Diversity: Phylogenetic Approaches* / Ed. by R. Mace, C. Holden, and S. Shennan, pp. 69–85. Publisher: UCL Press.

Burke A. M. *et al.* 2010. Periodic Scarred States in Open Quantum Dots as Evidence of Quantum Darwinism. *Physical Review Letters* 104(17801): 1–4.

Campbell D. 1960. Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes. *Psychological Review* 67: 380–400.

Camus A. 1942. *Le mythe de Sisyphe: essai sur l'absurde*. Paris: Les Éditions Gallimard.

Chaisson E. 2001. *Cosmic Evolution: The Rise of Complexity in Nature*. Cambridge: Harvard University Press.

Christian D. 2004. *Maps of Time: An Introduction to Big History*. Berkeley, CA: University of California Press.

Craig W. L. 1979. *The Kalam Cosmological Argument*. New York: Macmillan.

Davies P. 1995. *The Last Three Minutes*. London: Phoenix.

Dawkins R. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins R. 2006. *The God Delusion.* New York: Houghton Mifflin.

Deacon T. 2011. *Incomplete Nature: How Mind Emerged From Matter*. London: W.W. Norton & Company.

Dennett D. 1996. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Touchstone.

Duff M. 1998. The Theory Formerly Known as Strings. *Scientific American* 278(2): 64–69.

Emiliani C. 1995. *The Scientific Companion: Exploring the Physical World with Facts, Figures, and Formulas*. New York: John Wiley.

Farmelo G. 2009. *The Strangest Man: The Hidden Life of Paul Dirac, Mystic of the Atom*. London: Faber and Faber.

Fodor J. 2012. What are Trees About? *London Review of Books* 34(10): 31–34.

Fracchia J., and Lewontin R. 2005. The Price of Metaphor. *History and Theory* 44: 14–29.

Goodenough U., and Deacon T. 2006. The Sacred Emergence of Nature. *The Oxford Handbook of Religion and Science* / Ed. by P. Clayton and Z. Simpson, pp. 853–871. Oxford: Oxford University Press.

Greene B. 2004. *Fabric of the Cosmos: Space, Time, and the Texture of Reality*. New York: Alfred A. Knopf.

Guth A. 1997. Was Cosmic Inflation the 'Bang' of the Big Bang? *The Beamline* 27(14): 14–21.

Guth A. 2007. Eternal Inflation and its Implications. *Journal of Physics A: Mathematical and Theoretical* 40: 6811–6826.

Harrison E. R. 1995. The Natural Selection of Universes Containing Intelligent Life. *Quarterly Journal of the Royal Astronomical Society* 36: 193–203.

Hawking S. 2001. *The Universe in a Nutshell*. New York: Bantam.

Hawking S., and Mlodinow L. 2010. *The Grand Design*. New York: Bantam Books.

Hazen R. *et al.* 2008. Mineral Evolution. *American Mineralogist* 93: 1693–1720.

Kauffman S. 1993. *The Origins of Order: Self-Organisation and Selection in Evolution.* Oxford: Oxford University Press.

Kauffman S. 1995. *At Home in the Universe: The Search for Laws of Complexity*. Oxford: Oxford University Press.

Kauffman S. 2006. Beyond Reductionism. Reinventing the Sacred. *The Edge*: *Third Culture.* URL: http://www.edge.org/conversation/beyond-reductionism-reinventing-the-sacred.

Lake M., and Venti J. 2009. Quantitative Analysis of Macroevolutionary Patterning in Technological Evolution: Bicycle Design from 1800 to 2000. *Pattern and Process in Cultural Evolution /* Ed. by St. Shennan, pp. 147–162. Berkeley, CA: University of California Press.

Mace R., Holden C. J., and Shennan St. (Eds.) 2005. *The Evolution of Cultural Diversity: A Phylogenetic Approach*. London: UCL Press.

Nazaretyan A. **2010.** *Evolution of Non-Violence: Studies in Big History, Self-Organization, and Historical Psychology*. Saarbrucken: LAP.

Richerson P. J., and Boyd R. 2005. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: University of Chicago Press.

Ritt N. 2004. *Selfish Sounds and Linguistic Evolution: A Darwinian Approach to Language Change*. Cambridge: Cambridge University Press.

Quaedacker E. 2014. To See the World in a Building: A Little Big History of Tiananmen *Teaching & Researching Big History: Exploring a New Scholarly Field* / Ed. by L. E. Grinin, D. Baker, E. Quaedackers, and A. Korotayev, pp. 90–111. Volgograd: Uchitel.

Shennan St. 2002. *Genes, Memes, and Human History: Darwinian Archaeology and Cultural Evolution*. London: Thames and Hudson.

Shennan St. (Ed.) 2009. *Pattern and Process in Cultural Evolution*. Berkeley, CA: University of California Press.

Smolin L. 1997. *The Life of the Cosmos*. London: Phoenix.

Spier F. 2010. *Big History and the Future of Humanity*. Chichester: Wiley-Blackwell.

Ziman J. (Ed.) 2000. *Technological Innovation as an Evolutionary Process*. Cambridge: Cambridge University Press.

Zurek W. 2003. Decoherence, Einselection, and the Quantum Origins of the Classical. *Reviews of Modern Physics* 75: 715–765.

5
The Chandra Multiverse

Tom Gehrels

Abstract

Equations of Planck and Chandrasekhar lead to our universe being a member of a quantized system of universes, the 'Chandra Multiverse'. It is a trial-and-error evolutionary system; all universes have the same critical mass and finely tuned physics that our universe has. The origin and demise of our universe are described. In our astronomical environment everything ages and decays; even the proton may have a limited half-life. The decay products of all the universes expand into the inter-universal medium (IUM), clouds form in the IUM, from which new universes are started. When the density at the center of our proto-universe cloud reached proton density, photons, protons and neutrons were re-energized. A Photon Burst marks the beginning of our universe, 10^{-6} s, i.e. 10^{37} Planck times, later than a Big Bang; the evolution of forces, sub-atomic particles, and finely tuned physics, occurs in the Chandra Multiverse. This paper is based on 30 observations, 8 previous papers, and 2 books; the multiverse makes the identification of dark energy and dark matter possible.

Keywords: *Universe, Chandra Multiverse, Planck Time.*

To leave the earth's surface was the desire of Leonardo da Vinci – in 1783, the first balloonists left the earth! There were *four* flights that year, from the squares of Paris, in full sight of thousands of people who had come out to see the miracle: 'If we can do that, what can we *not* do?' The French Revolution began soon after that. Of the next step, to the moon, President John F. Kennedy spoke to the US Congress in 1961. It sounds like poetry when you read it aloud:

I believe that this nation
should commit itself
to achieving the goal
before this decade is out,
of landing a man on the Moon
and returning him safely to Earth.

It was done, and had the same effect on the world population as on the people of Paris before. The trials and failures, and ultimately the flights and walks on the moon were shown on international television, in newspapers and magazines, watched and discussed by millions on Earth: '*If we can do that, we can do anything!*' The astonishing feat of people leaving Earth may have spurred the surge of activity and questioning that took place during the 'Great Sixties' all over the world.

These two examples of broadening the horizons stimulate the Chandra Multiverse[1] to solve old problems and open new disciplines ('Chandra' was used by his colleagues and for the Chandra X-Ray Observatory spacecraft). There was an inkling of high expectations in 1954 when I was a student in the class of Subrahmanyan Chandrasekhar (1910–1995) at the Yerkes Observatory of the University of Chicago. There were many equations and derivations, but his cosmic-mass equation was exceptional because of its unified operation of quantum, relativity, gravity, and atomic physics. He hinted the possibility of deeper relations between atomic theory and cosmogony.

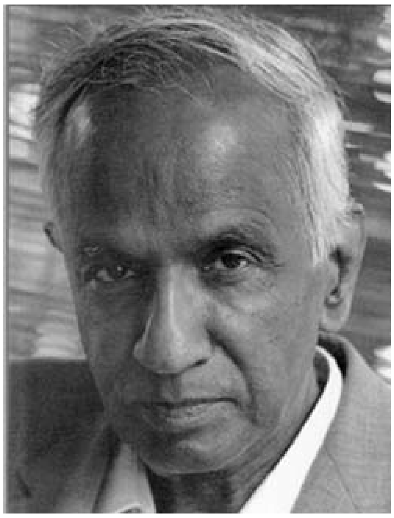


Fig. 1. Chandra

It took me years of preparation, but when I finally dared to use that equation, it pointed to our universe not being alone but being a member of the Chandra Multiverse with many others, each with the same mass and physics. I published that in 2007, and thought I was done, because others would surely jump to such a basic change in our worldview. No one stirred, so that I proceeded with various papers[2] and two books (Gehrels 2011, 2012). This is a beginning, a sketch to show possibilities, while expansion by specialists is needed for all parts.

[1] This notion is explained further in this article.
[2] URL: www.lpl.arizona.edu/faculty/gehrels2.html.

Better understanding of the vast dimensions of the Chandra Multiverse follows from changing scales, using various measuring sticks. We use centimeters or inches for small distances on the desk, but it would be silly to use the large number of centimeters for going home, so that we switch to kilometers or miles. Beyond the Earth and Moon, the astronomical unit (AU) is used, which is the average distance of the Earth from the center of the Sun. To the outer edges of the solar system that becomes a large number again – 100,000 AU. So, beyond that to the stars and galaxies we use the light-year, which is the distance light travels in a year, $\sim 10^{13}$ km.

In order to help you visualize the multiverse, consider the next larger grouping after planets and stars, which is a galaxy, the ensemble of at least a thousand million stars, and then the next step, another thousand million galaxies to populate our universe. Having come to the end of our universe, as far as we can observe it, we ask: 'Why would our universe be the only one?' Some theoreticians have already made models regarding the existence of other universes, but this article is about a specific set of universes. It is based on the mass equation of Chandrasekhar (1951) and is therefore called the 'Chandra Multiverse'; it has nothing to do with any previous usage of the word 'multiverse'.

This paper is all assembled for historians who call themselves 'Big Historians' for wanting to go back all the way in time. The evolution in the Chandra Multiverse is now the earliest we can go back to – there never was a Big Bang – and the early evolution is now properly taken care of in the multiverse.

Big Bang, Planck Time, Inflation and Strings

The words 'Big Bang' are commonly used for the beginning of our universe. As it is conceived, a supercharged explosion created the basic matter and energy that evolved over the course of 13.7 thousand million years into what we perceive to be our universe today. These Big Bang words are also used in a presently accepted theory called the 'Standard Model'. In this model, our universe is generally believed to be the only universe... this is *it* for the Cosmos.

There are problems with this concept. The usual thinking is to reverse the *expansion* of our universe as a *contraction* back in time. A problem arises when one continues the contraction, without stopping, until all of the mass of our universe is in nearly zero space and compressed into nearly infinite density. That infinite condition is called a 'singularity', which has caused theoreticians to write a large amount of literature in papers, books, and encyclopedias. The Big Bang is believed to be confirmed by three observations, but these were made at ages of minutes and much later, applying nicely to the Chandra-Multiverse model.

The Big-Bang starting time for the Standard Model is $t = 10^{-43}$ sec and all of *our basics* would have had to be produced during the very beginning of our universe. The earliest stages after the Big Bang would have had to be able to evolve quantum mechanics, relativity, gravity, and atomic physics, as well as fundamental forces.

The Standard Model is that the volume of the universe must have stayed small for a while because of the condition that *interaction must have flourished*; in other words, everything could interact with everything else in such confinement. This requirement emerged when the background radiation between the stars was observed by the COBE spacecraft to have the same temperature of 1.728 K in all directions – the universe is *uniform* on its largest scales. However, this leads to that impossible early condition, because in order to interact, the components must be close enough. All of the quantum fluctuation components, with a total mass equivalent to 10^{21} solar masses, would have had to be confined in a volume so small that all their components could interact with all others at the velocity of radiation. Their density would have had to be as high as the Planck density of 10^{96} kg/m^3, which does not seem realistic.

At this point of reasoning, *inflation theory* seemed to solve the problems in the 1970s, with a complex homogenizing and processing between 10^{-43} sec and 10^{-32} sec, removing a variety of uncertainties in the understanding of the early stages for our universe. (That interval is hard to imagine, and looks brief, while it is actually long for the fast subatomic actions. The usage of the *second* as a unit of time is better replaced here for sub-atomic action by the *Planck time*, which is about 10^{-43} sec. That interval is then $[10^{-32} - 10^{-43}] / 10^{-43} \sim 10^{11}$ PT.) Anyway, the inflation theory has in that interval a fast increase in the size of the universe, an inflationary expansion, with thorough interaction of components, explaining the uniformity of our universe found by COBE. This fast expansion was confirmed in later observations by the WMAP spacecraft (Spergel *et al*. 2007).

Such confirmation has not yet happened for *string theories*, which have no observations supporting them; they may be too small and energetic. String theories consider sub-atomic particles as one-dimensional curves called ‘strings’. The strings all differ in order for their vibrations to represent the variety of properties of sub-atomic particles. They facilitate storage of information by having those particles replacing the infinitely small point-particles of quantum theory. The theories are mathematically expressed and are powerful because they can describe atomic forces and fields; the strings are imagined to be embedded in space-time. They come in various sizes and shapes for storing the various properties; they may be curled and imperceptibly small. The variety of particle properties is a reflection of the various ways in which a string can *vibrate*: electrons vibrate in one way, quarks in another, *etc.* Several *dimensions* are added, at least seven, to the four we are used to (which are up-and-down, close-and-far, left-and-right, and time). The seven higher dimensions are not defined; no one seems to be able to describe them other than by making a comparison with the four (Randall 2005, 2007; Smolin 2007).

There are many string theories, Smolin gives their number as 10^{500}, and the name 'M-theory', M for Many, is therefore used for the ensemble. There is some collaboration of inflation and string theorists, while there seems to be a general acceptance that some combination of inflation and strings is what happened at the beginning of our universe. Large physics departments have at least one inflationist and a string-theorist in their faculty; hundreds of physicists work on these theories in international collaboration.

Physicist Brian Greene has provided an overview of string theories and their history, and he revels in an analogy with vibrational patterns of music, how the strings orchestrate the evolution of the world into a cosmic symphony (Greene 1999).

Physicists Paul Steinhardt and Neil Turok criticize inflation theories, while presenting a string theory that has our universe pulsating, expanding ↔ contracting; they do not believe inflation theories (Steinhardt and Turok 2007). Physicists Peter Woit and Lee Smolin do not believe in string theory (Woit 2006; Smolin 2007). Steinhardt specifies why he finds inflation theory deeply flawed (Steinhardt 2011). Such is the state of affairs for understanding the early stages of our universe before $t \sim 10^{-6}$ sec, the beginning time of our universe in the Chandra-Multiverse model, which has the earlier sub-atomic evolution in its multiverse.

Parallel Universes and Anthropic Principles

In the large literature about other multiverses, some of it seems to be *anthropic* (from the Greek *anthropos*, human being), *i.e.* based on human interpretations of physics. The anthropic principle has the idea that nature is the way it is because we are here to observe it – it was created for people. For example, physicist Hugh Everett (1971) adopted the wave interpretation of quantum mechanics held by Schrödinger and others (Bitbol 1995), but he also felt that this picture makes sense only when observation processes are treated within the theory. He favored an interpretation of quantum theory by which reality is brought through observation or measurement. Some of the literature about the multiverse is therefore based on an interpretation of quantum mechanics in which *observations* bring *reality*. I cannot help thinking of that when crossing a major street on the bike, 'Perhaps if I just do not observe an oncoming car, it would not exist', but I have not tried it yet. Books and articles describe various multiverses; some have 'parallel' universes with identical copies of us reading this paper at this time (Everett 1971; Vilenkin 2006). In contrast, the Chandra-Multiverse model is free of anthropics.

Expansion of 100 % Space

The word 'space' is a most frequently used word, while we actually do not know what it is, in a physical sense. The best feel for space can be obtained by a thought experiment: Imagine that the outer perimeter of the hydrogen atom with its electron shell is as large as the longest outer dimension of an Olympic stadium. How large is then the proton nucleus of the atom on center field? As the size of the stadium is roughly three times the length of a soccer or football field (100 meters or 300 feet), take the long dimension to be about 1,000 feet, which is 10^3 x 12 (in/foot) x 2.54 (cm/in) = 3 x 10^4 cm. Divide that by the known factor between the radii of the hydrogen atom and the proton, 10^5, and it follows that the nucleus would be 0.3 cm, the size of a pea. All around inside that stadium is modeled as space. Furthermore, the proton is a structure of space as well. So, the answer to the question of how much of this atom is 'space' is 100 %. The most astonishing conclusion is that this applies to all atoms, to our entire visible world consisting of 100 % space.

You might protest that the objects around us are solid and we experience our mass and weight – how can that be if we are 100 % space? The answer is that space has an abundance of particles with their properties. The word 'particle' is confusing because it reminds us of a solid grain. When physicists use the word, however, they do not mean anything like the Greeks did (an inert grain of dull matter), but rather that *a particle is a point in space that produces a phenomenon, like mass.* By what it does, *a particle is something that provides action and an effect, like electricity.* Both sentences together may help us understand the particles as forms of space and waves by which they produce and interact with gravity, electromagnetism, *etc.* It is because of the properties of particles and photons that we experience weight and all other characteristics of nature.

In this paper there is also much discussion about the expansion of space. It is simple to understand by the galaxies having been pushed apart and still moving because there is little to stop them. It may be helpful to imagine a cake in the oven, which expands as the temperature increases. Imagine yourself to be one of the raisins and you see the others drifting away... that is the expansion. A raisin twice as far away will seem to expand at twice the speed.

The First Minute of Our Universe

As of the notable event in the Standard Model that we mentioned at 10^{11} PT after the Big Bang, high-energy particle physicists begin to understand the physics involved at this time, just barely, and make proposals for the forces and subatomic particles that subsequently evolved. A feature of the theories indicates that particles formed symmetrically with their anti-particles of opposite properties. The two would then have largely annihilated each other. However, it was apparently not a total annihilation; one of the two types would have prevailed a little, and that survivor is what we now call a 'particle'. It will be interesting to see if this feature of the theory survives in the evolution of the Chandra-Multiverse.

The next milestone came at 10^{31} PT with the emergence of four nuclear forces: the weak, gravitational, electromagnetic and nuclear forces. The forces and their laws began to be recognized in physicists' modeling of the temperature and pressure of that time.

At 10^{37} PT, protons and neutrons were made; a large amount of radiation was produced. We will see later that this is the beginning time for the Chandra-Multiverse model, t = 0 for that model, and from here onwards the Chandra-Multiverse model adopts and follows the Standard Model. The evolution during the earlier 10^{37} PT is however in the multiverse; that is not an *ad hoc* assumption, but based on the realization that the early evolution, the natural selections, could not have happened in a single universe with fast changes in their environment, while in the multiverse there are many samples and long times like Darwin had for his finches and people. An example of natural selection is in Fig. 2, that is of course for millions of years later and still happening today.

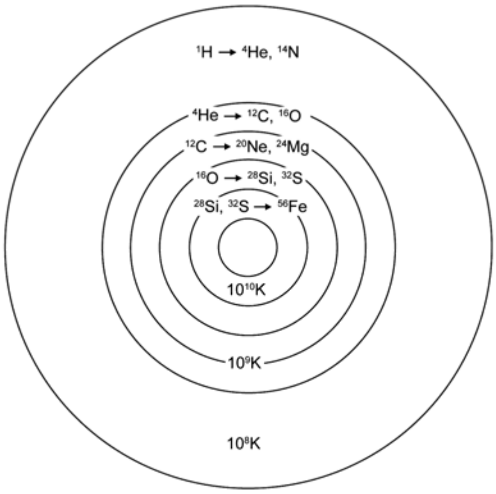


Fig. 2. An example of *natural selection* but at a much later time in our evolution. Increasingly heavier atomic nuclei are formed in increasingly energetic environments in massive stars of spectral types O and B

Neutrinos appear at t = 1 sec, electrons 15 seconds later, and helium nuclei one minute later. At about that one minute of age, the Standard Model shows a spell of a few minutes during which the conditions were right for the assembly of the nuclei of helium, plus traces of heavier nuclei. However, the change towards less dense conditions still happened fast in that expanding universe; such that the conditions of pressure and density for the combination of nuclei heavier than protons lasted just a few minutes. Therefore, only a limited number of helium nuclei were formed. However, the numbers computed for these conditions are confirmed by the numbers of helium nuclei that are presently observed in the universe. Here lies a success for the later times in the Standard Model.

About the following millennia we know little, other than that the universe consisted of photons being scattered around by protons, electrons, *etc.* – a noisy and energetic mass in expansion, too dense for light to escape (except, perhaps, an early Photon Burst, we shall see). Not much was happening excluding this *multiple scattering* of light, somewhat similar to what happens inside the Sun, where it takes a million years for a newly made photon at its center to be bounced to the outer levels and then onwards to Earth. All we can do with the present models is to consider the plasma 380,000 years later.

The Universe at Age 380,000

Finally, at t = 380,000 years, came the last remarkable event of the Standard Model for the Big Bang. A great transition took place when the temperature and pressure reached a level at which the scattering diminished, resulting in less density such that the photons were knocked about far less violently and frequently. As a result, the nuclei (of protons and neutrons) and the electrons no longer had enough collision energy to stay free, so they coalesced into hydrogen and helium *atoms*, which are widely spacious to let the photons pass through them, as was mentioned two sections before.

By the time the expanding universe had reached the temperature of about 3,000 Kelvin, it resulted in the separation of matter and radiation. The radiation from that stage is still observed today. It is all around us, observed far away between the stars to where it has expanded since $t \sim 10^{-6}$ sec. By now, the cloud has cooled to almost absolute zero. It is called *the 3-degree-Kelvin radiation*, 3 K being its approximate temperature. The COBE and WMAP spacecraft confirmed this. WMAP also derived that the detailed formation of galaxies and their stars began some 200 million years after the beginning of the universe. The galaxy formation peaked about 5,000 million years after the beginning, which is at 5×10^9 years of age for the universe, compared with the present age of 13.7×10^9 years (Spergel *et al*. 2007). The epoch of 5×10^9 years will return in the history of our universe, when the *acceleration of the expansion* was discovered.

Galaxies

Here is another thought experiment. Imagine that you are a galaxy in space... rather lonesome because space is so large and dark... but still, your gravity will make you move towards whichever galaxy wins over the gravitational tugging by all the others, because it may be closer or more massive. That is how the Andromeda Galaxy is moving towards our Milky Way Galaxy. Such gravitational *interaction* causes the galaxies' motion, groupings, clusters, and collisions.

Although nearly all galaxies expand away from us, they themselves do not expand (nor do we): galaxies have their own gravitational regime. The expansion takes place between the galaxies, and is therefore properly referred to as *expansion of intergalactic space*.

With the naked eye we can see at a certain place in the sky a 'Milky Way' band encircling the sky with the multitude of stars in the galaxy (Fig. 4). We see that because our solar system lies in the flat galaxy that from the outside would be seen in the shape of a discus (Fig. 3). The solar system is also about halfway from the center to the outer rim, and so we see more stars in the direction towards the galactic center than in the opposite direction. Now, if we look in the third direction, away from the plane of the Milky Way, up and down in the paper of Fig. 3, we see fewer stars that belong to our own galaxy and more of the dark sky because we look deep into space where there are few disturbing stars in our own galaxy or its halo. Now we can see other galaxies in the distance.

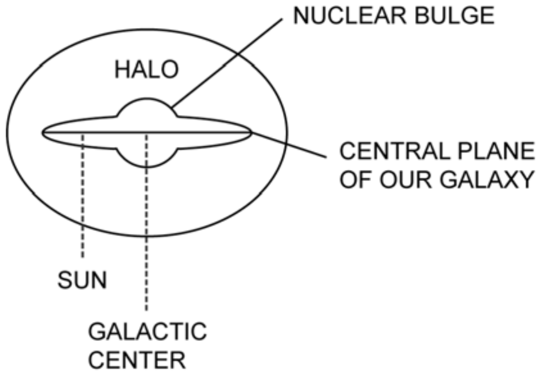


Fig. 3. An edge-on sketch of our Galaxy. The central plane is at the middle of the Milky Way seen in Fig. 4. The Halo is exceedingly faint

The detailed studies of our galaxy require a rich variety of *observational facilities* such as infrared instruments and radio telescopes, along with *theoretical studies* for the formation of star systems and for the dynamics of moving objects. Our galaxy has about 10^{11} bright and faint stars and a great variety of clouds consisting of gas and dust; it all moves through endless turbulence and dynamics.

How large is our galaxy? At the speed of light, it takes about 26,000 years to travel to its center, from the sun to the right in Fig. 3. We are roughly 24,000 lightyears away from the outermost visible stars in the other direction, to the left; this is approximate, because there is no sharp edge to the galaxy. We therefore are located a little more than halfway from the center to the edge of our galaxy; the diameter of the galaxy is about 100,000 (10^5) lightyears.

Galaxies occur in groupings. Outside of our Milky Way, we first encounter others of our *Local Group,* which has about 30 members and is about three million lightyears in size. The local groups are members of *clusters*, which have a million or so groups, and they in turn may be parts of *superclusters*, having some ten-thousand clusters.

The Interstellar Medium (ISM)

Experts study the characteristics of whatever floats through interstellar space by observing its interaction with starlight, and they find scattering properties in different colors at different wavelengths. Classical studies have made use of such observations with special instruments such as spectrometers and polarimeters at large telescopes.

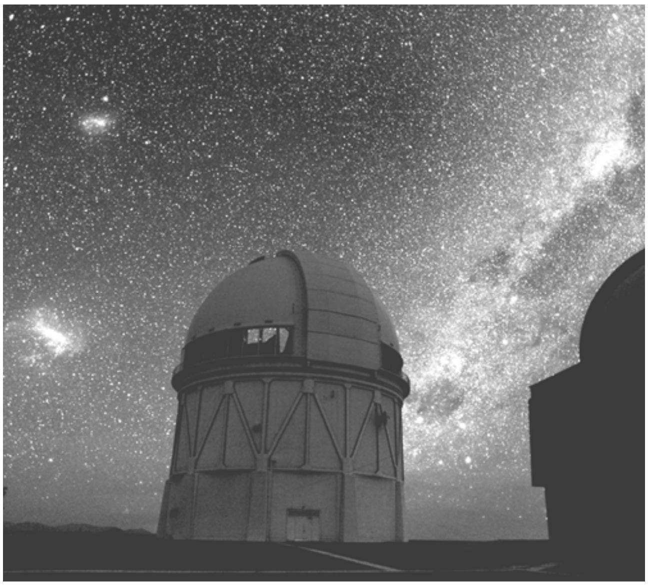


Fig. 4. View of the Milky Way with its dark dust clouds, the two Magellanic Clouds, and telescope domes at the Cerro Tololo International Observatory in Chile (Copyright: Roger Smith, NOAO/AURA/NSF)

Interplanetary and perhaps even interstellar grains have been collected by high-altitude aircraft and balloons, and made available for investigation with microscopes and other instruments in the laboratory. In this way, the 'Stardust' spacecraft flew through interplanetary space to collect samples, as well as grains near Comet Wild 2 in 2004, dropping them off in a capsule for a parachute landing on the Utah desert in 2006. It was a great success; other comets had been studied from spacecraft, but this was the first return of a sample.

Another branch in astronomical science studies molecules that are either drifting freely in space or are attached to interstellar grains. Faint glows emitted by hydrogen molecules sometimes show clouds of interstellar gas. The reason for emphasis on this interstellar medium of gas and dust is that they play a primary role in the formation of the stars. The reason for mentioning the ISM here is that we use it for the study of the inter-universal medium (IUM).

Introduction to Dark Energy and Dark Matter

There is a special challenge in astrophysics and cosmology, which is to understand the physical nature of dark energy and dark matter observed and called by these names, but not understood until now. They are the dominant contents of our universe. You will see the miracle occurring when we switch from our sole universe to considering the multiverse; both problems will then be solved in an elegant manner.

Observable matter, which is called ‘baryonic’, from the Greek word *barys*, or heavy, accounts for only 4.6 % of the total mass in the universe, while neutrinos equal less than 1 %. The rest is ‘dark’ or not observed: 23 % of dark matter, while 72 % of the total is in some form of dark energy, believed to be the cause of the acceleration of the intergalactic space expansion.

It was not always so. The abundances were not the same when the universe was young as they are now, and that may be a clue for origins. When the universe was young – age 380,000 years – the observable universe was made up of 12 % atoms, 15 % photons, 10 % neutrinos. What was not observable but otherwise derived to be present was 63 % dark matter and a small amount of dark energy.

The Cosmos Has a Single Unified Physics

Young Chandrasekhar showed such promise at the age of 19 at school in Madras (now Chennai), India, that he was accepted for PhD studies at Cambridge University in England. During his long sea voyage to Britain, he did much of his work on stellar structure, which he would further develop throughout the 1930s, and for which he would receive the Nobel Prize in 1983. He moved to the Yerkes Observatory of the University of Chicago in 1937 and to the Chicago campus in 1956.

His equation for masses I referred to at the start of this paper is

$$M(\alpha) = (hc/G)^{\alpha}\, H^{1-2\alpha}, \qquad \text{(Eq. 1)}$$

in which h is the Planck constant, c is the velocity of radiation, G is Newton's gravitational constant, and H is the mass of the proton; positive exponents α identify the objects shown in Table 1. The derivation takes seven pages, and there was some awe about it, that Chandrasekhar (1951) was aware of – he did publish the equation four times – but he thought it was too early to use it to explore '...deeper relations between atomic theory and cosmogony...' Shu (1982) referred to it in his classical textbook as '...one of the most beautiful and important formulae in all of theoretical astrophysics...' Max Planck had already expressed the possibility for *extraterrestrial* application of constant h:

> …the possibility is given to establish units for length, mass, time and temperature, which, independent of special bodies or substances, keep their meaning for all times and for all cultures, including extraterrestrial and non-human ones, and which therefore can be called 'natural measurement units'… (Planck 1899).

He wrote that more than a century before the present time, when we had not as yet discovered any society on a planet of another star. What he probably surmised is that the Cosmos has a single physics that is nearly perfect; it is that physics we consider in the Chandra-Multiverse model.

At the time, however, no one *did* anything with Chandra's equation. The time was not ripe; one had to wait for large telescopes and powerful spacecraft to deliver essential information about our own universe. This began in 1989 with the Cosmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (WMAP) that followed (Spergel *et al.* 2007).

I started working with Chandra's equation in 2001. The only purpose of the term $H^{1-2\alpha}$ is to be able to use any unit for the mass, such as kilograms and solar masses. But these would be useless in a multiverse; other planetary cultures would have their own units. For a truly *universal* mass, the obvious choice of unit is that of the proton. If all masses are expressed in terms of proton mass H = 1 and $H^{1-2\alpha} = 1$ for all values of α, such that:

$$M(\alpha) = (hc/G)^{\alpha}. \qquad \text{(Eq. 2)}$$

It appears to be a concise and powerful combination in terms of the Cosmos' physics, which would encompass perfect versions of quantum, relativity, gravity, and atomic physics (the latter because all masses are expressed in terms of the proton mass, H).

The next excitement was to see what happened for various values of exponent α, and that is in Table 1. Chandra had already published these values for the mass for the 'primordial stars' that provide our heavier atomic nuclei (spectral type O) at $\alpha = 1.5$, and the mass of our universe at $\alpha = 2.0$. The values under 'units shown' were computed with his Eq. 1; s.m. = solar masses. Each step of $\Delta\ \alpha = 0.50$ is a factor of 3.3×10^{19}. The last column has observations presented in values of α, for comparison with the first column. It is remarkable that the proton mass, determined in laboratories, would occur in Table 1, alongside the stupendously large structures of the cosmos, all by using the same equation. Observations became the predominant feature of this model, in the end as many as 30 have been used; this is unusual in cosmology.

Table 1. Computed and observed masses

Computed α	Proton mass	Units shown	Type of object	Observed α
2.00	1.13×10^{78}	9.52×10^{20} s.m.	Primordial universe	1.998–2.008
1.50	3.47×10^{58}	29.179 s.m.	Primordial stars	1.49–1.53
0.50	3.26×10^{19}	5.46×10^{-8} kg	Planck mass	0.500
0.00	1	1.67×10^{-27} kg	Proton	0.000

At this point, I was not rushing forward, but checking the quantization and the enormity of its consequences, because this modeling was clearly going outside of our universe. A correlation with Max Planck was found. On the last pages of his classical paper that initiated quantum mechanics, he has a derivation of units. For mass it is:

$$M(\alpha) = (hc/G)^{0.5}, \quad \text{(Eq. 3)}$$

which is what we now call the 'Planck mass'. He found the units from dimensional analysis of the cosmological constants *h, c,* and *G.* Chandra's analysis now provides the numerical calibration of the Planck mass. Eq. 3 looks like the Planck mass with a variable exponent; it might be called 'the universal Planck mass', and it is used in Table 2, below.

It is easy to do a dimensional analysis as a trial-and-error until getting a mass; for instance the Planck mass in kilograms is $(m^2 \text{ kg s}^{-1} \text{ m s}^{-1} \text{ m}^{-3} \text{ kg s}^2)^{0.5}$. The numerical values of the constants are being improved all the time (except for *c*); modern values for it are published yearly by Mohr and Taylor (2005). In 2005, their numerical values for the constants and their dimensions were $h = 6.626{,}0693(11)\times10^{-34}$ m^2 kg s^{-1}, c = 299,792,458 m s^{-1} (in a vacuum, exact by definition), $G = 6.6742(10) \times 10^{-11}$ m^3 kg^{-1} s^{-2}.

The treatment of the present paper indicates that the Planck constant is *h*, not ħ = h/2π, as is now used in most of the literature; it is easily seen that the values in the Tables would be off unrealistically for ħ. Three of Planck's units are:

$$\text{Planck mass} = (hc/G)^{0.5} = 5.455{,}55(40) \times 10^{-8} \text{ kg}, \quad \text{(Eq. 4)}$$

$$\text{Planck length} = (Gh/c^3)^{0.5} = 4.051{,}31(30) \times 10^{-35} \text{ m}, \quad \text{(Eq. 5)}$$

$$\text{Planck time} = (Gh/c^5)^{0.5} = 1.351{,}38(10) \times 10^{-43} \text{ s}. \quad \text{(Eq. 6)}$$

The Planck temperature and charge are also considered members of the basic set, and there are derived units; the entire set is called 'the Planck domain'.

Theoreticians prefer to have quantization in steps of whole units rather than fractions, and Table 2 shows a form of Eq. 3 for statistics with whole numbers for *N*. The key equation of cosmic masses is then,

$$M(N) = (hc/G)^{0.5N}, \quad \text{(Eq. 7)}$$

where the exponent gives the values of *N* shown in Table 2. *N* = 2 may be for the mass of the planetesimal; this needs checking by an expert. The constant quantization factor of 0.5 in the exponent yields the same factor 3.3×10^9 as before; for instance, the Local Group has 3.3×10^{19} universes. Note the arrow pointing up as *α* is not restricted so that there may be other steps of universes; the quantization topic needs expert study too. A theoretical challenge for future work may be to explain the mechanism of quantization. Such a solution seems to resort to the basic studies of quantum physics, with its connection to Planck's constant *h*. The cosmos has a quantized distinctness, without which it would be an amorphous brew.

Table 2. Masses in steps of Planck mass

No	Proton mass	Type of object
5	3.7×10^{97}	Local Group of universes
4	1.1×10^{78}	Universe
3	3.5×10^{58}	Original star
2	6.7×10^{38}	Planetesimal
1	3.3×10^{19}	Planck mass
0	1	Proton

The numbering of the alphas in *M(α)* and of *N* in *M(N)* is, however, an *anthropic curiosity* because we think of our universe as having α = 2.00, while another culture in another universe will at first do the same for its own universe. Imagine another culture elsewhere, with its intelligent beings seeing a hierarchy as we do, within and outside of their universe. They also start with their symbols for *h, c, G,* and *H,* which are measured in their laboratories, and they consider that their 'α values' are from 0 for the proton mass to 2.00 for their 'universe', and onward out into 'the multiverse'. Their universe is then imagined by them as we do for ours, as a member of an assembly of 3.3×10^{19} universes at α = 2.50, which is a member of 3.3×10^{19} assemblies at α = 3.00, *etc.* However, while their universe is at their α = 2.00 and is followed by their 2.50 outside, for us these two may be, for example, at our α = 7.50 and 7.00.

The fact that the equations are open – that *α* in Eq. 2 and N in Eq. 7 are unrestricted – has brought us the multiverse; I believe that an expert study is needed to establish whether it is an infinite multiverse. In any case, the equations set this multiverse apart from other models in having several characteristics of the universes. To begin with, we know their number at each step, with 3.3×10^{19} in the Local Group, the next step in the quantization being 3.3×10^{19} larger, *etc.* The *h, c, G,* and *H* cosmic physics is the same for all; the universes therefore are the same in their grand design, except for age (having started from inter-universal clouds at different times), and perhaps for some small effects of evolution.

This is such an important point that it bears repeating in different words. If our universe resulted from, and is decaying back into the inter-universal medium, the IUM must have the *h, c, G, H* physics we know for our universe. The IUM has uniformity through mixing of debris from a large number of universes, and they all emerged from the IUM as well, such that all universes surviving from the medium have that same physics.

Another characteristic is the *critical mass.* 'Critical' is defined as being not too large, or the universe would gravitationally collapse under its own weight; and not too small, or lack of gravitational pull would let the universe quickly expand into nothingness. The mass for N = 4 is indeed near that critical mass for our universe, and thereby for all others in the multiverse. Any universe with a mass different from this critical limit would simply not have survived; it would have collapsed or blown apart.

The interaction between the universes is by their expansion. The effect is seen from a bridge over a quiet pond on which raindrops fall. The rings on the water expand and travel through each other. The result in three dimensions is a thorough mixing of the debris from old universes. The resulting new universes will then be the same in composition because the mixing brings material together from many old universes.

The multiverse *endowed the universes with their mass, energy, physics, and evolution.* Each universe is thereby capable of doing all that we observe in our universe. That may include evolution going as far as having cells and chromosomes for its tools, and making beings who claim to be intelligent as one of its species, depending on local conditions such as the right distance of their planet from the central star. Comet and asteroid impact is also an uncertain factor.

The time scale of the multiverse is estimated by using the scaling factor of the Tables, 10^{19}, times the lifetime of our universe, $\sim 10^{11}$; the resulting 10^{30} seems consistent with the observations of the half-life of the proton. This is probably a real average, as some universes may have longer times, and others shorter, to come to the epoch of their formation through random cloud accretion.

The predominant characteristic of the multiverse is, however, that it is an *evolutionary* system (Gehrels 2012). It has the abundance of time and possibilities for change, just as Darwin and Wallace's evolutionary model for life had long times and much trial and error. The physics, the forces and components more basic than photons and protons are established and maintained during the long times and through many universe-samples in the multiverse. Universes that do not succeed end back in the inter-universal medium.

This intense and extensive evolution explains the fine-tuning of the nuclear transactions in stars that Fred Hoyle used to point out, how the selections and combinations could not have occurred if the physical constants of the elements had been even slightly different (Hoyle 1999, originally 1987, is actually on a discrepancy he had with Darwin's theory). The fine tuning, or any evolution, could not have happened if ours had been the *single* Big-Bang universe unfolding *fast* in the times before t $\sim 10^{-6}$ sec, 10^{37} PT on the standard clock.

The Aging and Decaying of Our Universe

We now begin a history of our universe, discussing its decay and demise first because we are in that stage at present so that we can identify the components of debris that are spreading into the multiverse. The major concepts of evolution include birth and demise; that we must die is therefore an evolutionary predicament.

Not only people and animals die, but inorganic substances age and decay as well. Even the protons may have a limited half-life, predicted at greater than 10^{50} years, but they are observed to be at least 10^{33} years old. This model then becomes useful right away because these two large numbers have been puzzling in comparison to the lifetime of our universe, which is some 10^{11} years. Why would the proton have evolved to live that long? The large numbers now are nicely in perspective with the timescale of the multiverse, $\sim 10^{30}$ years.

In summary, our universe is slowly dying with aging stars and galaxies. The dying of universes is a basic process in evolution, because the debris of radiation, atomic particles and other masses are used in the accretion for new universes.

Debris of Photons, Protons and Everything Else

Photons emerge from stars, supernovae, gamma-ray bursters and other energetic sources. Their aging is in terms of moving out, cooling steadily. Their radiation is observed 3,000 K at age 380,000 and 3 K at present. On the time scale of 10^{30} years, the old photons must be near zero degrees during most of their time in the multiverse. The situation near 0 Kelvin, –273.15 degrees Celsius, 'absolute zero', needs to be studied by experts, including the common understanding that all thermal activity stops. I use the term 'near 0 K' for what seems to be the ground state of photons, protons, *etc.*

The *protons* cool in the expansion as the photons do; other particles such as neutrons and electrons, they all are part of our universe's decay debris, as are old stars and brown dwarfs. Whole galaxies expand outwards too – and clusters of galaxies – while they internally keep their gravitational ensemble together for whatever is aging inside; these mass configurations, as decayed as they are, appear conserved in the multiverse because they are recognized with the Wilkinson Microwave Anisotropy Probe (WMAP) in our early universe. Dark matter and dark energy must be included in the discussions too because they are abundant in our universe, 23 % and 72 %, as we saw above.

The Acceleration of the Expansion

In the late 1990s, two teams of observers went to two different but large telescopes on the prevailing prediction that the expansion of our universe is decelerating, such that a shrinking and collapse would be its fate. Because of their competition and interesting goal, the astronomical world seemed to be travelling with them. The two used the same techniques and made the same opposite discovery: of an accelerating universe as of age 5,000 million years after the beginning, which is at 5×10^9 years of age for our universe (Goldhaber and Perlmutter 1998; Riess *et al.* 1998).

The discovery is crucial for the Chandra-Multiverse model because a decelerating expansion, leading towards collapse of our universe, would have made it impossible. We speak of *accelerated* expansion of intergalactic space, and now see how all the debris of a decaying universe proceeds on the accelerated expansion into the inter-universal medium in which our universe is imbedded.

Furthermore, the cause of the acceleration was interpreted to be dark energy (Riess *et al.* 1998), and we will use this as the key to the physical interpretation of dark energy below.

The Inter-Universal Medium (IUM)

The description of the IUM uses the extensive knowledge and understanding of the interstellar medium (ISM) between our stars and galaxies. The ISM is tenuous in most places while in others there are extended clouds of hydrogen and other molecules. The ISM clouds have lifetimes of $\sim 10^9$ years, there is always motion in the universe so that they encounter, combine and grow. Concepts from the ISM may be used in this modeling because the same types of processes are bound to happen in the IUM, albeit over much larger cosmological scales of time, $\sim 10^{30}$ years. The comparisons must, however, be made with care because the conditions of size, density, radiation environment, *etc.* of the two media differ greatly.

Anyway, the space density in the IUM will be as uneven as it is in the ISM, with huge clouds accreting. The clouds grow by sweeping the material up during their motion through space. Eventually, self-gravitation will become dominant by its increasing gravitational cross-section, speeding the sweeping and contraction of the cloud towards making a new universe.

The principal difference of the IUM is in its *composition*, now no longer the young and active material of hydrogen and other atoms but, instead, the above old and cold decayed debris objects that had come long before from dying universes. This is *energy seeking* material. Gravitational energy of the contracting cloud is now being used internally. The growing proto-universe can therefore increase in mass without getting as hot as a proto-star would do, because the gravitational energy is used to re-energize the old cold photons and to re-energize and re-constitute the atomic particles into regular photons, protons, and neutrons.

That this actually happens is seen in observations of the preservation of characteristics, for instance the clustering of galaxies coming from old universes being recognized for our young universe, as was mentioned above. In other words, the characteristics have *not been melted away.*

Now we are ready to consider the cloud of the IUM that made our new universe. The spherical gravitational mass is of overall *uniform* composition because the debris from many universes is mixed together, with all the above components of old photons, subatomic parts of protons and neutrons, and of dark mass and dark energy. This is the uniformity found by the COBE spacecraft till the third decimal of 2,728 K.

Characteristics of Photons

Cooling is the prime characteristic of this decay, and it is here connected with the expansion whereby the density of the material diminishes as a cooling mechanism. We then speak of old cold photons, of old cold protons, *etc.* COBE discovered them to have 3,000 K at age 380,000 and 3 K at present. At the 10^{30}-years time-scale of the multiverse, they all spend most of their existence near absolute zero. Their *ground state* is therefore near 0 K; the energy equation of photons has a zero-point energy term (Lamb 1995). Only occasionally may they be scooped up to proto-universal duty, and only a very few will serve the most complex and capable form of evolution known to us as *life*.

Photons are even more important than the protons because there are at least 10^9 more of them; this enormous ratio may become better understood with the role the photons play in the birth of universes in the present model. The characteristic most needed here is their radiation pressure. It depends on the *fourth power of the temperature*, such that it will be strongest for re-energized photons, rather than for the old cold ones.

The Proto-Universe and Photon Burst

We are now ready to pursue the cloud that started our universe, the *universe's nebula*, towards the more advanced form that is called the *proto-universe*. We recall that the IUM *composition* is totally different from that in the ISM, namely of decayed energy-seeking debris. This is the *first reason* why the growing proto-universe did not become impossibly hot and the galaxy clustering was observed; 10^{13} K will be its maximum temperature, as the standard models indicate. The mass of our baryon universe is equivalent to 10^{21} solar masses, but it consisted at this time mainly of energy-absorbing material, and may therefore not have collapsed into a black hole during accretion.

The *second reason is the radiation pressure* of photons, which provides counter-action to gravity. Imagine a box in IUM space, with radiation from many photons pushing in all directions, but at the periphery of the box the pressure is outward. That results in direction opposite to that of gravity in the proto-universe cloud; it will have negative sign if and where one uses a positive sign for gravity.

It is noted that the maximum temperature of 10^{13} K makes it unnecessary to do a computer modeling at this time. Before a detailed modeling, an outline like this is needed, so we proceed with the history of our universe on first principles.

The rate of accretion was first controlled by the geometrical cross-section of the cloud, but gradually the larger gravitational cross-section would accelerate it. Already then, the cloud would gain outward force due to radiation pressure. During the accretion this build-up of pressure *lagged behind* the gravitational force of the cold photons. It brought about the appropriate temperatures, from near 0 to 10^{13} K, as density increased from the start at $\sim 10^{-28}$ to 10^{18} kg m^{-3}, but that would take time for the heat to penetrate, and cause a delay. The gravity thereby prevailed and the accretion continued.

The fourth-power dependence of the radiation pressure on the temperature must have caught up to end the accretion phase. The remarkable epoch would be reached for the completion of re-energizing the photons, namely at 10^{18} kg m^{-3} and 10^{13} K, at the time of t $\sim 10^{-6}$ sec on the old clock at which we know these three conditions from the Standard Model for Big Bang as well as the one for atomic physics.

An increasingly enormous number of old photons reached their full radiation; in other words, an *inferno* flared out. It must have increased from a small to a maximum central volume; the maximum depends on the total mass and on the percentages below. That central region of sizeable volume was converted from old photons to re-energized radiating photons in a Photon Burst. The one-but-last paragraph of this paper has more information about this stage.

The Photon Burst may have been detected by WMAP as a shell beyond that of the 3-K radiation; *i.e.* of greater radius than that of the 3-K radiation. Analysis of that WMAP observation would confirm that a flaring occurred like thousands of Super Novae!

Another observation may support this reasoning, in addition to the percentages observed uniformly and of a WMAP verification, namely that we do not observe a center for our universe. This confirms that the enormous Photon Burst brought passage and thereby *mixing* of baryons with the non-baryons throughout the whole universe. This explains that '12 % atoms, 15 % photons, 10 % neutrinos, 63 % dark matter and little dark energy' were observed by COBE and WMAP *everywhere the same* in our early universe.

Continuing in the central volume is the re-energizing of the protons *etc.*, which took longer than that of the photons because they were more complicated. They however encapsulated the inferno by holding the photons back through multiple scattering, which sustained the expansion; now the appropriate temperature from 10^{13} K to 3,000 K was *lagging behind* the decrease of the density from 10^{18} to 10^{-19} kg m^{-3}, because the cooling took time. The multiple scattering thereby prevailed and the expansion continued.

At t = 380,000 years, the space density had become low enough for the electrons, protons and neutrons to combine, to make atoms. Atoms have internal space to let photons pass through them; remember the 'stadium experiment' in the beginning parts of this paper. This blast would also have a pressure blowing the IUM material further away from the completed universe. The expansion continued for there was nothing to affect it, until age 5 x 10^9 y when the acceleration occurred.

The Identity of Dark Energy

We have seen in the previous section that at age 380,000 years there was only a small amount of dark energy left. The dark energy was apparently mostly used up in the accretion and birthing processes and so were most of the old, cold photons. A discovery thereby occurs: *dark energy is the energy of old photons*; *causing* the expansion has the same physical action as *accelerating* the expansion, which was mentioned above as being due to dark energy (Riess *et al.* 1998). In other words, the conclusion is that the original expansion was caused by dark energy, which also caused the later acceleration.

There appears an indirect but powerful reasoning to confirm this interpretation of dark energy. The evolution in the multiverse could not have left unused the major part of what each decaying universe contributes to the IUM, the old, cold photons (or, in the second interpretation, the 72 % dark energy). They must have been fitted into the evolution of universes and their survival. If they would have been left unused, the ever-increasing number of old-photon debris (*i.e.* dark energy) would have overwhelmed the cosmos.

The Identity of Dark Matter

The Chandra-Multiverse model might also provide a confirmation of the solution to the problem of the physical nature of dark matter, for which there is already a large literature. Astronomer Bernard Carr (2001) has a critical discussion of both baryonic and non-baryonic dark matter. Physicist Padmanabhan (2002) concludes: '*both baryonic and non-baryonic dark matter exist in the universe, with non-baryonic being dominant*'. His general rule is: '*There is not an a-priori reason for the dark matter in different objects to be made of the same constituent*'. He also discusses baryonic and non-baryonic matter such as protons, WIMPs, axions, neutrinos and massive astrophysical halo objects. Their results support that old, cold protons and other particles, such as neutrons and electrons, are part of our universe's decay debris, as are whole galaxies (each gravitationally holding its debris), clusters of galaxies, and whatever other debris such as old stars. As a short name for all these we use 'protons, *etc*'.

The dark-energy actions for multiple scattering and expansion have a counterpart in the dark-matter actions – the one gives and the other receives the kinetic energy of the photon. In other words, while dark energy is the kinetic energy of photons, old or new, dark matter may be the name for the ensemble with which these photons interact, old protons, *etc.*

This view of dark matter has the same simple but firm confirmation as dark energy. The evolution in the multiverse could not have left unused the major part of what all decaying universes contribute to the IUM, namely their 23 % dark matter (old protons, *etc.*). If this matter had been unused, the ever-increasing amount of dark matter would have overwhelmed the Cosmos.

The Cosmological Constant

The cosmological constant is sometimes held responsible for the acceleration of expansion at the universe's age of 5×10^9 years. The previous scenario applies just as well as the cosmological constant for the physical acceleration action by old photons. There is a certain form of logic in it: That the multiverse evolved the properties of the photons, in order to get the cycle of decay and rebirth accomplished, for expansion and survival of the system. In the equations, that is expressed as a cosmological constant.

A refined support of the model appears in the fact that the cosmological constant is exceedingly small, because its theory entails *fine-tuning* of physics and nature to 'bizarre' accuracy (Padmanabhan 2002). Hoyle (1999 [1987]) used similar strong words alike 'bizarre' for extreme fine-tuning of the nuclear constants inside stars. The evolution of the universes in the multiverse has indeed bizarre long times: $\sim 10^{30}$ years, and numbers of samples, $>10^{19}$ universes. That is how the Chandra Multiverse accomplishes extreme fine-tuning for all parameters and characteristics of the universes.

Schwarzschild's Confirmations

In conclusion of this paper is a warning that there was no Big Bang by Schwarzschild (1873–1916), which also confirms the Chandra-Multiverse model. He was a well-known and versatile physicist in Germany who had volunteered for duty during World War I. He caught a serious illness on the Russian Front, but worked on elucidating one of Einstein's equations until the end of his life (Schwarzschild 1916). It included a discussion that can be illustrated as follows.

Light can escape only when its kinetic energy is greater than its potential energy,

$$GMm\, R_S^{-1} < \tfrac{1}{2}\, mc^2, \qquad \text{(Eq. 8)}$$

such that there is a limiting radius of the body,

$$R_S > 2GM\,c^{-2}, \qquad \text{(Eq. 9)}$$

which is 3×10^8 lightyears; that is the radius of our universe at age 380,000 when the radiation did escape. Table 3 shows the comparison of R_S with radius $R = (3M/4\pi\sigma)^{1/3}$ for a sphere with uniform space density σ; $M = 1.13{\times}10^{78}$ proton masses.

The first line applies for the above case when the universe's radiation did escape, $R/R_S = 1$. The Schwarzschild limit does not seem precise because standard theories predict that density to be $\sim 10^{-19}$ kg m^{-3} at that time, not 10^{-23}. However, the precision of these predictions is low and the effect is small, because if 10^{-19} were used in the calibration of R_S, the following ratio is still rounded off to 10^{-14}, and the next is 10^{-38} instead of 10^{-40}.

Table 3. Radii and Schwarzschild radii

R/R_S	t	σ	R(ly)
1	380,000 y	10^{-23}	10^8
10^{-14}	10^{-6} sec	10^{18}	10^{-5}
10^{-40}	0	10^{96}	10^{-31}

The second line is for the above starting time of our universe, $t \sim 10^{-6}$ sec, using proton density of 10^{18} kg m^{-3} to derive R. Because $R/R_S = 10^{-14}$ is so very negative, the Schwarzschild radius indicates a giant black hole, but our entire universe is not a black hole. Furthermore, for photons to escape at age 380,000, they must have been *generated much earlier*. In the case of the Sun, it takes a million years for a photon generated at its center to escape; physically that is a different situation, but it serves this comparison. The shorter time of 380,000 years may be appropriate, even though the body is much larger than the Sun, because the medium is *expanding*. The photons' scattered journey was increasingly speeded up because it went through diminishing density, eventually as low as the above 10^{-19} kg m^{-3} (which was the density of the whole universe).

In the third line, the Planck density of 10^{96} kg m^{-3} is used for obtaining R and thereby R/R_S. Under these conditions, the Schwarzschild limit calls resoundingly for a giant black hole of our whole universe. Our universe was not a black hole, and neither radiation pressure nor dark energy is available to save the modeling (this is not to say that our universe would not have smaller black holes, of course). The Table confirms the Chandra-Multiverse model.

Acknowledgements

It is a pleasure to acknowledge the support and suggestions of Kees de Jager, Leonid Ksanfomality, and Barry Rodrigue.

References

Bitbol M. (Ed.) 1995. *The Schrödinger Interpretation of Quantum Mechanics*. Woodbridge: OxBow Press.

Carr B. J. 2001. Baryonic and Non-baryonic Dark Matter. *Birth and Evolution of the Universe* / Ed. by K. Sato, and M. Kawasaki, pp. 55–70. Tokyo: Universal Academy Press (earlier version at URL: http://arxiv.org/abs/gr-qc/0008005).

Chandrasekhar S. 1951. The Structure, the Composition, and the Source of Energy of the Stars. *Astrophysics, a Topical Symposium* / Ed. by J. A. Hynek, pp. 508–681. New York: McGraw-Hill.

Everett H. III 1971. The Theory of the Universal Wave Function. *The Many-Worlds Interpretation of Quantum Mechanics* / Ed. by B. S. DeWitt, and N. Graham, pp. 3–140. Princeton: Princeton University Press.

Gehrels T. 2011. *Can We Do without the Big Bang?* Tucson, AZ: University of Arizona BookStores..

Gehrels T. 2012. *Survival through Evolution, from Multiverse to Modern Society*. 2nd ed. Tucson, AZ: University of Arizona BookStores.

Goldhaber G., and Perlmutter S. 1998. A Study of 42 Type Ia Supernovae and a Resulting Measurement of Omega(M) and Omega(Lambda). *Phys. Letters* 307(1–4): 325–331.

Greene B. 1999. *The Elegant Universe, Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory.* New York: Random House.

Hoyle F. 1999[1987]. *Mathematics of Evolution.* Memphis, TN: Acorn Enterprises.

Lamb W. E. Jr. 1995. Anti-Photon. *Applied Physics B* 60: 77–84.

Mohr P. J., and Taylor B. N. 2005. CODATA Recommended Values of the Fundamental Physical Constants. *Reviews of Modern Physics* 80: 377–470, Table XXVI.

Padmanabhan T. 2002. *Theoretical Astrophysics.* Vol. 3. *Galaxies and Cosmology.* Cambridge: Cambridge University Press.

Planck M. 1899. Über irreversible Strahlungsvorgänge. *Sitzungsber. preusz. Akad. Wissenschaften* 5: 440–490.

Randall L. 2005. *Warped Passages.* New York: Harper-Collins.

Randall L. 2007. The Case for Extra Dimensions. *Physics Today* 60(7): 80–81.

Riess A. G. *et al.* 1998. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astron. J.* 116: 1009–1038.

Schwarzschild K. 1916. Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie. *Sitzungsberichte Königlich. Preus, Akad. Wiss. Berlin, Phys.-Math. Klasse* 189: 189–196; translated in 2003, On the Gravitational Field of a Mass Point according to Einstein's Theory. *Gen. Rel. Grav.* 35(5): 951–959.

Shu F. H. 1982. *The Physical Universe*. Mill Valley: University Science Books.

Smolin L. 2007. *The Trouble with Physics.* New York: Houghton Mifflin Co.

Spergel D. N. *et al.* 2007. Wilkinson Microwave Anisotropy Probe (WMAP) Three-Year Results: Implications for Cosmology. *Astrophys. J. Suppl*. 170(2): 377–470.

Steinhardt P. J. 2011. The Inflation Debate. *Scientific American* 304(4): 36–43.

Steinhardt P. J., and Turok N. 2007. *Endless Universe, Beyond the Big Bang.* New York: Doubleday.

Vilenkin A. 2006. *Many Worlds in One, the Search for Other Universes*. New York: Hill and Wang.

Woit P. 2006. *Not even Wrong, the Failure of String Theory and the Search for Unity in Physical Law*. New York: Basic Books.

6
Was There a Big Bang?

Leonid E. Grinin

Abstract

The idea that our Universe emerged as a result of the extraordinary power of the Big Bang from singularity (i.e., a state of an infinitely small quantity and infinitely high concentration of matter) is still very popular today. It was one of the main postulates of the Big Bang theory that completely formed in the 1960s–1970s. However, at present this idea as well as the Big Bang theory is outdated, although it is still shared by many scientists. Being widespread since the end of the 1970s the Inflation theory appears more modern. The main reason for the emergence of the Inflation theory was that the Big Bang theory could not satisfactorily explain a number of the contemporary parameters of the Universe.

The Inflation theory makes still widespread views of the Big Bang theory archaic as regards the following points: 1) the history of the Universe started with the Big Bang; 2) it started with the singularity. According to the Inflation theory, the Big Bang was not the beginning and the moment of the origin of the Universe, but it was preceded by at least two epochs: inflation and post-inflationary heating. That is, the Big Bang or precisely the hot Big Bang is just a phase transition from the state of cold inflation to the hot phase. Since the Inflation theory does not consider the Big Bang as the initial phase there emerges an intricate problem of the role of the Big Bang in the process of the formation the Universe as a whole. The paper considers the confusion with the Big Bang notion, a number and sequence of 'bangs' and why the theory can dispense easily without the notion the Big Bang. We will also discuss some advantages and disadvantages of the Inflation theory.

Keywords: *Big Bang, hot Big Bang, inflation, Universe, inflaton, false vacuum, the Inflation theory, singularity, quantum fluctuation, post-inflationary heating.*

I. Introduction: Big Bang Theory's Limitations and the Emergence of Inflation Theory

It is supposed that our Universe emerged about 13.82 billion years ago from the unknown state and substance. The history of the Universe, particularly its initial stages, is a scientific reconstruction whose most points are still considered to be scientific hypotheses. The latter sometimes may seem unbelievable ones. It is not surprising that there is no agreement among physicists and cosmologists about many issues of the initial stages of the history of the Universe. Completely formed in the 1960s–1970s the Big Bang theory has been popular long enough (about history of the term 'Big Bang' see Wood 2018: 1–4). At present it is outdated, although it is still shared by many scientists. Being widespread since the late of the 1970s the Inflation theory appears more modern.

The main reason for the emergence of the Inflation theory was that the Big Bang theory could not satisfactorily explain a number of the contemporary parameters of the Universe, in particular why the Universe is so homogeneous, isotropic, 'large' (spatially flat) and hot (Gorbunov and Rubakov 2010: 341; Guth 1997, 2002, 2004).

Due to the emergence of the Inflation theory many problems of the Big Bang theory can be eliminated. The Inflation theory makes still widespread views of the Big Bang theory archaic. For our topic the following points of this theory are more important: 1) the history of the Universe started with the Big Bang; 2) it started with the singularity (*i.e.*, a state of an infinitely small quantity and infinitely high concentration of matter). Below we will consider the main ideas of the Inflation theory. At present one should note that according to this theory the *Big Bang was not the beginning and the moment of the origin of the Universe, but it was preceded by at least two epochs*: *inflation and post-inflationary heating.* That is, the Big Bang or precisely the hot Big Bang is just a phase transition from the state of cold inflation to the hot phase. Furthermore, due to the fact that the Inflation theory does not regard the Big Bang as the initial phase there emerges an intricate problem of the role of the hot Big Bang as a whole. In the words of Alan Guth, it is not explained what 'exploded', how it 'exploded' and what caused the 'explosion' (Guth 1997). Thus, it is not surprising that (as we will see below) as a rule there is no clear description of the phase of the Big Bang in the modern researches.

II. Main Ideas of Inflation Theories

Historically, Alexei Starobinsky's model of inflation (1979) was the first model developed in detail. But the best-known Inflation theory was first formulated in the famous article by Alan Guth in 1981, which continues to promote it extensively. Within the framework of this theory 'the fundamental properties of our Universe (*i.e.*, it is homogeneous, isotropic, spatially flat and hot) appear as the consequences of extremely unnatural initial conditions' (Gorbunov and Rubakov 2010: 341).

The Universe before the hot Big Bang. The inflationary stage. According to the Inflation theory our Universe's origin was the result of *quantum fluctuation* (*i.e.*, negligible small fluctuation but still having certain spatial parameters[1]). This fluctuation has put the forces of the so-called false vacuum in motion. A false vacuum is a hypothetical state of matter, in which, matter is repulsed and space is expanded due to negative pressure. That is why this stage is called *the inflationary stage* (*i.e.*, inflation of the Universe). The Universe has reached enormous proportions in the smallest fractions of the second. One should mention that a false vacuum had constant temperature. That is why the inflation is defined as cold. The heating had begun due to the processes described below. Vacuum-like energy as well as false vacuum itself is now often called the *inflaton*.

Completion of the inflationary stage, post-inflationary heating. The false vacuum is an unstable state of matter, so it started to decay quickly. On the whole, the inflationary period (as well as all the initial stages of the early Universe) was very short. Nevertheless it is important for the theory that it must not be smaller than an extremely short period of time, measured in the smallest units, so-called Planck units (from 70 to 100 such units within the smallest fraction of a second).[2] This duration in terms of the Inflation theory was called the slow-roll of a scalar (inflaton) field. During this process the potential energy of this field decreased, transforming into the kinetic one. It is supposed that this leads to the formation of the so-called boson condensate. Eventually, the potential energy of the inflaton (inflaton field) reaches a minimum at a certain moment. This means that the conditions necessary for exponential increase are violated and the inflationary stage ends.

Thus, this leads to a rather rapid heating of the Universe. There comes *the stage of post-inflationary heating*, in which the boson condensate decays due to the vibrations (oscillations) of the inflaton field, which has reached its minimum energy. During the oscillation of the inflaton field one can observe the beginning of the formation of different particles about the nature of which there are different assumptions. The energy of the inflaton transforms into the energy of the emerging particles as a result of their interaction with the rapidly changing inflaton field. Figuratively speaking, one can observe 'pumping out' of the energy that led to the rapid heating of the Universe and the formation of elementary particles of ordinary matter. In other words, the entropy that was previously low in the false vacuum increased sharply. At the same time there was a rapid expansion of the Universe. And the inflationary equation of state of matter transforms into the powdered one. And later, when the heating had reached its peak, the powdered equation of state transformed into a radiation dominated equation. In other words, when reaching an ultrahigh temperature, the matter passed into the state of 'super-hot plasma consisting of free quarks,

[1] It differs much from the Big Bang theory which regards the starting point of the Universe with the singularity (see above).

[2] 100 Planck times is something like a period of time from $5 \times 10^{-44} – 5 \times 10^{-42}$ s.

gluons, leptons and high-energy quanta of electromagnetic radiation' (Levin 2010).[3] *Hence, within a fraction of a second there took place successively equations of state of a false vacuum – powdered – radiation dominant* (for more details see Grinin 2013).

Was there the hot Big Bang Phase? There are discussions about the temperature which the post-inflation Universe achieves as a result of these processes. In any case, it was very high, although, most likely, it was lower than it was expected in the Big Bang theory.[4] According to cosmologists and physicists, this leads to a kind of 'boiling up' of the vacuum, which, it should be noted, has already occupied plenty of space by that time. Postnov (2001) notes that during the period of 10^{-34} s. the stage of inflation 'prepares' the primary very hot substance in a very small area, and it expands by inertia.

Thus, in the previous description one can see a phase transition from the state of cold inflation to the hot phase. Just in this point of the Inflation theory there emerges a problem of identification of the place, role, and even reality of the Big Bang. The Inflation theory does not give any definition of this concept. On the one hand, they maintain that heating had resulted in a hot Big Bang which further dispersed the expansion of the Universe. One cannot find any great explosion in the hot Big Bang phase unlike the picture drawn in the classical Big Bang's scenario. Of course, one can call heating of the Universe the Big Bang, but the heating is a process, it was not momentary as a sudden explosion assumed by the Big Bang theory. We will return to the problem of the Big Bang concept below.

III. Comments on the Inflation Theory

The advantages of the Inflation theory from a philosophical point of view in comparison with the Big Bang theory are as follows: 1) the existence of matter and the Universe before the phase of inflation is supposed; 2) anyway, the process of the formation of the Universe looks exactly like a process (although very fast), but not as an act of creation from nothing; 3) the original size of the Universe, although small, but it is still more verisimilar than the singularity (the latter is an artifact of outdated cosmology); 4) the introduced hypothetical substance – the inflaton field – explains the processes as a whole with the help of physics, and not simply by the assumption of an explosion.

In the Big Bang theory, as the beginning of everything which emerged from the singularity, it was believed that the classical space-time started to form immediately in the course of explosion, because the Universe began to expand

[3] An ordinary matter that had appeared as a result of boiled vacuum and then a hot 'bang' had been remaining in a state of hot plasma for hundreds of thousands years (until the process of hydrogen recombination).

[4] Although there are no direct experimental indications that there were temperatures above several MeV in the Universe (*i.e.*, several tens of billions of degrees) (Gorbunov and Rubakov 2012).

at once after the hypothetical state of singularity and also acquire the related characteristics. As Hawking (2001) wrote, Einstein's general relativity theory concludes that space-time arose at the singular point of the Big Bang. However, we proceed from the fact that the hot Big Bang was preceded by the inflationary phase, during which the Universe significantly expanded. Thus, a very rapid expansion of the Universe during the given period leads to the origin of classical space and time.

Disadvantages of the Inflation theory and its Physical Fatalism

Now let us consider the disadvantages of the Inflation theory. They are as follows: 1) Introduction of the hypothetical substance. Inflation requires the introduction of some powerful repulsive force for its explanation (*i.e.*, the inflaton field or a false vacuum with negative pressure), the nature of which is not clear in many respects (see May *et al.* 2007: 38–39). In the inflaton field the laws of ordinary gravitational physics change, because 'matter becomes not a source of attraction, but a source of repulsion' (Sazhin 2002: 38). Filling gaps with hypothetical kinds of matter is a form of science development. In this case it seems as a too bold idea.

2) The assumption that initially the Universe had very small (almost Planck) size and a huge Universe could arise from that size. We have no example of evolutionary processes when something very large would have turned out from one tiny unit. The process always proceeds as either the coexistence of the mass of small units, which then form a new macrosubstance (system), or the gradual acquisition by a certain number of small units of the ability to grow and, as a result, the emergence of large units.[5]

3) The Universe's origin time is too short. In the end, although the Inflation theory significantly withdrew from the concept of the 'act of creation', as opposed to the Big Bang theory, but the generic features of this approach are still visible.

4) As in the case with the Big Bang theory an issue about the origin of the Universe which appeared to be 'almost from nothing' raises many questions (Rubin 2004). It is also unclear, 'where does the material come from in the first state of the world' (Cherepashchuk and Chernin 2004: 278).[6]

[5] To a certain extent the assumption of multiple of multi-faceted universes also implies such a variant of gaining the ability to grow, but the idea of multitude of universes is too speculative to be associated with evolutionary processes.

[6] Postnov (2001) points out that the exponential growth in the sizes of the area with constant density means the growth of mass (energy) inside the area 'out of nothing', which might seem strange at first sight. However, there is no violation of energy conservation law – the growth of the positive energy is exactly compensated by the negative energy of the gravity field, which is created by the 'emerging' positive energy inside the expanding area. Therefore, the total energy remains the same in the course of inflationary expansion (see also Sazhin 2002).

5) In general, both the Big Bang theory and the Inflation theory proceed from the fact that they must explain the present observed states of the Universe, including Hubble's law, the spatial homogeneity of the Universe, its flatness, *etc.* How far is such predeterminancy possible in terms of evolution?[7]

Why should these states be explained by the very initial conditions? Why could not they arise later under the influence of any factors? Apparently, this is connected both with the desire of cosmologists and physicists to see a complete picture that would explain everything, and that otherwise, if the theory of the origin of the Universe does not explain the present observed circumstances, then it is easily refuted and, in fact, not even considered. As a result, the emergence of Hubble's law is included in the Inflation theory, although why should not this expansion (if the redshift would not be later explained in another way) emerge later? The expansion of the Universe having arisen at the very first moment does not change by inertia. It looks rather fatalistic.[8] Moreover, the entire subsequent large-scale structure of the Universe was therefore predetermined by the smallest density fluctuations, which already appeared at the inflationary stage within extremely short fractions of milliseconds. It is very sad to realize that everything was decided in such a short period and in such a small amount (from the Planck size to 1 cm^3) of matter. Although the Inflation theory aims to withdraw from the concept of singularity with its full uncertainty in the physical realm, nevertheless the original dimensions are difficult to perceive.

One should mention that quantum dimensions of the original Universe in comparison with the singularity from the point of view of physics is a principally different state, since it allows using already known or at least formulated hypothetical laws and forces. But from the perspective of the ideas of evolution the differences between singularity and quantum dimensions are not considerable.

IV. Confusion and Problems with Big Bangs

As we have seen from the above discussion, the stage of the hot Big Bang succeeded the post-inflationary heating stage. However, there are still a number of scientists who, just as before, consider the Big Bang as the moment of the origin of the Universe followed by inflation. However, this disagreement can be explained not only by differences in points of view but also by the confusion in terminology. The question is that when speaking about the Big Bang as

---

[7] Even the proponents of such views have to admit that 'according to modern ideas, space-time in the Planck scale is a fantastic figure, more like a monster from horror films than the object of physical research. Future research will show whether this picture is correct' (Sazhin 2002: 81). About evolutionary approaches in respect of Cosmic phase of Big History see Grinin 2014, 2018.

[8] Not to mention the fact that this contradicts the fact according to which the speed of the receding galaxies not only decreases but increases.

an event preceding the beginning of inflation, it is often meant not the hot Big Bang (*i.e.,* classical Big Bang), but another one, *i.e.* the pre-inflation Big Bang.

Thus, today speaking about the Big Bang, it is necessary to specify which explosion is mentioned. The fact is that there is no common terminology concerning the Big Bang in physics and cosmology, which study the early Universe: there is considerable confusion here.

Sometimes the followers of the Inflation theory mention the Bang that preceded the inflation stage. *They might regard this bang as above mentioned quantum fluctuation or another hypothetical event of uncertain origin. Sometimes they talk about such Big Bang as a special phase of early history of the Universe. Unfortunately they do not clarify whether this Big Bang was the trigger for the quantum fluctuation, or it is just the beginning of the inflationary stage. In any case this Big Bang was definitely cold but not hot.* However, some researchers do not identify the pre-inflation cold Big Bang or do not mark it out as a special stage because such a variety in approaches implicitly creates a great confusion in our understanding of the notion of the Big Bang. Were there two Big Bangs or was only one or none at all? And after what stage it occurred? The confusion is growing because the Big Bang theory also implies the inflationary stage. But the sequence of stages differs from that of the Inflation theory. According to the Big Bang theory, the Big Bang was the first to occur and led to great inflation. And the Inflation theory suggests that the hot Big Bang resulted from the inflation. As we will see below such a shift of the processes' order makes the Big Bang unnecessary stage in the sequence of events that occurred in the Universe. One should also note that not all researchers distinguish the stage of post-inflationary heating.

As a result this situation seems paradoxical. On the one hand, practically, there are no scientists who would definitely reject the Big Bang. On the other hand, a number of researchers who use this concept as something conventional, but indefinite, increases. It appears that implicitly or even explicitly they understand that the theory can easily avoid using the Big Bang notion. However, because the direct negation of the Big Bang may cause difficulties, probably they think that the best way to avoid problems is the indistinct mentioning of this moment. Thus, one should mention that the Big Bang seems to become a kind of metaphor, an indicator of fidelity to the mainstream, playing a role similar to that of the incomprehensible god in deism philosophy. We recall that the situation is greatly complicated.

Among many followers of the Inflation theory there is an implicit assumption that there could be two trigger events which can be described as 'bangs', one of which preceded the inflation, and the other – followed it.

But the description and characteristics of the pre-inflation Big Bang are even more obscure than those of the hot Big Bang. It also does not have any common term; there are references to the Planck era of the Big Bang, the early

Big Bang stage, the real Big Bang, *etc.* One should mention that due to this terminological and theoretical confusion it is extremely difficult to understand whether one or two explosions are meant, as well as to describe the real sequence of stages.[9] If there were two Big Bangs then the origin of the Universe would schematically look like this: the pre-inflationary Big Bang – inflation (expansion of the Universe) – post-inflationary heating of the Universe – the hot Big Bang. But such a reconstruction is not presented anywhere because perhaps as was mentioned above it is easier to avoid difficulties. Most commonly mentioning of the Big Bang among physicists simply looks like a tribute to a tradition they dare not to violate, and therefore such mentions are rather ritual than filled with specific content.[10] In general, it appears that the early history of the Universe may well do without using the concept of the Big Bang, using the scheme: fluctuation (whatever it may have been caused by) – inflation – post-inflationary heating.

Thus, among a large number of astrophysicists the very idea of the Big Bang has been losing not only its substantiality and uniqueness, but the need in general. However, at the same time among others and especially among those who popularizes the early history of the Universe one can observe dominating desire to see something extremely real and apocalyptic in the Big Bang. Perhaps, it would be too strong to call the Big Bang 'a misleading, ugly and trivializing name' (as Timothy Ferris did; see Wood 2018: 2). Nevertheless in the light of modern points of view it is very necessary to regard the Big Bang not as a real huge explosion but rather as a metaphor that still exists due to tradition.[11] The matter is that 'the Big Bang, just as we imagined it traditionally, most likely did not occur at all' (Mukhanov and Orlova 2006). At the same time, most initial conditions that determine the most important characteristics of the modern Universe are also referred to the inflationary stage, rather than the hot Big Bang.

---

[9] It is difficult to understand also from Guth's article with the title 'Was Cosmic Inflation the "Bang" of the Big Bang?' which is relative to our topic.

[10] Though one can find the following arguments. It is shown that such an event as the hot Big Bang is not a necessary stage in the Inflation theory. Now it is clear that the inflationary stage played a role of the 'bang' (Postnov 2001). The moment when the Universe is heated up is *now called the Big Bang* (*Ibid.*). The boundaries between inflating and thermalized regions play the role of the Big Bang for the corresponding thermalized regions (Garriga and Vilenkin 2001; Vilenkin 2006, 2010; in all cases, emphasis added by me. – *L. G.*).

[11] Perhaps, it also has some sense from pedagogical point of view. In the paper by Wood (2018) one can see a discussion on possibilities of using the notion of Big Bang for pedagogical and other purposes as well as the author's suggestion to use as synonym of 'Big Bang' term 'the big beginning' as beginning of TIME, SPACE, MATTER and ENERGY as well as other initial substances. He clarifies his goal: 'Assuming "big beginning" as a non-contentious synonym for big bang, the task of communication must be redefined: How can this incomprehensible event when time began, space unfolded, matter appeared, and energy bifurcated into various forces be formulated as imaginative narratives that will broaden and deepen its meaning and significance in harmony with discoveries over the past half century?' (*Ibid.*: 3).

V. Conclusion

The importance of the Inflation theory. The theory of the Big Bang could not explain very much, which could be explained precisely with the help of the Inflation theory. *At present, the Big Bang theory is firmly integrated with the Inflation theory*. From the point of view of cosmology and physics the introduction of the stages preceding the hot Big Bang more or less successfully solves all the problems related to the initial data of the hot Big Bang epoch, and eventually explains the flatness, homogeneity and isotropy of the observed Universe. The inflationary era is very important for modern cosmological and cosmophysical concepts. Alan Guth explains with enthusiasm, 'Inflation is not just a theory of the initial (ultimate) beginning, but it is a theory of evolution that explains essentially everything that we see around us, starting from almost nothing' (Guth 2002; about creation from nothing see our comments above).

However, one should understand that the emergence of the Inflation theory is the result of searching for such physical conditions under which it would be possible to explain the characteristics of the modern Universe. For modelling such initial conditions some scientists introduced hypothetical states of matter and energy. Therefore, it is absolutely normal that there are dozens of competing models of the inflationary stage, as well as the fact that nearly all parameters of this stage are unclear. What is really surprising is that science can put forward well-structured and reasonable hypotheses about such distant and extremely short periods.

Thus, on the one hand, the Inflation theory is a triumph of possibilities of modern cosmology and physics, but on the other hand, it perfectly demonstrates the limits of our knowledge, and the extent to which these hypotheses can be exotic and strange to explain things near these limits.

One should not forget that in these cases we are talking not about even theories but paradigms (Guth 1997), not the proved facts but *hypothetical* events and substances. It is also worth agreeing with Guth's (1997) statement that if it is true that the Universe arose from inflation, we cannot regard the quantum fluctuation as the cosmic origins. This idea leads us closer to its beginning. However, it is absolutely unclear, whether there was something before the inflation and what it was. There is a number of theories about these topics.

In particular there is a wide variety of very original theories according to which our Universe is not the only one, but just one of the myriad universes of the Multiverse.[12] According to some theories, these universes do not contact with each other, according to others the collisions of universes cause Big Bangs. Anyway, such approaches show that the origin of our Universe 13.82 billion years ago is

[12] *E.g.*, according to A. Linde's theory, 'the area of the Universe that we are observing now occupies a part of one of these "bloated" clusters' (Rubakov and Gorbunov 2010: 357).

an ordinary event in Multiverse. However, it is a very important event for humanity and the starting point for Big History because any history must have a beginning.

References

Cherepashchyuk A. M., and Chernin A. D. 2004. *The Universe, Life, Black Holes (Science for Everybody).* Fryazino: Vek 2. *In Russian* (Черепащук А. М., Чернин А. Д. *Вселенная, жизнь, черные дыры (наука для всех).* Фрязино: Век 2).

Garriga J., and Vilenkin A. 2001. *Many Worlds in One.* URL: http://cds.cern.ch/record/486136/files/0102010.pdf.

Gorbunov D. S., and Rubakov S. A. 2010. *Introduction to the Theory of the Early Universe. Cosmological Perturbations and Inflationary Theory.* Moscow: URSS. *In Russian* (Горбунов Д. С., Рубаков С. А. Введение в теорию ранней Вселенной. Космологические возмущения. Инфляционная теория. М.: УРСС).

Gorbunov D. S., and Rubakov S. A. 2012. Introduction to the Theory of the Early Universe. The Theory of Big Bang. Moscow: LKI. *In Russian* (Горбунов Д. С., Рубаков С. А. *Введение в теорию ранней Вселенной. Теория горячего Большого взрыва*. 2-е изд. М.: ЛКИ).

Grinin L. E. 2013. *Big History of the World Development: Cosmic Evolution.* Volgograd: Uchitel. *In Russian* (Гринин Л. Е. *Большая история развития мира: Космическая эволюция*. Волгоград: Учитель).

Grinin L. E. 2014. The Star-Galaxy Era of Big History in the Light of Universal Evolutionary Principles. *Teaching & Researching Big History: Exploring a New Scholarly Field* / Ed. by L. E. Grinin, D. Baker, E. Quaedackers, and A. V. Korotayev, pp. 163–187.Volgograd: 'Uchitel' Publishing House.

Grinin L. E. 2018. Evolution of the Early Solar System in Terms of Big History and Universal Evolution. *Journal of Big History* 2(1): 15–26. URL: http://dx.doi.org/10.22339/jbh.v2i1.2252.

Guth A. H. 1997. Was Cosmic Inflation the 'Bang' of the Big Bang? *Beem Line* 27(3). URL: http://ned.ipac.caltech.edu/level5/Guth/Guth1.html.

Guth A. 2002. *The Inflationary Universe.* URL: http://www.edge.org/conversation/the-inflationary-universe-alan-guth.

Guth A. 2004. Inflation. *Carnegie Observatories Astrophysics.* Series 2: *Measuring and Modeling the Universe* / Ed. by W. L. Freedman. Cambridge: Cambridge University Press. URL: http://www.astro.caltech.edu/~george/ay21/ readings/guth.pdf.

Hawking S. 2001. *A Brief History of Time. From the Big Bang to Black Holes.* St. Petersburg: Amfora. *In Russian* (Хокинг С. *Краткая история времени. От большого взрыва до черных дыр.* СПб.: Амфора).

Levin A. 2010. A Trillion Years before the Big Bang. *Popular Mechanics* 6. URL: http://elementy.ru/lib/431131?page_design=print. *In Russian* (Левин А. За триллион лет до Большого взрыва. *Популярная механика* 6).

Mukhanov V. F., and Orlova O. 2006. Big Bang Became a Victim of Inflation. *In Russian* (Муханов В. Ф., Орлова О. *Большой взрыв стал жертвой инфляции.* URL: http://polit.ru/article/2006/12/19/vzryv).

Postnov K. A. 2001. *Lectures on General Astrophysics for Physicists.* Moscow: Moscow State University. *In Russian* (Постнов К. А. *Лекции по общей астрофизике для физиков*. М.: Изд-во МГУ. URL: http://www.astronet.ru/db/msg/1170612/index.html).

Rubin S. 2004. The World Born out of Nothing. *Vokrug Sveta* 2(2761): 56–65. *In Russian* (Рубин С. Мир, рожденный из ничего. *Вокруг света* 2(2761): 56–65).

Sazhin M. V. 2002. *Modern Cosmology in Popular Presentation*. Moscow: URSS. *In Russian* (Сажин М. В. *Современная космология в популярном изложении.* М.: Едиториал УРСС).

Vilenkin A. 2006. *Many Worlds in One: The Search for Other Universes.* New York: Nill and Wangs.

Vilenkin A. 2010. *The World of Many Worlds. Physicists in Search for Other Universes*. Moscow: Astrel. *In Russian* (Виленкин А. *Мир многих миров. Физики в поисках иных вселенных.* М.: Астрель).

Wood B. 2018. Imagining the Unimaginable: Narratives of the Big Bang Time, Space, Matter, Energy. *Journal of Big History* 2(1)*:* 1–13. URL: http://dx.doi.org/10.22339/jbh.v2i1.2251.

III. BIOLOGICAL ASPECTS

7
Ecological Darwinism or Preliminary Answers to Some Crucial though Seldom Asked Questions

Edmundas Lekevičius

Abstract

The author asserts that evolutionary regularities might be deduced from principles of life's functioning. First of all, the latter should describe the part-whole relationships and control mechanisms. The author suggests supplementing the concept of struggle for existence with the concept of functional hierarchy: no solitary individual or species is functionally autonomous, life as we know it can exist only in the form of a nutrient cycle. Only two purely biotic forces – 'biotic attraction' and 'biotic repulsion' – act in the living world. The first one maintains and increases diversity and organizes solitary parts into systems integrated to a greater or lesser degree. The second one, in the form of competition, lessens biodiversity but at the same time provides life with necessary plasticity. On that basis, tentative answers to the following questions are given: (1) Why does life exhibit such a peculiar organization: with strong integration at lower levels of organization and weak integration at the higher ones?; (2) Why did particular species and guilds appear on the evolutionary stage at that particular time and not at any other?; (3) Why was the functional structure of ecosystems prone to convergence despite a multitude of stochastic factors?

Keywords: *ecological Darwinism, biology, selection, ecosystems.*

Introduction

As a matter of fact, Darwin's theory on natural selection consists of two constituent parts: the ecological and the genetic one. The first of them ('struggle for existence') deals with a surplus in offspring and unfavourable environmental factors, which cause mortality of the former. The part of genetics focuses on undirected variability and inheritance of selected traits. Geneticists of the 20th century specified and elaborated the latter part of the theory. Meanwhile, the first part of the theory fell into the hands of ecologists and also underwent elaboration. However, ecologists did not restrict themselves to 'struggle for existence' and created something that was new in principle, *i.e.* the ecosystem conception. Many of its statements are still valuable to date. Strange as it is, until recently evolutionists have hardly made any use of this part of ecology, and it lingered where it was created. It is strange because when referring to any hypothetico-deductive theory (Darwinism is undoubtedly such a theory), it is advisable from time to time to revert to its original postulates to verify if they are in agreement with new data (Popper 1959). The ecological part of the natural selection theory deals with the way organisms react to the environment. If these relations are not restricted just to 'struggle for existence', it is not only possible but advisable to supplement the premises of the theory with the new ones. In turn, conclusions following from the original statements will change. So, if we want to have a more profound evolutionary theory which better corresponds to the present-day achievements, we must revert to Darwin's original premises and reassess them not only from the viewpoint of genetics but also from that of ecology.

I am pretty sure that the need for a new evolutionary theory is felt most strongly by those who cannot find answers to the questions concerning the 'essence' of life and the main regularities of its functioning and evolution. It is quite possible that the majority of biologists believe that all the questions of fundamental significance have already been answered. I assume such a viewpoint to be excessively superficial. I think that it is only with the help of ecological theory that it is possible to give an answer to many fundamental questions which traditional biology did not even raise. For instance:

- Why does life exhibit such a peculiar organization: with strong integration at lower levels of organization and weak integration at the higher ones?
- Why did particular species and guilds appear on the evolutionary stage at that particular time and not at any other?
- Why was the functional structure of ecosystems prone to convergence despite a multitude of stochastic factors?

The material presented in this survey raises hope that answers, tentative as they are, to these and the like questions may be perceived in the near future. Such 'ecologizing' of Darwinism is likely to benefit not only this theory but ecology itself as well.

The Possible Methodological Framework for the Future Evolutionary Biology

Let us start by formulating the main methodological principles, *i.e.* particular 'recipes' which should be followed if we want to guarantee success in devising a more extensive synthesis.

It is usually claimed that populations and ecosystems are complicated and difficult to investigate objects; subjects of study of molecular biology and especially biophysics are less complicated by comparison, therefore progress in these spheres is greater. My opinion is slightly different: the complexity of life phenomena is largely the creation of our own minds and is a consequence of research methods applied that are not entirely adequate. At first glance it seems that life is objectively complex only if we measure complexity in terms of heterogeneity or the variety of structures. Biologists who talk about the complexity of life very often appeal to the abundance of links as well. And in this regard they are right again: multicausality is the result of this abundance. However, it is quite possible that life will lose most of its complexity when the new method of logical simplification that has been ignored heretofore is applied. I will try to propose the guidelines for such simplification.

I would define *Recipe 1* in the following way: *the deductive method, especially thought experiments, should not be avoided while pursuing wider biological synthesis*. Biologists have largely ignored the deductive method. Undoubtedly, this has led to adverse consequences and it is hard to explain exactly why this has happened. It is possible that the majority of biologists identify the deductive method with axiomatic methods, which are unacceptable to most biologists, and maybe not without good reason. I am also sure that almost all biologists associate deduction with mathematical methods, which is also a real misunderstanding. We use deduction in our everyday lives, and without it we would be simply unable to understand each other. Although Darwin generated a number of ideas through deduction and without using mathematics, most biologists understand and appreciate their value.

Recipe 2 is as follows: *while proceeding with the development of evolutionary theories (deriving evolutionary regularities from functional ones) initiated by Darwin, it is necessary to revise not only the conception of the 'struggle for existence' but also the attitude towards the nature of intra-organism links adopted in Darwin's lifetime.* The main question is: what is the nature of the 'part-whole' relationship at every level of organization, starting with macromolecules and ending with ecosystems. Such union of functional biology and evolutionary biology makes it possible to explain the maximum number of phenomena on the basis of the minimum number of statements. This is the main purpose of any logical simplification.

Recipe 3: while analyzing the causes of biological phenomena, it is recommended that the widest implications of cause-effect relationships be given consideration. Of course, this recipe can be useful only to a theorist pursuing synthesis. The well-established tradition, which encourages interest only in direct relations is potentially disastrous to theoretical work aimed at synthesis. However, it must be noted that a physicist or chemist would hardly admire this recipe, and many may claim that it would make the biological view of the world even more complex, but we should not be concerned with that; as it becomes clear through causal analysis of this type, that biotic connections are 'neatly built' and characterized by a particular hierarchy. Using this recipe, it has been established, for example, that it is not only plants that participate in the process of photosynthesis (as it is usually considered) but almost all the local ecological community (Lekevičius 1985).

Recipe 4 recommends using a qualitative or conceptual method of modeling. Even though this method is used in biology quite widely, I suppose I should describe it in greater detail. This modeling can be viewed as intermediate between verbal and mathematical modeling. Darwin's theory of natural selection can be considered as a typical verbal model. To transform it into a qualitative model, it is necessary to formulate and define its original terms and statements (premises) strictly. Qualitative models would include graph diagrams indicating only trends and various kinds of diagrams displaying connections between objects and phenomena and the like. The disadvantage of this modeling is its insufficient precision. However, there are particular advantages to using this modeling also: it does not restrict the modeler to any particular mathematical apparatus, the researcher therefore has much more freedom to raise questions and suggest hypotheses than he/she would have if mathematical modeling was employed. This kind of modeling additionally offers the opportunity to 'cover the uncoverable' (see *Recipe 3*). Furthermore, it allows adapting the method to the aims and research objects rather than the other way round, which is often the case, especially in ecology and evolutionary biology when mathematical methods are employed. Mathematical modeling could even follow from qualitative modeling as it usually happens in physics.

According to the instructions of *Recipe 5, one of the main ways to engage in logical simplification is to adopt the functional point of view.* This rule is based on the fact that the variety of structures (macromolecules, sub-cell organelles, cells, organs, genotypes, phenotypes, and species) is far richer than the variety of roles or functions that these structures perform. From the structuralist point of view, every enzyme is fairly complicated, and in order to describe this variety in detail, an exhaustive and difficult research effort is required. Meanwhile, its function is comparatively easy to identify and can be defined in a single sentence. Additionally, simplicity can be seen in the fact that the organization of life in its entirety is based on a certain hierarchical system: general functions, such as local nutrient cycling, can be fulfilled only through partial functions that are performed by individual guilds and species of the ecosystem. To attain simplicity, it is necessary to abstract from details. In our situation abstraction is easy to achieve because nature seems to have already provided it for us through the manifestation of a few functions ('emergent' features) present at the highest levels of organization. This sharply contrasts with the abundance of partial functions found at lower levels.

The Nature of 'Part-Whole' Relations

Let us conduct a thought experiment: let us imagine an organ in isolation and try to find an answer to the question as to how long it could survive without the appropriate biotic context, in this case – the organism (1). Let us now do the same with an individual animal or plant (2) and any population (3). The results are going to be more or less as follows. The organ will cease to function very soon, the individual will, however, survive for a longer period of time, and the population will survive still longer. It does not matter if you isolate an individual plant from its biotic environment, the whole population, or all 'autotrophs' of a particular ecosystem. The result in all cases is going to be the same – loss of life. The only difference, when compared with animals, might lie in the fact that some 'autotrophs', when isolated from detritivores (= decomposers) will be able to survive for a longer time – up to a few years or a few decades, depending on the amount of supply of inorganic nutrients available at the beginning of the experiment.

This fact illustrates that functional autonomy is not characteristic of any of these structures. If the biotic environment is eliminated they cannot be considered to be alive, in a sense. Following similar logic, biologists have claimed that viruses are not live organisms since they can only reproduce using the nuclear apparatus of a host's cell. This verdict does not seem to be controversial, but then, using this same logic, we may ask why we consider a deer or a lion, for example, to be alive.

Having conducted these experiments you will be forced to admit that in some sense the main feature of being alive, *i.e.* functionally autonomous, is only typical of an ecosystem, since life cannot last independently without nutrient cycling. The formula 'only an ecosystem is living' should be interpreted in the following way: a nutrient cycle is an emergent feature shaped by the coordinated activity of the whole ecological community (at least that of 'autotrophs' and detritivores). Let us call the local cycle and the energy flow that follows it the global function. The activities of individual guilds or species, then, could be treated as partial, or minor, functions, those of separate individuals – as even smaller functions, and so on, until we come to the functions of macromolecules. Eventually, we arrive at what systemologists refer to as a functional hierarchy. This concept might be more convenient to use, but it essentially means the same as functional dependence. It follows then that it is not simply integrity that is characteristic of life, but integrity based on mutualism, or links of reciprocal dependence. A biological species is not an aim *per se*, as it is usually assumed, but also a means.

This can still be expressed in a different way, by adopting the concept of labour division that Darwin (1872) was so very fond of: nutrient cycles are the outcome of common activity of individual species that are involved in labour of a narrower kind. Each of them performs a different operation. Again, specialization in reproductive or any other function is available within the population framework. This is directly analogous to the division of labour typified by the arrangement of organs, cells and macromolecules in a single living being.

There might be, in fact, several varieties of hierarchy. One kind of hierarchy is typical of clockwork mechanisms, for example, another kind of hierarchy – of a multi-cellular organism, and still another one – of a population or a community/ecosystem. Clock-parts have no capacity for reproduction. It might be claimed that the structure of a clock is therefore inflexible and completely inert, and that its parts therefore lack the 'freedom' to pursue their own self-interest. A multi-cellular organism has a more or less different hierarchy of functions, with cells of an organism capable of proliferating and therefore having some freedom to pursue their self-interest, although minimally. An organism is flexible and can adapt to the ever changing environment (physiological and biochemical adaptation), the freedom to pursue self-interest at the level of sub-individual structures being a prerequisite for this. Cells might even compete with each other while pursuing their individual 'goals'; experiments with chimeras largely contributed to this conclusion (McLaren 1976; see also the review by Lekevičius 1986: ch. 3.4). As far as individuals and species are concerned, they possess even greater freedom. This freedom is so great that the majority of ecologists still conceive of nature as being governed by competition ('biotic repulsion') and still argue that interdependence ('biotic attraction') does not exist at all at the levels of population and ecological community; and even if it does, this interdependence can be ignored. Extensive biological data show that individuals and species use this freedom for their own 'purposes' which are usually related to generating even larger populations.

To my mind, the analogy of the two-faced Janus can be used (Koestler 1967) to reveal the essence of the part-whole relationship. The side of his face that is turned upwards, towards the higher level of organization, shines with obedience and devotion, whereas the one turned downwards is the face of an individual who recognizes only his own objectives. Biosystems can be regarded not only as multilevel, but also as multigoal systems (Mesarović *et al.* 1970) in the sense that the goals of individuals and species do not necessarily have to be the same. Their respective objectives might even be in conflict with one another, which is what we usually observe in nature. On the other hand, the fairly great freedom of action that is noticeable at these levels seems to be very useful to ecosystems when they have to adapt to drastic and unpredictable environmental changes. So, the functional hierarchy in nature is not rigid or stiff. From a functional point of view, biotic components, starting with cells and ending with species, do not only depend on each other, but are also conditionally independent, as they cooperate and compete with each other simultaneously. The interaction of these two opposite forces, 'biotic attraction' and 'biotic repulsion', determines the behaviour of life and its evolution.

The functional hierarchy cannot be realized without an adequate hierarchy of control. However, there exists no control device at these levels of organization, which is similar to that of multi-cellular organisms. Many people may consider this situation absurd, but this is nothing new for experts in systems theory. This type of control has been termed as diffusive or passive (Novoseltsev 1978; Lekevičius 1986). It is achieved through the interaction of subsystems, whose behaviour towards control is the same. During these interactions, certain constraints (positive or negative feedbacks), helping the whole to control its constituent parts, emerge. These constraints usually evolve because not all combinations of subsystems or their activities can ensure stability. Populations and ecosystems therefore adjust on their own, without any external contribution. It means that joint efforts help ecological communities not only to support local nutrient cycles, but also to ensure their conditional independence from various kinds of inner and external perturbations. In other words, global parameters, vitally important for the whole biota, are homeorhetic because of self-organization and self-regulation. Nutrient cycles are the most highly buffered features of life (Lekevičius 1986).

It might be even easier to understand how the ecosystem's functioning is controlled by considering an analogous example of capitalist economy, the laissez-faire mode in particular. The forces of 'repulsion' and 'attraction' in capitalist economy are almost equal in power, their counterbalance being nearly the same as that in nature. The initiative and the right to decide belong to individuals. Although, as a rule, they pursue self-interest rather than the interest of society, the society, guided by an 'invisible hand' (in fact – by the market), inevitably tends towards the universal well-being. This conception became popular in England as the paradox of private vices and public benefits.

Incidentally, Darwin was probably the first to notice parallels between the organization of economy and that of nature. They have also recently been discussed by Salthe (1985) and Lekevičius (2009a). Naturally enough, using analogies is not an appropriate way to explain something. However, I do think that it may be beneficial for the clarification of statements.

Why does life exhibit such a peculiar organization: with strong integration at lower levels of organization and weak integration at the higher ones? To answer this question, let us think what animate nature would look like if individuals of the same and different species only cooperated, *i.e.* if competition as a phenomenon completely disappeared. A preliminary answer to this and some other questions of a similar kind has been provided by GAT, the general adaptation theory (Conrad 1983; Lekevičius 1986, 1997). According to the theory, if this hypothetical situation came into being, we would probably have both ecosystems and nutrient cycles. In fact, these would not be typical ecosystems; they would have a much greater degree of integration – somewhat comparable to today's coral reef ecosystems. These ecosystem-superorganisms would perform their vital functions incredibly effectively, but would fall to pieces like a giant with clay feet as soon as the first unusual environmental change took place. Specialization and integration allow maximizing the degree of adaptation, but that is incompatible with maximum adaptability. The capacity for disintegration and the conditional freedom of subsystems are essential attributes of life on this planet, where environmental conditions are continually changing to a great degree and are very often unpredictable.

What would happen in the opposite situation, *i.e.* if these relationships were only of a competitive type? I think that the most likely final outcome would be that only one species would exist in any given location at any given time; *i.e.* the one that would have replaced all the rest species, those that are not so well adapted to struggling for existence. And within this species, a single ('wild') genotype that has replaced all the genotypes of lower adaptive value would exist. Naturally, there would be no ecological communities or ecosystems under these circumstances. However, as it has already been mentioned, this sort of life would have no chances of survival since none of the species can maintain a nutrient cycle on its own. To summarize, it might be claimed that life has chosen a compromise between two incompatible strategies – to be maximally efficient and not to compromise adaptability. This compromise must have conditioned the long-term existence of life. However, the problem with this kind of an answer lies in its teleological implications. We could arrive at a far better answer if we discovered the evolutionary processes through which this form of life could emerge.

On Evolutionary Interdependence of Individuals and Species

Having applied the methodological recipe advocated by us (evolutionary regularities can be deduced from principles of functioning – *Recipe 2*), we come to the conclusion that even when evolving, individuals and species cannot have autonomy. Functional dependence inevitably leads to evolutionary dependence. It is clear that the most obvious manifestation of this regularity is likely to be observed in cases of cooperation and mutualism. For instance, it is clear that organs of a multicellular organism can evolve only in a coordinated manner. Otherwise, the integrity and vitality of an organism will be destroyed. Similarly, the evolution of members of the same population, which are bound by relations of interdependence, cannot be uncoordinated. For instance, an uninterelated evolution of males and females of the same population is difficult to imagine. In these cases, loss of coadaptation is equivalent to population extinction. It is also obvious that evolutionary changes in species bounded by mutualistic relationships cannot be uncoordinated either. For example, such coevolution should be characteristic of flowering plants and insects pollinating them. The same holds true for the relations between producers and detritivores of the same ecosystem: both these ecological groups should inevitably affect the evolution of each other, as they are mutually dependent. In short, coevolution or coordinated evolution is the inevitable outcome of functional dependence.

Populations of prey and its predators like those of hosts and their parasites also coevolve. For instance, extensive available evidence shows that prey/host populations accumulate features reducing exploiter-induced mortality. In length of time, the latter, in their own turn, acquire features enabling them to continue the exploitation of their usual prey/host. It is clear that the initiative should not necessarily come from exploited populations. Such coevolution usually leads to moderate exploitation. And only in case of moderate exploitation, we have the right to assert that both partners are coadapted. In this context, I suppose, I do not violate the terminological discipline, as, to my knowledge, the terms 'coevolution' and 'coadaptation' are treated in this way by the majority of users.

In my opinion, the evolution of most species was and still is restricted from every side, as the ecospace nearest to them was and still is occupied by other species well adapted to their environment. Species do not exist in some kind of ecological vacuum – both their functioning freedom and that of evolution are restricted. Therefore prohibitions have always outnumbered permissions. Stabilizing selection and evolutionary stasis are daily routine of animate nature. Many of the evolutionarily old species can be treated as living fossils, which is not because they lack variability, but because other species (most often those that have emerged later) did not leave free ecospace for the new variations to penetrate. This approach, in my opinion, considerably differs from the opinion that has been dominant for a long time. According to that view, the rate and success of evolution are predetermined only by genetic variability; and maybe also by climatic and edaphic conditions and geographical isolation.

In this context, permission is understood as a vacant niche, and, more exactly, as a vacant environmental niche. Two terms are used in ecology: an ecological or Hutchinsonian niche, on the one hand, and an environmental niche, on the other. The first one is understood as the totality of needs. The ecologists using the term 'an environmental niche', first of all, have environmental conditions in mind, which, in their opinion, can exist and exist independently of organisms. It is only they who use the term 'a vacant niche'. They understand a vacant niche as potentially usable resources. Solar energy having no consumer, some organic or inorganic substance as a potential source of energy, electrons or carbon can be taken as examples of such resources. Of course, a live organism also can be treated as a vacant niche if it does not have consumers (predators or parasites) (for more information about the vacant niche concept see the survey by Lekevičius [2009b]).

An occupied niche can be viewed as prohibition for another species to occupy it. However, this prohibition can be overcome in cases of successful innovations or immigration of stronger competitors. Sometimes it is more expedient to replace this term (prohibition) by the term 'constraints', which may sound less categorical. Of course, in addition to constraints associated with the availability or absence of resources, there are other types of constraints, such as thermodynamic, climatic or edaphic constraints. Their evolutionary impact is quite well-known and we are going to discuss them as well. In essence, prohibition can be associated with the fact that not all evolutionary trajectories ensure stability of living systems. For instance, ecosystem-level constraints are, first of all, negative feedbacks, which do not permit species to evolve in such a way that the local nutrient cycle should be disturbed. So, it is possible then to view permissions as positive feedbacks and prohibitions as negative ones.

Evolutionary Assembly of Ecosystems

Ecosystem 'assembly rules' may be formulated in the following way (Lekevičius 2002). It is quite possible that since the very moment of life appearance there existed quite a simple mechanism by which ecosystems and nutrient cycles were formed – end products of some organisms' metabolism turned into waste, *i.e.* into resources potentially usable though used by nobody. Such vacant niches provoked the evolution of organisms capable of exploiting those resources. The final result was that end products of detritivores' metabolism became primary materials for producers. The formation of ecological pyramids should have followed a similar pattern: producers provoked the evolution of herbivores, the latter – that of primary predators, and so on and so forth until eventually the evolution produced common to us pyramids with large predators at the top.

So, vacant niches not only stimulate diversification, but also determine its direction. And this fact, most probably, witnesses causality. This idea can be viewed as a keystone of evolutionary theory because it is not so difficult to explain, and, at least partly, predict results of diversification from data on vacant niches.

In order to clarify the vacant niche concept and its usage, I have constructed a *table* demonstrating some steps in ecosystem evolution.

The first terrestrial organisms should have probably been heterotrophs. The main shortcoming of the first ecosystem was that decomposition was carried out far more intensively than the chemical synthesis of organic matter. This misbalance might have caused the very first in the history of life ecological crisis, which finished with the rise of the first producers. The latter could have been green and purple non-sulfur bacteria, which carried out anoxygenic photosynthesis. They used organic compounds as a source of hydrogen (electrons).

Along with these bacteria, detritus-decomposing ones, too, are likely to have been involved in local nutrient cycles of that time. Their emergence and diversity was determined by the diversity of organic substances present in detritus. Already at that time cycles must have been non-waste, and decomposition was carried out to the very inorganic nutrients.

Table. Some of the vacant niches/adaptive zones that existed in the Archean and Proterozoic, and their occupants

Description of vacant niches/adaptive zones	Hypothetic occupants
Organic substances as donors of energy, electrons and carbon. Organic substances as final electron acceptors	Protobionts
Light as an energy donor, H_2S/H_2O as an electron donor and CO_2 as a donor of carbon	Green and purple sulphur bacteria, cyanobacteria
Detritus as an energy, electron and carbon donor. S^o and SO_4^{2-} as final electron acceptors	Sulphur- and sulphate-reducing bacteria
Fe^{2+}, Mn^{2+}, H_2S, CO, H_2, CH_4, NH_4^+ as energy and electron donors, CO_2 as a carbon donor. O_2 as a final electron acceptor	Aerobic chemolithoautotrophs
Detritus as an energy, electron and carbon donor. NO_{3-} as a final electron acceptor	Denitrifying bacteria
Detritus as an energy, electron and carbon donor. O_2 as a final electron acceptor	Aerobic decomposers
Biomass as an energy, electron and carbon donor. O_2 as a final electron acceptor	Protists as ‘herbivores’ and decomposer-eaters
Biomass (‘herbivorous’ and decomposer-eating protists)	Protists as primary predators
‘Herbivores’ and primary predators	Multicellular organisms as secondary predators

In the Table, attempts are made to list events in chronological order, from the appearance of protobionts to that of secondary predators. Take note of the fact that some vacant niches/adaptive zones preexisted the emergence of organisms, while others were presumably created by organisms themselves (compiled from Lekevičius 2002).

As biomass accumulated, sooner or later aquatic resources of free organic compounds had to be depleted. That could have caused the rise of true autotrophs (photolithoautotrophs). The latter could have been green or purple sulfur bacteria, which used H_2S and H_2 as a source of hydrogen (electrons). Those bacteria accumulated sulfur and sulphates as waste, so after a while evolution should have brought about organisms reducing sulfur and sulphates. The vacant niche was occupied to make the cycle become non-waste again. After some time, however, the resources of H_2S and H_2 must have run short, that should have resulted in the appearance of cyanobacteria carrying out oxygenic photosynthesis. The merit of that kind of photosynthesis is in that it uses water molecules as a source of hydrogen (electrons). However, when oxygen turned into waste, it began to accumulate in water. As a result, the evolution of oxygen resistance was triggered off. Still after a while, presumably some 2.0–2.5 billion years ago, cyanobacteria and detritivores accompanying them became aerobes. It must have been at that time that all modern aerobic chemolitotrophs came into existence. The motives for their rise were very simple: oxygen accumulating in the environment reacted by itself with the dissolved in water ferrous iron and manganese, hydrogen, carbon monoxide, sulfur, hydrogen sulfide, ammonia, and methane. The energy produced during oxidation was lost. Naturally, those vacant niches became factors stimulating and directing evolution. Thus, after a while all those niches were occupied.

The nitrogen cycle was presumably assembled in the following way (for details see in Lekevičius 2002). At the dawn of life, nitrogen compounds, especially ammonia and ammonium ions, might have apparently been much more abundant in the atmosphere and waters. Thus, selection pressure, forcing organisms to acquire the ability of nitrogen fixation, might have been absent for a while. Yet there are reasons to believe that later the amount of ammonia and ammonium ions in the environment decreased to minimum, and not only because part of it converted to organic nitrogen, the biomass. Due to the presence of cyanobacteria, oxygen began to accumulate in the environment and, affected probably by lightning, reacted with ammonia and molecular nitrogen, thereby producing oxides. Besides, as it has been mentioned above, soon thereafter originated nitrifying bacteria oxidizing ammonia and ammonium ions to nitrates. I think that could have given rise to selection pressure, which induced diversification of nitrogen fixing organisms and their spread. Nitrates immediately created a vacant niche that provoked the rise of denitrificators. The latter used nitrates as unchangeable under anoxic conditions glucose oxidizers, final acceptors of electrons. Due to nitrate respiration nitrates converted to free nitrogen. The global nitrogen cycle became closed. Accumulating in the environment nitrates might have soon become an additional source of nitrogen to cyanobacteria. Thus, we obtain the following picture of the evolution of the nitrogen cycle (see Fig. 1). I understand that this scenario of the changes in the nitrogen cycle is rather speculative, though it, seemingly, is in accordance with the one proposed by experts (Falkowski 1997; Raven and Yin 1998; Beaumont and Robert 1999). The difference lies merely in some details of secondary importance.

There are sound reasons to believe that 2 billion years ago all modern global cycles – carbon, oxygen, nitrogen, sulfur – had been already formed. From the point of view of chemistry, they have not changed until nowadays (for details see Lekevičius 2002).

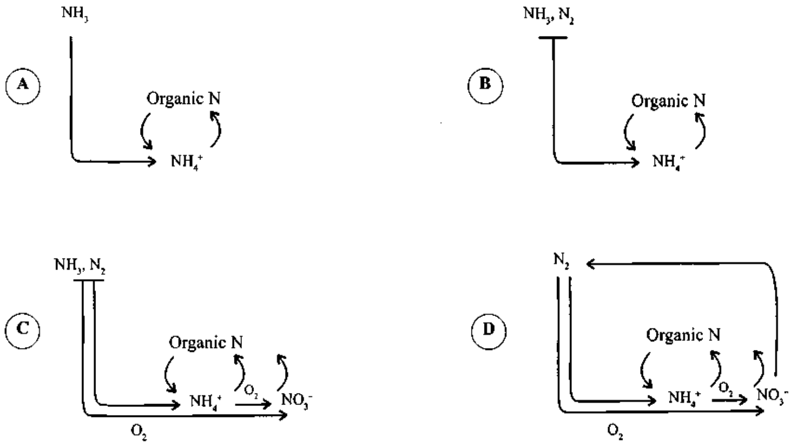


Fig. 1. Assembly of the nitrogen cycle. A – local cycles are formed; B – biological nitrogen fixation appears; C – nitrates are produced in large quantities; D – denitrification arises

Two billion years ago, ecosystems were still composed of only two 'functional kingdoms'– producers and detritivores. For quite a long time, some organisms exploited others not before the latter died. Accordingly, there must have been a huge adaptive zone. Its exploitation presumably started about 1.5 billion years ago, after the emergence of protozoans, although fossil records do not evidence the existence of organisms that could be called the first biophages. Hence, we are speaking about the appearance of the 2nd trophic level. Another possibility, *i.e.*, the emergence of parasitism as a phenomenon at that particular time should not be ruled out either. The only certainty is that immediately upon emergence, organisms representing the 2nd trophic level automatically became an adaptive zone for the future 3rd level representatives, *i.e.*, primary predators (see Table). In their own turn, the latter became prey for the future bigger predators, *etc.* This self-inducing process, as a matter of fact, ended in the appearance of top predators in the Ordovician. It is quite probable that in length of time, in addition to predators, the newly emerging species acquired a set of parasites exploiting them. So, it seems that at the end of the Ordovician, a supply of vacant niches suitable for biophages was depleted.

The stages of terrestrial ecosystem development and its mechanisms did not differ much from those of marine ecosystems (for details see Lekevičius 2002): appearance of producers (1), vegetative detritus (2), detritivores and local cycles (3), herbivores and organisms feeding on detritivores (4), primary predators (5), and so on up to the top-level predators. The latter came into existence in the late Carboniferous, approximately 300 million years ago. When the formation of the block of biophages finished in the seas and 135 million years later on land, there were almost no vacant niches left in ecosystems. Therefore cases of competitive exclusion, preconditioned by migration and the emergence of new forms, became more frequent. However, species diversification continued: life was penetrating into new territories, and what is more, the process of niche splitting was going on (Lekevičius 2002).

One may ask what there is new in such explanations of the well-known facts. In general, it is not customary in modern evolutionary biology to raise a question and to look for an explanation as to why certain guilds, let us say, aerobic chemolithoautotrophs or primary predators appeared on the evolutionary stage at that particular time and in that particular place. This can be probably explained by the fact that to find answers to questions of this kind, it is necessary to employ the deductive method, which is not popular with biologists. It has been only in this decade that somewhat wider, but still tentative use of the vacant niche term in the evolutionary theory has been started (*Idem.* 2009b). In case of failing to provide an explanation, a phenomenon itself is somewhat ignored. Another thing that makes my approach to evolution unconventional is that in respect of a population I emphasize external factors influencing the course of species evolution. Meanwhile, the conventional approach focuses all the attention on inner mechanisms. That does not mean, of course, that these approaches cannot be reconciled; they perfectly complement each other.

How Selection has Made Ecosystems Converge

The functional convergence of ecosystems was discovered quite long ago. Darlington (1957) wrote in his book *Zoogeography: The Geographical Distribution of Animals*:

> Neither the world nor any main part of it has been overfull of animals in one epoch and empty in the next, and no great ecological roles have been long unfilled. There have always been (except perhaps for very short periods of time) herbivores and carnivores, large and small forms, and a variety of different minor adaptations, all in reasonable proportion to each other. Existing faunas show the same balance. Every continent has a fauna reasonably proportionate to its area and climate, and each main fauna has a reasonable proportion of herbivores, carnivores, *etc.* This cannot be due to chance (Darlington 1957).

Here I would like to draw the reader's attention to one important, in my opinion, episode from the history of general ecology. It is known that the ecosystem conception was developed on the basis of empirical data obtained in the 1960s of the last century. It was discovered, for instance, that neither the number of trophic levels, nor the ecosystem structure in general is dependent on primary productivity, which is known to vary within very great limits on a world scale. Fortunately for ecologists, nature turned out to be undivided, in that respect. Otherwise, it would have been necessary to develop individual conceptions for individual ecosystems. Thus, ecosystem convergence was a trivial fact for ecologists of that time.

Time passed and ecologists of the older generation retired one after another to be replaced by young people interested in other problems. That was possibly due to the fact that in those times it was not easy to explain facts of the functional convergence of ecosystems, since they were hardly within the framework of the neo-Darwinian paradigm. It was difficult, or, according to somebody, impossible to build a bridge between a change in gene frequency in a population and a global phenomenon such as ecosystem convergence. It was 'common knowledge' that each species is affected by a multitude of internal and external factors and that its fate depends not only on an accidental genetic variability, but also on gene drift, climatic changes that are usually difficult to forecast, the impact of other species, and other difficult-to-define events. In the course of millions of years, these numerous factors must have produced such chaos of consequences in living nature that none of theorists was able to explain it. In a word, the opinion, which, by the way, persists to date, was formed that evolution is controlled by accidental forces and that it cannot be predicted. That is why the phenomenon of ecosystem convergence was and is out of place in the neo-Darwinian conception. On the contrary, facts of convergence contradicted the neo-Darwinian experience rather than supported it. However, it is known that facts do not necessarily refute theories. It is often the other way round – facts contradicting the generally accepted theory are simply ignored. Thus, it is no wonder that in the course of time the interest in that phenomenon gradually abated.

I propose using the notion of the functional convergence of ecosystems in a somewhat wider sense than that used by my colleagues some 20–30 years ago. I have in mind the invariability of ecosystem functions both in time and space. By this, I do not mean that ecosystems were not changing over time. I am inclined to take the view that approximately 2 billion years ago ecosystem metabolism finally became settled and since then nutrient cycles have been just replicated. The shape of production (energy) pyramids characteristic of local ecosystems seems to have undergone no considerable changes over the last several hundred million years despite all internal changes followed by numerous extinctions and adaptive radiations. Geographical invariance is also characteristic of these pyramids. Their form almost does not depend on the primary production, which may differ at least several ten-fold (the 10 % rule). Besides, when using the term 'functional convergence of ecosystems', it is necessary to have in mind the convergence at the level of individual species, too, *i.e.* a great abundance of ecological equivalents (species that have no consanguinity and live in different locations but have converged due to the fact that they occupy similar ecological niches).

As distinct from the traditional approach, I believe that all evolutionary processes are quite rigidly canalized. That role of canalization is performed by species interaction, which always and everywhere directs species evolution onto a few invariant ways. The raw material, from which evolution moulds a community, may differ. However, the final result, *i.e.* what the structure and function of that community is going to be like, is easier to predict because it often recurs both in time and space. God does not dice, so evolution could be predicted. But for this purpose, of course, one should have sufficient information not only about ancestral forms, but also about constraints. However, this information is as a rule lacking, because until today, in my opinion, evolutionists have not paid proper attention to factors constraining the evolution of species.

It is well-known that ecosystems may be assembled in two ways: via migration (ecological succession) and/or evolution. However, the final result does not depend on the mode of assembly, and that is evidenced by the fact of functional convergence. Probably, the same ecosystem-level constraints operate both in succession and evolution, although mechanisms are different. As a matter of fact, there are some similarities. Primary succession as a rule starts with the settlement of herbaceous plants (sometimes lichen). Then vegetative detritus is formed, niches suitable for the settlement of herbivores and detritivores (bacteria, protists, fungi and invertebrates) appear. As a result, necessary conditions for the appearance of soil are created (Olson 1958). In its own turn, the formation of soil stimulates the emergence of niches for new plants, woody plants among them. The latter change their surroundings, thus facilitating the settlement of still other plants and animals (facilitation theory – Connel and Slatyer 1977). The sequence of events is presumed to have been similar in the Palaeozoic when life occupied land (see above). However, then occupants came into existence mainly as a result of evolution *in situ*. So, I maintain that ecological succession may be interpreted as a process of niche filling as well, and it should not differ much in its course and final result (having in mind functional properties of ecosystems) from what is observed in cases of adaptive radiation and evolutionary recovery after extinction.

Odum (1969) put forward a hypothesis according to which ecological succession and evolution are characterized by the same trends of variation in ecosystem parameters (species diversity, primary production, total biomass, production and biomass ratio, efficiency of nutrient cycle). Although later this hypothesis was used as a target by many critically disposed opponents, it seems to be enjoying popularity among some ecologists and evolutionists (*e.g.*, Loreau 1998; Solé *et al.* 2002; Lekevičius 2002, 2003) to date. In the opinion of these authors, forces directing the evolution of ecosystems are in fact the same as those controlling their routine action. Consequently, in both cases trends cannot differ much. This idea, that 'ontogeny' of ecosystems may recapitulate their 'phylogeny', I think, is quite attractive.

What is Selected *vs.* What is Making Selection

Extremist neo-Darwinians suggest that only the gene ('selfish gene') can be a unit of selection. Still others maintain that this role is more suitable for the genotype. Some evolutionists have claimed that differential survival may involve entire populations (species) and even ecosystems. Thus, there have been attempts not only to reveal mechanisms of individual features' evolution, but also to explain how parameters specific to populations and ecosystems could have evolved. So, there was hope to finally clarify how nature creates and maintains biodiversity and, on the basis of the latter, communities and nutrient cycles. Still others suggested combining all these ideas rejecting the mentality of 'either-or'. Thus, the idea of hierarchic, or multilevel, selection emerged (*e.g.*, Williams 1966; Gould 1982; Wilson 1997; Gould and Lloyd 1999). According to this idea, differential survival involves all or almost all structures ranging from single genes to entire ecosystems. As far as I understand, these evolutionists do not doubt that evolving are not only individual features, but also populations, ecosystems, and even the biosphere. However, they believe that adaptation at any level requires the process of natural selection operating at that level. I think that here they make an essential mistake for they restrict the problem of selection to the question of what is being selected. What is more, they seem to be little interested in what is making that selection. Because of that, the problem becomes quasi-complicated and, unfortunately, insoluble. I am inclined to think that Darwin, however, was right in assuming that it is an individual that should be regarded as the major unit of selection.

As far as I understand, the problem of selection units has become so complicated and intricate because it has not been associated with functional biology. The imaginary wall between biological time and biological space hinders researchers from finding a solution to this problem. If this wall was demolished, the problem would immediately become quite simple and clear. The greater the integration of constituent parts of a biosystem is, the greater the possibility is that selection will affect the whole system as a unit. And on the contrary, if constituent parts of a system are functionally autonomous, they will be involved into the ever-lasting 'struggle for existence' and each of them will become a selection unit. Even ecosystem selection would be possible, if ecosystems functioned as real superorganisms. However, this is inconceivable either for populations, or for ecosystems. By the way, the question of selection units was already solved in a similar way by Rosen (1967), but his point, apparently, has been overlooked.

In my opinion, natural selection is a 'black box' turning non-directional inheritable variability into a more-or-less directed evolutionary development (see Fig. 2). This is an essential attribute of selection. Differential survival and that kind of reproduction are merely external and most obvious features of selection. Quite possibly, selection may have another external form, too, but anyway it is the force constraining inheritable variability in a specific way.

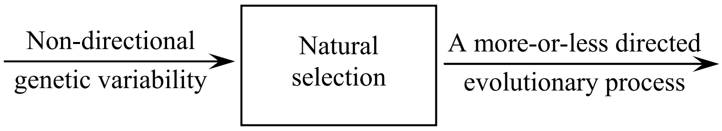


Fig. 2. Natural selection as functional constraints. Mutations and recombination create a field of potential evolutionary possibilities, whereas functional requirements constraint it in a specific way

Intraindividual constraints ('internal selection'), constraints emerging from the interactions of individuals of both the same and different species and from their interactions with abiotic surroundings are under discussion. Evolution at the levels of species and ecosystems progresses through genetic variability and differential survival and reproduction of individuals. Neither species selection nor that of entire ecosystems is necessary for the evolution to occur. Prohibitions and permissions that stimulate or suppress the spread of certain heritable variations emerge in the course of these interactions.

How does new genetic information in the form of a mutation or recombination become an attribute that changes the functioning of an individual, population and the entire ecosystem? Even the pathway of an especially successful mutation/recombination always begins from a single change in one of the cells. In the case of its success, novel genetic information passes several stages of strengthening. This may be done by means of the following mechanisms (Lekevičius 1986):

– transcription and translation of the newly emerged gene, increasing in concentration of mutant (recombinant) macromolecules in a cell;

– mitosis of cells carrying the gene;

– growth in the frequency of mutants (recombinants) in a population;

– growth of the population carrying the evolutionary novelty and widening of the species range.

Additionally, the variation has at least one more theoretical chance to be consolidated, which is to become the property of numerous species in the course of speciation.

As the novel genetic information is reinforced, an ecosystem reacts to it as to a gradually increasing internal disturbance. Individual – biochemical and physiological – mechanisms of adaptation are the first to respond. Mutants/recombinants are incorporated into adaptational and coadaptational processes at the population level after they pass barriers of internal selection. In case of success, new characters spread, but they have to prove they meet the requirements for constituent parts of an ecological community. If such co-adaptation happens, evolutionary diversification might follow and novel genetic information is disseminated among several or more species. In summary, the spread of evolutionary novelty always evokes feedbacks from individual, population and biocenotic mechanisms of adjustment, individual mechanisms being the first to react.

To sum up, traditional approach emphasizes selection units and cares about *what* is selected, whereas I propose taking interest in what *is making selection.* Differential survival and reproduction of individuals are merely external attributes and thus are the first impacts of adaptation to be noticed. It is functional constraints coordinating routine activities of individuals, populations and ecosystems that perform selection. They convert undirected hereditary variability into the far more directed evolutionary development. It is the individual that dies or produces fewer offspring, whereas structures, which may range from those of macromolecules to those of ecosystems, change and evolve. Moreover, competition is not necessary for the process of selection: it might be even more intense in the case of cooperation (*e.g.*, features disturbing inner balance of an organism are done away via internal selection, or variations reducing the co-adaptation of sexual partners are also successfully eliminated). The difference is that in case of cooperation, only the characters beneficial to all cooperating partners are selected, while in case of competition, only the characters that enhance the adaptedness of particular competitors are selected. Of course, any novelty that is beneficial for the whole population or ecosystem, must be primarily beneficial for its possessor, only then it can be spread and, in this way, strengthened.

Concluding Remarks

During the past decade, strong nihilistic trends, far stronger than before, appeared in evolutionary biology. This is how one of the most authoritative evolutionists has summarized his approach:

> Natural selection is a principle of local adaptation, not of general advance or progress. The history of life is not necessarily progressive; it is certainly not predictable. The earth's creatures have evolved through a series of contingent and fortuitous events (Gould 1994).

So, it turns out that Darwinism is suitable for the description of local phenomena of adaptation only. In this context it is worth remembering the previously published article by Gould and Lewontin (1979) expressing the authors' doubts regarding the whole adaptationist paradigm.

What way out do these authorities propose? Stephen J. Gould and Richard C. Lewontin seem to expect much from the theories of chaos, catastrophe, and complexity.

To describe that situation I could find no better word than 'crisis'. My opinion regarding the question is somewhat untraditional: biologists should reconcile themselves to the idea that no one else will propose a suitable methodology for the description of their subjects of study. A new methodology should take root in the depth of biology itself. It should be sodden with biologists' sweat and experience. None of the chaos, catastrophe, or complexity theories can or will take root, like dozens of other exotic matters, for they have originated in a different medium. If we do not want strange methods to dictate strange to us objectives and world outlooks, we should assume the responsibility for the future of biology. I disapprove of a further depreciation of mind and reasoning, entrusting the function of thinking to the computer, being simply afraid to form daring and audacious hypotheses that do not follow directly from the data available. I dare to claim that the naked empiricism combined with scientism raises monsters, *i.e.* young people who, for the sake of solidarity, cut their own wings and burden themselves with weights and lead in order to make their thinking as standard as possible. I doubt whether Francis Bacon, the father of empiricism, would like the scientific society so prone to standardize, but for me it is not very appealing – it is my civic position if you like. I am for the balance of induction and empiricism with deduction and rationalism rather than the counterbalance between them as it is usually the case. I think that the method of hypotheses advanced by Popper (1959) will be vindicated sooner or later. Biologists should do this as soon as possible.

Acknowledgements

I am grateful to Michael Woodley and Laimutė Monkienė who helped me in translating and amending the text.

References

Beaumont V., and Robert F. 1999. Nitrogen Isotope Ratios of Kerogens in Precambrian Cherts: A Record of the Evolution of Atmosphere Chemistry? *Precambrian Research* 96: 63–82.

Connell J. H., and Slatyer R. O. 1977. Mechanisms of Succession in Natural Communities and their Role in Community Stability and Organization. *American Naturalist* 111: 1119–1144.

Conrad M. 1983. *Adaptability: The Significance of Variability from Molecule to Ecosystem*. New York – London: Plenum Press.

Darlington P. J. Jr. 1957. *Zoogeography: The Geographical Distribution of Animals*. New York: Wiley.

Darwin C. 1998 [1872]. *The Origin of Species by Means of Natural Selection*. 6th ed. Middlesex: Tiger Books Intern.

Falkowski P. G. 1997. Evolution of the Nitrogen Cycle and Its Influence on the Biological Sequestration of $C0_2$ in the Ocean. *Nature* 387(6630): 272–275.

Gould S. J. 1982. Darwinism and the Expansion of Evolutionary Theory. *Science* 216(4544): 380–387.

Gould S. J. 1994. The Evolution of Life on the Earth. *Scientific American* 271(4): 63–69.

Gould S. J., and Lewontin R. C. 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 205(1161): 581–598.

Gould S. J., and Lloyd E. A. 1999. Individuality and Adaptation across Levels of Selection: How shall We Name and Generalize the Unit of Darwinism? *Proceedings of the National Academy of Sciences, USA* 96(21): 11904–11909.

Koestler A. 1967. *The Ghost in the Machine*. New York: Random House.

Lekevičius E. 1985. Explanation in Biology: An Attempt to Synthesize Functional and Causal Approaches. *Scientific Knowledge: Systematic Aspect* / Ed. by R. Pavilionis, pp. 119–135. Vilnius: Institute of Philosophy, Sociology and Law. *In Russian* (Лекявичус Э. Объяснение в биологии: Попытка синтеза функционального и причинного подходов. *Научное знание: системный аспект* / Ред. Р. Павилионис, с. 119–135. Вильнюс: Институт философии, социологии и права).

Lekevičius E. 1986. *Elements of the General Theory of Adaptation*. Vilnius: 'Mokslas'. *In Russian* (Лекявичус Э. *Элементы общей теории адаптации.* Вильнюс: «Мокслас»).

Lekevičius E. 1997. *A Model of Adaptive Rearrangements on Community, Population and Organism Levels.* Thesis for Habilitation. Vilnius: Institute of Ecology.

Lekevičius E. 2002. *The Origin of Ecosystems by Means of Natural Selection*. Vilnius: Institute of Ecology.

Lekevičius E. 2003. Evolution of Ecosystems: Main Stages and Potential Mechanisms. *Zhurnal obshchei biologii* 64(5): 371–388. *In Russian* (Лекявичус Э. Эволюция экосистем: основные этапы и возможные механизмы. *Журнал общей биологии* 64(5): 371–388).

Lekevičius E. 2009a. On Some Analogies between Ecosystems' Evolution and Economic Development: From A. Smith and Ch. Darwin to the Newest Ideas. *Evolution: Cosmic, Biological, Social* / Ed. by L. E. Grinin, A. V. Markov, and A. V. Korotayev, pp. 226–259. Moscow: LIBROKOM. *In Russian* (Лекявичюс Э. О некоторых аналогиях между эволюцией экосистем и развитием экономики: от А. Смита и Ч. Дарвина до новейших идей. *Эволюция: космическая, биологическая, социальная* / Ред. Л. Е. Гринин, А. В. Марков, А. В. Коротаев, с. 226–259. М.: ЛИБРОКОМ).

Lekevičius E. 2009b. Vacant Niches in Nature, Ecology, and Evolutionary Theory: A Mini-Review. *Ekologija* (Vilnius) 55(3–4): 165–174.

Loreau M. 1998. Ecosystem Development Explained by Competition within and between Material Cycles. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 265: 33–38.

McLaren A. 1976. *Mammalian Chimeras*. Cambridge: Cambridge University Press.

Mesarović M. D., Macko D., and Takahara Y. 1970. *Theory of Hierarchical Multilevel Systems*. New York: Academic Press.

Novoseltsev V. N. 1978. *Theory of Control and Biosystems. Analysis of Conservation Properties.* Moscow: Nauka. *In Russian* (Новосельцев В. Н. *Теория управления и биосистемы. Анализ сохранительных свойств*. М.: Наука).

Odum E. P. 1969. The Strategy of Ecosystem Development. *Science* 164: 262–270.

Olson J. S. 1958. Rates of Succession and Soil Changes on Southern Lake Michigan Sand Dunes. *Botanical Gazette* 119: 125–170.

Popper K. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

Raven J. A., and Yin Z.-H. 1998. The Past, Present and Future of Nitrogenous Compounds in the Atmosphere, and their Interactions with Plants. *New Phytologist* 139: 205–219.

Rosen R. 1967. *Optimality Principles in Biology*. New York: Plenum Press.

Salthe S. N. 1985. *Evolving Hierarchical Systems.* New York: Columbia University Press.

Solé R. V., Montoya J. M., and Erwin D. H. 2002. Recovery after Mass Extinction: Evolutionary Assembly in Large-Scale Biosphere Dynamics. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 357: 697–707.

Williams G. C. 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought.* Princeton, NJ: Princeton University Press.

Wilson D. S. 1997. Biological Communities as Functionally Organized Units. *Ecology* 78(7): 2018–2024.

8
Relationship between Genome Size and Organismal Complexity in the Lineage Leading from Prokaryotes to Mammals

Alexander V. Markov, Valery A. Anisimov, and Andrey V. Korotayev

Abstract

The lack of a strict relationship between genome size and organismal complexity (level of organization) is largely due to size variability of the facultative part of the genome. However, there is a direct relationship between the level of organization and the minimal genome size (MGS) in the lineage leading from prokaryotes to mammals, in which the tendency towards increasing complexity is especially clear. The dynamics of MGS in this lineage can be adequately described by the model of hyperexponential growth. This implies the existence of nonlinear positive feedbacks that account for the acceleration of MGS growth. The nature of these feedbacks is discussed, including the formation of new genes by means of recombination of the fragments of existing genes, formation of 'niches' for new genes in the course of evolution of gene networks, and the expansion of regulatory regions. Hyperexponential growth of different variables related to the level of organization of the biosphere and society (biodiversity, MGS, size and complexity of organisms, world population, technological development, urbanization, etc.) suggests that the evolution of the biosphere and humanity in the direction of increasing complexity is a self-accelerating (autocatalytic) process.

Keywords: *genome size, evolution, level of organization, complexity.*

Introduction

The problem of a possible connection between genome size and level of morphophysiological organization (complexity) of organisms has been a focus of attention of biologists for a while. A hypothesis of a positive correlation existing between these parameters seems logical because it appears that to create more complex organisms in the course of ontogeny a larger 'developmental program' should be necessary. However, until very recently no strict correlation of this kind could be revealed (Gregory 2005). Attempts to reveal regular growth of a genome in isolated evolutionary lineages were usually unsuccessful (Thomas 1971; Gregory and Hebert 1999; Gregory 2005, 2008). On the other hand, the size of a genome was noted to considerably increase over the transition from prokaryotes to eukaryotes and from unicellular eukaryotes to multicellular eukaryotes. There are also indications of the existence of a positive relationship between the size of unique genome sequences and level of organization. In addition there is general growth of the genome in the most 'progressive' evolutionary lineage leading from prokaryotes to mammals (this lineage most clearly shows the trend toward morphophysiological progress) (Raff and Kaufman 1986; Patthy 1999; Sharov 2006). This paper aims at the analysis of the relationship between the genome and organismal complexity in this evolutionary lineage. The minimal genome size (MGS) within a large taxon was used as a measure of genome complexity. It is shown below that this evolutionary lineage displays an increase of MGS, which is adequately described by a biexponential variety of the model of hyperexponential growth. Previously models of hyperexponential (primarily hyperbolic) growth have been successfully used to describe the dynamics of some demographic and macrosociological parameters (population, levels of technological and economic development, urbanization, literacy, *etc.*) (Korotayev, Komarova, and Khaltourina 2007; Korotayev, Malkov, and Khaltourina 2006, 2007; Grinin and Korotayev 2009). In addition, it has been shown that the hyperbolic model can be used to describe the dynamics of taxonomic diversity of the Phanerozoic biota (Markov and Korotayev 2008, 2009). The hyperexponential growth usually suggests the presence of complex nonlinear positive feedback facilitating growth acceleration of a parameter under consideration.

Minimal Genome Size as a Measure of Necessary (Nonredundant) Amount of Genetic Information

We consider minimal (as opposed to mean, or maximum) genome size in each group because genomes of most organisms are known to contain a large amount of so-called 'junk DNA', for which no function has yet been identified. This amount may greatly vary even within a class or order. The amount of 'junk DNA' is largely determined by factors not directly linked to life and adaptations of an organism. A large amount of 'junk DNA' consists of mobile genetic elements (MGE). For instance, mobile elements compose roughly 50 % of primate genomes (Xing *et al.* 2007). Presence of 'extra' DNA in the genome places an additional load on the organism, which has to use more resources for its replication; the genome size may influence the cell size, rates of replication and cell division, *etc.* (Gregory 2005; Gregory and Hebert 1999). While there is a conventional hypothesis on the redundancy of MGE, introns and other noncoding sections, MGE are sources of genetic variability, while fragments of these selfish elements are actually dynamic reservoirs for new cellular functions ('domesticated elements') (Miller *et al.* 1999; Volff 2006). MGE are shown to play a significant role in the evolution of eukaryotes including evolution towards increased complexity (Bowen and Jordan 2002; Muotri *et al.* 2007).

A dynamic equilibrium between the trends toward longer and shorter 'junk DNA' sections is established in the course of evolution. The former trend results from spontaneous self-duplication of MGE, while the latter – from deletions (Gregory 2004b). MGE loss is generally advantageous because of the economy of resources used for synthesis and upkeep of 'extra DNA' in each cell. If one of two trends prevails, the genome 'inflates' or 'shrinks'. The prevalence of the first trend can be related to the appearance of a new form of MGE with a higher replication rate or with loosening of cellular systems of MGE control. The second trend may prevail if the loss of extra fragments gives a significant adaptive advantage. A typical example is the advantage of having a reduced genome in birds and bats compared to flightless tetrapods, because of the reduced body weight. 'Extra' DNA is present in each cell and its upkeep and replication requires numerous 'extra' proteins, which results in positive correlation between the genome size and cell size (Organ *et al.* 2007). Therefore, selection favored loss of 'junk DNA' in flying vertebrates, which led to the reduction in genome size (Hughes and Hughes 1995). It is noteworthy that neither increased complexity, nor increased genome size are uniform evolutionary tendencies. For instance, the evolution of prokaryotes is dominated by a reduction in genome size rather than by its increase (Ochman 2005). A similar pattern was apparently present in the evolution of Saurischia and birds (Organ *et al.* 2007). The physiology of a particular organism may affect the genome size. For instance, in prokaryotes, with their imperfect systems of DNA repair and distribution (absence of mitosis), the genome cannot grow beyond some maximum limit (Sharov 2006): the maximum genome size in bacteria is 13.03 Mb (*Sorangium cellulosum*), in archaea 5.75 Mb (*Methanosarcina acetivorans*).[1] Genome growth in prokaryotes may be restricted by large population size, which slows down genetic drift (Lynch and Conery 2003). On the other hand, the existing level of complexity of an organism suggests that the genome cannot be reduced below a particular minimum level. We suggest that MGS in a large group of organisms can be used for an approximation of the amount of essential (non-redundant) genetic information necessary for the existence of representatives of a taxon under consideration. A more precise proxy is difficult to obtain because there is no reliable means to differentiate genuinely redundant sections in DNA from functional ones (*e.g.*, from non-coding sequences performing regulatory functions).

Materials and Methods

We compared MGS in nine successive groups of organisms that 'nest one inside the other' (see Table 1). The succession of groups corresponds to the evolutionary lineage from the earliest prokaryotes to mammals. The choice of an organism with minimal genome in each group was made without taking into account intracellular symbionts and parasites, which are often subjected to genetic simplification and to some extent lose the right to be called independent

[1] URL: http: www.ncbi.nlm.nih.gov/genomes/lproks.cgi.

organisms (see *e.g.*, Nakabachi *et al.* 2006). Intracellular symbionts exploit the host organism's genes instead of their own lost genes, allowing them to survive without many genes which are absolutely essential for free-living organisms in the same group. Genes of symbionts can be transferred to the host organism's genome where they continue functioning to the mutual benefit of the symbiont and the host (as happened in the symbiosis of early eukaryotes with the future mitochondria and plastids) (Stegemann *et al.* 2003; Markov and Kulikov 2005).

Each of the successive groups under consideration is a subset of the previous one, and is the subset within which organisms achieved the highest level of complexity. It is characteristic that within each group, the smallest genome size was recorded for those members of the group, which were not included in the subsequent subset. For instance, group 6 (tetrapods), the smallest genome, is characteristic for a representative of amphibians, *i.e.*, lower tetrapods, rather than for some members of a higher subset 7 (amniotes), although amniotes are included in tetrapods. This alone shows that a correlation exists between MGS and organismal complexity.

Table 1. Minimal (nonredundant) genome size (MGS) in nine nested groups of organisms

Group	MGS (Mb)	Approximate time of appearance (Ma)	Species with the smallest genome (apart from intracellular parasites)
1. All living beings	1.3	4000	Marine free-living bacterium *Pelagibacter ubique*, strain HTCC1062
2. Eukaryotes	9.2	2000	The sac fungus *Ashbya gossypii* with the smallest genome among free-living eukaryotes
3. Animals (Metazoa)	19.6	1250	Nematode *Pratylenchus coffeae*
4. Chordates	68.6	575	*Oikopleura dioica appendicularium*, representing the subphylum Tunicata in the phylum Chordata
5. Vertebrates	342	540	Bony fish *Tetraodon fluviatilis*
6. Tetrapods	931	375	Frog *Limnodynastes ornatus*
7. Amniotes	951	315	Pheasant *Phasianus colchicus*
8. Mammals	1695	220	Bat *Miniopterus schreibersi*
9. Primates	2215	65	Collared Titi monkey *Callicebus torquatus*

Note: Data on the size of the minimal genomes in groups are based on the following sources: Pellicciari *et al.* 1982; Gregory 2004a, 2008; Dietrich *et al.* 2004; Complete … 2008; Eukaryotic genome sequencing projects 2008. Approximate dating of the appearance of groups is based on the molecular and paleontological data from: Marshall and Schultze 1992; Shu *et al.* 1999; Heges and Kumar 2003; Battistuzzi *et al.* 2004; Fedonkin 2006; Falcon-Lang *et al.* 2007. Some data suggest earlier appearance of eukaryotes, probably 2.7 Ga (Rozanov 2003), although one of the most important facts indicating this (presence of eukaryotic biomarkers in the Archean rocks) was recently questioned by Rasmussen *et al.* (2008). Therefore, we use a more conservative estimate of 2 Ga, agreed among most authors, and supported by molecular data.

We had to exclude some intermediate levels which lack reliable data. Following the logic used above, it was possible to place an intermediate group of gnathostomes between subset 5 (vertebrates) and 6 (tetrapods). This was not done because of the insufficient data on agnathans, *i.e.*, the vertebrates that are not included in gnathostomes. Agnathans were diverse in the Paleozoic, but in the recent biota they are represented by only two highly specialized groups – lampreys and hagfishes, among which species with small genomes have not been yet identified (Gregory 2008). However, this does not necessarily mean that in Paleozoic agnathans genomes were as large as in extant lampreys and hagfishes (based on data presented in this paper, we suggest that this was not the case). Nevertheless, it has to be acknowledged that taking into account such relicts or poorly studied groups could to some extent obscure patterns discussed in this paper. One important question that inevitably emerges in any study using regression analysis is the selection of adequate models. An oversimplified model may occasionally not reveal essential details in patterns studied, whereas an overcomplicated model may accidentally focus on a background constituent of the experimental data (noise). Taking this into consideration we studied two classes of models: biparametric and triparametric ones. Biparametric models are generally more reliable and less affected by measurement errors, whereas triparametric models are more informative (when the given data are sufficiently precise).

We studied two families of biparametric models: exponential and hyperbolic. The exponential models can be generalized as: $L = Ae^{(-BT)}$, where *L* – MGS of a taxon (in Mb); *A*, *B* are adjustment parameters, *T* – time from a supposed appearance of a taxon (in Myr). Hyperbolic models are described as $L = A/(B + T)$. Because the value of *L* in the lineage from prokaryotes to mammals changes by more than three orders of magnitude, the use of logarithmic scale is reasonable (we used natural logarithms of MGS). We have also considered three families of triparametric models: power exponential, power hyperbolic, and biexponential. The power exponential model is described by a formula $L = Ae^{(-BT^N)}$, where *N* is the third adjustment parameter. The power hyperbolic model is described by a generalized formula $L = A/(B + T)^N$. The biexponential model is described as $L = Ce^{Ae^{(-BT)}}$, where the third parameter is the coefficient *C*. Optimal parameters of the models were chosen directly by selection of numbers using the least square method.

Table 2. Data previously used to substantiate the exponential growth of nonredundant genome size

Group	Nonredundant genome size (Mb)	Time of appearance
1. Prokaryotes	0.5	3500
2. Eukaryotes	2.9	2000
3. Worms	72.8	1000
4. Fishes	133.3	500
5. Mammals	480	125

Source: Sharov 2006.

Results

Of the biparametric models considered, the given data (Plate 1) are best described by an exponential model with values of parameters A = 1.002 × 10^9, B = 0.00191; the Pearson's correlation coefficient is relatively high (R^2 = = 0.914). Of triparametric models the best fitted is a biexponential model with the values: A = 8.41, B = 7.98 × 10^{-4}, C = 1.069 × 10^8. The correlation with the observed data is higher than for a simple exponential model (R^2 = 0.979) (which is understandable because triparametric models generally allow a more precise approximation of the existing data than the biparametric models). As a control, the parameters of the exponential and biexponential models were calculated based on Sharov's data (Sharov 2006; Table 2; see Discussion). For the exponential model the following optimal values of parameters were calculated: A = 4.37 × 10^8, B = 0.00206, for biexponential: A = 11.75, B = 2.88 × × 10^{-4}, and C = 6.077 × 10^3. Note that Sharov, using an exponential model, obtained similar results.

The MGS growth dynamics in the evolutionary lineage considered is shown in Fig. 1 (black diamonds). Sharov's data are shown as grey squares. Graphs corresponding to optimal biparametric models are shown by solid lines, black (our data) or grey (Sharov's data), whereas curves corresponding to triparametric models as moire black strips (for our data) and grey (for Sharov's data) lines.

Discussion

Exponential or hyperexponential growth?

The analysis of size dynamics of a ‘non-redundant’ genome in a given evolutionary lineage (Sharov 2006) leads to a conclusion that this parameter grows exponentially. Our analysis partly repeats Sharov's work; however, we estimate the essential genome size differently and also use more detailed data. Sharov built his exponential graph based on five points that in their biological sense correspond to groups 1, 2, 3, 5, and 8 considered herein (see Tables 1, 2). Why, then, did Sharov suggest exponential growth, while we more or less confidently suggest more accelerated, hyperexponential growth? This is firstly because Sharov's paper did not consider any other models except for exponential, and secondly because of different methods for estimating the non-redundant genome. For groups 1 and 2 Sharov used genome size of parasitic microorganisms, whereas we use minimal genomes of free-living species. However, these differences do not much influence the correlation of data with either exponential or biexponential models. The essential genome size in group 3, according to Sharov, is 72.8 (this is the genome size of the nematode *Caenorhabditis elegans* excluding 25 % that are supposedly non-functional). Our estimate (19.6) is more realistic, because this is the genome size of another nematode (*Pratylenchus coffeae*). Obviously, the genome of *C. elegans*, even excluding ‘nonfunctional’ regions, has many regions that are absent in other round worms. For fish and mammals, Sharov's estimates are lower than ours because Sharov did not consider those parts of genome that are supposedly non-functional (65 % of the fish genome and 85 % of the mammal genome). However, criteria used by Sharov to estimate the size of the ‘nonfunctional’ genome regions apparently do not take into account that many noncoding regions, with function as yet unidentified, may perform important regulatory functions, or code functional RNA. For instance, in mammal genomes many regions previously considered as nonfunctional are transcribed. In addition, these regions have recently been found to contain a whole class of previously unknown genes coding large RNA molecules with regulatory functions (Guttman *et al.* 2009). There are many indications of a very important role of MGE in eukaryote evolution, including evolution towards increased complexity (Miller *et al.* 1999; Bowen and Jordan 2002; Muotri *et al.* 2007). In our opinion, the adequate interpretation of the minimal necessary genome in vertebrates can be obtained from organisms in which a decrease in genome size has adaptive signifycance. Apparently flying vertebrates (birds and bats) can provide this evidence (Hughes A. and Hughes M. 1995; Organ *et al.* 2007). We considered their genomes as ‘minimally necessary’ for groups 7 (amniotes) and 8 (mammals). Sharov's data also differ from ours because, according to Sharov, mammals appeared ca. 125 Ma (Early Cretaceous). We cannot agree with this estimate because mammals are known as early as the Late Triassic (*ca.* 220 Ma), whereas 125 Ma is the time of the earliest find of placental mammals (Ji *et al.* 2002).

Comparison of our Biparametric Model with Sharov's Results

Despite the above disparities of the initial data, the slopes of the exponential lines on the resulting logarithmic graph (Fig. 1) are very similar. This indirectly supports the adequacy of rough estimates given by exponential models. Simpler biparametric models apparently give rougher local estimates, but global estimates obtained that way are more reliable. Therefore, it is possible that estimates of the appearance of life in the Universe (at least 7 Ga) obtained by Sharov from extrapolation of models into the past, can be considered rather seriously (but only if such extrapolation itself is assumed to be reliable).

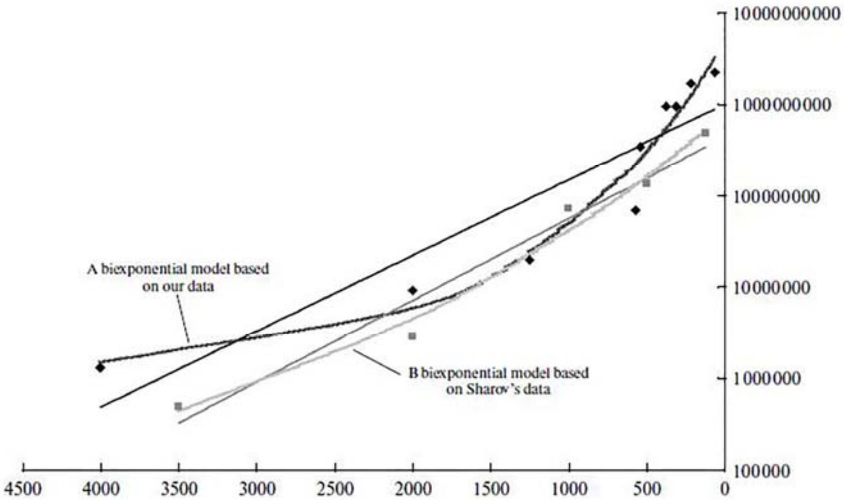


Fig. 1. Biexponential growth of the minimal genome size (MGS) in the lineage from prokaryotes to mammals. The horizontal axis shows the time of the group appearance in Ma. The vertical axis shows MGS in Mb in the logarithmic scale. Markers correspond to the groups in Table 1. (A) biexponential model (moire black line) in our data describes the observed dynamics better than the exponential (solid black line). (B) biexponential model (moire grey line) according to Sharov's (2006) data also described the observed dynamics better than the exponential model (solid grey line)

Genome Size and Organismal Complexity

The absence of a direct correlation between organismal complexity and genome size is well substantiated. For instance, unicellular eukaryotes include taxa with a genome that exceeds all studied genomes of multicellular animals. Amphibians include species with genomes larger than in mammals, *etc.* (Thomas 1971; Gregory and Hebert 1999; Gregory 2005). Some reasons for high variability of genome size are mentioned above. However, in the evolutionary lineage from prokaryotes to mammals, in which the trend toward increased complexity was the strongest, positive correlation between the nonredundant genome size and organismal complexity is clearly displayed (Patthy 1999; Sharov 2006). This agrees with theoretically expected results based on an interpretation of genome as a 'program' of the development and function of an organism. It is natural to expect that more complex organisms would have a more complex and hence larger 'program'. For instance, it has been proposed that the size of the functional nonredundant regions of the genome can be considered as a measure of the biological complexity of organisms (Adami, Ofria, and Collier 2000). We think that MGS within a large group of organisms is a good approximation to the size of the nonredundant genome, which is difficult to calculate. Thus, the hypothesis of relationships between MGS and organismal complexity does not contradict conventional wisdom.

Vendian – Cambrian Acceleration of MGS Growth

Despite a generally good approximation of the MGS dynamics from the biexponential formula, the graph shows a considerable discrepancy between the observed data and the approximating curve in the latest Proterozoic (Vendian) and Cambrian. A sharp acceleration of MGS growth in the Vendian – Cambrian coincides with the adaptive radiation of Metazoa (Vendian – Cambrian explosion). This time apparently corresponds to the appearance of the first chordates (point 4 on the graph), and then vertebrates (point 5). The most obvious reason for the sharp increase in MGS growth at the time of the appearance of vertebrates is the occurrence of two whole-genome duplications; this was a key event in the early evolution of the vertebrates (Putnam *et al.* 2008).

It should be taken into account that in the period under consideration the biosphere underwent a fundamental restructuring. A sharp increase in MGS, perhaps, reflects the transition of MGS from one stable (exponential?) growth trajectory to another, which could be to some extent a consequence (or demonstration) of this global change. Study of complex systems with nonlinear positive feedbacks developing over a large time scale hyperexponentially (and hyperbolically) shows that in many cases such global evolutionary motion of a system when studied at a smaller scale becomes fragmented into a number of stages separated by phase transitions, the succession of which forms large-scale hyperexponential dynamics. A system usually has so-called 'attractors', near which it exists in the condition of a local optimum and can persist in this way for a long time until the external pressure or changes gradually accumulating in it push the system away from the attractor. After that, the system again evolves rapidly until after the next phase transition it is trapped by the next attractor. Recent papers (Korotayev 2006, 2007; Korotayev and Grinin 2007; Korotayev, Komarova, and Khaltourina 2007a; Korotayev, Malkov, and Khaltourina 2007b; Grinin and Korotayev 2007, 2009) give examples of similar behavior in the social World System, which was until very recently evolving towards the hyperexponential (including hyperbolic) growth in its basic macrosociological parameters (size of populations, levels of technical, economic, and sociocultural development, degree of urbanization and political complexity, *etc.*). It can be suggested that the sharp change in the MGS dynamics in the Vendian–Cambrian might have resulted from the fact that the system (biosphere) at that time was leaving its then attractor (the world of prokaryotes and unicellular eukaryotes) and beginning to transform in a process of phase transition into the world of multicellular organisms. Interestingly, at the end of the Proterozoic – beginning of the Paleozoic, the maximum size of living organisms sharply increased (by many orders of magnitude), which is related, among other things, to the sharp growth in the concentration of free oxygen (Payne *et al.* 2009). We mentioned above that a positive correlation can be traced between the size of an organism and genome size.

Positive Feedback Providing Hyperexponential MGS Growth

Similar to the human population dynamics and other macrosociological parameters, and the biodiversity dynamics (Markov and Korotayev 2009), the hyperexponential genome growth is supposedly provided by complex positive feedbacks. The hypothesis that the genome growth in evolution was governed by positive feedbacks is discussed by Sharov (2006), who has suggested the following mechanisms of genome growth based on positive feedback:

1) A genome can be considered as an assemblage of symbiotic self-replicating elements, or as a hypercycle (Eigen and Schuster 1977). For instance, the gene responsible for higher precision of DNA replication facilitates more precise copying of all other genes in the cell, and this advantage involves not only genes that are already present, but also those that will appear in the future. Thus, already existing genes can facilitate the invasion and affixation of new genes into the genome.

2) New genes are often produced by duplication or recombination of already existing genes. Therefore, a large genome represents more initial material for the formation of new genes.

3) Large genomes support a higher diversity of metabolic networks and morphological elements and thus provide more potential niches for new genes (Sharov 2006).

To understand which of the supposed positive feedbacks can provide hyperexponential MGS growth, they should be discussed in greater detail. Firstly, the exponential model corresponds to cases when the variable under consideration grows with a rate proportional to the variable's value (dL/dT ~ L), whereas in biexponential models the growth is proportional to its current value multiplied by its logarithm (dL/dT ~ Lln(L)). The presence of the second factor determines the hyperexponential growth of the parameter under consideration. Even greater acceleration of the variable's growth is observed in another kind of hyperexponential dynamics, *i.e.*, hyperbolic growth. In that case the variable's growth rate is proportional to the square of its value: dL/dT ~ L^2 (the solution of this differential equation is a hyperbolic function, see *e.g.*, Kapitsa 1992, 1999; Korotayev 2006: 119–120). Note that the biexponential dynamics is an intermediate between the exponential and hyperbolic dynamics.

The analysis of existing data on the mechanisms of genome growth suggests that some positive feedbacks governing MGS growth could, with the course of time, give rise to exponential dynamics, while others could result in the hyperbolic dynamics, whereas their joint action leads to the intermediate result, *i.e.*, biexponential dynamics. Let us consider duplication of the DNA fragments, as one of the major mechanisms of genome growth. In the simplest case it can be assumed that the probability of duplication of a DNA fragment of fixed length is a constant value. Because the number of such fragments is proportional to the genome length, its growth rate due to random duplications should also be proportional to its length. With the course of time, this should lead to exponential genome growth. However, the triparametric model indicates a considerable deviation from the exponential law, especially after the appearance of the metazoans. Hyperexponential growth can be related to the formation of new functional (coding and regulatory) regions of DNA based on the combinatory principle. New genes are often formed due to recombination of fragments of existing genes (Patthy 1999). It is easy to demonstrate that the number of potentially possible new combinations of fragments (*i.e.*, new genes that can potentially be formed in such way) is approximately proportional to the squared number of existing genes. In an idealized situation each gene consists of two domains (functional blocks); new genes are formed by merging of copies of two domains, originating from two different genes. In this situation, each pair of genes can potentially give rise to four new genes (if the order of the domain arrangement in a new gene is not taken into account). Hence, the total number of potentially possible new genes can be calculated as $2(N^2 - N)$, where N is the number of genes in a genome. Considering that genomes of free-living organisms contain quite a high number of genes (from thousands to tens of thousands), the formula can be simplified as $2(N^2 - N) \approx 2N^2$. Assuming that the growth rate of the number of genes is proportional to the number of potentially possible new genes, we obtain the following expression: $dN/dt = kN^2$, which corresponds to the hyperbolic growth of the number of genes. A similar deduction can also be applied to noncoding regulatory sequences, which apparently can also be formed by recombination of fragments of existing regulatory sequences. As noted above, the combined action of factors, some of which facilitate exponential, and some hyperbolic MGS growth, can lead to an intermediate kind of dynamics, for example, biexponential growth. It is important that the above mechanism of formation of new genes by recombination of domains (or exons) of old genes is found in Metazoa much more often than in other organisms. Most new genes of animals that formed in such a way (*i.e.*, module, multidomain genes), appeared early in the evolution of Metazoa (*Ibid.*), which coincides with the period of sharp acceleration of MGS growth at the end of the Neoproterozoic – beginning of Paleozoic. It is noted that most of these genes are related to specific features of animals such as cell and tissue differentiation, intercellular interactions, and other aspects of function of a metazoan animal organism as an entity. In other words, these genes are directly related to coding of organismal complexity in Metazoa (Patthy 1999). The predominance of this mechanism capable to provide hyperexponential genome growth in animals (as opposed to other multicellular organisms) explains rather well why this growth type is distinctly recognized in the evolutionary lineage considered. The hyperexponential MGS growth may also be connected with patterns of the evolution of gene networks (Kolchanov *et al.* 2000; Kolchanov, Suslov, and Shumnyi 2003; Kolchanov, Suslov, and Gunbin 2004). According to the principle of combinatory coding of complexity, the increase in complexity of gene networks proceeds not only due to the increase of the number of genes involved in their work but also through the increased complexity of mechanism of their interactions (mutual regulation). Evidently, the potential number of intergene interactions, direct or indirect, should grow proportionally to the number genes squared (because potentially any gene can interact with any other gene in the genome). Regulatory regions of DNA (various enhancers containing binding sites of transcriptional regulators) are used to perform intergene interactions. It is known that the increased complexity of metazoans is largely provided by the appearance of new regulatory sequences (Mikkelsen *et al.* 2007). Thus, growth in the number of genes should create new niches for regulatory sequences, and the number of niches should grow proportionally to the number of genes squared. The picture will become more complicated if it is taken into account that the regulatory sequences themselves can interact with each other, for example, due to competition for the same regulatory proteins, and the same gene may be regulated by the coordinated work of the entire complex of various regulatory sequences and transcription factors (Kolchanov *et al.* 2000; Takahashi *et al.* 2008). It is important that new regulatory sequences often appear by 'domestication' of MGE, which thus are effectively transformed from 'junk DNA' to functionally important genome components (Miller *et al.* 1999; Mikkelsen *et al*. 2007). More so, new regulatory proteins (transcription factors) can also be formed by domestication of MGE genes (Lin *et al.* 2007). This mechanism of growth of the functional regions of the genome due to the transformation of 'junk DNA' into functional DNA, apparently should lead to a situation when growth of nonfunctional regions of the genome in a long term should facilitate the acceleration of MGS growth.

Note that organism size growth (partly co-ordinated to MGS growth) leads to reduction in population size and hence to an increase in genetic drift (*i.e.*, to higher probability of random fixation of neutral and slightly harmful genetic changes). Therefore, new copies of reproducing MGE have more chance of being fixed in a small population than in a large one. This can lead to passive growth of 'junk' regions of the genome in organisms with small populations, for which large size and large genomes are characteristic (Lynch and Conery 2003). This is partly supported by the well-known fact that as the genome's (and organism's) size grows, the proportion of noncoding regions grows as well. For instance, in the prokaryotes to mammals lineage considered, the genome size grows approximately by three orders of magnitude (from a few million to a few billion Mb), whereas the number of protein-coding genes increases only by one order of magnitude (from a few thousand to a few tens of thousands). On the other hand, it is known that the genome of complex organisms contains many functional non-coding regions, which play an important role in the evolution of complexity (Miller *et al.* 1999; Bowen and Jordan 2002; Volff 2006; Mikkelsen *et al.* 2007; Muotri *et al.* 2007). In other words, an increase in the complexity and size of organisms can lead to an increase in both 'junk' and functional noncoding regions.

Apparently, an increase in complexity in organisms can precede the genome growth and stimulate it due to the mechanism of 'escape from adaptive conflict' during duplication of genes. As complexity increases, many genes can acquire additional functions, *i.e.*, becoming multifunctional. Such genes are in a state of adaptive conflict: selection cannot efficiently optimize them to perform one of the functions because that would result in the reduction in efficiency of other functions performed by the gene. New gene copies which appeared as a result of gene duplications can specialize to perform various functions. This considerably lowers the probability that the new copies that appeared as a result of duplications will be redundant and will be lost (Des Marais and Rausher 2008). This mechanism played an important role in the early evolution of vertebrates, when after two whole-genome duplications, many newly formed extra gene copies remained in descendants and acquired new functions (Putnam *et al.* 2008). It is possible to assume the connection between the increase in biodiversity and genome size in the most highly organized representatives of the biota. Growth of biodiversity leads to increased complexity and heterogeneity of the biotic environment. This creates predisposition to the development of complex adaptations and hence, complex organisms. Increased complexity of organisms in turn facilitates further genome growth. Computer simulations have shown that in organisms evolving in an 'information-rich' (complex and heterogenous) environment, the genome grows because it embraces information about the environment and of how to function most efficiently in this environment. In contrast, in organisms evolving in an information-impoverished environment the genome size decreases (Adami *et al.* 2000; Ofria, Adami, and Collier 2003).

It is noteworthy that the dynamics of the hyperexponential growth of biodiversity and MGS in the lineage from prokaryotes to mammals are essentially different. In the former, a period of explosive growth occurs in the last

100–150 Ma (Late Mesozoic – Cenozoic). The first signs of the beginning of the end of the blow-up regime are recorded for the second half of the Cenozoic (Markov and Korotayev 2009). In the MGS dynamics, the period of explosive growth corresponds to the Paleozoic and essentially finished in the Triassic with the entry of mammals. Thus, the MGS growth dynamics shows the end of the regime with hyperexponential acceleration as early as the beginning – middle of the Mesozoic, 100–200 Myr earlier than in biodiversity growth. It is possible that genome growth to some extent contributed toward biodiversity growth by creating additional levels of freedom for genetic transformations (although it is necessary to remember that the explosive diversity growth in the Mesozoic and Cenozoic occurred mainly due to comparatively simply organized animals, such as mollusks and insects, and to a lesser extent due to teleosts, birds, and mammals).

The morphological complexity of organisms in the lineage under consideration probably also grew at least exponentially (or even hyperexponentially), although it is difficult to check because of the absence of reliable estimates of the level of morphological complexity in the groups considered.

The decrease in the MGS growth rate after the appearance of mammals does not mean that the increase of complexity slowed down to the same extent. If the phenotype is understood in the wide sense, including not only morphology but also behavior, and extrasomatic adaptations (beaver dams, bird nests, *etc.*) (Dawkins 1982), it becomes apparent that the phenotype complexity growth rate did not slow down in the Mesozoic and Cenozoic. Growth of encephalization quotients of mammals in the Cenozoic was accompanied by an increase in ability to learn and increased complexity of behavior (including social behavior). More so, this increase gradually created a basis for biological evolution to transform into cultural and social evolution (Grinin, Markov, and Korotayev 2008). These phenomena can be considered as stages of one accelerating global process of extraction of information from the environment by the biota and its preservation on an 'external carrier'. DNA initially worked as such a carrier, but after the appearance of the more efficient means of processing, transmission and storage of information in a complex nervous system (particularly the mammalian brain with a developed neocortex), and then speech and writing, the evolutionary pressure towards the increase of the informational capacity of the genome apparently weakened. Although the subsequent progressive changes were not accompanied by an accelerated MGS growth, some of them apparently required the development of more complex mechanisms of gene regulation (Mikkelsen *et al.* 2007).

Alternative splicing, a process of editing of matrix RNA molecules through which a cell can synthesize more than one different protein based on the same gene is an example of a mechanism allowing an increase in the 'useful' complexity and informational capacity of the genome without increasing its size. It has been shown that about 94 % of human genes undergo alternative splicing, whereas in lower animals alternative splicing is found in the minority of genes (*e.g.*, about 15 % in *C. elegans*) (Wang *et al.* 2008; Pan *et al.* 2008). This discovery answers the intriguing question of why the human genome contains approximately the same number of genes as a much less complexly organized worm *C. elegans* (about 20,000). It has been shown that the diversity of proteins in the human organism is in fact (as would be expected) much higher than in the worm, although this diversity is achieved not by genome growth, but by the development of alternative splicing.

The existing data on mechanisms and rates of genome growth are still insufficient to build adequate mathematical models of this process. The main difficulty is the absence of rigid quantitative evaluations of the relative contribution of different mechanisms of genome growth in the total dynamics of this growth.

The discussion of possible extrapolation of the model curves onto the past has remained beyond the scope of this paper. The results of such a procedure were interpreted by Sharov (2006) as evidence of the extraterrestrial origin of life. The validity and methods of such extrapolation and the conclusions that can be made based on it require detailed discussion, which will be presented in a separate paper. Here we shall only note that there are arguments both 'for' and 'against' the hypothesis that the sharp discontinuity of the curve in the left side of the model graph (see Fig. 1), *i.e.*, at the time of the supposed appearance of the prokaryotic cell, can be used as an argument supporting the hypothesis of extraterrestrial abiogenesis.

Conclusions

In the evolution of the biosphere, as in the evolution of society, some parameters reflecting the general level of the development or complexity of the system changed in time in accordance with the hyperexponential (and often hyperbolic) mode. In the biosphere such parameters include biodiversity and nonredundant genome size in the most complex organisms, and also apparently maximum size of organisms and maximum level of complexity of their organization. The hyperexponential growth of these parameters suggests that the evolution of the biosphere towards general increased complexity, like social and cultural evolution of mankind, is regulated by nonlinear positive feedback and is a self-accelerating process. In other words it is possible that complexity itself is the reason for the progressive increase in complexity of biological and social systems.

Acknowledgments

The authors are grateful to A. D. Panov, A. S. Rautian, and A. A. Fomin for discussion of ideas considered in this paper and valuable recommendations.

References

Adami C., Ofria C., and Collier T. C. 2000. Evolution of Biological Complexity. *Proceedings of the National Academy of Sciences* 97(9): 4463–4468.

Battistuzzi F. U., Feijao A., and Hedges S. B. 2004. A Genomic Timescale of Prokaryote Evolution: Insights into the Origin of Methanogenesis, Phototrophy, and the Colonization of Land. *BMC Evolutionary Biology* 4: 44.

Bowen N. J., and Jordan I. K. 2002. Transposable Elements and the Evolution of Eukaryotic Complexity. *Current Issues in Molecular Biology* 4(3): 65–76.

Complete Microbial Genomes. **2008.** URL: http: www.ncbi.nlm.nih.gov/genomes/ lproks.cgi. Date accessed: 11.10.2008.

Dawkins R. 1982. *The Extended Phenotype: The Long Reach of the Gene.* Oxford: Oxford University Press.

Dietrich F. S., Voegeli S., Brachat S. *et al.* 2004. The Ashbya gossypii Genome as a Tool for Mapping the Ancient Saccharomyces cerevisiae Genome. *Science* 304(5668): 304–307.

Eigen M., and Schuster P. 1977. The Hypercycle. A Principle of Natural Self-Organization. Part A: Emergence of the Hypercycle. *Naturwissenschaften* 64(11): 541–565.

Eukaryotic Genome Sequencing Projects. **2008.** URL: http: www.ncbi.nlm.nih.gov, genomes/leuks.cgi. Date accessed: 11.10.2008.

Falcon-Lang H. J., Benton M. J., and Stimson M. 2007. Ecology of Earliest Reptiles Inferred from Basal Pennsylvanian Trackways. *Journal of the Geological Society* 164(6): 1113–1118.

Fedonkin M. A. 2006. Two Records of Life: Experience of Comparison (Paleobiology and Genomics on the Early Stages of the Biosphere Evolution). *Problems of Geology and Mineralogy* / Ed. by A. M. Pystin, pp. 331–350. Syktyvkar: Geoprint. *In Russian* (Федонкин М. А. Две летописи жизни: опыт сопоставления (Палеобиология и геномика о ранних этапах эволюции биосферы). *Проблемы геологии и минералогии* / Ред. А. М. Пыстин, с. 331–350. Сыктывкар: Геопринт).

Gregory T. R. 2004a. Macroevolution, Hierarchy Theory, and the C-Value Enigma. *Paleobiology* 30(2): 179– 202.

Gregory T. R. 2004b. Insertion-Deletion Biases and the Evolution of Genome Size. *Gene* 324: 15–34.

Gregory T. R. 2005. The C-Value Enigma in Plants and Animals: A Review of Parallels and an Appeal for Partnership. *Annals of Botany* 95(1): 133–146.

Gregory T. R. 2008. *Animal Genome Size Database.* URL: http: www.genomesize. com.

Gregory T. R., and Hebert P. D. N. 1999. The Modulation of DNA Content: Proximate Causes and Ultimate Consequences. *Genome Research* 9(4): 317–324.

Grinin L. E., and Korotayev A. V. 2007. The Political Development of World-System: Formal and Quantitative Analysis. *History and Mathematics: Macrohistorical Dynamics of the Society and State* / Ed. by S. Yu. Malkov, pp. 49–101. Moscow:

KomKniga/URSS. *In Russian* (Гринин Л. Е., Коротаев А. В. Политическое развитие Мир-Системы: формальный и количественный анализ. *История и Математика: макроисторическая динамика общества и государства* / Ред. С. Ю. Малков, Л. Е. Гринин, А. В. Коротаев, с. 49–101. М.: КомКнига).

Grinin L. E., and Korotayev A. V. 2009. *Social Macroevolution: The Genesis and Transformations of the World System.* Moscow: LIBROKOM/URSS. *In Russian* (Гринин Л. Е., Коротаев А. В. *Социальная макроэволюция: Генезис и трансформации Мир-системы.* М.: Либроком/УРСС).

Grinin L. E., Markov A. V., and Korotayev A. V. 2008. Macroevolution in Living Nature and Society. Moscow: LKI/URSS. *In Russian* (Гринин Л. Е., Марков А. В., Коротаев А. В. *Макроэволюция в живой природе и обществе*. М.: ЛКИ/УРСС).

Guttman M., Amit I., Garber M. *et al*. 2009. Chromatin Signature Reveals over a Thousand Highly Conserved Large Non-Coding RNAs in Mammals. *Nature* 458 (7235): 253–227.

Heges S. B., and Kumar S. 2003. Genomic Clocks and Evolutionary Timescales. *Trends in Genetics* 19(4): 200–206.

Hughes A. L., and Hughes M. K. 1995. Small Genomes for Better Flyers. *Nature* 377(6548): 391.

Ji Q., Luo Z.-X., Yuan C.-X. *et al.* 2002. The Earliest Known Eutherian Mammal. *Nature* 416(6883): 816–822.

Kapitsa S. P. 1992. A Mathematical Model of the World Population Growth. *Matematicheskoye modelirovaniye* 4(6): 65–79. *In Russian* (Капица С. П. Математическая модель роста народонаселения мира. *Математическое моделирование* 4(6): 65–79).

Kapitsa S. P. 1999. *General Theory of the Growth of Humankind: How Many People Have Lived, Live, and Are to Live on Earth*. Moscow: Nauka. *In Russian* (Капица С. П. *Общая теория роста человечества: Сколько людей жило, живет и будет жить на Земле.* М.: Наука).

Kolchanov N. A., Anan'ko E. A., Kolpakov F. A. *et al.* 2000. Gene Networks. *Molecular Biology* 34(4): 533–544, 449–460.

Kolchanov N. A., Suslov V. V., and Shumnyi V. K. 2003. Molecular Evolution of Genetic Systems. *Paleontological Journal* 6: 58–71; 37(6): 617–629.

Kolchanov N. A., Suslov V. V., and Gunbin K. V. 2004. Biological Evolution Simulation: Genetic Regulatory Systems and Biological Complexity Encoding. *Vestnik VOGiS* 8(2): 86–89.

Korotayev A. V. 2006. The World System History Periodization and Mathematical Models of Socio-Historical Processes. *History and Mathematics: The Problems of Periodization of Historical Macroprocesses* / Ed. by L. E. Grinin, A. V. Korotayev, and S. Yu. Malkov, pp. 116–167. Moscow: KomKniga/URSS. *In Russian* (Коротаев А. В. Периодизация мировой системы и математические модели социально-исторических процессов. *История и математика: проблемы периодизации исторических макропроцессов* / Ред. Л. Е. Гринин, А. В. Коротаев, С. Ю. Малков, с. 116–167. М.: КомКнига / URSS).

Korotayev A. V. 2007. The World System Urbanization Macrodynamics: A Quantitative Analysis. *History and Mathematics: Macrohistorical Dynamics of the Society and State* / Ed. by S. Yu. Malkov, L. E. Grinin, and A.V. Korotayev, pp. 21–39. Moscow: KomKniga/URSS. *In Russian* (Коротаев А. В. Макродинамика урбанизации Мир-системы: количественный анализ. *История и математика: макроисторическая динамика общества и государства* / Ред. С. Ю. Малков, Л. Гринин, А. В. Коротаев, с. 21–39. М.: КомКнига / УРСС).

Korotayev A. V., and Grinin L. E. 2007. Urbanization and Political Development of the World System: Comparative Quantitative Analysis. *History and Mathematics: Macrohistorical Dynamics of the Society and State* / Ed. by S. Yu. Malkov, L. E. Grinin, and A. V. Korotayev, pp. 102–141. Moscow: KomKniga/URSS. *In Russian* (Коротаев А. В., Гринин Л. Е. Урбанизация и политическое развитие Мир-Системы: сравнительный количественный анализ. *История и математика: Макроисторическая динамика общества и государства* / Ред. С. Ю. Малков, Л. Е. Гринин, А. В. Коротаев, с. 102–141. М.: КомКнига / URSS).

Korotayev A. V., Malkov A., and Khaltourina D. 2006. *Introduction to Social Macrodynamics: Compact Macromodels of the World System Growth.* Moscow: URSS.

Korotayev A. V., Komarova N. L., and Khaltourina D. A. 2007. *Laws of History. Secular Cycles and Millennial Trends. Demography. Economics. War*. Moscow: URSS/KomKniga. *In Russian* (Коротаев А. В., Комарова Н. Л., Халтурина Д. А. *Законы истории. Светские циклы и тысячелетние тенденции. Демография. Экономика. Война*. М.: УРСС / КомКнига).

Korotayev A. V., Malkov A. S., and Khaltourina D. A. 2007. *Laws of History: Mathematical Simulation of the Development of the World System: Demography, Economics, Culture.* Moscow: URSS/KomKniga. *In Russian* (Коротаев А. В., Малков А. С., Халтурина Д. А. *Законы истории: Математическое моделирование развития мировой системы: Демография, экономика, культура*. М.: УРСС / КомКнига).

Lin R., Ding L., Casola C. *et al.* 2007. Transposase-Derived Transcription Factors Regulate Light Signaling in Arabidopsis. *Science* 318(5854): 1302–1305.

Lynch M., and Conery J. S. 2003. The Origins of Genome Complexity. *Science* 302(5649): 1401–1404.

Markov A. V., and Kulikov A. M. 2005. Origin of Eukaryota: Conclusions Based on the Analysis of Protein Homologies in the Three Superkingdoms. *Paleontological Journal* 4: 3–18; 39(4): 345–357.

Markov A. V., and Korotayev A. V. 2008. Hyperbolic Growth of Marine and Continental Biodiversity through the Phanerozoic and Community Evolution. *Zhurnal obshchey biologii* 69(3): 175–194. *In Russian* (Марков А. В., Коротаев А. В. Гиперболический рост разнообразия морской и континентальной биот фанерозоя и эволюция сообществ. *Журнал общей биологии* 69(3): 175–194).

Markov A. V., Korotayev A. V. 2009. *Hyperbolic Increase in Living Nature and Society*. Moscow: URSS. *In Russian* (Марков А. В., Коротаев А. В. *Гиперболический рост в живой природе и обществе*. М.: URSS).

Marshall C., and Schultze H.-P. 1992. Relative Importance of Molecular, Neontological, and Paleontological Data in Understanding the Biology of the Vertebrate Invasion of Land. *Journal of Molecular Evolution* 35(2): 93–101.

Mikkelsen T. S., Wakefield M. J., Aken B. *et al.* 2007. Genome of the Marsupial Monodelphis Domestica Reveals Innovation in Non-Coding Sequences. *Nature* 447(7141): 167–177.

Miller W. J., McDonald J. F., Nouaud D., and Anxolabehere D. 1999. Molecular Domestication – More than a Sporadic Episode in Evolution. *Genetica* 107(1–3): 197–207.

Muotri A. R., Marchetto M. C., Coufal N. G., and Gage F. H. 2007. The Necessary Junk: New Functions for Transposable Elements. *Human Molecular Genetics* 16 (Spec. 2), R159–R167.

Nakabachi A., Yamashita A., Toh H. *et al.* 2006. The 160 – Kilobase Genome of the Bacterial Endosymbiont Carsonella. *Science* 314(5797): 267.

Ochman H. 2005. Genomes on the Shrink. *Proceedings of the National Academy of Sciences USA* 102(34): 11959–11960.

Ofria C., Adami C., and Collier T. C. 2003. Selective Pressures on Genomes in Molecular Evolution. *Journal of Theoretical Biology* 222(4): 477–483.

Organ C. L., Shedlock A. M., Meade A. *et al.* 2007. Origin of Avian Genome Size and Structure in Non-Avian Dinosaurs. *Nature* 446(7132): 180–184.

Pan Q., Shai O., Lee L. J. *et al.* 2008. Deep Surveying of Alternative Splicing Complexity in the Human Transcriptome by High-Throughput Sequencing. *Natural Genetics* 40(12): 1413–1415.

Patthy L. 1999. Genome Evolution and the Evolution of Exon-Shuffling – a Review. *Gene* 238(1): 103–114.

Payne J. L., Boyer A. G., Brown J. H. *et al.* 2009. Two-Phase Increase in the Maximum Size of Life over 3.5 Billion Years Reflects Biological Innovation and Environmental Opportunity. *Proceedings of the National Academy of Sciences, USA* 106(1): 24–27.

Pellicciari C., Formenti D., Redi C. A., and Manfredi Romanini M. G. 1982. DNA Content Variability in Primates. *Journal of Human Evolution* 11(2): 131–141.

Putnam N. H., Butts T., Ferrier D. E. K. *et al.* 2008. The Amphioxus Genome and the Evolution of the Chordate Karyotype. *Nature* 453(71198): 1064–1071.

Raff R. A., and Kaufman T. C. 1986. *Embryos, Genes, and Evolution: The Developmental Genetic Basis of Evolutionary Change*. New York: Macmillan 1983; Moscow: Mir.

Rasmussen B., Fletcher I. R., Brocks J. J., and Kilburn M. R. 2008. Reassessing the First Appearance of Eukaryotes and Cyanobacteria. *Nature* 455(7216): 1101–1104.

Rozanov A. Yu. 2003. Fossil Bacteria, Sedimentogenesis, and the Early Biospheric Evolution. *Paleontological Journal* 6: 41–49; 37(6): 600–608.

Sharov A. A. 2006. Genome Increase as a Clock for the Origin and Evolution of Life. *Biology Direct* 1: 17.

Shu D. G., Luo H-L., Conway M. S. *et al.* 1999. Lower Cambrian Vertebrates from South China. *Nature* 402(6757): 42–46.

Stegemann S., Hartmann S., Ruf S., and Bock R. 2003. High Frequency Gene Transfer from the Chloroplast Genome to the Nucleus. *Proceedings of the National Academy of Sciences USA* 100(15): 8828–8833.

Takahashi K., Hayashi N., Shimokawa T. *et al.* 2008. Cooperative Regulation of Fc Receptor Gamma-Chain Gene Expression by Multiple Transcription Factors, Including Sp1, GABP, and Elf-1. *Journal of Biological Chemistry* 283(22): 15134–15141.

Thomas C. A. J. 1971. The Genetic Organization of Chromosomes. *Annual Review of Genetics* 5: 237–256.

Volff J. N. 2006. Turning Junk into Gold: Domestication of Transposable Elements and the Creation of New Genes in Eukaryotes. *Bioessays* 28(9): 913–922.

Wang E. T., Sandberg R., Luo S. *et al.* 2008. Alternative Isoform Regulation in Human Tissue Transcriptomes. *Nature* 456(7221): 470–476.

Xing J., Witherspoon D. J., Ray D. A. *et al.* 2007. Mobile DNA Elements in Primate and Human Evolution. *Yearbook of Physical Anthropology* (Suppl. 45): 2–19.

9
A Mathematical Model of Influence of the Interaction between Civilization Center and Barbarian Periphery on the World System Development[*]

Andrey V. Korotayev and Leonid E. Grinin

Abstract

This article offers an analysis and mathematical modeling of the influence of one of the major factors of the World System macrodynamics throughout most part of its history (since the 'urban revolution') – the factor of interaction of civilizations with their barbarian periphery. The proposed mathematical model is intended to describe possible influence of interaction between civilizational core of the World System and its barbarian periphery on the formation of the specific curve of the world urbanization dynamics. It simulates completion of the phase transition, behavior of the system in the attraction basin and beginning of the phase transition to the attraction basin of the new attractor and is aimed to identify the role of the factor of interaction between the civilizational core and barbarian periphery in the formation of attractor effect during the completion of phase transition, that is for clarification of the reason why there was observed not only slowdown of growth rates of the main indicators of the World System development after completion of phase transitions during its development, but also their falling with the subsequent temporary stabilization near some equilibrium level. Achievements of modern barbarology, including the understanding of complexity of the barbarian periphery itself and its heterogeneity are considered. The basic principle of the proposed dynamic model is that sizes, power and level of complexity in realization of external policy functions in nomadic unions (empires) closely correspond to sizes, power and level of political culture and activity of the core states with which nomads constantly had to do (this point has been established in works of the known experts in nomadic studies). Various alternatives are shown in the model, when depending on power and size of one of the two components of the

[*] The research is supported by the Russian Foundation for Basic Research (Project No. 17-06-00464).

system 'civilization – barbarian periphery' studied by us, another one also changes significantly as it has to respond to the challenge properly, or can make less efforts feeling no threat or resistance. This principle is observed throughout the long period of the history of the World System. It is shown that interaction between the civilizational center and barbarian periphery really can explain some characteristic features of the World System dynamics in the 4th millennium BCE – the 2nd millennium CE. The ways of further development of the model are outlined.

Keywords: *dynamic modeling, the World System, civilizations, barbarian periphery, urbanization, asabiyyah, technology, warfare.*

The emergence and development of the world urban network was one of the main components of the World System's evolution that accelerated its development and increased its integration. Not without reason V. Gordon Childe focused on the urban revolution (Childe 1952: Chapter 7; Childe 1956). It is also quite clear that the processes of functional differentiation, social stratification and class formation proceeded in many ancient agricultural societies under a considerable influence of the 'urban revolution' (Alekshin 1986: 22). The city also implies a complex concentration of geographical, social, political, and sacral resources and assets. 'The city is a direct territorial concentration of multiple heterogeneous forms of human activities' (Akhiezer 1995: 23). One can see the closest connection between urbanization, on the one hand, and the formation and development of civilizations and statehood, on the other.[1]

As is known, relatively large pioneer settlements (like Jericho in Palestine) vaguely resembling cities, emerged more than 9,000 years ago. In particular, about 7200 BCE Jericho was surrounded by a stone wall three meters thick, and four meters high (Lamberg-Karlovsky and Sablov 1992: 75). In the 7th – 6th millennia BCE, a number of settlements with estimated population of about 2,000 people appeared in Western Asia (Ain Ghazal, Beisamoun, Beida, Abu Hureira, Çatal Hüyük). However, it is undisputable that the first real cities appeared only in the 5th – 4th millennia BCE. And, finally, the first period of a rapid urban growth within the World System occurred between the second half of the 4th and the first half of the 3rd millennium BCE.

Our previous research (Korotayev 2006c, 2007a, 2007b; Korotayev, Komarova, and Khaltourina 2007: 169–177; Korotayev and Grinin 2012, 2013; Grinin 2017a, 2017b) has shown that the curve describing the dynamics of the world urban population has a rather peculiar form (see Figs 1 and 2).

---

[1] See, *e.g.*, Korotayev and Grinin 2006, 2007; Grinin 2007a; Masson 1989. This situation was typical for many regions: Ancient Greece (Gluskina 1983: 36; see also Frolov 1986: 44; Andreev 1979: 20–21); Mesopotamia, in particular in the late 4th millennium and the 3rd millennium BCE (Dyakonov 2000: 46), a number of African regions; *e.g.*, in South East Madagascar in the 17th century a few small Betsileo states originated in this way (Kottak 1980; Claessen 2000, 2002, 2004).

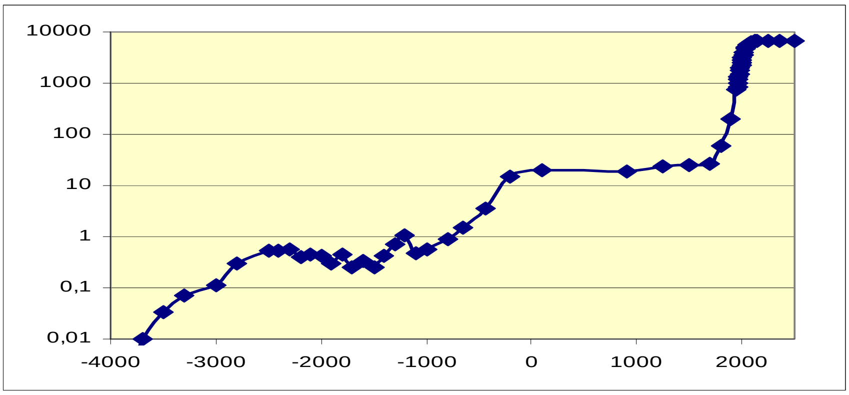


Fig. 1. The World Urban Population Dynamics (in millions), for cities with > 10,000 inhabitants (logarithmic scale)

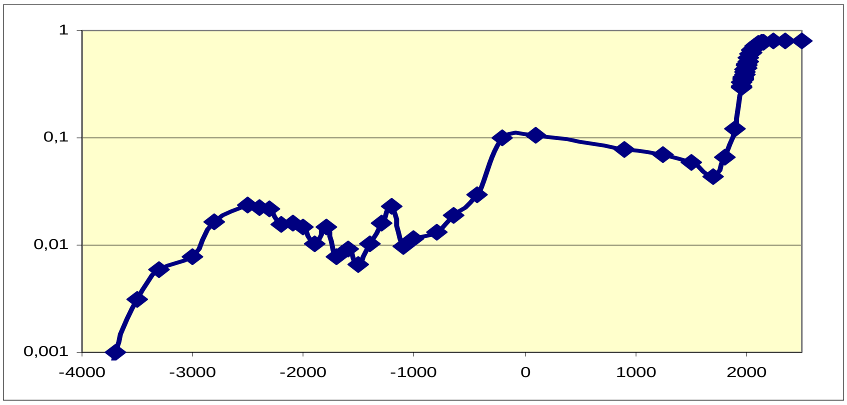


Fig. 2. Dynamics of the World Urbanization Index (proportion of population living in the cities with population more than 10,000 inhabitants in the in the overall population of the world, with a projection of modern trends (logarithmic scale)

As can be seen, one can single out here three rather distinct periods of relatively fast world urban population growth: (A1) from the mid-4th millennium BCE to the mid-3rd millennium BCE, (A2) the 1st millennium BCE (roughly corresponding to the 'Axial Age'); and (A3) the 19th – 21st centuries CE. Moreover, one can note two periods of relatively slow growth of the world urban population (including long phases when the urban population and the world urbanization level would hardly grow or could even considerably fall): (B1) from the mid-3rd millennium BCE to the late 2nd millennium BCE and (B2) between the 1st and 18th centuries CE. Two other periods turn out to be essentially close to these epochs: Period (B0) immediately preceding the mid-4th millennium (when the world urban population did not grow simply because cities had not appeared yet and no cities existed on the Earth), and Period (B3) that is expected to begin in the 22nd century, when, according to forecasts, the world urban population will again stop growing in any significant way (since the World System urbanization is supposed to reach the saturation level along with the stabilization of the world population) (see, *e.g.*, Korotayev, Malkov, and Khaltourina 2006a, 2006b, 2007; Korotayev, Komarova, and Khaltourina 2007; Korotayev 2008, 2009, 2012, 2013; Grinin 2006a, 2006b).

Note that the detected world urbanization dynamics correlates rather well with the dynamics of the World System political organization (Grinin and Korotayev 2006, 2007; Korotayev and Grinin 2007, 2012, 2013; Grinin 2016a, 2016b; Grinin, Ilyin, and Andreev 2016). Moreover, the above mentioned synchronous phase transitions to the new levels of the world urbanization and new complexity levels of the World System political organization temporally coincide with phase transitions to higher levels of the World System political centralization that were detected by Taagepera and that took place, according to his calculations, during periods A1, A2 and A3 (Taagepera 1997: 485).

Similar phase transitions appear to be observed with respect to the world literacy macrodynamics. In fact, during Period A1 we observe the emergence of the first literate people whose share in the world population by the end of this period had reached the level of decimals of a percent and fluctuated at this level during Period B1. And it is no coincidence. Although literate people could live not only in cities, nevertheless, their number in cities was incomparably larger than in rural areas. During Period A2, the world literacy rate grew by an order of magnitude and amounted to several percent of the world total population; it fluctuated at this level during Period B2 till the late 18th century when Period A3 started. During that period the world literacy level reached the level of dozens percent, and by the beginning of Period B3 (presumably in the 22nd century) it is likely to stabilize at the hundred-percent level (see, *e.g.*, Korotayev 2006d; Korotayev, Malkov, and Khaltourina 2006a, 2006b).

In fact, the above-mentioned phase transitions can be regarded as different aspects of a series of unified phase transitions: Phase Transition A1 from medium complexity agrarian societies to complex agrarian ones, Phase Transition A2 from complex agrarian societies to supercomplex ones, and, finally, Phase Transition A3 from supercomplex agrarian societies to postindustrial ones (within this perspective, the period of industrial societies turns out to be a period of phase transition B2 – B3). These phase transitions are also exceptionally strongly connected with production revolutions and transitions from one principle of production to another (see in more detail: Grinin 2003a, 2007a, 2007b; Korotayev and Grinin 2006, 2007). The period of the first attractor (the first phase), in particular, is connected with the first variant of the intensive phase of agrarian revolution (transition to irrigation agriculture); the second attractor/phase – with the second variant of the intensive phase of agrarian revolution (transition to intensive plow non-irrigation agriculture). From the 16th to the first half of the 20th century (especially the 19th – the first half of the 20th century) the phase transition was connected with the transition to the industrial principle of production. The period from the end of the 20th century and (presumably) the whole 21st century is connected with the transition to scientific and information/cybernetic principle of production (for more details see Grinin 2006a, 2006b; Grinin L., Grinin A., and Korotayev 2017a, 2017b).

The proposed mathematical model aims at analyzing the possible impact of the interaction between the World System's civilizational core and its barbarian[2] periphery on the formation of a specific curve of the world urbanization dynamics. It describes the completion of the phase transition, system's behavior in the basin of attraction, and the start of the phase transition to a new basin of attraction. This model also identifies the role of the interaction between civilizational core and barbarian periphery in the formation of the attractor effect during the completion of phase transition. In other words, we try to find out why after the phase transitions were completed there was observed not only the slow-down in the growth rates of the main indicators of the World System development but also their decline with subsequent temporary stabilization at some equilibrium level (let us note that the offered model cannot describe the fluctuations observed at the respective levels).

The issues of coexistence, interaction and struggle between civilization and barbarian periphery are extremely important for understanding of the evolution of the World System over the last five thousand years after the emergence of the first states and civilizations. This also remained relevant up to some extent for the Modern Period up until very recent times. In some regions, like, for example, the Middle East and North Africa, the non-state tribal and chiefdom forms of political organization continued coexisting, competing or cooperating in the military field with the states up to the early 20th century. In particular, even in the territory of Egypt, one of the most ancient civilizations, in the second half of the 18th century the Bedouin raids were a great threat for sedentary populations (as evidenced by the famous Egyptian historical chronicler of the late 18th and the early 19th centuries Abd ar-Rahman al-Jabarti [1978]). In the middle of the 20th century within the territory of another most ancient, Chinese, civilization one could find a true internal barbarian periphery which made regu-

[2] Needless to say that, following Lewis H. Morgan (1877) tradition, we use the term 'barbarian' as a purely technical one – devoid of any pejorative implications (note that Morgan himself in no way despised the 'barbarians' – he – in company with Friedrich Engels [1884/1978] – rather admired them).

lar incursions into the center.[3] One can recollect a similar situation in the Caucasus in the 19th and even in the 20th centuries.[4]

As is known, there are great differences in the definition of civilization (see, *e.g*., Grinin 1997, 1998a, 1998b, 1998c; Grinin and Korotayev 2008: Introduction). In this paper we operationalize civilization or civilizational core (center) of the World System as the societies of the World System core with urban settlements; while the peripheral communities without urban settlements are defined as 'barbarian'. Within the framework of the present mathematical model the existence of cities is assumed to be the only formal characteristic of civilization.[5]

As Vera P. Budanova (2002: 168) notes, there are some directions in modern barbarology which study general relations between barbarian world and civilization (see, *e.g*., Budanova 1990, 1994, 2000, 2002; Masson 1986, 1989; Barfield 2006, 1991; Kradin 1992, 2001a, 2001b; Kradin and Bondarenko 2002; Pershits and Khazanov 1978; Khazanov 1975, 2002, 2006; Sannikov 2002, 2003, 2005; Kradin, Bondarenko, and Barfield 2003; Grinin and Korotayev 2013, 2014, 2018). Nevertheless, it is a very broad subject, and many of its aspects are hardly well scrutinized. Still, the following conclusions drawn by barbarologists are especially important for our subject: 1) the center and barbarian periphery are considered as closely related elements of a single pan-

[3] *E.g.*, the *Yi* people (or *Nuosu people*) in the high mountain region of Liangshan of China's Sichuan province. There were four 'classes' in this society, one of which (actually [*Nuosu*] – 'black') contrary to subordinate 'whites' was superior, noble, and therefore it did not participate in productive labor. The rest three classes were in different degree of dependence – from semi-bond to slavish against the background of the absence of any developed political structure (Its and Yakovlev 1967; Kubbel' 1988: 241–242). Such a situation developed from the 7th – 9th centuries CE after the cattle breeding tribes had subordinated farmers (Its and Yakovlev 1967: 79). Slavery was widespread in this society. Herewith the Nuosu often attacked and captured the Han people, reducing them to slavery. Thus, in 1919 the Liangshan Yi people captured and took away more than 10,000 people from neighboring counties to their highlands. In the early 19th century the total population of the Liangshan Yi was rather small, numbering about 10,000 people. But in 1838 it amounted already 40–50,000, and in 1910 – about 200–300,000. It continued to increase, having reached 630,000 people in the mid-1950s, among whom non-assimilated Han slaves totaled 50,000–60,000 (*Ibid.:* 79–80).

[4] *E.g.*, the General Anton Denikin in his *History of the Civil Strife in Russia* (Denikin 1993: 122) speaks about the Ingush as the most organized among the Caucasian peoples who took advantage of the anarchy during the Civil War and systematically plundered and terrorized all the neighbors. In fact, the most recent events in the Caucasus in Russia still remind us 'living vestiges' of such relationships between civilization and barbarian peripheries.

[5] The identification of this characteristic as a working criterion of civilization within our mathematical model should, of course, be treated just as an assumption and is explained by the necessity to determine in an operationalizable way within the present formal/mathematical research the societies forming the World-System core and having urban settlements as distinct from peripheral societies (designated here as 'barbarous') lacking those settlements. Let us note that it does not contradict some researchers' rather fair statements that in the context of their research this characteristic can be substituted, *e.g.*, by the presence of monumental buildings (see, *e.g.*, Masson 1989).

Oecumene system (= the World System) in which peoples with different levels of socio-cultural complexity interact (see, *e.g.*, Pershits and Khazanov 1978: 4; Budanova 2002: 168); 2) within the barbarian periphery itself a certain center[6] – a 'core' of the barbarian world – can be formed in relation to civilization, which in many respects defines relationships between civilization and barbarians and with that part of barbarians who inhabited territories that were distant from the civilization core (the emergence of such centers, as a result of the civilizational core pressure quite often led to the growth of collective solidarity [*asabiyyah*] of barbarians that we try to account for in our model); 3) the complexity level of the barbarian alliances (especially among the nomads) closely corresponds to the size and level of political culture of states with which they were in contact (see, *e.g*., Barfield 2006); and 4) foreign policy and economic (trade) interests played a significant role in the relations between barbarian world and civilization, however, military contacts prevailed there (Budanova 2000, 2002; Kradin 1992, 2001a; Barfield 1991, 2006).

Elsewhere we have already analyzed the possible role of interaction between the civilizational core and barbarian periphery (see Korotayev, Malkov, and Khaltourina 2007: 189–208; Grinin 2003b, 2004a, 2004b, 2011; Grinin and Korotayev 2013, 2014, 2018; Grinin *et al.* 2004, 2006) and considered the reasons of an essential decline (up to negative values) of the growth rates of the main indicators of the World System development in the 1st millennium CE after the completion of A2 phase transition to supercomplex agrarian societies (Korotayev, Malkov, and Khaltourina 2006b; Zinkina, Ilyin, and Korotayev 2017). Note that the above-mentioned analysis allowed identifying that factor as one of very important causes (but not the only one) of the considered phenomenon. Thus, we have come to the following preliminary conclusions:

> The fact that the regime of hyperbolic growth changed after the World System's political centralization had reached critically high level of hyperbolic rates (in the early 1st millennium CE the absolute majority of World System's inhabitants turned out to be under control of only four empires – Roman, Parthian, Kushan and Han) is not accidental also for some other reasons. The rapid growth of political centralization in the 1st millennium BCE was driven by the diffusion of iron metallurgy (for more details see Grinin and Korotayev 2008: Ch. 6; Korotayev and Zinkina 2017; Zinkina, Ilyin, and Korotayev 2017), which not only considerably increased the Earth's carrying capacity, but also led to the development of production of rather cheap and effective weapons which promoted the formation of numerous armies without which the emergence of the world empires would be almost impossible. However, this process had important side effects. The politically centralized systems

[6] At the same time in those barbarian polities such leaders would constantly change (Budanova 2002).

> quite often secure military superiority through the development of specialized military subsystems – rather small but well trained and professional armies. However, to preserve this superiority there is necessary to have monopoly on certain effective types of weapons (war chariots, bronze weapons, *etc*.). If the revolution in production of means of violence takes place and the monopoly on them cannot be efficiently supported (*e.g.*, in case of emergence of iron weapons), the less politically centralized societies with a high proportion of military active population get considerable advantage and in military terms can become stronger than politically centralized societies. This was the case in many parts of Oecumene of the Old World in late antiquity. Moreover, less politically centralized societies with a greater share of military active population could considerably increase their military efficiency without noticeable increase in their political centralization or internal differentiation, for example, through nomadization, growth of specialization on cattle breeding since the herder's everyday work and the character of his socialization make him a combat-effective warrior. Nomadic cattle breeding with a widespread use of herders-riders could considerably increase military potential of such societies without additional political centralization and functional differentiation. In this context it is important for us that the side effect of the technological shifts of the first millennium BCE was strengthening of the barbarian periphery's military potential in general and nomadic socio-political systems, in particular... As a result, the nomads got a consistent military superiority over the settled societies throughout most part of the 'Junior Hyperbole' epoch (additionally strengthened by the invention and diffusion of stirrups and sabers); this led to an additional reduction in the World System's demographic growth rates not only due to mass depopulations resulting from recurring nomadic invasions, but also as a result of some decrease in the Earth's carrying capacity in many important zones of the World System due to the pressure of barbarian (and, in particular, nomadic) peripheries (here we could recollect the Russian 'bread-basket' – Black Earth region which through the most part of the 2^{nd} millennium was known as the Wild Field since the lands in this region were almost not cultivated because of the threat of nomadic raids) (Korotayev, Malkov, and Khaltourina 2007: 207–208).

Let us note that a systematic military superiority at a certain phase of the World System evolution does not mean a constant superiority. China, for example, defeated the Xiongnu a few times, carrying out deep raids in their lands (see, *e.g*., Gumilev 1993; Kradin 2001a), similar as the Russian dukes did with respect to the lands of Cumans (see, *e.g.*, Rybakov 1966a: 561–562). Therefore, we actually deal with an unsteady balance of forces between the barbarian periphery and civilizational center but this balance could change under some circumstances. In a certain situation, the barbarian periphery's pressure would come over civilization or, *vice versa*, civilization would invade the barbarian periphery. Thus, *one can speak about certain cycles when phases of civilizational center's expansion and the barbarian periphery retreat are alternated by the phases of barbarian periphery expansion and civilizational center's retreat.*

It is worth mentioning that depending on the power and scale of one of the two components of 'civilization – barbarian periphery' system, the other element would also significantly change to give an adequate response to the amplified (or changing in some other way) challenge; otherwise, feeling no threat or resistance it can make less efforts. Anyway, it was noticed, that in nomadic alliances (empires) the size, power and complexity level of foreign policy functions correlated closely with the size, power and level of political culture and activity of the states with which the nomads constantly interacted (see, *e.g.*, Barfield 2006: 429). This can also explain the situation described below in the model, when a civilization expands to the barbarian periphery rather actively, while the latter is unable to actively resist the former. It may happen because the barbarian periphery turns unable to adapt to the power and size of the advancing civilization yet. After absorbing the part of the barbarian periphery which is less capable to resist, especially the territories with environment, suitable for the civilization's economic expansion (and with peoples who are somehow ready to become a part of civilization), civilization can face more persistent representatives of the periphery especially those living under marginal conditions. As a result, the above-mentioned dynamic equilibrium can be established for certain periods (sometimes for a rather long time).

Although a rather long coexistence of civilization and barbarian periphery is obvious, each part of this dynamic system tries to weaken or even destroy the other at every opportunity, so there emerges a situation of 'interdiffusion' when various innovations are borrowed (mostly by the barbarians from civilization, but sometimes *vice versa*) and also civilization uses the barbarians for its own needs. As a result one can observe an accelerated development of the barbarian periphery which in order to have advantages and resist civilization tries to develop similar political and social forms. This usually aims primarily at achieving a military balance or military superiority, and also at achieving the parity of prestige. This also involves ideology which can be rather developed among the barbarians. The latter is important for understanding of the *asabiyyah* concept which was developed by an outstanding medieval Arab thinker Abd ar-Rahman Ibn Khaldūn as a scientific category (Ibn Khaldūn 1958, 2004) and introduced into the scientific vocabulary of modern Cliodynamics by Peter V. Turchin (2003, 2007) who, in our opinion, quite reasonably interprets this concept as 'collective solidarity'.

It is a peculiar ideology of tribal solidarity, which allows uniting barbarian people into powerful military force, for example, when putting together groups of tribes. Therefore, Morton Fried (1967) has reason to state that tribes are the secondary non-primitive formations emerging under the influence of neighboring communities with significantly higher level of sociocultural complexity (see also Korotayev 1997a, 1997b, 1998, 2000a, 2000b, 2003, 2004, 2006b, 2006c; Grinin 2007a). In fact, many analogue forms of polities of the 'main sequence' are often secondary phenomena associated with the impact of civilizational center, or their development is significantly modified being effected by more developed neighbors (see Grinin 2007a). For example, such a modification might have happened in the development of the Scythian polity since the Scythians actively communicated with the Medians and Persians, and later with Greeks (see, *e.g.*, Dyakonov 1956; Khazanov 1975). Moreover, such forms quite often emerge just because they best fit the marginal environment, while civilizations, as a rule, emerge in the environment more favorable for the development of intensive production. It is natural that the analogues in barbarian periphery possessing certain environmental, economic and demographic features could get along without cities. Only some barbarians had a developed system of cities as it was in Gaul where there were up to a thousand of 'genuine cities', and in some of them population reached tens thousands people (Shkunaev 1989: 134, 143). The size of some cities was 100 and more hectares, and they were secured by powerful walls (see Filip 1961: 116–129; Mongait 1974: 248–253).

Thus, due to interaction of different kind between civilization and barbarian periphery: a) the World System expanded and became more and more complex[7]; b) the socio-political, economic and cultural level of the barbarian periphery generally increased; c) the civilizational level, including urbanization, could temporarily decrease due to generally increasing size of the World System, and temporal 'barbarization' of extensive territories as it was repeatedly observed in the 1st millennium CE (especially in Europe).[8] One can apply here Adolf Leo Oppenheim's idea (1990: 88) about constant counteraction between anti- and pro-urbanistic trends in ancient Mesopotamia and in the ancient world in general, while the barbarian periphery was the most important agent of the former trend.

It is also worth mentioning that when speaking about each certain barbarian onslaught towards the civilizational zone, as well as about definite periods of such mass movement, we can hardly know the exact reasons that launched such migrations. For example, Budanova (2000: 5–6) writes that there is still no definite answer to the question what triggered the migration engaging territories

---

[7] This may have involved the integration of a number of civilized societies by barbarian conquerors. We think that the Mongolian amalgamation is one of the most significant in this context since for a certain period it strengthened the relations within the World System from the Pacific Ocean to the Atlantic (see, *e.g.*, Abu-Lughod 1989, 1990).

[8] One can also assume that the World System expansion rate could be sometimes inversely proportional to the rate of quality growth of its particular parts and processes (such as urbanization).

from Scandza to Mauritania, from China to the Pyrenees in the Great Migration Period (the 3rd – 7th centuries CE).[9]

Here we present the model which is founded on the above-described ideas, and also on our earlier general models of development of the World System (Korotayev 2005, 2006c, 2006d, 2007a, 2008, 2009, 2012, 2013; Korotayev, Malkov, and Khaltourina 2006a, 2007; Korotayev and Malkov 2012; Korotayev and Zinkina 2017; Grinin 2006a, 2006b, 2007a, 2010; Grinin and Grinin 2015, 2016) and some ideas of the theory of dynamics of community solidarity (*asabiyyah*) formulated by Peter Turchin (2003, 2005, 2007).

In the proposed model the World System is assumed to be divided into three main geographical zones: (1) small (1 mln km^2) and highly productive zone; (2) a larger size zone with average producing capacity (24 mln km^2) which surrounds Zone 1; and (3) the largest in size (96 mln km^2) and the least productive zone surrounding Zone 2 (see Fig. 3).

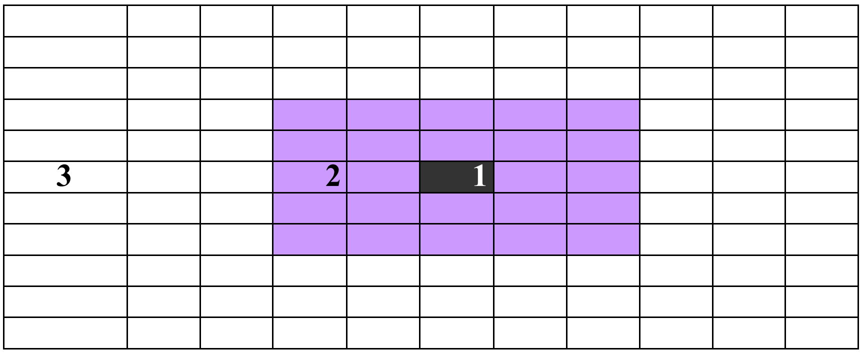


Fig. 3. Spatial structure of the World System assumed in the model

It can be assumed that the first cities originated in Zone 1 (see, *e.g.*, Korotayev and Grinin 2006, 2012, 2013; Grinin and Korotayev 2009a, 2009b) which therefore, can be identified as the 'civilizational center'. It is assumed that the initial level of technological development in this center (T_{c0}) is significantly higher than that for the barbarian periphery (T_{b0}) coinciding with Zone 2 at the start of computer simulation. At this point, Zone 3 with the lowest initial level of technological development (T_{h0}) is considered as the World System's hinterland.

At the first stage of computer simulation the model's major scenario describes the initial vigorous territorial expansion of the civilizational center supported by its more developed technologies, which in combination with signifi-

[9] However, the possible general reasons of the barbarian periphery pressure on civilizational center are quite clear: the demographic pressure, the shortage of resources (land, pastures) correlated with this and other (natural in the first turn) factors; aspiration to the spoils of war; and pressure of enemies (*i.e.*, conflicts within the barbarian periphery) and other similar factors.

cantly denser population of the civilization zone results in a significantly higher military potential. In the proposed model, the civilization's territorial dynamics is mathematically described by means of the following differential equation:

$$\frac{dA_c}{dt} = a(M_c - M_b), \qquad \text{(Eq. 1)}$$

where A_c is the territory controlled by the civilizational core; M_c is the military potential of the civilizational core; M_b is the military potential of the barbarian periphery; a is the constant which determines the rate of transformation of military superiority into territorial acquisitions (the calculation pattern for M_c and M_b values will be described below, see Eqs 2 and 3).[10]

However, after a while this expansion is exhausted in the major scenario of the model and the barbarian periphery's counterattack unfolds.[11] Note that in the suggested model (as well as in historical reality) less numerous and technologically backward barbarians can put pressure on more numerous and technologically advanced 'civilized' enemies. This effect may be produced by the following factors:

1) A higher military participation ratio that was characteristic of the barbarians. It is proved by written, ethnographic and even archaeological sources. For example, in some territories occupied by the German tribes before the Great Migration epoch about 80 % of males were buried with iron weapons (see Gurevich 1999: 44). One should also mention the early military training for boys among many barbarian (especially nomadic) peoples, for example, among Huns, Mongols or Turks when they were nomads (see, *e.g.*, Nefedov 2008[12]).

---

[10] In real history, it could be just the result of demographic pressure of migrants who would absorb numerically insignificant aboriginals, or the result of combination of demographic dissipation and military superiority. Thus, many barbarians just disappeared as ethno-social entities and were assimilated by civilized peoples. But those who survived became ethnicities of a new generation capable of both military and cultural opposition, and development of their own complex political systems which led to creation of analogues of the state among barbarians (see, *e.g.*, Grinin 2001–2006, 2003b, 2004b, 2007a).

[11] In real history it was most often observed that civilization reached the limits of the natural zones, suitable for its economic pattern (and apparently, it would be worth taking into account this effect in future mathematical models). Note, also, that 'barbarous counterattack' in reality might start both after the period of established power balance, and sometimes at once without intermediate period of balance. Khan Konchak's campaign against the Russians in 1185 following Duke Igor Novgorod-Seversky's unsuccessful campaign on the Cumans can serve a classic example here. Moreover, the Cuman troops moved toward Rus' in three directions: toward deserted Igor and Vsevolod's principalities, toward Pereyaslavl and Kiev, 'where Konchak was attracted by the memories of Khan Bonyak knocking with sabre on Kiev's Golden Gate' (Rybakov 1966b: 595).

[12] 'Training of the Turkish archers does not appear to have been inferior to Mongolian archers, and similar to the Mongols, constant trainings also promoted the development of muscles of arms. 'From eight, or even seven years old they began to shoot at a target, – the imperial ambassador Ghiselin de Bousbecq wrote, – and for ten or twelve years they would practice in archery. This continuous training strengthened muscles of their arms and gave them such a skill that they could hit the smallest targets with their arrows' (Nefedov 2008).

The most important role of this factor in explaining the cases of successful advance of the barbarian periphery on civilizational center was described in the 16th century by the Ethiopian monk Bahrey in his well-known *History of the Galla*. Bahrey tried to explain why the politically centralized Ethiopian state was constantly defeated by the politically less centralized and less developed Galla (Oromo) tribes ('How is it that the Galla defeat us though we are numerous and well supplied with arms?' Bahrey 1976 [1593]: 140). The answer which Bahrey proposed is very interesting and convincing: just because the Ethiopian society was much more developed and socially differentiated (*i.e.*, actually more 'civilized'), it suffered continuous defeats in the fight against less developed 'barbarians', the Galla. In this case the high level of internal differentiation ('civilization') becomes a source of military weakness:

> How is it that the Galla defeat us, though we are numerous and well supplied with arms?.. It is because our nation is divided into ten classes, nine of which take no part whatever in war, and make no shame of displaying their fear; only the tenth class makes war and fights to the best of its ability. Now, although we are numerous, those who can fight in war are few in number, and there are many who do not go to war. Of these classes, the first is that of the monks, of whom there are vast numbers. Among them are those who become monks at an early age, drawn thereto by the other monks while they are studying, as indeed was the case with him who has written this history, and others like him. There are also others who become monks because they fear war. A second group is composed of those who are called *dabtara*, or clerks; they study the holy books and all works relating to the occupations of the clergy; they clap their hands and stamp their feet during divine service, and have no shame for their fear of going to the wars. These people take as their models the levites and priests, namely, the sons of Aaron. The third group is that of the people called *Jan Hasana* and *Jan Maasare*, who look after the administration of justice, and keep themselves from war. The fourth group is formed by those who escort the wives of dignitaries and the princesses; they are vigorous, brave, and strong men who nevertheless do not go to war, for they say, 'We are the protectors of the women'. The fifth group calls itself *ema gelle*, 'elders'; they are the lords and hereditary landowners: they share their land with their laborers, and are not ashamed of their fear. The sixth group is that of the laborers in agriculture, who live in the fields and have no thought of taking part in war. The seventh group is composed of those who engage in trade and gain profit thereby. The eighth group is that of the artisans, such as the smiths, scribes, carpenters, and such-like, who know not the art of war. The ninth group is that of the wandering singers, those who play the *qanda kabaro* [a small drum] and the *bagana*, whose profession is to beg, to collect money. They invoke blessings on those who reward them, flattering them with vain praises and idle panegyrics; while those who refuse to give them presents they curse, though they are not blamewor-

> thy for this, for, as they say, 'This is our custom'. Such people keep themselves as far as possible from war. The tenth group, finally, is composed of those who carry the shield and spear, who can fight, and who follow the steps of their king to war. It is because these are so few in number that our country is ruined. Among the Galla, on the contrary, these nine classes which we have mentioned do not exist; all men, from small to great, are instructed in warfare, and for this reason they ruin and kill us (Bahrey 1976 [1593]: 140–141).[13]

In our model the higher coefficient of military participation that was typical for barbarians is mathematically described in Eq. 3 by giving b coefficient (representing the military participation ratio of barbarians here, *i.e.*, the percentage of barbarian population participating in military operations) *significantly* higher value than that of c coefficient (representing the military participation ratio for the 'civilized' population) in Eq. 2. For example, in the computer si-

[13] However, it is worth noticing that military forces of small polities were quite often comparable with military forces of large ones in those political entities of civilizations where the level of military participation of inhabitants was high (*e.g.*, in some civil communities, in particular in Greek *poleis*, Roman *civitas*, some medieval cities). The best known example is Greek-Persian wars when an alliance of civil communities with a high military participation ratio defeated a low military participation ratio empire. Note also that the suggested mathematical model does not consider the following factors of barbarians' military superiority (which it would be worth considering in future generations of similar mathematical models): a) a high mobility of some barbarian peoples in comparison with settled farmers which is quite often defined by their own way of life, as well as a low specific useful biomass output per unit of economically exploited territory (that causes the need of moving in order to increase their zone of economic exploitation). Especially it refers to nomads and the sea peoples, as well as the inhabitants of those places where rivers constituted the main communication lines. From the very beginning water transport was the main means of long distance connections (McNeill 1995). Therefore we suggested an idea that it is necessary to multiply the number of inhabitants among herders and seamen by the coefficient of their mobility for considering their potential for the intensification of political complexity growth processes (Korotayev 1991; Grinin 2007a). Such mobility often secures possibilities of the rapid advance in huge territories where civilizations are located; b) Higher prestige (concerning the whole population) of military activities. In other words, in a number of civilizations military professionals had no such prestige as priests or officials. For example, the founders of the Song Dynasty in China (960–1279) significantly downgraded and changed the position of military elite in order to prevent the possibility of the 'military coups' that undermined the stability of the political system of their predecessors (Wright 2001). But even where military estate stood high (as, *e.g.*, in medieval Europe or Japan), monopolization of military affairs in their hands led to the fact that most of population were specialized in peaceful occupations and as a result their military potential was close to zero. So, for example, in the 8th and 9th centuries in Charlemagne's empire, especially in France, the hardships that came along with military service actively induced peasants to pass voluntarily under the protection of large secular and religious landowners, thereby even by sacrificing their civil freedom (see, *e.g.*, Gurevich 1970: 145–183). In Russia, it was not uncommon when voluntary transfer of noble children to serf status took place in order to save them from military service. At the same time participation in military affairs was very honorable among barbarians (especially in state analogues), and very often volunteers' participation was sufficient to support major military actions in such affairs (see, *e.g.*, Fenton 1978: 127 on the Iroquois).

mulations whose results are given below for the main scenario of our model, the value of *c* is 0.05, whereas the value of *b* coefficient is 0.2.

$$M_c = cN_cT_cH_c, \qquad \text{(Eq. 2)}$$

where N_c is the size of 'civilized' population; T_c is the level of technological development of the civilizational core (for simplicity it is assumed that the level of development of military technologies of civilization is proportional to the general level of its technological development; therefore, within this model it is not identified as a separate variable); H_c is the level of asabiyyah of civilized population (we will dwell upon this variable below);

$$M_b = bN_bT_{mb}H_b, \qquad \text{(Eq. 3)}$$

where N_b is the number of inhabitants of the barbarian periphery; T_{mb} is the level of development of military technologies in the barbarian periphery (it is assumed that this variable is not identical with the general level of technological development of the barbarian periphery; the implications of this assumption will be considered below); and H_b is the level of barbarians' asabiyyah.

2) Borrowing of military technologies by barbarians happened at higher rates, than borrowing of non-military technologies (the chosen by us way to model this assumption mathematically will be described below). For example, due to the fact that Mongols borrowed siege equipment and technology from China, they were able to take a lot of cities successfully. Thus, in this case borrowing of military innovations was one of important reasons of mass destruction of cities of the World System in the 13th century. Nevertheless, the abovementioned facts about borrowings refer not only to weapons, but also to the strategy, tactics, and organization of the army. Quite often barbarians just imitated the structure of armies (or separate military institutes) of the neighboring civilizations. For example, the German leader Marobod (the late 1st century BCE – the early 1st century CE), having united Marcomanni with the Lugians, the Mugilones, Ghots and other Germanic tribes, created a large army on the Roman pattern which numbered 70,000 of infantry and 4,000 of cavalry (SIE 1966: 123).[14]

[14] By the way, such borrowings became the main impulse for transformation of the non-state systems into the early state society quite often. This happened, *e.g.*, as a result of the borrowing of iron weapons, and later fire arms (in particular, one could see examples of the latter in Madagascar in the 17th century [Deshan 1984: 353; Ratzel 1902, vol. 1: 445], or in Tahiti and Hawaii in the 18th century [Service 1975; Earle 2002: 86; for more details about similar cases see Grinin 2007a]). Note some more points here which are not considered in the present version of our model, but which it would be worth considering in the next generation of the models of interaction between the civilizational center and barbarian periphery. A) Barbarians themselves could be inventors of important military innovations. It is fair enough since many people considered war as the most important issue and became professionals of military attacks and robberies. Sometimes such inventions helped some barbarian chiefdoms to defeat others. A classic example

3) The beginning of forceful expansion of the civilizational core upon the barbarian periphery can be interpreted as the formation of a metaethnic border between the civilization and the barbarian world. As was clearly demonstrated by Peter V. Turchin (2003, 2005, 2007), the formation of such a metaethnic border tends to lead to a significant increase in collective solidarity (asabiyyah) in that party that turned out to be under pressure.[15] As a result, if at the beginning of its forceful expansion civilization faced scattered groups of barbarians incapable to produce any effective resistance, further on these groups began to cooperate more and more among themselves for putting up resistance, and civilization had to deal with more and more united and large coalitions of barbarians (which were formed in many respects as a reaction to forceful expansion and were able to show more and more effective resistance, and further to start

was that Shaka, the leader of the Zulu people, who applied a new type of cold weapon that in many respects promoted progress of his army and formation of the empire (Ritter 1968; Ratzel 1902, 2: 116). As a result, the Zulu polity was transformed from the pre-state level to the one of the state. There were cases when such inventions promoted expansion of barbarians against civilization. East Germans in the 5th century CE probably invented some kind of huge backsword with the straight sharpened blade (*scramasax*) which was up to 80 cm long. It was a typical saber weapon capable of giving terrible wounds which increased the power of a horse soldier. Therefore, it was borrowed by Huns, and then Goths and Francs (Kardini 1987: 263–264). The striking example of such innovations of barbarians were ships and naval tactics of Vikings who 'were second to none at sea' and whose sea advantage was often absolute (Gurevich 2005: 41 and ff.). One can also mention, *e.g.*, military tactical and organization innovations used in Genghis Khan's army, which undoubtedly played a great role in the Mongol victories. Thus, if barbarians and civilization were incomparable by cultural level, they could be quite comparable as regards their military-strategic levels, and quite often barbarians also had superiority, but at the same time kept such forms of organization of society which, according to to the well-known expert of nomadic studies William Irons, were real alternatives to state organization (Irons 2002, 2004) and could reproduce themselves without cities (though they could control cities populated by conquered peoples). B) In the process of weakening collective solidarity (~ asabiyyah) of civilizations and states, conflicting parts of civilization begin to use barbarians as allies, which gives them a chance to interfere with affairs of the civilization core. One can recollect that the author of *The Song of Igor's Campaign* wrote that dukes began 'to forge feuds for themselves', and 'to draw the pagans onto the Russian land'. The late Roman and Byzantine history gives a lot of examples of the 'integration' of barbarians into policy of civilization. A classic example is a tragedy of post-Roman Britain. After the withdrawal of the Roman troops from Britain in 410 CE, the Britons (Romanized British Celts) searching for the defenders from attacks of the Irish and Scottish barbarians invited the Saxones and gave them some land (thereby having exercised a certain social innovation, which was, however, repeatedly used in the Roman world with its practice of 'fighting against barbarians with barbarians' hands'). But having seen weakness of the Britons, the Saxones ceased to obey the local authorities and together with the Angles and the Jutes became eventually the owners of the country. And, despite their prolong and persistent resistance, the Britons were partly expelled, partly enslaved, and partly destroyed. Therefore, Anglo-Saxon barbarous kingdoms emerged in Britain in place of the 'Briton' state (*e.g.*, Blair 1966: 149–168; Chadwick 1987: 71). Thus, military opportunities of barbarians could significantly increase with their involvement into military-political affairs of civilization.

[15] Those, who withstood it, found adequate responses to the challenge that finally led to the selection of types of barbarian communities most adapted to the fighting against civilization.

successful counterattacks). As has already been mentioned above, Turchin suggests using for denoting 'collective solidarity' the term *asabiyyah* that was introduced into the scientific discourse by Abd ar-Rahman Ibn Khaldun[16] (1332–1406).

In the model, the dynamics of barbarian *asabiyyah* (H_b) is described mathematically by means of the following equation:

$$\frac{dH_b}{dt} = e \times \frac{dA_c}{dt}, \qquad \text{(Eq. 4)}$$

where e is a constant. It means that the higher the rates of forceful territorial expansion of civilization, the higher the growth rates of barbarians' *asabiyyah*.[17]

Respectively:

$$\frac{dH_c}{dt} = -e \times \frac{dA_c}{dt}, \qquad \text{(Eq. 5)}$$

where H_c is *asabiyyah* of civilized population.

Note that it means that *asabiyya* of civilization begins to grow under the pressure of barbarians, and the stronger this pressure is, the quicker it grows (for more details see Turchin 2005).

While describing population dynamics, we base ourselves upon the simplified version of the compact model of demographic, technological and economic development of the World System (Kremer 1993; Korotayev 2005, 2006d, 2007, 2008, 2009, 2012, 2013; Korotayev, Malkov, and Khaltourina 2006a, 2006b, 2007; Korotayev and Malkov 2012; Zinkina, Malkov, and Korotayev 2014; Korotayev and Malkov 2016; Korotayev and Zinkina 2017; Grinin 2003a, 2012; Grinin L. and Grinin A. 2015, 2016; Grinin A. and Grinin L. 2015; Grinin and Korotayev 2016; Grinin L., Grinin A., and Korotayev 2017a). We make a Malthusian assumption that throughout the most part of the period of existence of the humankind, the human population was limited by the level of development of life-supporting technologies. As in simplified Kremer's model (Kremer 1993: 685), we assume that population comes to technologically determined level of the Earth's carrying capacity instantly (or, in other words, instantly fills the ecological niche expanded as a result of technological growth).[18] Besides, we take into account the fact that territory with a higher

---

[16] See, *e.g.*, Ibn Khaldun 1958, 2004; Batsieva 1965; Ignatenko 1980; Alekseev and Khaltourina 2004; Turchin 2003, 2007; Korotayev and Khaltourina 2006; Korotayev 2006e, 2007d; Inan 1933; Mahdi 1937.

[17] We also assume that with the increase in barbarians' asabiyyah the rate of borrowing of military technologies of civilization increases (this assumption is modeled by the Eq. (8'''). We also assume that variable *H* cannot have negative values.

[18] Let us note that it deprives us of an opportunity to describe cyclical dynamics of the system in the basin of attraction (see, *e.g.*, Korotayev, Komarova, and Khaltourina 2007) that would bring dynamics generated by the model considerably closer to actually observable one, but at the same

natural productivity can support the existence of a larger population at the same level of technological development, than the territory with smaller natural producing capacity, and otherwise under equal conditions a larger territory can support a larger population than a smaller territory. Thus, the size of population *(N)* of some zone with productivity *F* and area *A* at the level of development of life-supporting technologies *T* will be described mathematically by means of the following equation:

$$N = gFTA, \qquad \text{(Eq. 6)}$$

where g is a constant.

As a result, the mathematical description of the population for year i for a hinterland of the World System (Zone 3) appears to be the simplest one in our model, since we have initially assumed that the territory occupied by it throughout the modeled period remains constant, and the level of technological development is the same for the whole zone:

$$N_{3i} = gF_3T_{3i}A_3. \qquad \text{(Eq. 7)}$$

The situation with the civilizational core and barbarian periphery of the World System is a little more complicated. The matter is that throughout the most part of the modeled period the civilization zone is divided into two subzones with different natural productivity, *i.e.* the core of the civilization zone with high natural productivity (~ Zone 1) and the periphery of the zone corresponding to the part of less productive Zone 2 taken by the civilization from 'barbarians'. Thus,

$$N_{ci} = N_{1ci} + N_{2ci}, \qquad \text{(Eq. 8)}$$

where N_{ci} is population of the civilization core for year i; N_{1ci} is the 'civilized' population of Zone 1 for year i; N_{2ci} is the 'civilized' population of Zone 2 for year i.

At the same time:

$$N_{1ci} = gF_1T_{ci}A_{1ci}, \qquad \text{(Eq. 9)}$$

where A_{1ci} is the area of the part of Zone 1 controlled by civilization for year i;

$$N_{2ci} = gF_2T_{ci}A_{2ci}, \qquad \text{(Eq. 10)}$$

where A_{1ci} is the area of the part of Zone 2 controlled by civilization for year i.

Respectively,

$$N_{bi} = N_{2bi} + N_{1bi}, \qquad \text{(Eq. 11)}$$

where N_{bi} is population of the barbarian periphery for year i; N_{2bi} is the 'barbarian' population of Zone 2 for year i; N_{1bi} is the 'barbarian' population of Zone 1 for year i.

---

time this considerably simplifies the suggested model, which made us dwell on this simplified version of description of dependence of population on the level of technological development.

Herewith,

$$N_{2bi} = gF_2T_{bi}A_{2bi}, \quad \text{(Eq. 12)}$$

where A_{2bi} is the area of the part of Zone 2 controlled by 'barbarians' for year i;

$$N_{1bi} = gF_2T_{bi}A_{2bi}, \quad \text{(Eq. 13)}$$

where A_{1bi} is the area of the part of Zone 1 controlled by 'barbarians' for year i. The way of calculation of A_{1c}, A_{2c}, A_{2b} and A_{1b} variables employed by us in this model is described below (see Table 1 and Eqs 9, 10, 24, 25).

The total population of the World System for year i (N_{wi}) is calculated by means of the following equation:

$$N_{wi} = N_{ci} + N_{bi} + N_{3i}. \quad \text{(Eq. 14)}$$

Similarly to our general model of the World System development, mathematical description of technological dynamics is based upon the equation for technological growth proposed by Michael Kremer[19] (Kremer 1993: 686):

$$\frac{dT}{dt} = hNT, \quad \text{(Eq. 15)}$$

where h is a constant (~ coefficient of technological innovative activity of population).

We assume that the diffusion of innovations proceeds from the civilization center of the World System to its barbarian periphery and from it to hinterland (1). Though in reality the diffusion of innovations from hinterland to periphery (2), from periphery – to center (3), as well as between various subzones of periphery (4) and hinterland (5) was also observed, after all the main flow of technological diffusion went in the first of the abovementioned directions (see, *e.g.*, Chubarov 1991; Grinin A. and Grinin L. 2015, 2016) and we decided to refrain from the modeling of diffusion of technological innovations in other directions for the sake of simplicity of our model.

Thus, the following system of difference equations has been used in our model to model the technological development of the World System:

$$T_{ci} = T_{ci\text{-}1} + hN_{ci\text{-}1}T_{ci\text{-}1}, \quad \text{(Eq. 16)}$$

where T_{ci} is the level of technological development of the civilization core of the World System for year i;

$$T_{bi} = T_{bi\text{-}1} + hN'_{bi\text{-}1}T_{bi\text{-}1} + k(T_c - T_b), \quad \text{(Eq. 17)}$$

where T_{bi} is the level of technological development of the barbarian periphery for year i; N'_b – population of the barbarian Zone; k – a constant;

$$T_{3i} = T_{3i\text{-}1} + hN'_{3i\text{-}1}T_{3i\text{-}1} + l(T_{bi\text{-}1} - T_{3i\text{-}1}), \quad \text{(Eq. 18)}$$

[19] Note that quite independently from Michael Kremer this equation was proposed by Rein Taagepera (1976, 1979), A. V. Podlazov (2000, 2001, 2002) and S. V. Tsirel (2004).

where T_{3i} is the level of technological development of hinterland (Zone 3) for year i; N'_3 is population of hinterland; l is a constant.

As has already been mentioned above, we introduce an additional equation for description of dynamics of development of 'barbarians'' military technologies to take into account the effect of the more rapid borrowing by 'barbarians'' of military technologies in comparison with peaceful technologies:

$$T_{mbi} = T_{mbi\text{-}1} + mN'_{bi\text{-}1}T_{mbi\text{-}1} + nH_{bi\text{-}1}(T_{ci\text{-}1} - T_{mbi\text{-}1}), \quad \text{(Eq. 19)}$$

where m and n are constants.

The effect that is of the most interest for us can be described by means of Eqs 17 and 19 by giving a higher value to n coefficient in Eq. 19 in comparison with the value of k coefficient in Eq. 17. This equation also describes an assumption that the rates of barbarians' borrowings of military technologies grow along with the growth of their asabiyyahs.

We assume that all the urban population of the World System is concentrated in its civilizational core. For mathematical description of urbanization dynamics in the model the following equation is used:

$$u_{ci} = pT_{ci}, \quad \text{(Eq. 20)}$$

where u_{ci} is the index of urbanization of the civilizational core (a share of urban population in the total population of civilization core), and p is a constant. The possibility of approximation of $u_{ci} \sim pT$ follows from the equations of our compact model of the general World System development (Korotayev, Malkov, and Khaltourina 2007; Korotayev, Komarova, and Khaltourina 2007; Korotayev 2012, 2013). At the same time an empirical testing of this approximation was not carried out, and this testing was done by us. We use the World System Technological Development Index proposed by us earlier (Korotayev 2006a) for an empirical test of this hypothesis. Let us recollect that this index was calculated on the basis of Hellemans – Bunch database (Hellemans and Bunch 1988). In this database Hellemans and Bunch tried to record in chronological sequence all the main inventions and discoveries that had been made by the 1980s. As a value of the World System Technological Development Index for the moment X we use the total number of inventions and discoveries which were made in the World System up to that moment.

The correlation between this World System Technological Development Index and the World System Urbanization Index calculated by us earlier (Korotayev, Malkov, and Khaltourina 2007: 122–127; Korotayev 2007; Grinin and Korotayev 2008: Ch. 4) looks as follows (see Fig. 4):

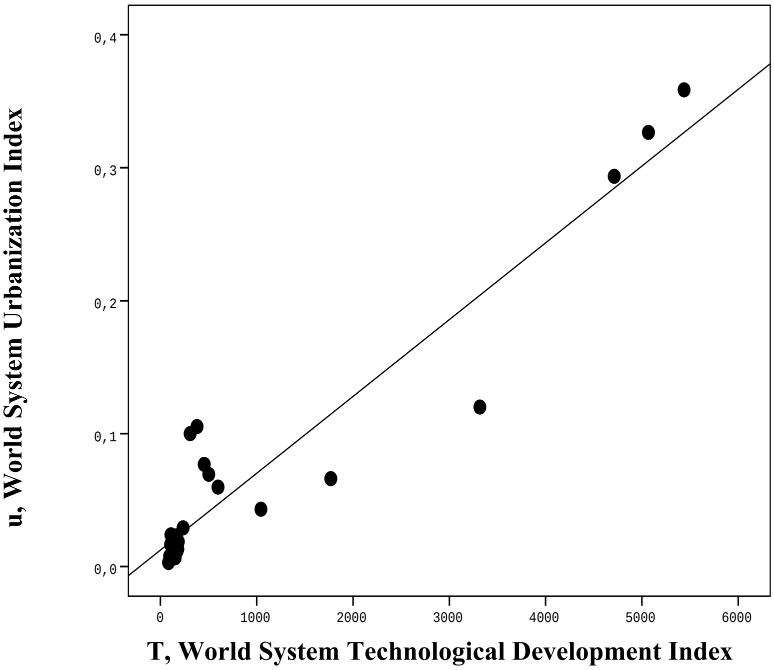


Fig. 4. Correlation between the World System Technological Development Index (*T*) and the World System Urbanization Index (*u*) (3500 BCE – 1970 CE): scatterplot with a fitted regression line

Note: $R = 0.95$; $R^2 = 0.903$; $p = 1.08 \cdot 10^{-15}$.

Thus, we find a rather strong and statistically significant correlation between these indices.

The total number of urban population in the model is defined by the following equation:

$$U = u_c N_c. \quad \text{(Eq. 21)}$$

Finally, the World System Urbanization Index (a share of urban population in a total number of the World System population) u_w is defined by the following equation:

$$u_w = \frac{U}{N_w}, \quad \text{(Eq. 22)}$$

where N_w is the total population of the World System.[20]

[20] Thus, the World System urbanization appears here in our model as a purely dependent variable. Perhaps, it would make sense to consider its influence on some other key variable models (*e.g.*,

Table 1 gives a summary description of the model:

Table 1. Compact mathematical model of influence of interaction of the civilizational center and barbarian periphery on the development of the World System (a detailed description)

Variable symbol	Meaning	=	Value for year *i*	Equation number
A_1	The territory of Zone 1	=	Constant, in computer simulations the results of which are given below, has the value of 1 mln km^2	–
A_2	The territory of Zone 2	=	Constant, 24 mln km^2	–
A_3	The territory of Zone 3	=	Constant, 96 mln km^2	–
F_1	'Index of natural fertility' of Zone 1	=	Constant, in computer simulations, the results of which are given below has the value 10	–
F_2	'Index of natural fertility' of Zone 2	=	Constant, 3	–
F_3	'Index of natural fertility' of Zone 3	=	Constant, 1	–
A_c	The territory of 'civilization zone'	=	$A_{ci-1} + a(M_{ci-1} - M_{bi-1})$; $A_{c0} = A_1 = 1$ mln km^2. This variable cannot have negative values	(1)
A_b	The territory of the 'barbarian periphery'	=	$A_{bi-1} + a(M_{bi-1} - M_{ci-1})$; $A_{b0} = A_2 = 24$ mln km^2. This variable cannot have negative values either	(23)
A_{1c}	The territory of the part of Zone 1, controlled by civilization	=	It is described by a version of Eq. 1; $A_{1c0} = A_{c0} = A_1 = 1$ mln km^2; it does not change while there is an expansion of civilization; if as a result of counterattack of barbarians they completely return Zone 2 to themselves, then $A_{1ci} = A_{1ci-1} + a(M_{ci-1} - M_{bi-1})$ till A_{1c} (= A_c) reaches zero value (it is interpreted as a complete conquest of civilization by barbarians) or returns to value of 1 mln km^2 (it is interpreted	(9)

on the rates of technological growth which was already made by M. Artzrouni and J. Komlos [Artzrouni and Komlos 1985] and that, in our opinion, might allow us to give a more exact description of technological dynamics of the World System in the basins of attraction of attractors B_1 and B_2), but in order to avoid excessive complication of the model we opt to refrain from this, though the action of this factor may be taken into account in future models.

Variable symbol	Meaning	=	Value for year *i*	Equation number
			as a full expulsion of barbarians from Zone 1).[21] This variable cannot have negative values	
A_{2c}	The territory of the part of Zone 2 controlled by civilization	=	$A_c - A_1$ if $A_c > A_1$; 0 if $A_c \leq A_1$.	(10)
A_{2b}	The territory of the part of Zone 2, controlled by 'barbarians'	=	A_b if $A_b \leq A_2$; A_2 (= in our case 24) if $A_b > A_2$	(24)
A_{1b}	The territory of the part of Zone 1, controlled by 'barbarians'	=	0 if $A_b \leq A_2$; $A_1 - A_c$ if $A_b > A_2$.	(25)
M_c	Military potential of civilization	=	$cN_cT_cH_c$	(2)
M_b	Military potential of the 'barbarians'	=	$bN_bT_cH_c$. It is assumed that the value of the military participation ratio of 'barbarians' *(b)* is *significantly* higher than that for 'civilized' population. In computer simulations whose results are presented below, the value of *c* is assumed to be equal to 0.05, and the value of coefficient *b* is assumed to be equal to 0.2	(3)
H_b	Index of barbarians' collective solidarity (*asabiyyah*)	=	$H_{bi\text{-}1} + e(A_{ci} - A_{ci\text{-}1})$; $H_{bi} \geq 0$	(4)
H_c	Index of collective solidarity (*asabiyyah*) of 'civilized' popula-tion	=	$H_{ci\text{-}1} - e(A_{ci} - A_{ci\text{-}1})$; $H_{ci} \geq 0$	(5)
N_c	Population of the civi-lizational core	=	$N_{1ci} + N_{2ci}$	(8)
N_{1c}	'Civilized' population of Zone 1	=	$gF_1T_{ci}A_{1ci}$	(26)
N_{2c}	'Civilized' population of Zone 2	=	$gF_2T_{ci}A_{2ci}$	(27)
N_b	Population of the bar-barian periphery	=	$N_{2b} + N_{1b}$	(11)
N_{2b}	'Barbarian' population of Zone 2	=	$gF_2T_{bi}A_{1bi}$	(28)

[21] It is obvious that the easiest way to model the dynamics of this variable is to give it the value of A_1 (that is 1 in our computer simulations) when $A_c \geq A_1$ and value A_c when $A_c < A_1$. This method was also applied by us in real computer simulations for this and other similar variables (A_{2c}, A_{2b} и A_{1b}).

Variable symbol	Meaning	=	Value for year *i*	Equation number
N_{1b}	'Barbarian' population of Zone 2	=	$gF_1T_{bi}A_{1bi}$	(29)
N_3	Population of Zone 3	=	$gF_3T_{3i}A_3$	(7)
N_w	Total population of the World System	=	$N_{ci} + N_{bi} + N_{3i}$	(14)
T_c	Level of technological development of the World System civilizational core	=	$T_{ci\text{-}1} + hN_{ci\text{-}1}T_{ci\text{-}1}$	(16)
T_b	The level of technological development of the barbarian periphery	=	$T_{bi\text{-}1} + hN'_{bi\text{-}1}T_{bi\text{-}1} + k(T_c - T_b)$	(17)
N'_b	Population in one subzone of 'Barbarian Zone' (with conditionnal area of each subzone being equal to 1 mln km^2)	=	N_{bi}/A_{bi} (note that the area of zones in our model is measured in mln km^2 therefore this division gives the population of 'barbarians' per 1 mln km^2)	–
T_3	The level of technological development of the World System hinterland (= Zone 3)	=	$T_{3i\text{-}1} + hN'_{3i\text{-}1}T_{3i\text{-}1} + l(T_{bi\text{-}1} - T_{3i\text{-}1})$	(18)
N'_3	Population in one subzone of Zone 3 (with conditional area of each subzone being equal to 1 mln km^2)	=	N_3/A_3	–
T_{mb}	The level of development of military technologies of 'barbarians'	=	$T_{mbi\text{-}1} + mN'_{bi\text{-}1}T_{mbi\text{-}1} + nH_{bi\text{-}1}(T_{ci\text{-}1} - T_{mbi\text{-}1})$	(19)
u_c	Index of urbanization of the civilizational core (a share of urban population in the total population of the civilizational core)	=	pT_c; $0 \le u_c \le 0.9$	(9)
U	Total urban population	=	u_cN_c	(21)
u_w	Index of the World System urbanization (a share of urban population in the total population of the World System)	=	U / N_w	(22)

A typical dynamics generated by the model with average values of parameters and initial conditions is presented in Figs 5–7:

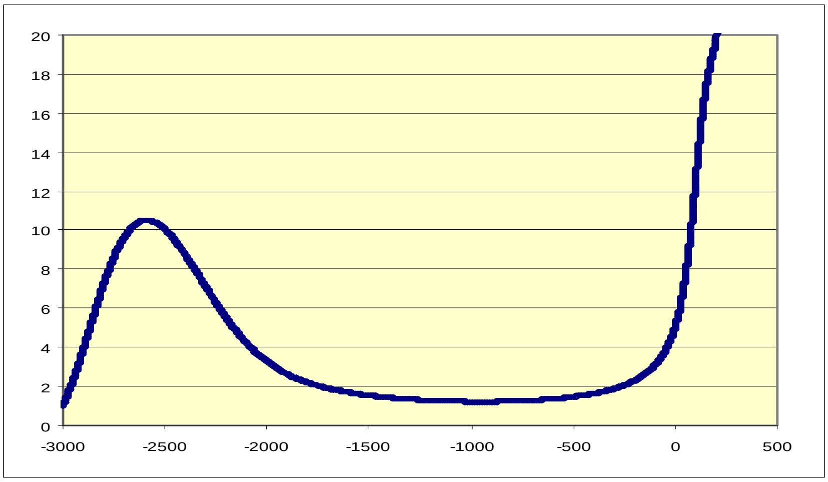


Fig. 5. Dynamics of the territory of the World System civilizational core generated by the main scenario of the model (mln km^2)

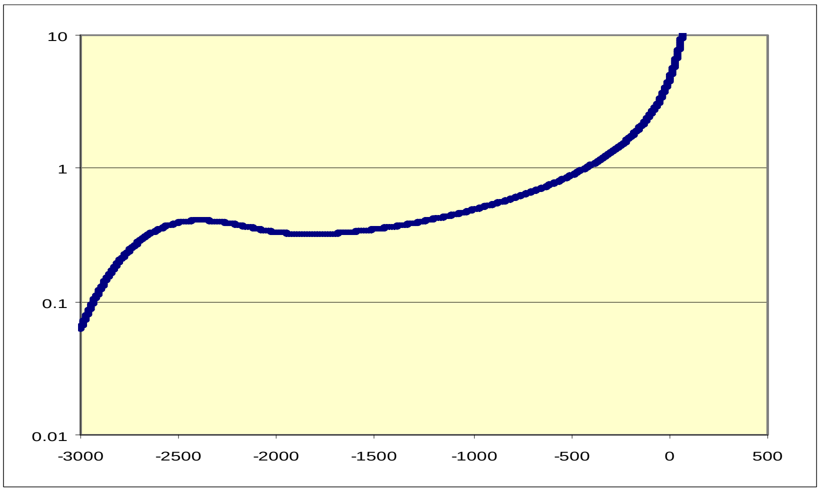


Fig. 6. Dynamics of the World System urban population (millions) generated by the model, logarithmic scale

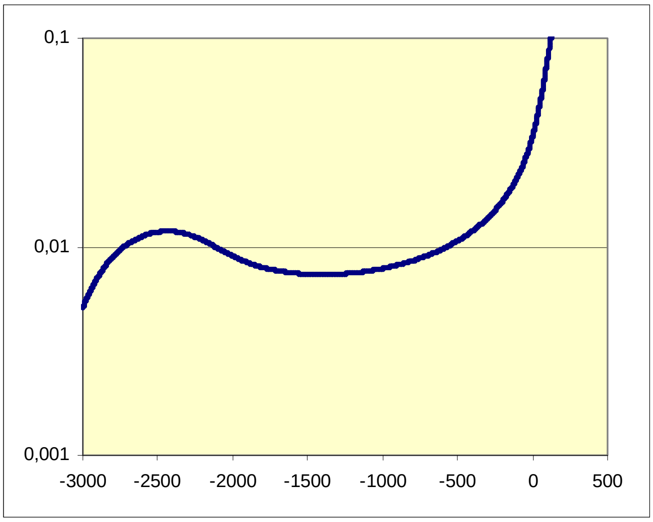


Fig. 7. Dynamics of the World System urbanization index (proportion of urban population in the total population of the World System) generated by the model, logarithmic scale

Note. Figs 5–7 show results of computer simulation with the following values of parameters and initial conditions: t_0 = 3000 BCE = –3000; A_1 = 1 mln km^2; A_2 = = 24 mln km^2; A_3 = 96 mln km^2; F_1 = 10; F_2 = 3; F_3 = 1; A_{c0} = 1 mln km^2; A_{1c0} = = 1 mln km^2; A_{2c0} = 0; A_{b0} = 24 mln km^2; A_{2b0} = 24 mln km^2; A_{1b0} = 0; T_{c0} = 10; T_{b0} = 2; T_{mb0} = 3; T_{3_0} = 0.2; H_{c0} = 1; H_{b0} = 0.1; a = 0.012; b = 0.2; c = 0.05; e = 0.052; g = 0.05; h = m = 0.0000315; k = l = 0.000504; n = 0.00504; p = 0.00125.

Within this computer simulation one can distinguish the following phases:

Phase 1 (years 0–130 of the computer simulation). *Vigorous accelerating expansion of civilization*.

Accelerating expansion of civilization in this phase is generated by the following system of positive feedbacks (see Fig. 8):

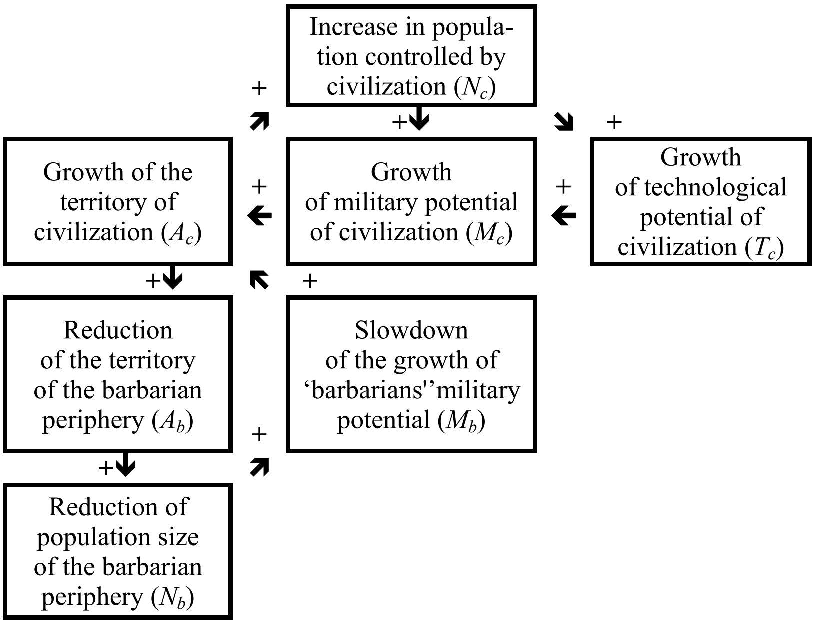


Fig. 8. System of positive feedbacks generating the accelerating territorial expansion of civilization during the first phase of the computer simulation

Thus, at this phase the growth of the civilization territory leads to the increase in its population, which results in increase in its military potential both directly (the size of the army [*i.e.* the number of soldiers] increases along with the increase in population size), and through acceleration of technological growth rates (allowing to supply the soldiers with more effective weapons); the increase in military potential of civilization leads to further increase in its territory which results in further acceleration of growth of its population, *etc.*; on the other hand, acceleration of the growth of the territory of civilization leads to a substantial reduction of the territory of barbarian periphery and consequently, decrease of population size and military potential of 'barbarians' that promotes further acceleration of growth of the territory of civilization, reduction of the territory of the barbarian periphery, *etc.*

At this phase one can observe the accelerated growth of population of the World System[22], the World System urbanization index[23] and the urban population.[24]

Phase 2 (years 130–340 of the simulation). *Slowdown of expansion of civilization.*

The following system of negative feedbacks comes to the foreground during this phase: the growth of the civilization territory leads to the growth of asabiyyah of 'barbarians', which is expressed in increase in the level of their political culture and organization[25], and leads to the growth of their military potential both directly and through the acceleration of rates of borrowing of military technologies of civilization (including military and organizational and tactical innovations) which results in reduction of rates of growth of the civilization territory which, until it slows down to zero level, continues to lead (through the mechanisms mentioned above) to the growth of barbarians' military potential and further slowdown of rates of territorial expansion of civilization (see Fig. 9).

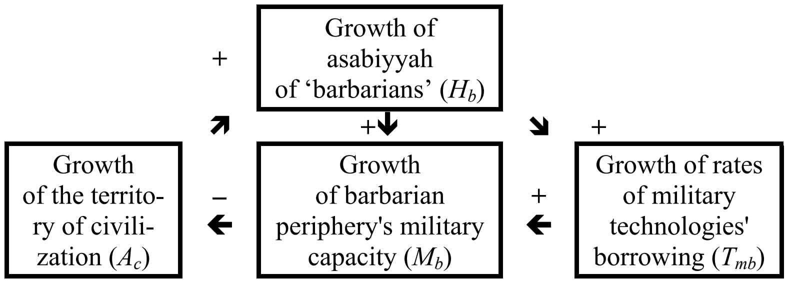


Fig. 9. System of negative feedbacks generating slow-down of the territorial expansion of civilization at the second phase of computer simulation

[22] First of all as a result of increasing diffusion of high technologies of civilization in the territories of Zone 2 subordinated by it, the growth of the carrying capacity there, and, therefore, the population.

[23] In connection both with accelerating technological growth of civilization and with the growth of 'civilized' population percentage in the total World System population.

[24] In connection both with the growth of urbanization in civilizational zone, and with the accelerated growth of its population as a result of the territorial expansion.

[25] In historical reality we find the corresponding situation: as A. M. Khazanov notes, though nomads may seem barbarians for settled contemporaries , these 'barbarians' may be quite sophisticated in a political sense (Khazanov 2002: 54).

Nevertheless, at this phase the expansion of civilization proceeds at rather rapid (though more and more slowing down) rates; besides, a rather rapid (though slowing down) growth of population of the World System also continues. However, since year 154 of our computer simulation the absolute growth rates of population of the World System begin to decrease, but up to the end of Phase 2 they remain rather high. Since year 209 of the computer simulation the absolute growth rates of urban population also begin to decrease (remaining nevertheless rather high). During this phase, the growth rates of the World System urbanization index decrease almost twice (nevertheless remaining rather high if compared with the subsequent two phases).

Phase 3 (years 340–510 of the simulation). *Expansion of civilization is exhausted and stops. Approximate power balance. The barbarian periphery begins its counterattack.*

During this phase, the territory of civilization in comparison to the territory of its barbarian periphery changes rather slowly, no more than 0.01 million km^2 per year (reaching at the inflection point, in year 408 of simulation, 48 km^2 per year). At the first stage of this phase the action of the above-mentioned mechanism of negative feedback leads to its logical conclusion – military potentials of civilization and barbarian periphery become equal to each other, and the rates of expansion of civilization reduce to zero level. However, the process of rather fast borrowing of military technologies of civilization by 'barbarians' continues. As a result military potential of the barbarian periphery begins to exceed that of civilization, and 'barbarians' start their counterattack. At the beginning it develops extremely slowly (83 km^2 during the first year); but the beginning of 'barbarian counterattack' leads to the formation of the system of positive feedbacks giving more and more noticeable results every year – acceleration of the growth of barbarian periphery territory leads to the acceleration of the growth of population of barbarian periphery, which in turn, leads to the increase in the military potential of 'barbarians' and even greater increase in the territory of barbarian periphery and consequently, to a greater increase in 'barbarian' population, *etc*.; on the other hand, acceleration of growth of the territory of barbarian periphery leads to a substantial reduction of the territory of civilization and consequently, to the decrease of civilization population, and military potential of civilization which promotes further acceleration of the growth of the territory of barbarian periphery, reduction of the territory of civilization, *etc*. (see Fig. 10).

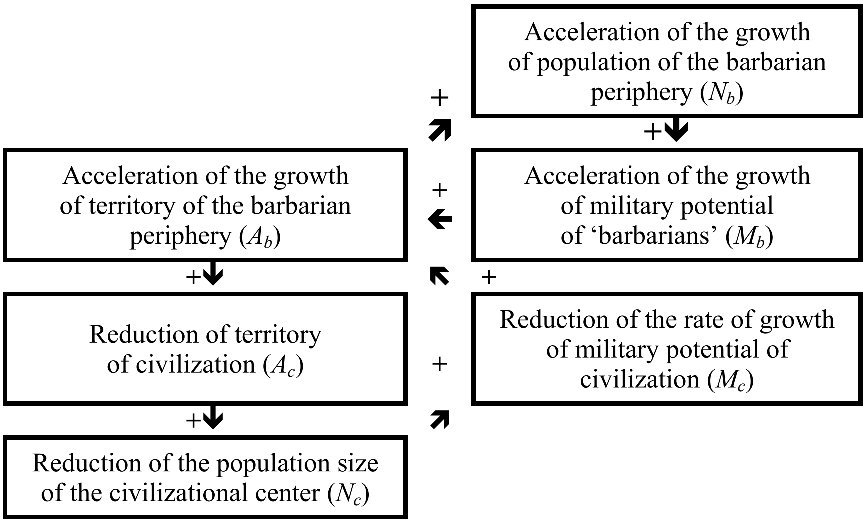


Fig. 10. System of positive feedbacks generating the accelerating territorial counter-expansion of the barbarian periphery at the 4th and 5th phases of computer simulation

It is interesting to note that in our computer simulation by the time of the beginning of counterattack of barbarian periphery the total number of 'barbarians' (7.9 million) is almost three times less than the number of 'the civilized population' (23.4 million), and the index of their general technological development (3.7) is much lower than the level of technological development of civilization (12.1). At the same time the counterattack of barbarians appears to be possible due to a higher military participation ratio peculiar to them, and also due to the point that at the beginning of counterattack their asabiyyah is higher than the civilization asabiyyah, and their military technology is much higher (7.15) than the general level of their technological development (3.7).

The growth rates of the total population of the World System decrease in this phase almost three times from moderate 0.125 to 0.043 % per year. At the end of this phase there begins the slowdown in rates of technological growth of civilization. Urban population growth falls to 260 people per year, and growth of the index of urbanization – to 0.0002 % per year.

Phase 4 (years 510–680 of the simulation). *Accelerating expansion of the barbarian periphery.*

At this phase the mechanism of the positive feedback accelerating the counterattack of the barbarian periphery works at full capacity. The territory controlled by the civilizational center is reduced by 2.5 million km^2. The population of the civilizational center is reduced from 24.1 to 21.7 million people. This reduction is only partially compensated by the increase in population of the barbarian periphery and hinterland of the World System; as a result, the growth rate of total population of the World System falls from 0.043 to 0.01 % per year. The rates of technological growth of civilization are reduced from 0.076 to 0.068 % per year. Since year 614 of our computer simulation the continuing growth of urban population in the territory unoccupied by barbarians ceases compensating the reduction of the urban population as a result of counterattack of the barbarian periphery, and the total number of urban population begins to decrease. Even earlier (since year 577 of the simulation) the World System urbanization index begins to decrease.

Phase 5 (years 680–935 of the simulation). *Slowdown of the expansion of the barbarian periphery.*

During this phase a system of feedbacks reducing the vigor of counterattack of barbarian periphery comes to the foreground.

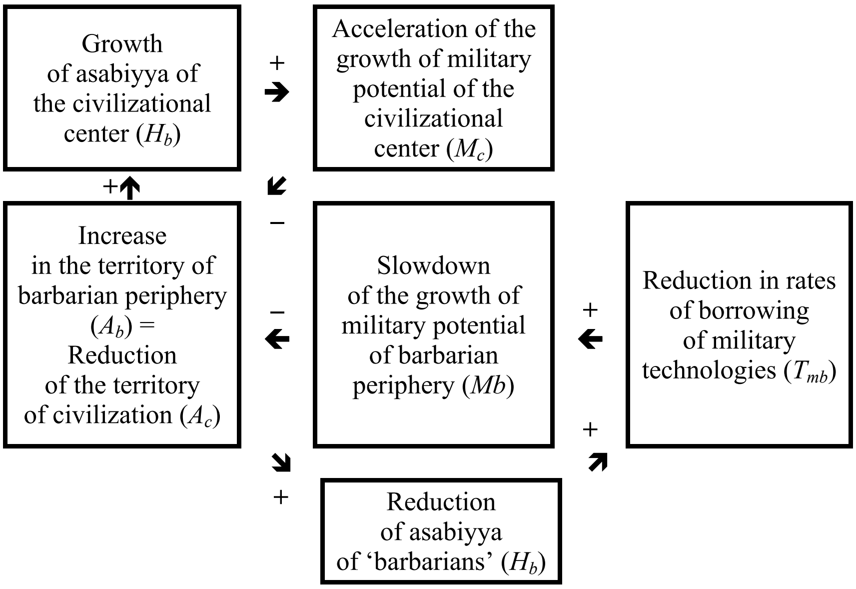


Fig. 11. The system of negative feedbacks, generating the slowdown of territorial expansion of the barbarian periphery during Phase 5 of the computer simulation

Despite its slowdown, the continuing counterattack of the barbarian periphery throughout this phase is able to lead to very significant consequences. The area of the territory controlled by the civilizational center is reduced almost twice. The population of the civilizational center falls from 21.73 to 15.97 million people. Until year 730 of the simulation this reduction is compensated in a lesser degree than earlier by the increase of population of the barbarian periphery and hinterland of the World System; therefore the growth rate in the World System total population reduces almost to zero. Later, the effect of the slowdown of expansion of the barbarian periphery begins to manifest itself, which against the background of continuing acceleration of growth rates of population of the barbarian periphery and hinterland leads to the renewal of the increase in growth rates of the total population of the World System (though it restarts growing very slow, showing the growth, say, only by 0.002 % between years 730 and 789 of our simulation, and even in 935 population growth rates of the World System remain extremely low, *i.e.*, 0.027 % per year, while at the beginning of the first phase they were 0.37 % per year, *i.e.* they were ten times higher). Technological growth rates of civilization continue to decrease (from 0.068 to 0.05 % per year). The urban population declines from 403,000 to 344,000 people (at the same time the effect of the decline of the vigor of the barbarian counterattack begins to manifest itself – the absolute rates of decrease in urban population reach maximum in years 817–838, and then begin to decline). The World System urbanization index falls from 0.0116 to 0.0095 (though reduction rates of this indicator since year 865 of our simulation begin to decline too).

Phase 6 (years 935–2885 of the simulation). *Expansion of the barbarian periphery reaches its peak. Approximate balance of forces. Civilization launches the counterattack.*

During this phase, the territory of the barbarian periphery as compared to the territory of the civilizational core changes rather slowly, no more than 0.01 million km^2 per year (reaching at the inflection point just 0.5 km^2 per year in year 2047 of our simulation). At the first stage of this phase (years 935–2047), the action of the negative feedback described in Fig. 10 produces its logical conclusion: military potentials of civilization and barbarian periphery become equal to each other, and expansion rates of the barbarian periphery decline to zero. However, the rates of technological development (including the growth rates of military technologies) of civilization continue to outpace those for the barbarian periphery. As a result, the military potential of civilization begins to exceed that of the barbarian periphery, and the civilization begins its counterattack (with the level of 1.23 million km^2) in year 2048 of our simulation. At first it proceeds very slowly (only 1.5 km^2 during the first year of the 'counterattack'); but the beginning of civilization counterattack leads to the formation of a system of positive feedbacks (described in Fig. 8), giving more and more noticeable results every year, – growth of the civilization territory leads to the increase in its population which leads to the growth of military potential both directly and through acceleration of technological growth rates; the growth of military potential of civilization leads to further increase in its population, which leads to further acceleration of growth of its territory, *etc.* As a result, the territory of civilization grows from the level of 1.23 million km^2 in year 2048 to 2.94 million km^2 by the end of this phase at increasing (but still, in general, rather slow) rates.

The growth rates of the total population of the World System continue to increase throughout all the 6th Phase, but by very slow rates, increasing from 0.027 to 0.12 % per year, still remaining lower than the rates characteristic of the beginning of Phase 1. The population of the civilizational center continues to fall until year 1438 of our simulation, declining from 15.97 to 12.73 million, and then its growth resumes, and, by the end of the phase, the civilization population reaches 38.5 million, considerably exceeding the level reached at Phases 1–3. The reduction of technological growth rate of civilization continues till year 1439 of the simulation (declining from 0.05 to 0.04 % a year), and then this rate begins to grow rapidly, reaching the level of 0.12 % per year by 2885 (*i.e.*, it considerably exceeds the level reached at Phases 1–3). The World System urban population continues to decline till year 1169 of the simulation (decreasing from 344,000 to 320,000 people, then its growth is resumed, gradually accelerating, and, by the end of Phase 6, the urban population of the World System reaches 2,340,000). The World System urbanization index continues to fall much longer – till year 1600 of our simulation, declining from 0.0095 to 0.0074; and then it begins to grow with gradual acceleration, reaching the level of 0.0206 by the end of Phase 6.

Phase 7 (years 2885–3209 of the simulation). *As a result of a vigorous counterattack civilization completely subordinates Zone 2, absorbing the whole barbarian periphery*.

During this phase, there is a rapid growth of all the modeled indicators of the level of development of the World System. By year 3065 of the computer simulation the urban population of the World System reaches the level of 10 million, and the World System urbanization index exceeds 10 % in 3123. In reality at this level, civilization already have to contact with the extensive hinterland of the World System (which during our simulation managed to achieve rather high levels of population and technological development), whereas the World System hinterland would transform into the new barbarian periphery of civilization, which, with certain values of parameters, could lead to a new counterattack of the barbarian periphery at a higher level. However, it would lead to additional complication of the model from which we have decided to refrain at this stage.[26]

It is interesting that correlation between technological development and urbanization of the World System generated by this model is surprisingly similar to what we have seen above (see Fig. 4) for the empirical estimates of the level of technological development of the World System, on the one hand, and the level of its urbanization, on the other (see Figs 12–14):

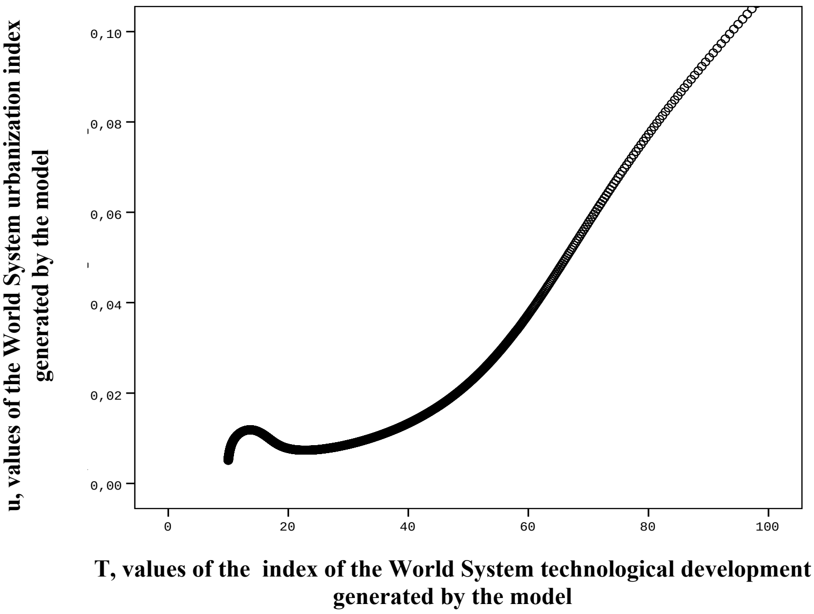


Fig. 12. Correlation between values of the World System technological development index (*T*) and the World System urbanization index generated by the model (*u*)

[26] This model also does not describe the withdrawal of the World-System from the blow-up regime. In theory, in our case it might be possible, having described basic population dynamics by means of the following equations: $dN/dt = r \cdot dT/dt \cdot (1-l)$; $dN/dt \leq 0{,}04$; $dl/dt = s \cdot dT/dt \cdot (1-l)$ (where *l* is a proportion of literate population, and *r* and *s* are constants), and basic urbanization dynamics – by means of the following equations: $du/dt = v \cdot dT/dt \times (u_{lim} - u)$ (where u_{lim} is a maximum possible share of urban population, and *v* is a constant); justification of the equations of this type, see, *e.g.*, in the following works: Korotayev 2006a; Korotayev, Malkov, and Khaltourina 2007; Korotayev, Komarova, and Khaltourina 2007. However, we decided not to do it in order to avoid the excessive complication of the model especially since it does not include the description of withdrawal of the World-System from the blow-up regime, and mathematical models with such a description have been already offered and published by us earlier (see *e.g.*, Korotayev 2006; Korotayev, Malkov, and Khaltourina 2007; Korotayev, Komarova, and Khaltourina 2007).

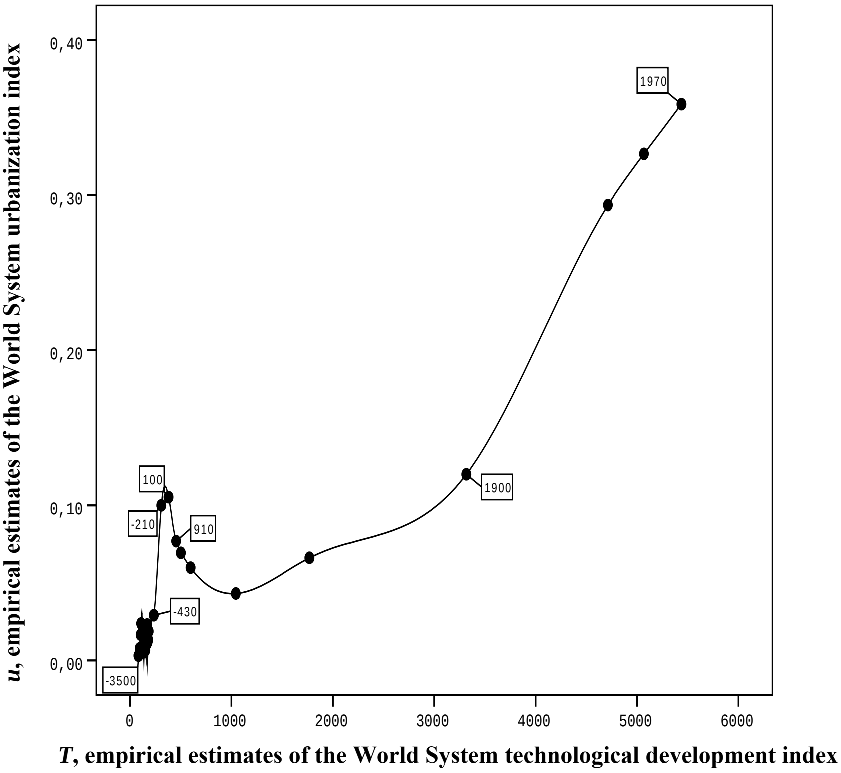


Fig. 13. Correlation between the empirical estimates of the index of the World System technological development (*T*) and the empirical estimates of the World System urbanization index (*u*) (3500 BCE – 1970 CE)

In this figure, most part of the curve up to 430 BCE looks like a solid black spot. However, after 'zooming' in this spot, it is possible to see that the curve in this sector has a form that is surprisingly similar to the one of the whole graph – a sort of fractal effect (see Fig. 14):

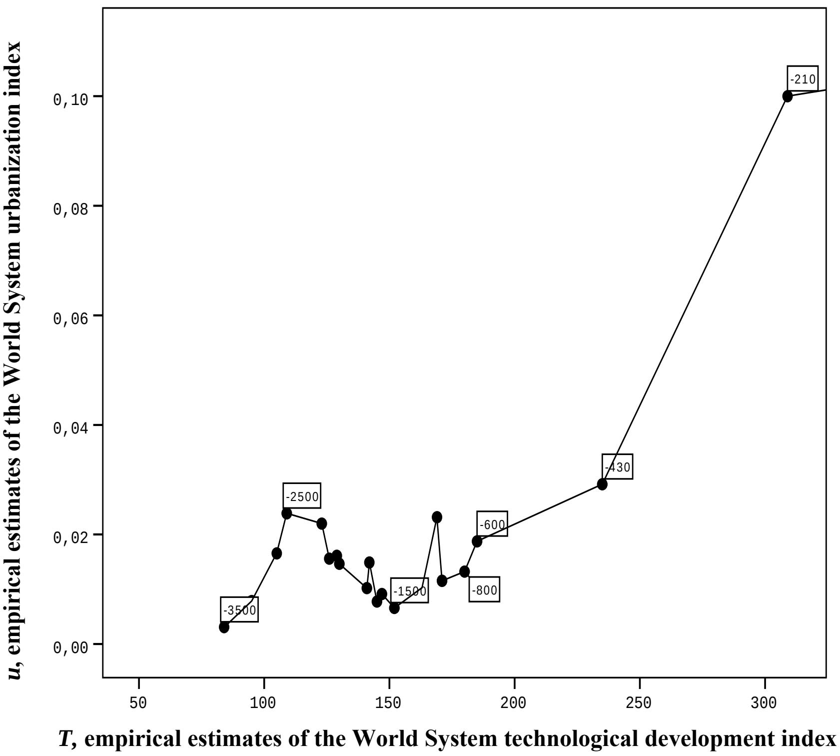


Fig. 14. Correlation between the empirical index of the World System technological development (*T*) and the World System urbanization index (*u*) (3500 BCE – 210 BCE)

As we see, the shapes of all the three curves are amazingly similar: in the initial part of the figure even rather small technological growth is followed by very noticeable growth of urbanization of the World System. Then it is followed by a pronounced interval when further technological growth is accompanied by decrease of the urbanization level, which is changed by an interval where technological growth is accompanied by slow growth of urbanization which is followed by a stretch where technological growth is accompanied by rapid growth of urbanization with a subsequent new interval of relative slowdown.

It is worthy of note that formal indicators of correlation for model values of these two variables (R = 0.95, R^2 = 0.903) are almost identical to those for correlation calculated by us (see note to Fig. 4) for empirical values of these variables.

Perhaps, it is not a pure coincidence, as this correlation both in our model and in reality in many respects was generated by similar mechanisms. Thus, the impact of the barbarian periphery on the central ('civilized') territories of the World System in the 1st millennium CE led to noticeable barbarization and de-urbanization of many of these zones as a result of which the World System urbanization index considerably decreased. At the same time, the rates of technological growth (both in our model and in reality) decreased, but technological growth did not stop completely since at that time new inventions and discoveries continued to be made (especially in those zones of the World System which underwent de-urbanization least of all). As a result, at that time the decrease in the level of urbanization was followed by some (albeit decelerated) growth of the level of its technological development, which caused an apparent negative correlation reducing general level of the general positive correlation between the variables in question.

Note that, in our computer simulation, the final phases of conquest of the barbarian periphery by civilization are not deprived of some dramatism. The matter is that, not long before the end of the full conquest of the barbarian periphery by civilization, its asabiyyah falls to zero ('dizziness with success'?) which leads to a fast and powerful counterattack of the barbarian periphery, burst of civilization asabiyyah and even more rapid final submission of the barbarian periphery (Fig. 15):

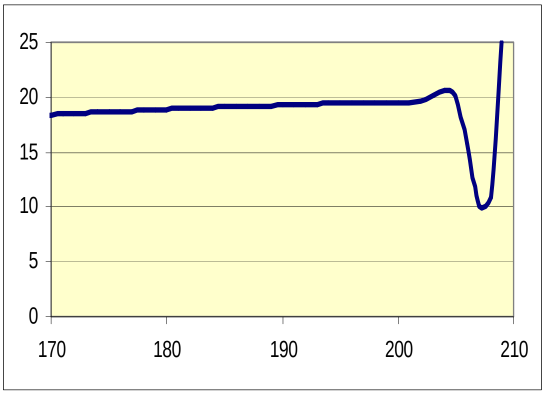


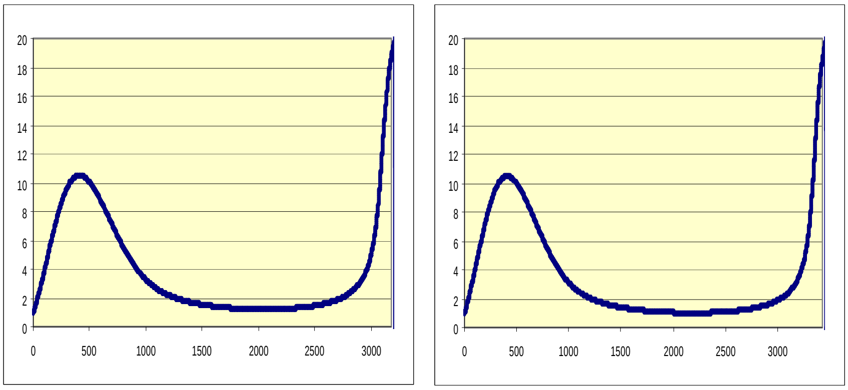
Fig. 15. Dynamics of the civilization territory before the final conquest of barbarian periphery (years 3170–3209 of our computer simulation)

In our opinion, quite a probable scenario of events is modeled here; however, fortunately, it does not appear to have been ever realized in real history of the World System (though the events of 9/11, perhaps, can serve here as a slightly resembling analogue).

* * *

The numerical study of influence of parameter values on the dynamics of our model shows that the key parameters determining the length of phases are as follows: the coefficient of innovative activity (*h* in Eq. 8) defining the rates of technological growth; the coefficient of territorial expansion (*a* in Eq. 1); the coefficient of borrowing technologies of civilization by 'barbarians' (*k* in Eq. 17) and in particular military technologies (*n* in Eq. 19); the coefficient of dynamics of *asabiyyah* (*e* in Eqs 4 and 5), and also the relationship between the coefficient of military participation of barbarians (*b*) and civilization (*c*).

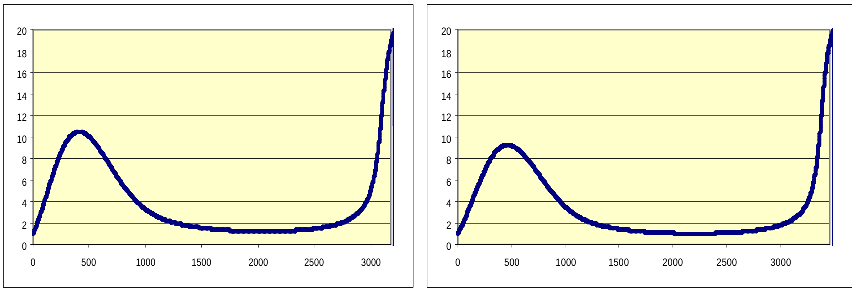
Small reduction of the value of coefficient *h* with respect to the value mentioned in the note to Figs 6–8 leads to some reduction of duration of phases 1–3 and significant increase in duration of other phases, but first of all of Phase 6 ('relative equilibrium of forces'); with reduction in coefficient of technological development (*h*) the rates of historical development are slowing down (see Fig. 16).



a) h = 0.0000315 **b)** h = 0.0000310

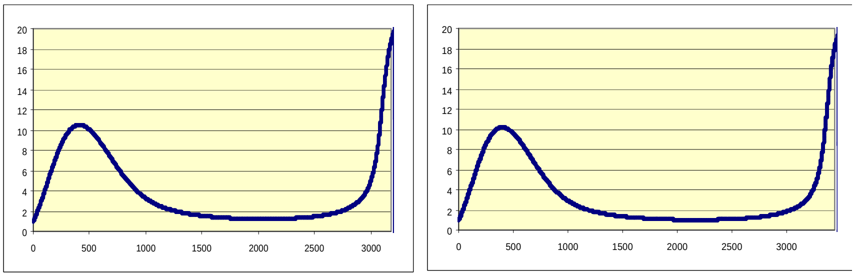
Fig. 16. Dynamics of the territory of the World System civilizational center generated by model (millions km^2) with a small decrease in the coefficient of technological growth (*h*)

Small reduction of the coefficient of territorial expansion (*a*, Fig. 17), or small increase in values of the coefficient of asabiyyah dynamics (*e*, Fig. 18), coefficient of borrowing of civilization technologies by 'barbarians' (*k*, Fig. 19) and, in particular, military technologies (*n*, Fig. 20), and also small increase in gap between the coefficients of military participation of population of the barbarian periphery and civilizational core (*b/c,* Fig. 21), small reduction of initial values of the level of technological development of civilization (T_{c0}, Fig. 22), its territories and asabiyyah, and also small increase in the initial values of the level of technological development, territory, and asabiyyah of the barbarian periphery produce similar results.



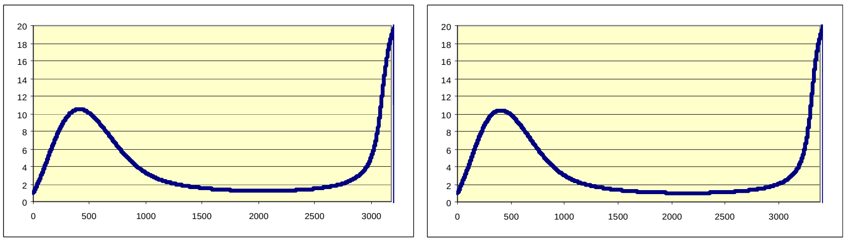
a) $a = 0.0120$ **b)** $a = 0.0094$

Fig. 17. Dynamics of the territory of the World System civilizational center generated by model (millions km^2) with a small decrease in the coefficient of the territorial expansion (a)



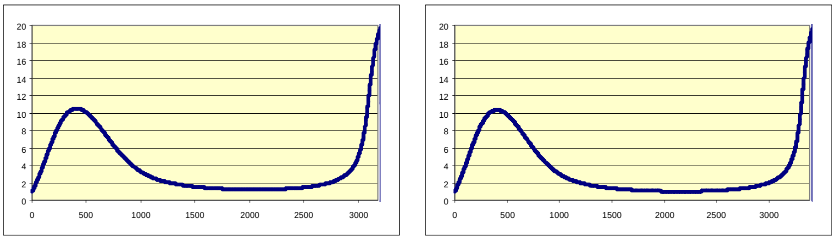
a) $e = 0.0520$ **b)** $e = 0.0528$

Fig. 18. Dynamics of the territory of the World System civilizational center generated by the model (in millions km^2) with a small increase in the coefficient of asabiyyah dynamics (e)



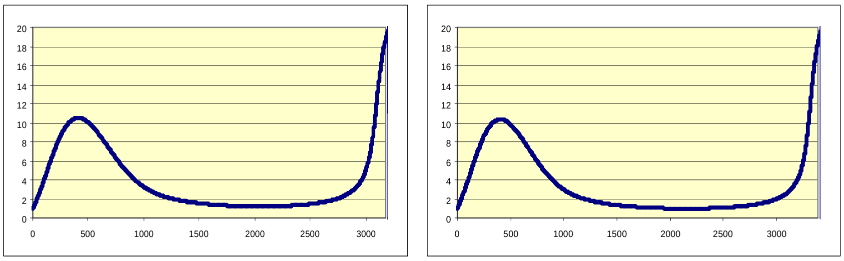
a) $k = 0.000504$ **b)** $k = 0.000517$

Fig. 19. Dynamics of the territory of the World System civilizational center generated by the model (in millions km^2) with a small increase in the value of overall coefficient of the borrowing of civilization technologies by 'barbarians' (k)



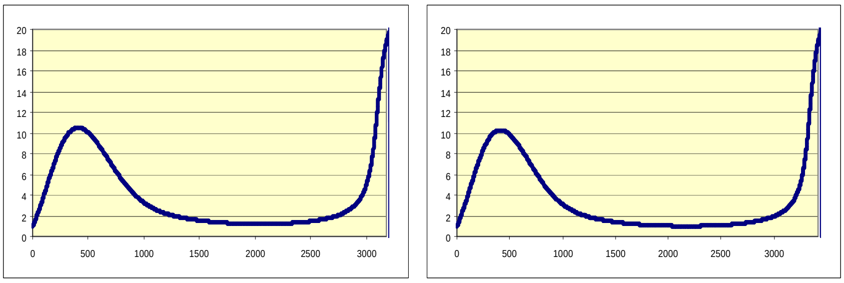
a) $n = 0.000504$ **b)** $n = 0.000525$

Fig. 20. Dynamics of the territory of the World System civilizational center of generated by the model (in millions km^2) with a small increase in the value of the coefficient of borrowing of civilization military technologies by 'barbarians' (n)



a) $c = 0.05$; $b = 0.2000$ **b)** $c = 0.05$; $b = 0.2027$

Fig. 21. Dynamics of the territory of the World System civilizational center generated by the model (in millions km^2) with a small increase in the gap between the military participation ratio of barbarian periphery and civilizational core (b/c)

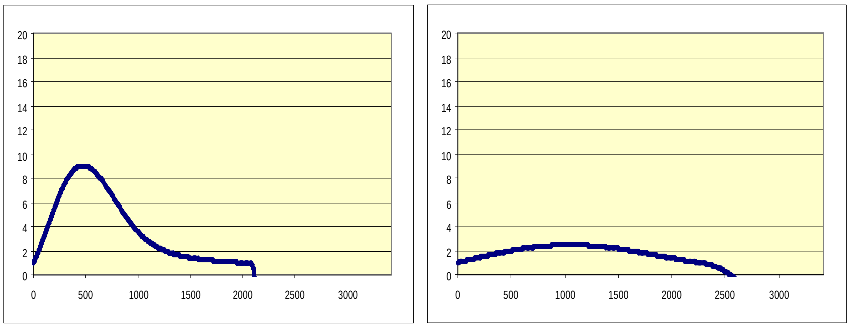


a) $T_{c0} = 10$ **b)** $T_{c0} = 9.9$

Fig. 22. Dynamics of the territory of the World System civilizational center generated by the model (in millions km^2) with a small reduction of initial value of the level of civilization technological development (T_{c0})

The above described small changes of parameters and initial conditions of the model lead to the reduction of sizes of territory to which control of civilization extends during the first wave of its expansion (Phases 1–2.5), and also to increase in duration and intensity of counterattack of the barbarian periphery (Phases 2,5–5,5). As a result, the zone remaining under control of civilization at the maximum of barbarian counter-expansion is reduced. The further change of parameters and initial conditions in this direction leads to significant changes of the overall picture of dynamics and implementation of a significantly different scenario. If (as a result of the barbarian expansion) the civilization zone is reduced to the level below 1 million km^2, it means that in the respective model simulation the barbarian periphery manages to conquer a part of the nuclear civilization zone (Zone 1) with an especially high natural productivity. As one can see in the model simulation, this leads to a very pronounced strengthening of barbarians (even at the phase when their counterattack approaches its exhaustion) and to a very serious weakening of civilization. In our simulations, the civilization could only launch a counterattack if barbarians managed to take no more than 1–2 % of Zone 1. Otherwise sharp strengthening of barbarians together with sharp weakening of civilization leads to a new acceleration of barbarian expansion and rapid final conquest of civilization by barbarians. Thus, a new phase is added (Phase 6' – the phase of the new acceleration of barbarian expansion and final conquest of the civilization by barbarians), and Phases 6.5–7 disappear.

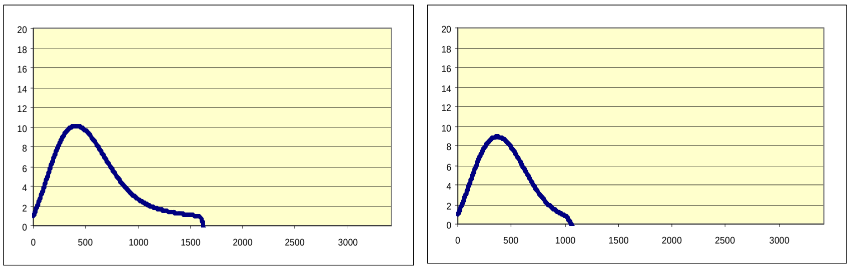
Further reduction of the coefficient of territorial expansion (a) reduces the size of the territory which is under control of civilization at the peak of its territorial expansion, but at the same time postpones 'barbarian occupation', prolonging the life of civilization (see Fig. 23).



a) $a = 0.009$ **b)** $a = 0.001$

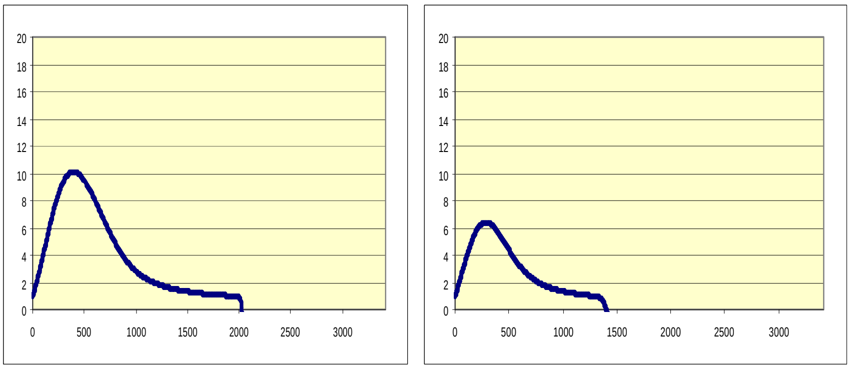
Fig. 23. Dynamics of the territory of the World System civilizational center generated by the model (in million km^2) at considerable decrease in the coefficient of territorial expansion (a)

On the other hand, further reduction of the coefficient of innovative activity (*h*) or increase in the values of the coefficient of asabiyyah dynamics (*e*), the coefficient of borrowing technologies of civilization by 'barbarians' (*k*) (and, in particular, military technologies (*n*)), as well as further increases in the gap between coefficients of military participation of population of the barbarian periphery and civilizational core (*b/c*), further reduction of the initial values of the level of technological development of civilization (T_{c0}, Fig. 22), its territory and asabiyyah, and also further increase in initial values of the level of technological development, territory, and asabiyyah of the barbarian periphery lead to reduction of the 'life of civilization', to its more rapid conquest by barbarians (Figs 24–30):



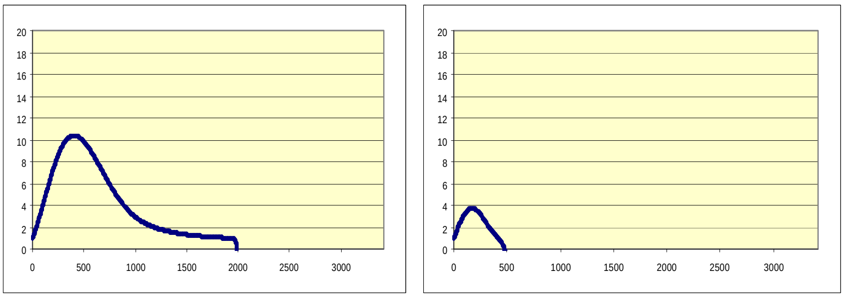
a) *h* = 0.000030 **b)** *h* = 0.000010

Fig. 24. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a considerable decrease in the coefficient of technological growth (*h*)



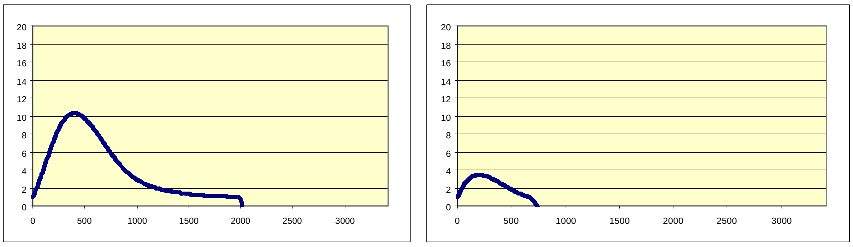
a) *e* = 0.0530 **b)** *e* = 0.07

Fig. 25. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a significant increase in the coefficient of asabiyyah dynamics (*e*)



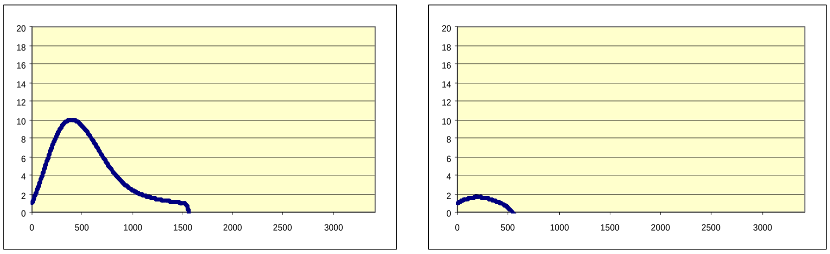
a) $k = 0.000520$ **b)** $k = 0.005000$

Fig. 26. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a significant increase in the value of general coefficient of borrowing of civilization technologies (*k*) by 'barbarians'



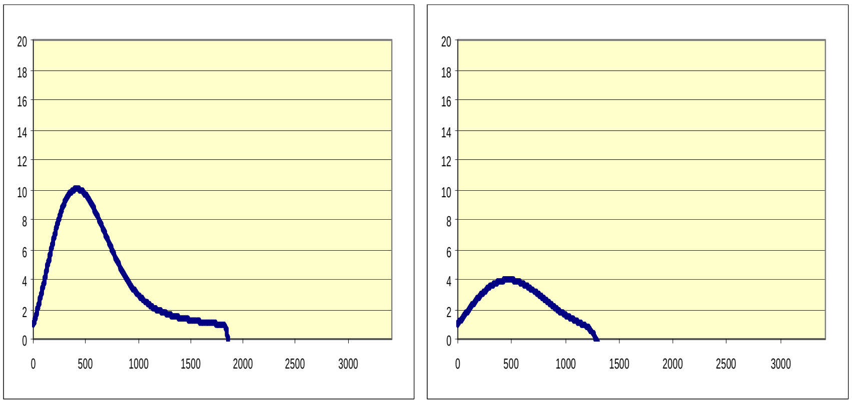
a) $n = 0.00053$ **b)** $n = 0.01000$

Fig. 27. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a significant increase in the value of coefficient of borrowing of military technologies of civilization by 'barbarians' (*n*)



a) $c = 0.05$; $b = 0.21$ **b)** $c = 0.02$; $b = 0.25$

Fig. 28. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a significant increase in the gap between military participation ratios of the population of barbarian periphery and civilizational core (*b*/*c*)



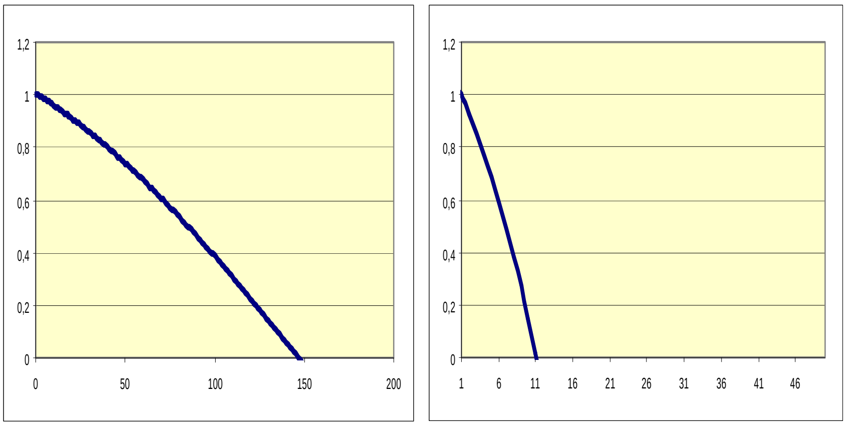
a) T_{c0} = 9,8 **b)** T_{c0} = 7,0

Fig. 29. Dynamics of the territory of the World System civilization center (in millions km^2) generated by the model with a considerable reduction of the initial value of the level of technological development of civilization (T_{c0})

It is obvious that with such values of parameters the model describes quite a real scenario. Indeed, with a certain set of parameters the expansion of civilization could create such a powerful barbarian periphery that its counterattack could be able to destroy this civilization (the classic example here is the conquest by barbarians of Rome, expansion of which in many respects gave barbarians that very strength which eventually helped them to break down their formidable opponent [see, *e.g.*, Turchin 2005]).

Beyond a certain limit we get a scenario of more and more rapid conquest of civilization by barbarians already without a phase of the initial civilizational expansion (see Fig. 30).[27]

[27] We should also note that the change of parameters of the model in this direction is after all meaningful only to a certain degree. Say, the coefficient of military participation of barbarians (*b*), cannot by definition exceed 1.0, but even 1.0 describes an unrealistic scenario because the entire population of any society (including, as we know, newborn babies and very old men and women) cannot actually take part in combat. Similar restrictions also exist for all other parameters and initial conditions of the model.



a) $c = 0.01$; $b = 0.4$ **b)** $H_{b0} = 2$

Fig. 30. Dynamics of the territory of the World System civilization center generated by the model (in millions km²) with a large gap in coefficients of military participation and very high initial value of 'barbarians' asabiyyah

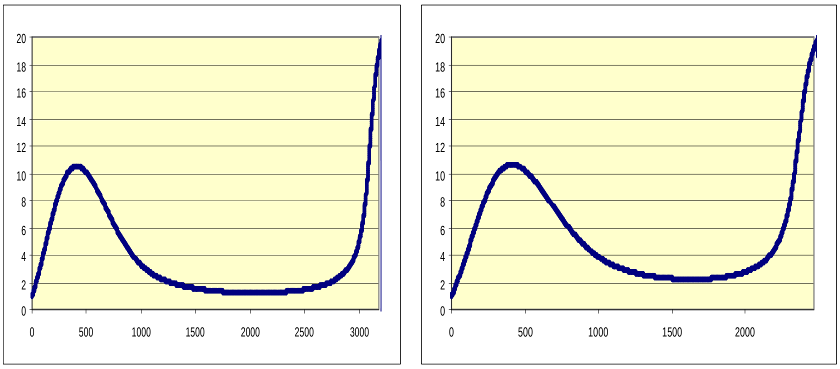
In this case we have already a rather banal scenario of the final conquest of civilization by barbarians whose military superiority has already been determined by initial parameters and initial conditions, while civilization (with the same parameters and initial conditions) has no chance to effectively resist the barbarian expansion.[28]

In case of little change of parameters and initial conditions in the opposite direction[29] (in relation to the parameters and initial conditions described in the note to Figs 5–7) we will see in our computer simulations some increase in du-

[28] At the same time, say, if the conquest of civilization by 'barbarians' within the model is brought about by assigning to 'barbarians' a higher initial value of asabiyya *(H)*, it can already be interpreted within Ibn Khaldun's tradition (see, *e.g.*, Ibn Khaldun 1958, 2004; Batsieva 1965; Ignatenko 1980; Alekseev and Khaltourina 2004; Turchin 2003, 2007; Korotayev and Khaltourina 2006; Korotayev 2006e, 2007c, 2007d; Inan 1933; Mahdi 1937) as the conquest of a low-assabiyyah civilization by high-asabiyyah barbarians, which does not 'end the history', but begins its new round, a new 'Khaldunian' dynastic cycle, but this will be quite a different model.

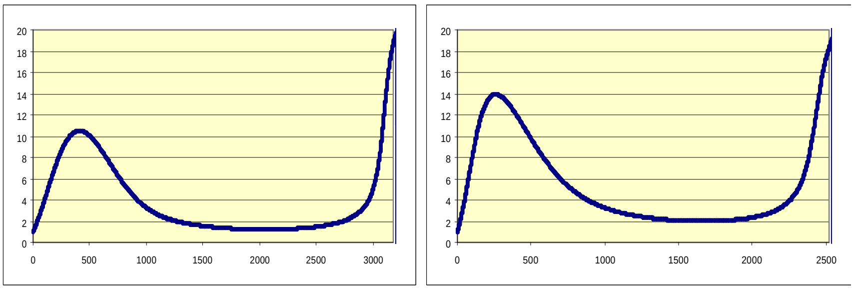
[29] That is with a small increase of the coefficient of technological growth (h) and the coefficient of territorial expansion (a), or with a small reduction of the values of the coefficient of asabiyyah dynamics (e), the coefficient of borrowing of civilization technologies by 'barbarians' (k) and, in particular, military technologies (n), and also with a small decrease of the gap between military participation ratios of the barbarian periphery and civilizational core (b/c), with a small increase in the initial values of the level of technological development of civilization (T_{c0}, Fig. 22), its territory and asabiyyah, and also with a small reduction of the initial values of the level of technological development, territory and asabiyyah of the barbarian periphery.

ration of phases of the initial expansion of civilization, the territory occupied by it during expansion (as well as the territory retained by it at the peak of 'barbarian' counterattack), and reduction of duration of all subsequent phases, and in particular Phase 6 ('relative balance of forces') (Figs 31–37):



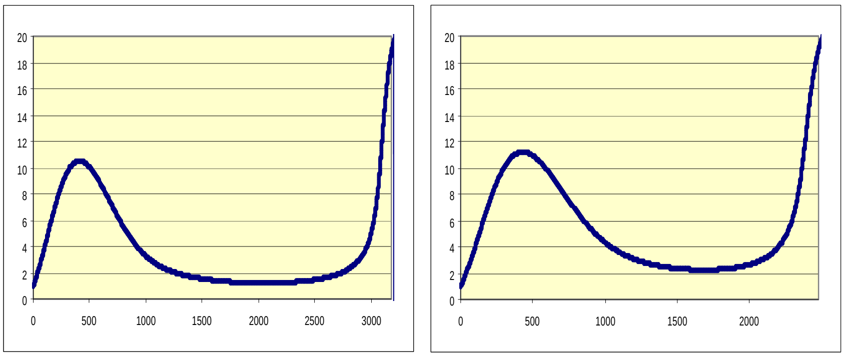
a) *h* = 0.0000315 **b)** *h* = 0.0000330

Fig. 31. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a small increase in coefficient of technological growth (*h*)



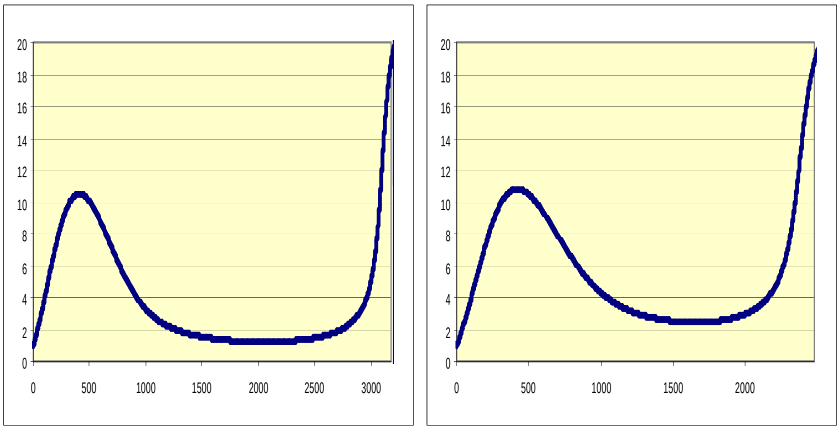
a) *a* = 0.0120 **b)** *a* = 0.0250

Fig. 32. Dynamics of the area of the territory of the World System civilization center generated by the model (in millions km^2) with a small increase in the coefficient of territorial expansion (*a*)



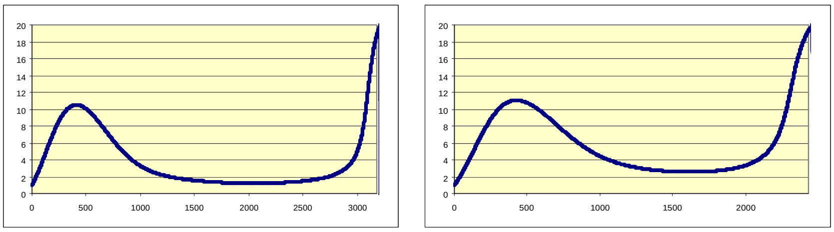
a) $e = 0.0520$ **b)** $e = 0.0505$

Fig. 33. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a small reduction of the coefficient of asabiyyah dynamics (*e*)



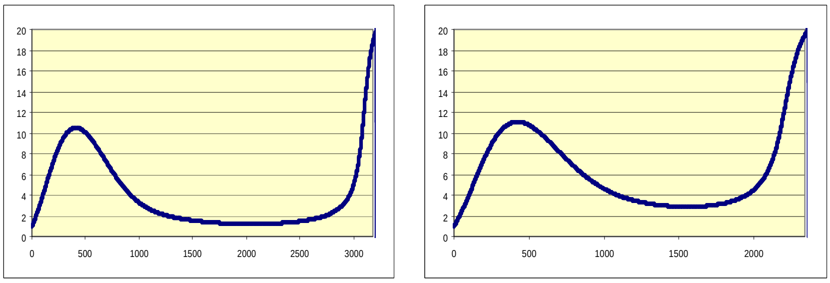
a) $k = 0.000504$ **b)** $k = 0.000465$

Fig. 34. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a small reduction of the value of the general coefficient of borrowing of civilization technologies by 'barbarians' *(k)*



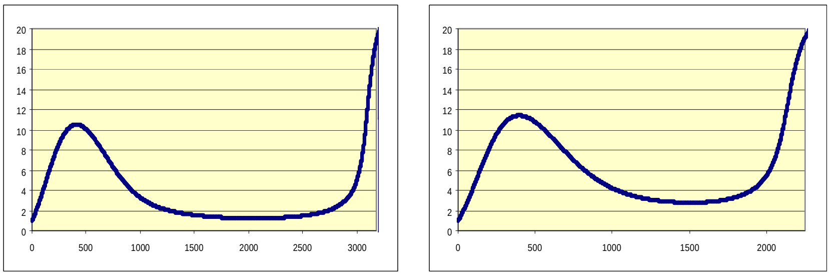
a) $n = 0.000504$ **b)** $n = 0.000440$

Fig. 35. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a small reduction of the value of the coefficient of the borrowing of military technologies of civilization by 'barbarians' (n)



a) $c = 0.05$; $b = 0.20$ **b)** $c = 0.05$; $b = 0.19$

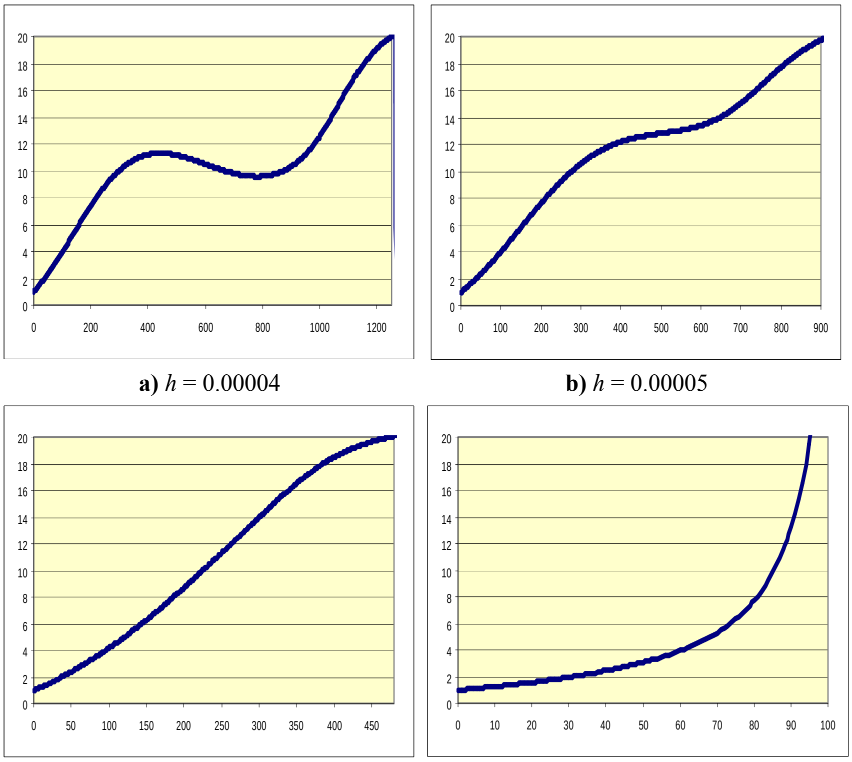
Fig. 36. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a small reduction of gap between military participation ratios of barbarian periphery and civilizational core (*b/c)*



a) $T_{c0} = 10$ **b)** $T_{c0} = 10.4$

Fig. 37. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a small increase in the initial value of the level of technological development of civilization (T_{c0})

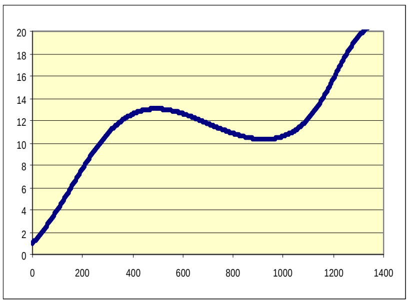
In case of further changes of parameters in this direction (and, in particular, with their combined changing) one can observe a further increase in duration of phases of the initial expansion of civilization, the area of the territory occupied by it during the first expansion phases (as well as the territory kept by it at the peak of 'barbarian' counterattack), and reduction of duration of all subsequent phases, and in particular Phase 6 ('relative balance of forces'). Gradually duration of phases 3–6 reduces to zero, and they disappear, whereas phases 1–2 eventually merge with phase 7. As a result, we have only phase 1, and further change of parameters in this direction leads only to reduction of the period of time which is required for civilization to conquer its barbarian periphery (Figs 38–41):

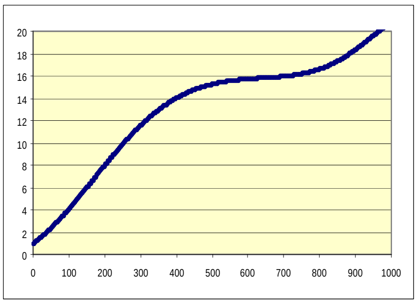


a) h = 0.00004 **b)** h = 0.00005

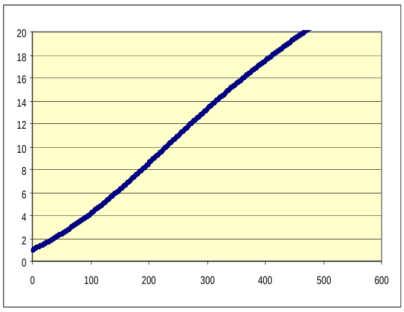
c) h = 0.00010 **d)** h = 0.00100

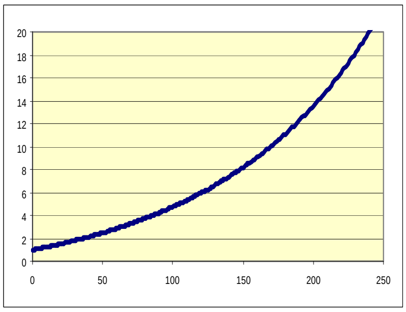
Fig. 38. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a significant increase in the coefficient of technological growth (h)





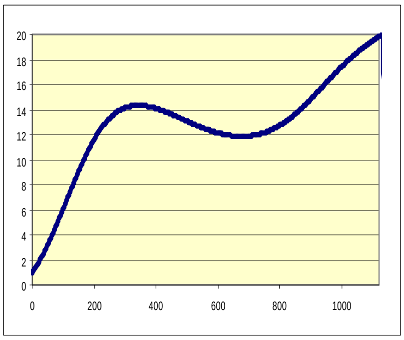
a) $e = 0.0475$ **b)** $e = 0.0450$

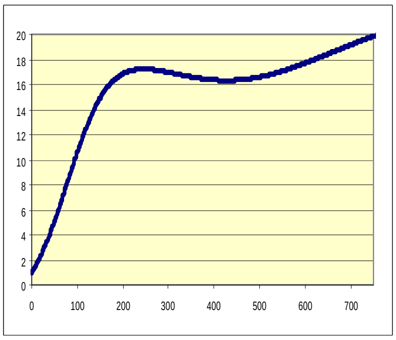




c) $e = 0.0400$ **d)** $e = 0.0100$

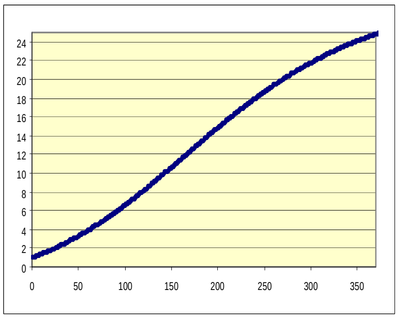
Fig. 39. Dynamics of the territory e of the World System civilization center generated by the model (in millions km^2) with a considerable reduction of the coefficient of asabiyyah dynamics (*e*)



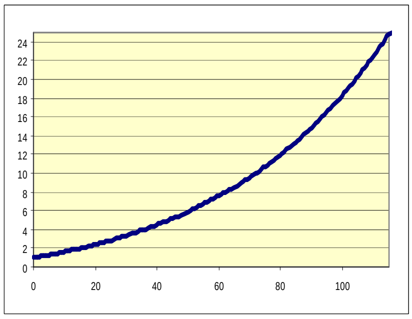


a) $T_{c0} = 12$ **b)** $T_{c0} = 15$

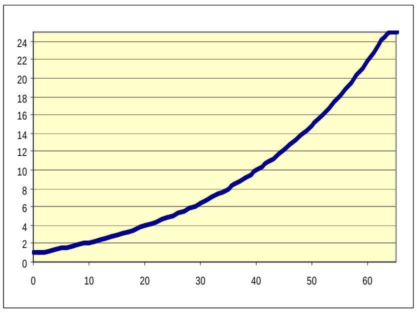
Fig. 40. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a large increase in the initial value of the level of technological development of civilization (T_{c0})



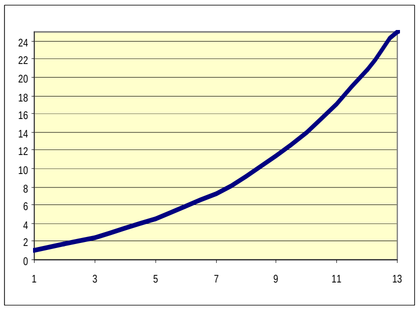
a) $e = 0.04$; $T_{c0} = 12$



b) $e = 0.01$; $T_{c0} = 15$



c) $e = 0.005$; $T_{c0} = 20$



d) $e = 0.001$; $T_{c0} = 50$

Fig. 41. Dynamics of the territory of the World System civilization center generated by the model (in millions km^2) with a simultaneous considerable reduction of the coefficient of asabiyyah dynamics (*e*) and a large increase in the initial value of the level of technological development of civilization (T_{c0})

Let us note that in all cases we deal with quite real scenarios of interaction between civilizational center and barbarian periphery. In fact if a technologically highly developed (and, in addition, rapidly developing) civilization came in contact with such a periphery which was extremely undeveloped technologically and incapable of borrowing technologies of civilization rather quickly, it led to the accelerated conquest by civilization of its periphery that was incapable to produce any effective resistance to advancing civilization (the British colonization of Australia presents a rather salient example of such a scenario).

Thus, depending on the given parameters and initial conditions the offered model can describe mathematically seven significantly different scenarios of the interaction between civilizational center and barbarian periphery (see Table 2):

Table 2. Scenarios of interaction between civilization center and barbarian periphery described by the model depending on the given parameters and initial conditions

Description of scenario	*Values of parameters and initial conditions*
1. Accelerated conquest of the barbarian periphery by civilization	Very high values of coefficients of technological growth (*h*), territorial expansion (*a*), very low values of coefficients of dynamics of asabiyyah (*e*), borrowing of civilization technologies by 'barbarians' (*k)* and, in particular, military technologies (*n*), very small gap between coefficients of military participation of population of the barbarian periphery and civilizational core (*b/c*), very high initial values of the level of technological development of civilization (T_{c0}), its territory (A_{c0}) and asabiyyah (H_{c0}); very low initial values of the level of technological development (T_{b0}), territory (A_{b0}) and asabiyyah (H_{b0}) of the barbarian periphery
2. The phase of the accelerated expansion of civilization is followed by the phase of its slowdown caused by more and more effective counteraction of 'barbarians' who, nevertheless, are not able to mount a counterattack	↓
3. The phase of accelerated and deep expansion of civilization and phase of its slow-down is followed by rather short-term and shallow counterattack of 'barbarians' which is succeeded by the phase of final submission of barbarian periphery	↓
4. The phase of still accelerating but less deep expansion of civilization and the phase of its slowdown is followed by the phase of a long and deep offensive of 'barbarians' which is succeeded by a long phase of relative balance of forces which is followed by the phase of final submission of the barbarian periphery by civilization	↓
5. The phase of less and less accelerating and deep expansion of civilization and the phase of its slow-down is followed by the phase of the long and deep offensive of 'barbarians' which is gradually slowed down by more and more effective counteraction of civilization which, nevertheless, is not able to prevent its complete conquest by barbarians	↓

Description of scenario	*Values of parameters and initial conditions*
6. The phase of rather shallow slowing down expansion of civilization is followed by the phase of accelerated advance of 'barbarians' which results in the conquest of civilization by the 'barbarians'	
7. As a result of a rapid, increasingly accelerating offensive, the barbarians conquer civilization	Very low values of the coefficient of technological growth (*h*) and the coefficient of the territorial expansion (*a*); very high values of the coefficient of dynamics of asabiyyah (*e*), the coefficient of borrowing of civilization technologies (*k*) by 'barbarians' and, in particular, military technologies (*n*); a wide gap between military participation ratios of the barbarian periphery and civilizational core (*b/c*); very low initial values of the level of technological development of civilization (T_{c0}), its territory (A_{c0}) and asabiyyah (H_{c0}); very high initial values of the level of technological development (T_{b0}), territory (A_{b0}) and asabiyyah (H_{b0}) of the barbarian periphery

Thus, the scenario displayed above in Figs 5–7 and most precisely describing the influence of interaction between the civilization center and barbarian periphery on the World System development is intermediate among the above-described scenarios. One can suppose that this is not a coincidence. In fact, there are some grounds to maintain that this interaction could have the historically attested impact only with intermediate values of parameters. Barbarians could not confront effectively the expansion of civilizational and launch a massive counterattack with very low values of coefficients *e, k, n* and *b*, *i.e.*, if their collective solidarity were not sufficiently amplified under the influence of pressure of civilization, if they had not been able to adopt vitally important technologies of civilization (including military ones) quickly enough, if they had not have a much higher military participation ratio. On the other hand, a too high value of these parameters would not have made the expansion of civilization of the World System possible, or would have even led (at especially high values of these parameters) to rapid conquest of civilization by barbarians before the civilization expansion could start. One can say the same about the values of all the other significant parameters of the model, and also about the values of the initial conditions (*e.g.*, if at the time of clash of civilization with the barbarian periphery the relative level of its technological development had been too low, it would not have been able to begin expansion at all, and with an extremely high value it would have rapidly subordinated the periphery that would be incapable of offering any effective resistance).

On the other hand, we can easily see that practically all the scenarios described above were observed in the history of the World System at the level of some specific civilization zones and their peripheries. In many respects this is precisely why the total dynamics of the World System appears to be the closest to the intermediate scenario of the model.

Of course, in our case the fit of theoretical curve with empirical estimates is still far from being ideal – which is not surprising since nobody would claim that the interaction between civilization core of the World System and its barbarian periphery was the only factor that defined the characteristic form of the curve of world urbanization dynamics.

What is really surprising for us, is the fact that the offered model after all appeared capable to describe the general form of this dynamics so precisely (though, of course, imperfectly). Hence, this suggests that the interaction between the civilization core of the World System and its barbarian periphery was really an important factor making a notable contribution (until very recently) to giving the peculiar form to the curve of the world urbanization dynamics.

The proposed model suggests that in the history of the World System development the important component of World System phase transitions A_1 and A_2 could be not only the movement of its core to a new level of complexity, but also the formation of barbarian periphery of an essentially new type capable to offer much more effective resistance to civilization expansion and to mount successful counterattacks that apparently could make a noticeable contribution to the formation of the 'attractor effect' during periods B_1 and B_2.

There are also some more points that seem to be capable to explain why this model which takes into account only one factor of the World System dynamics (and not always the most important one), was able to generate the curve so well describing empirically observed historical macrodynamics. The fact is that this model, most likely, describes general development logic of the World System within which the hyperbolic growth of its civilizational core creates powerful forces restricting this growth at certain phases of the World System development (or, to be more exact, moving the World System to a new, lower, hyperbolic trajectory) (see, *e.g.*, Korotayev 2006d, 2007; Korotayev, Malkov, and Khaltourina 2007; Grinin and Korotayev 2009b).

For example, one of such factors (which influenced dynamics of the World System in Phase B_2 probably, not lesser, if even not greater than the effect of interaction between civilization center and barbarian periphery) was earlier described by us as follows:

> The growth of the World System population by the end of the 1st millennium BCE up to 9-digit numbers produced a breeding ground that led to an almost inevitable appearance of a new generation of more lethal and epidemically destructive pathogens that could not reproduce themselves in smaller populations (Diamond 1999: 202–205; McNeill 1993; Koro-

> tayev, Malkov, and Khaltourina 2005: 105–113; 2006a), whereas the level of health care technologies achieved by the World System by the beginning of the 1st millennium CE turned out to be totally inadequate for the radically increased level of pathogen threat. Thus, the Antonine and Justinian's pandemics led to global depopulations of the 2nd and 6th centuries, contributing in a very significant way to the slowdown of the World System demographic growth in the 1st millennium CE. Note that due to this, since the early 1st millennium CE the role of health care technologies as a determinant of the carrying capacity of the Earth dramatically increases, which at least partly accounts for the change of the hyperbolic growth regime' (Korotayev, Malkov, and Khaltourina 2006b: 159; 2007: 206).

It is very important that the logic of action of this factor is extremely similar to the logic of action of the factor analyzed in this article. In the both cases, the hyperbolic growth of civilization creates powerful forces that block this growth. On the other hand, the pressure of the barbarian periphery was able to stimulate the growth of military potential of civilization, whereas pathogenic attacks on the World System eventually stimulated development of health care technologies which allowed the World System to repulse these attacks more successfully and renew its hyperbolic growth.[30]

The logic of this factor is similar to the logic of another factor (which, apparently, had an even greater influence on the World System dynamics World System in phase B_1,than the effect of interaction between the civilizational center and barbarian periphery) – the factor of environmental degradation under the influence of the hyperbolic growth of civilization. This factor is most evident in the history of ancient Mesopotamia (whose curve of urbanization dynamics in the 4th – early 2nd millennia BCE defined the general shape of the World System urbanization dynamics curve to a very considerable extent). As is well known, the explosive growth of civilization in this region led to catastrophic soil salination in its most developed zone, in Lower Mesopotamia, which, in turn, led to decrease in rates of demographic and urbanization growth here up to negative values in the middle of the 3rd millennium. On the other hand, environmental degradation stimulated the technological growth here which in the 1st millennium CE led to the restoration of the carrying capacity values here up to the levels of the beginning of the 3rd millennium BCE, and then to its noticeable expansion (Dyakonov 1983: 272, 330; Chubarov 1991; Roberts 1998: 175).

[30] It appears appropriate to recollect at this point that this growth continued up to the early 1970s when the World-System started to withdraw from the hyperbolic growth regime (*i.e.*, the blow-up regime) due to force created by its hyperbolic growth, but these were forces of another kind (Korotayev, Malkov, and Khaltourina 2006a, 2006b, 2007).

These facts explain to some extent a paradox that can be noticed above in Figs 1–2, 5–7, and 12–14. The point is that in order that the interaction of civilization center and barbarian periphery could produce an effect that is similar to the one that is actually observed, the territory under the control of civilization should be decreased manifold and remain at an extremely low level for a very long period of time, which apparently did not happen in reality. Apparently, what was not 'consumed' by 'barbarians' at 'sinks' in Figs. 1–2 corresponding to the phases B_1 и B_2 was mostly 'eaten away' additionally by pathogenic attacks and environmental degradation.

Finally, let us dwell upon some other factors that seem to be relevant for modeling of the long-term dynamics of the World System. One of these factors was described by us earlier as follows:

> Some hint here seems to be suggested by mathematical models (0.11)-(0.13)-(0.12) and (0.13)-(0.14)[31] described in the Introduction. According to these models any long-term decrease of per capita surplus (*S*) must lead to the decrease of population growth rates and, hence, the slow-down of technological growth. In the meantime, by the end of the Axial Age we seem to observe a World System trend towards the decline of precisely this indicator. This was connected not with decline of production, but rather with the growth of *m*, the per capita product that is necessary for the population reproduction with zero growth rate, the 'minimum necessary product' (MNP). In the 1st millennium BCE the rapid population growth sustained the hyperbolic growth of the complexity of sociopolitical infrastructures (on the other hand, of course, the hyperbolic population growth was also sustained by a hyperbolic growth of sociopolitical complexity – once more we are dealing here with the positive feedback phenomenon). However, the radical increase in sociopolitical complexity meant a radical increase in the MNP, as the substantial expenses necessary for the normal functioning of these sociopolitical infrastructures should be regarded, in this context, as a part of the minimum necessary product (rather than as surplus). Indeed, by the end of the 1st millennium BCE the World System population reached 9-digit numbers; even a simple reproduction (at zero growth rate) of so huge a population required maintenance of normal functioning of all those infrastructures (transportation, judicial, administrative and other such subsystems). Within such a context, if the product produced this year by a peasant is only sufficient to secure the survival of himself and his household, but not sufficient to pay any taxes, it is impossible to say that this peasant has produced this year the minimum necessary product. In fact, what he has produced this year is smaller than the MNP. Indeed, as the experience of post-Axial centuries showed on numerous occasions, in supercomplex agrarian societies the decrease of per capita

[31] These numbers refer to equations presented in the Introduction (Korotayev, Malkov, and Khaltourina 2006b).

production (usually as a result of relative overpopulation) down to a level that did not allow the population to pay taxes led to the disintegration of sociopolitical infrastructures and demographic collapse (see Korotayev, Malkov, and Khaltourina 2006b: Chapters 1–4[32]). There are grounds to maintain that the rapid growth of the MNP in the 1st millennium BCE exceeded the growth of the equilibrium per capita production, which resulted in the long-term decrease of real *S*, and, hence, the decrease of the World System population growth rates. On the other hand, it led to the decrease of sociopolitical system stability, and, hence, to the increase in the importance of the role of cyclical and chaotic components of the macrohistorical dynamics in comparison with the trend component (*Ibid.*: 159–160; 2007: 206–207).

Obviously, in this case we also deal with a force that was created by the hyperbolic growth of the World System and that blocked for some time the further hyperbolic growth of the World System.

In conclusion, we would like to make a few additional comments. We are well aware of the fact that many assumptions of our model simplify (sometimes even over-simplify) the situation observed in historical reality. It is worth pointing out some of such assumptions and points which should be developed, systematized and considered in future generations of models of interaction between civilization center and barbarian periphery.

First of all, one should note that one of the strongest simplifications of the model was that the World System consisted of only one civilization and only one barbarian periphery (though in history we naturally deal with a number of civilizations and barbarian peripheries surrounding them). Taking into consideration the multiplicity of civilizations may become one of the leading directions of further development of our model. It could also be possible to present some typology of both civilizations, and barbarian peripheries and to use this typology for explanation of characteristic features of the World System dynamics both within basins of attraction, and during phase transitions.

For example, it would be worth dividing barbarian periphery into two types. The first one is represented by barbarians-agriculturalists; the second one is by barbarians-herders (nomads). This point is important with respect to the level of the world urbanization, because nomadic population constituted a rather small fraction of the global population, whereas barbarians-farmers could constitute a substantially higher percentage (*e.g.*, according to some estimates, the population of Gaul before the Roman conquest was between 5 and 10 million [see, *e.g.*, Braudel 1995: 61–62]). Note that the transition of nomads, semi-nomads and extensive farmers to settled intensive agriculture appears to have greatly influenced the population dynamics of the World System. Apparently, period B_1 is

[32] See also, *e.g.*, Nefedov 2001, 2002a, 2002b, 2003, 2004, 2005, 2007, 2013; Turchin 2003: 121–127; Turchin and Nefedov 2009.

characteristic of such transition (in particular, in India among Indo-Aryans, in Iran, among the Dorians in Greece, in a number of other areas of Europe, *etc*.). In other words, transformations of barbarian periphery could significantly change proportions in the world population and its urbanization due to population growth of the barbarian periphery (which is considered only partially in the present version of the model).

At the initial stage (the first two thirds of the 1st millennium CE), period B_2 was connected with absorption by the civilization center of the huge number of barbarians who intruded in its territory. However, it appears necessary to take here into account the difference between complex agrarian societies of B_1 period and supercomplex agrarian systems of B_2 period in respect of the economic role of cities.

In typology of civilization centers it appears important to distinguish between irrigation and non-irrigation civilizations. It is very important due to significant differences in the processes of their urbanization. In particular, three functions of cities in the regions of river civilizations were more developed than in other places: economic, redistributive and sacral. And that fact made the cities actually a part of agrarian technology within civilization whereas cities could not act as such an integral element elsewhere. Moreover, the politically centralizing role of the cities in irrigation civilizations was also higher, and non-irrigation agricultural civilizations had more opportunities to remain decentralized without great losses for efficiency of functioning of their economy. As a result, the elite of irrigation civilizations was to a greater extent urban, and the elite of non-irrigation civilizations might be rural to a greater extent (*e.g.*, in the medieval Islamic Middle East the elite was generally urban, and in medieval Europe it was rural to a greater degree). Thus, the character of civilizations and barbarian periphery at different stages might influence the level of urbanization.

Distinguishing types of civilizations and barbarian peripheries could help us to achieve a significant clarification of typical reasons and variants of military rivalry between civilizations and barbarians. In particular, it is possible to preliminarily outline the reasons which forced civilizations and states to make the expansion to the barbarian periphery (the reasons of attacks of barbarians against civilizations will be discussed later):

1) Attempts to eliminate dangerous centers of constant concern and attacks. It was one of the main reasons of Chinese campaigns against Xiongnu, Russian campaigns against Cumans, *etc.*

2) Attempts to bring back territories occupied by barbarians.

3) Campaigns to seize slaves and booty (were characteristic of a number of African states). This type of campaigns includes the attacks for the purpose of getting tribute and necessary goods, strategic raw materials, *etc*. For example, the expansion of Russians to the North and Siberia was largely determined by their need for furs.

4) Campaigns with hegemonic purposes – 'in pursuit of power'. It was an important reason for which the Persian king Darius I tried to conquer Scythians.

5) Campaigns and conquests with some strategic aim (improvement of the situation, acquirement of favorable and convenient means of communication, elimination of potential danger, *etc*.). Among such cases one can mention conquests of Gaul and Dacia by Romans which were triggered by a special political situation, the involvement of these people into complex political game, attempts of their threat to Romans or their allies, *etc*. In particular, a cause for Caesar's campaign to Gaul, as we know, was that the Sequani, which suffered defeat from the Aedui, called Germans (Suevi and others) for help in campaign led by Ariovist who not only successfully defeated the Aedui, but also began to enslave the Sequani. These events served as an important pretext for Romans to interfere in affairs of the Gauls (see: Caesar. The Gallic War I: 31–39). Conquests which were made under such circumstances were also very characteristic of Europeans of the Modern Period.

6) Seizure of agricultural lands. One of examples are Charlenmagne's campaigns against the German and other barbarians, another one is the expansion of the German knights to the Baltic States (a motive to seize lands for agricultural purposes was especially salient in this case). One can add the expansion of Carthage to Corsica, Sardinia and to Spain. It appears appropriate to notice in all these cases we deal with the expansion to the territories occupied by agriculturalists. And therefore it is very important to take into account (as this is done in our model) that the territory of the barbarian periphery was divided, at least, into two types:

- the territory more economically attractive to the civilization;
- the territory less economically attractive.

It is obvious that expansion to the second type of territories was not important for civilization in itself (this only became important when this periphery disturbed it). Not without reason China could refrain from such expansion for a long time.

On the other hand, the systematic transition to intensive non-irrigation agriculture (that was observed just in the 1^{st} millennium BCE) might fully strengthen such expansion in its different types (for example, such was the case with Greek colonization which can be considered as one of the types of non-military or partly military expansion). In general, if the expansion of civilization to the barbarian periphery was in any way successful and prolonged, then as a rule it led to assimilation of barbarians.[33] One should note that in future models it would be worthwhile introducing mathematical description of processes of barbarians' assimilation in a more explicit way.

---

[33] However, civilization could not assimilate those barbarians who lived in marginal (unsuitable or almost unsuitable for agriculture) zones.

Probably, expansion of civilization to the periphery was not vital for a number of civilizations and played a subordinate role for them. However, such expansion was very important for the World System as a whole, which is shown in our model.

One should also define different types of expansion of civilization to the barbarian periphery. For example, it is possible to speak about assimilating expansion which might happen without essential resistance of the peripheral peoples. On the other hand, the more lands the civilization was bringing into its economic turnover, the more often it faced a more resistent and intractable (and at the same time less and less attractive economically) barbarian periphery.

Probably, in future models it would make sense to include mathematical description of the expansion of some barbarians to the territory of others. If the lands of nomads are often economically unattractive for civilizations, they are almost always economically attractive to other nomads. But nomads can also attack barbarians-farmers (as it happened with the Huns who attacked the Goths). And such kind of expansion quite often causes great changes by the 'domino' principle leading to general expansion of barbarians against the civilization center.

We have already mentioned above that the issue whether barbarians need or need not centralization for their successful expansion against the civilization has no unambiguous solution.

Successful wars and especially conquests of neighboring states by barbarians were quite often connected with successful centralization of barbarians (at least temporary) around some leader. It is relevant for the Xiongnu, the Mongols, the German tribes, the Huns, and many others. Sometimes as a result of these perturbations the center strengthened and a large chiefdom emerged. However, if this centripetal movement was insufficiently steady to become permanent, the life cycle of a new large polity was short. Such fragile formations as the Slavic Samo'state' (Lozny 1995: 86–87), the Germanic tribal unions of Maroboduus (among the Markomanni), Ariovistus (among the Suevi), Arminius (among the Cherusci), Claudius Civilis (among the Batavians) (Neusykhin 1968: 601–602; Oosten 1996); the Hunnish 'empire' of Attila (Korsunsky and Günther 1984: 105–116); the Getae and Dacians union led by the 'king' Burebista (Fedorov and Polevoy 1984) and others usually disintegrated after the death of the charismatic leader (and sometimes during his life as it happened to Maroboduus). Sometimes one could observe the decline of the supreme power in analogues even before the death of such a leader, especially if there was a strong and self-willed elite.

While considering the interaction between barbarians and civilization one should take into account the ambiguity of the solution of the issue of military and technological superiority of the latter over the former. Anyway, such a superiority was not always sufficient as barbarians could compensate it by suddenness or other advantages. Before the wide diffusion of iron, civilizations having expensive and highly technological weapons (*e.g.*, bronze arms) could not always successfully resist barbarians either, especially if civilizations were internally weakened.

It is also worth mentioning the features of barbarian political formations in comparison with the states of civilization zone.

Of particular interest are barbarian analogues of the early state where the political organism did not have rigidly fixed territory, more precisely, where socio-political organism can change its territory rather easily. Of course, it is much less characteristic for the states where the state control over a certain territory is almost obligatory. The states seldom change the core of their territory.[34] No matter how the borders of states change, the core usually remains the same, whereas barbarians, for example the Hungarians or the Goths (Shchukin 2005, *etc.*), were able to move thousands of kilometers 'in search for their home'. In some cases, an important cause for such migrations was constituted by population pressure. 'A sharp increase in population density in settled agricultural societies is well known, compared to the epoch of hunter-gatherers. It increased almost a hundredfold' (Masson 1976: 102–104, 189; 1980: 182–183). The importance of demographic growth in increasing role of wars in the relations between societies is great. And wars can also lead to the development of new political forms (*e.g.*, Grinin andKorotayev 2012; Turchin 2015). Not without a reason Robert Carneiro constantly emphasizes that increased population pressure can lead to wars and conquests, therefore state organization emerges in certain cases and under certain circumstances (Carneiro 1970, 1978, 2000a, 2002; 2006, 2012; see also: Lewis 1981). High population pressure quite often caused migrations and wars even among more populous state analogues. One of the most known examples is the huge polity of the Visigoths which suffered from relative overpopulation already under Hermanaricus in the middle of the 4th century CE (Shchukin 2005: 219), which was the most important reason for their migration to Byzantium (note that this episode is often considered as the beginning of the Great Migration Period).

One should also take into consideration the point that barbarian periphery being more poor and backward aspires to the robbery of civilization to a greater extent, than vice versa. Arnold J. Toynbee (1991) (somewhat simplifying the reality) said that 'external proletariat', *i.e.* barbarian peoples, posed threat for civilization having 'nothing to lose' just as Marx' proletariat, but they could gain lots of assets by attacking the civilization. This, of course, strengthened the barbarians' pressure.

[34] One of such extremely rare examples was represented by the South African Dutch who moved their states far away from the British in 1836–1839 (see, *e.g.*, *Büttner* 1981: 189–190).

It also appears appropriate to mention the role of wars in the life of barbarian societies. As some researchers note, wars among hunter-gatherers were observed a bit less frequently than among barbarians (Lesser 1968: 94; Korotayev, Malkov, and Khaltourina 2006b; Korotayev, Komarova, and Khaltourina 2007: 143, 148). There is also a certain significant correlation between extreme living conditions and a low level of aggression (Kazankov 2002). Therefore, nomadic hunter-gatherers who live under especially extreme conditions can be mostly characterized as relatively peaceful societies. It greatly differs from the behavior of many nomadic herders living under extreme conditions and having small population density. The latter are just distinguished by especially high levels of aggression. Thus, aggression considerably increases with the transition to barbarism. As Karl Marx and Friedrich Engels (who generally underestimate the role of wars in history) note in the *German Ideology* (2004 [1845–1846]: 89), 'with the conquering barbarian people war itself is still… a regular form of intercourse, which is the more eagerly exploited as the increase in population together with the traditional and, for it, the only possible, crude mode of production gives rise to the need for new means of production'.

It should also be noted that for barbarians (unlike for states/civilizations) external exploitation can play a significantly more important role than internal exploitation (see, *e.g.*, Kradin 1992; Grinin 1997, 2003a). Many researchers emphasize that very often exploitation begins not inside, but outside the society since the stranger is protected by neither tradition, nor custom. External exploitation strengthens inequality and, undoubtedly, promotes the development of politogenesis. For example, K. Pietkiewicz notes that before the formation of the Lithuanian state two basic strata were found among the Lithuanians: free farmers and warriors (nobles) who were called *kunigai*, *i.e.*, 'dukes', 'lords'. 'Dukes' attained material well-being and high positions in society through predatory wars, to a lesser extent through gathering tribute from their own population (Pietkiewicz 2006: 306).

Acts of violence often played a very important role in the life of culturally simple farmers and herders. Such actions were one of the most important ways for individuals to raise their social status. For example, N. A. Butinov writes the following about the Papuans:

> There were two ways of advance up to high rankers: peaceful and military; the second, probably, prevailed. The aspirant for a higher status brought together a group of men. Under his leadership they attacked neighboring villages, plundered, killed, subordinated survived to their power. The reason for attack was easy to invent (*e.g.*, black magic, theft of pigs, abduction of women, disputes on the territory *etc.*). Justification of murder of aliens was not necessary, as it was considered a good deed. Intercommunal wars took place very often (Butinov 1995: 62).

Wars played a significant role even in the formation of simple chiefdoms (see Carneiro 1970, 1981, 2004, 2012). On the whole the ideas to occupy neighboring settlements by force of arms; to take prisoners and to force them to work as slaves; to demand periodic payment of a tribute' were rather wide spread (though not universal) at that time (see, *e.g.*, Carneiro 2004; see, however, Zinkina *et al.* 2016).

Emergence or diffusion of some important technological innovations could lead to great changes in politogenesis and could become its accelerator. As a result, intensive politogenetic processes could begin in those places where politogenesis was strongly delayed or impossible at all before. As has been mentioned above, the role of such an accelerator might have been played by the diffusion of iron metallurgy, progress in using of mount or draught animals. Emergence of the cavalry, iron weapons, *etc.* also promoted 'intensification' of military operations, strengthened the role of wars in politogenesis.[35]

The role of wars in some respects is especially significant in the development and transformation of pre-state and state analogue barbarian societies. For most of relevant barbarian peoples, wars became the most important factor of their transformation into the state. A striking example was the state of the Zulu which at the beginning of the 19th century transformed from a conglomerate of chiefdoms into an empire very quickly (just in two or three decades) (see, *e.g.*, Davidson 1968: 5; 1984: 161; L'vova 1984: 47; Maquet 1974: 91; Potekhin 1954: 545; Gluckman 1987 [1940]: 29). Wars also played a prominent role in generating many important innovations which could be the source of 'energy' promoting the most powerful expansions and important reasons for fast military victories changing the World System map. In particular, there are some opinions that at the beginning of the second millennium CE evolution of nomadic military science reached such a level of development that essentially affected military art of other societies and civilizations of Eurasia (Khudyakov 1991: 3). The idea that at the beginning of the 13th century Mongols won in many respects due to some important innovations introduced in their military organization has become generally accepted (see, *e.g.*, Khrapachevsky 2005).

One should also note here that there are significant divergences concerning the role of wars in the processes of state formation. In the discussions over the state formation, as James Ambrosino notes, 'the role and influence of external social factors, *i.e.* such impetuses which are created by contacts with foreign societies, were practically ignored' (Ambrosino 1995: 54). Among modern researchers the theory of influence of wars on politogenesis and state formation was most systematically and consistently developed by Robert Carneiro

[35] *E.g.*, with the diffusion of horses in the Great Plains, theft of horses became the main reason of wars among the Amerindian peoples (such as the Omaha) in the North American prairies. Mobility of riding horses led to the point that the Utah, Apaches, Navaho, Shoshone and others began to raid peoples they never faced before the penetration of horses (Dennen 1995: 429).

(1970, 1978, 1981, 1987, 2000a, 2000b, 2000c, 2002, 2003, 2004, 2006, 2012; see also Godiner 1991; Turchin 2010, 2015; Grinin and Korotayev 2012). However, these ideas were not generally accepted. So, though the role of trade, cultural and other borrowings and influence in political anthropology is considered more adequately to some extent, the significance of wars for processes of politogenesis in general and formation of the state in particular is still underestimated by many researchers (if not by most of them).

It should be noted that in the late 19th century and the early 20th century most scholars estimated the role of wars in the state formation much higher than it was done later. For example, P. F. Preobrazhensky considered war as the inevitable concomitant of the government (2005: 154). K. Kautsky (1931) contrary to the Marxist doctrine had to eventually recognize that conquest was the most important cause for state formation. L. Gumplowicz and F. Oppenheimer were the most famous researchers who believed that the state formed due to a simple conquest of one nation by another. It is also worth mentioning Gustav Ratzenhofer.[36] Subsequently such approaches were rejected as too primitive not without some reason. Nevertheless, mistakes of a century ago do not mean that wars did not play an important role in politogenesis. Just on the contrary. At least, we do not know any case when the military factor was absent (at least in some form) during emergence and formation of the early state. By the military factor we mean a situation, connected both to waging wars in this or that way (aggressive or defensive), and with preparation for them, or with a direct conquest (submission) of some communities by means of military force (for more details see Grinin 2007a; Grinin and Korotayev 2012). At the same time the emergence of many states (even in terms of creation of truly new political and administrative forms) often took place on the basis of military structures, customs, institutes, for example military camps of young men, troops, personal guard, security structures, *etc*. (see, *e.g.*, L'vova 1995: 161; Orlova and L'vova 1978; Miller 1984: 191; Bocharov 1991: 70).

Probably it would make sense to include in the future models a mathematical description of the point that the diffusion of civilization happened not only due to military or settler expansion of civilizations, but also due to development of certain zones of barbarian periphery that quite often passed civilizational threshold directly in the course of offensive on the neighboring civilization or during defense against this offensive. So, A_2 period is characterized by transformation of the extensive part of barbarian agricultural periphery into civilization which was an important factor of the additional growth of urban population.

[36] The review of theories of power in respect to formation of the state, see, *e.g.*, in R. L. Carneiro's works (Carneiro 1970, 2006, 2012).

If we take into consideration these factors, as well as factors of cyclic and stochastic dynamics, this may allow us to develop such mathematical models that may be capable of giving more exact description of the long-term dynamics of the World System.

References

Abu-Lughod J. 1989. *Before European Hegemony: The World System A.D. 1250–1350.* New York, NY: Oxford University Press.

Abu-Lughod J. 1990. Restructuring the Premodern World-System. *Review* 13(2): 273–286.

Akhiezer A. S. 1995. The City is the Focus of the Urbanization Process. *City as a Sociocultural Phenomenon of Historical Process* / Ed. by E. V. Sayko, pp. 21–28. Moscow: Nauka. *In Russian* (Ахиезер А. С. Город – фокус урбанизационного процесса. *Город как социокультурное явление исторического процесса* / Ред Э. В. Сайко, с. 21–28. М.: Наука).

al-Dzhabarti, ʿAbd Ar-Rakhman. 1978. *Egypt on the Eve of Bonaparte's Expedition (1776–1798).* Moscow: Glavnaya redaktsiya vostochnoy literatury. *In Russian* (ал-Джабарти, ʿАбд Ар-Рахма̄н. *Египет в канун экспедиции Бонапарта (1776–1798)*. М.: Гл. ред. восточной литературы).

Alekseev I. L., and Khaltourina D. A. 2004. Ibn Khaldun and the Theory of Demographic Cycles in Modern Social Sciences. *School of the Young Orientalist – 2004* / Ed. by I. M. Steblin-Kamensky, I. M. Dyakov, A. A. Maslov, and Kh. R. Usoyan, pp. 129–133. St. Petersburg: St. Petersburg State University. *In Russian* (Алексеев И. Л., Халтурина Д. А. Ибн Халдун и теория демографических циклов в современных социальных науках. *Школа молодого востоковеда* – 2004 / Ред. И. М. Стеблин-Каменский, И. М. Дьяков, А. А. Маслов, Х. Р. Усоян, с. 129–133. СПб.: СПбГУ).

Alekshin V. A. 1986. Some Regularities of Development of Social System of the Ancient Agrarian Societies (According to Funeral Ceremonies). *Ancient Civilizations of the East* / Ed. by V. M. Masson, pp. 21–29. Tashkent: FAN. *In Russian* (Алекшин В. А. Некоторые закономерности развития общественного строя древнеземледельческих обществ (по данным погребальных обрядов). *Древние цивилизации Востока* / Ред. В. М. Массон, с. 21–29. Ташкент: ФАН).

Ambrosino J. N. 1995. Inter-Societal Contact and the Rise of the State: A Brief Note from Work in Progress. *Alternative Pathways to Early State* / Ed. by N. N. Kradin, and V. A. Lynsha, p. 54–59. Vladivostok: Dal'nauka.

Andreev Yu. V. 1979. Antique polis and the east city-states. *Antique Polis* / Ed. by E. D. Frolov, pp. 8–27. Leningrad: Leningrad University. *In Russian* (Андреев Ю. В. Античный полис и восточные города-государства. *Античный полис* / Ред. Э. Д. Фролов, с. 8–27. Л.: Издательство Ленинградского университета).

Artzrouni M., and Komlos J. 1985. Population Growth through History and the Escape from Malthusian Trap: A Homeostatic Simulation Model. *Genus* 41: 21–39.

Bahrey. 1976 [1593]. History of the Galla. *Some Records of Ethiopia (1593–1646).* / Ed. by C. F. Beckingham, and G. W. B. Huntingford. London : Hakluyt Society.

Barfield T. 1991. Inner Asia and Cycles of Power in China's Imperial History. *Rulers from the Steppe: State Formation on the Eurasian Periphery* / Ed. by G. Seaman, and D. Marks, pp. 21–62. Los Angeles, CA: Ethnographics Press.

Barfield T. J. 2006. The World of Nomads-Cattle Farmers. *Early State, its Alternatives and Analogues* / Ed. by L. E. Grinin, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 415–441. Volgograd: Uchitel. *In Russian* (Барфилд Т. Дж. Мир кочевников-скотоводов. *Раннее государство, его альтернативы и аналоги* / Ред. Л. Е. Гринин, Д. М. Бондаренко, Н. Н. Крадин, А. В. Коротаев, с. 415–441. Волгоград: Учитель).

Batsieva S. M. 1965. Historical and Sociological Treatise of Ibn Khaldun '*Mukaddim*'. Moscow: Nauka. *In Russian* (Бациева С. М. *Историко-социологический трактат Ибн Халдуна «Мукаддима»*. М.: Наука).

Blair P. H. 1966. *Roman Britain and Early England 55 B.C. – A.D. 871*. New York, NY – London: W. W. Norton and Company.

Bocharov V. V. 1991. Political Systems of Tropical Africa: From Tribe to State. Tribe and State in Africa / Ed. by Yu. M. Ilyin, V. A. Popov, and I. V. Sledzevsky, pp. 65–75. Moscow: Institut Afriki RAN. *In Russian* (Бочаров В. В. Политические системы Тропической Африки: от племени к государству. *Племя и государство в Африке* / Ред. Ю. М. Ильин, В. А. Попов, И. В. Следзевский, с. 65–75. М.: Ин-т Африки РАН).

Braudel F. 1995. What is France? Vol. 2. Part 1. Moscow: Izdatel'stvo imeni Sabashnikovykh. *In Russian* (Бродель Ф. Что такое Франция? Т. 2. Ч. 1. М.: Изд-во им. Сабашниковых).

Budanova V. P. 1990. *Goths during an Era of Migration of the Peoples*. Moscow: Nauka. *In Russian* (Буданова В. П. *Готы в эпоху Великого переселения народов*. М.: Наука).

Budanova V. P. 1994. *Barbarian World on the Boundary of Antiquity and Middle Ages*. Moscow: Nauka. *In Russian* (Буданова В. П. *Варварский мир на рубеже античности и средневековья.* М.: Наука).

Budanova V. P. 2000. *Barbarian World of an Era of Migration of the Peoples.* Moscow: Nauka. *In Russian* (Буданова В. П. *Варварский мир эпохи Великого переселения народов*. М.: Наука).

Budanova V. P. 2002. Migration of the Peoples as a Universal Model of Interaction of Civilization and Barbarity. *Civilizations. 5: Problems of Global Studies and Global History* / Ed. by A. O. Chubaryan, pp. 168–192. Moscow: Nauka. *In Russian* (Буданова В. П. Великое переселение народов как универсальная модель взаимодействия цивилизации и варварства. *Цивилизации.* 5: *Проблемы глобалистики и глобальной истории* / Ред. А. О. Чубарьян, с. 168–192. М.: Наука).

Bukhrey. 1976 [1593]. History of the Gaul. *Age groups of the People of East Africa* / Ed. by K. P. Kalinovskaya, pp. 133–143. M.: Nauka. *In Russian* (Бахрей. История галла. *Возрастные группы народов Восточной Африки* / К. П. Калиновская, с. 133–143. М.: Наука).

Butinov N. A. 1995. Papua of New Guinea: Power Structures. *Ethnic Aspects of the Power* / Ed. by V. V. Bocharov, pp. 51–78. St. Petersburg: St. Petersburg State University. *In Russian* (Бутинов Н. А. Папуа Новая Гвинея: властные структуры. *Этнические аспекты власти* / Ред. В. В. Бочаров, с. 51–78. СПб.: Санкт-Петербургский государственный университет).

Büttner T. 1981. *African History since the Ancient Times*. Moscow: Nauka. *In Russian* (Бюттнер Т. *История Африки с древнейших времен*. М.: Наука).

Carneiro R. L. 1970. Theory of the Origin of the State. *Science* 169: 733–738.

Carneiro R. L. 1978. Political Expansion as an Expression of the Principle of Competitive Exclusion. *Origins of the State* / Ed. by R. Cohen, and E. R. Service, pp. 205–223. Philadelphia, PA: Institute for the Study of Human Issues.

Carneiro R. L. 1981. The Chiefdom: Precursor of the State. *The Transition to Statehood in the New World* / Ed. by G. D. Jones, and R. R. Kautz, pp. 37–79. Cambridge, UK – New York, NY: Cambridge University Press.

Carneiro R. L. 1987. Cross-Currents in the Theory of State Formation. *American Ethnologist* 14: 756–770.

Carneiro R. L. 2000a. *The Muse of History and the Science of Culture*. New York: Kluwer Academic / Plenum Publishers.

Carneiro R. L. 2000b. Process *vs*. Stages: A False Dichotomy in Tracing the Rise of the State. *Alternatives of Social Evolution* / Ed. by N. N. Kradin, A. V. Korotayev, D. M. Bondarenko, V. de Munck, and P. K. Wason, pp. 52–58. Vladivostok: FEB RAS.

Carneiro R. L. 2000c. Process or Stages: False Dichotomy in Tracing the Rise of the State. *Alternative Ways to Civilization* / Ed. by N. N. Kradin, A. V. Korotayev, D. M. Bondarenko, and V. A. Lynsha, pp. 84–94. Moscow: Logos. *In Russian* (Карнейро Р. Л. Процесс или стадии: ложная дихотомия в исследовании истории возникновения государства. *Альтернативные пути к цивилизации* / Ред. Н. Н. Крадин, А. В. Коротаев, Д. М. Бондаренко, В. А. Лынша, с. 84–94. М.: Логос).

Carneiro R. L. 2002. Was the Chiefdom a Congelation of Ideas? *Social Evolution & History* 1(1): 80–100.

Carneiro R. L. 2003. *Evolutionism in Cultural Anthropology. A Critical History*. Boulder, CO: Westview.

Carneiro R. L. 2004. Was the Chiefdom a Congelation of Ideas? *The Early State, Its Alternatives and Analogues* / Ed. by L. E. Grinin, R. L. Carneiro, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 28–45. Volgograd: Uchitel.

Carneiro R. L. 2006. Theory of the Origin of the State. *Early State, Its Alternatives and Analogues* / Ed. by L. E. Grinin, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 55–70. Volgograd: Uchitel. *In Russian* (Карнейро Р. Л. Теория происхождения государства. *Раннее государство, его альтернативы и аналоги* / Ред.

Л. Е. Гринин, Д. М. Бондаренко, Н. Н. Крадин, А. В. Коротаев, с. 55–70. Волгоград: Учитель).

Carneiro R. L. 2012. The Circumscription Theory: A Clarification, Amplification, and Reformulation. *Social Evolution & History* 11(2): 5–30.

Chadwick N. 1987. *The Celts*. London: Penguin.

Childe G. 1952. *New Light on the Most Ancient East.* 4th ed. London: Routledge & Paul.

Childe G. 1956. *The Ancient East in the Light of New Excavation*. Moscow: Izdatel'stvo inostrannoy literatury. *In Russian* (Чайлд Г. *Древнейший Восток в свете новых раскопок*. М.: Издательство иностранной литературы).

Chubarov V. V. 1991. The Middle Eastearn Locomotive: The Developmental Pace of Technique and Technology in the Ancient World. *Archaic Society: Main Problems of Sociology of Development* / Ed. by A. V. Korotayev, and V. V. Chubarov, vol. 1, pp. 92–135. Moscow: Institut istorii of the USSR of AN of the USSR. *In Russian* (Чубаров В. В. Ближневосточный локомотив: темпы развития техники и технологии в древнем мире. *Архаическое общество: узловые проблемы социологии развития* / Ред. А. В. Коротаев, В. В. Чубаров, т. 1, с. 92–135. М.: Институт истории СССР АН СССР).

Claessen H. J. M. 2000. Problems, Paradoxes, and Prospects of Evolutionism. *Alternatives of Social Evolution* / Ed. by N. N. Kradin, A. V. Korotayev, D. M. Bondarenko, V. de Munck, and P. K. Wason, pp. 1–11. Vladivostok: FEB RAS.

Claessen H. J. M. 2002. Was the State Inevitable? *Social Evolution & History* 1(1): 101–117.

Claessen H. J. M. 2004. Was the State Inevitable? *The Early State, Its Alternatives and Analogues* / Ed. by L. E. Grinin, R. L. Carneiro, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 72–87. Volgograd: Uchitel.

Davidson A. B. 1968. About This Book. Introduction. *Shaka Zulu* / Ed. by E. Ritter, pp. 3–14. Moscow: Nauka. *In Russian* (Давидсон А. Б. Об этой книге. Вступление. *Чака Зулу* / Ред. Э. Риттер, с. 3–14. М.: Наука).

Davidson A. B. 1984. Countries of the South Africa. History of Africa in the 19th – the Beginning of the 20th century / Ed. by V. A. Subbotin, pp. 155–177. Moscow: Nauka. *In Russian* (Давидсон А. Б. Страны Южной Африки. *История Африки в XIX – начале XX в.* / Ред. В. А. Субботин, с. 155–177. М.: Наука).

Denikin A. I. 1993. Essays of the Russian Strife. Vol. 4. Ch. 10. *Questions of History* 10: 109–130. *In Russian* (Деникин А. И. Очерки русской смуты. Т. 4. Глава 10. *Вопросы истории* 10: 109–130).

Dennen J. M. G. van der. 1995. *The Origin of War. The Evolution of a Male-Coalitional Reproductive Strategy*. Vol. 1. Croningen: Origin Press.

Deshan Yu. 1984. History of the Big Island. *History of Tropical Africa (from the Ancient Times to 1800)* / Ed. by D. A. Olderogge, pp. 345–357. Moscow: Nauka. *In Russian* (Дешан Ю. История большого острова. *История тропической Африки (с древнейших времен до 1800 г.)* / Ред. Д. А. Ольдерогге, с. 345–357. М.: Наука).

Diamond J. 1999. *Guns, Germs, and Steel: The Fates of Human Societies.* New York, NY: Norton.

Dyakonov I. M. 1956. *History of Media from the Ancient Times up to the End of the 4th century BCE*. Moscow – Leningrad: Academy of Sciences of the USSR. *In Russian* (Дьяконов И. М. *История Мидии от древнейших времен до конца IV в. до н.* э. М. – Л.: Изд-во АН СССР).

Dyakonov I. M. (Ed.) 1983. *History of the Ancient East*. I. *Mesopotamia.* Moscow: Nauka. *In Russian* (Дьяконов И. М. *История древнего Востока*. I. *Месопотамия*. М.: Наука).

Dyakonov I. M. 2000. Cities-States of Sumer. *East History*. 1. *East in the Ancient Time* / Ed. by W. A. Jacobson, pp. 45–56. Moscow: Vostochnaya literatura. *In Russian* (Дьяконов И. М. Города-государства Шумера. *История Востока*. 1. *Восток в древности* / Ред. В. А. Якобсон, с. 45–56. М.: Восточная литература).

Earle T. K. 2002. Hawaiian Islands (800–1824). *Civilization Models of Politogenesis* / Ed. by D. M. Bondarenko, and A. V. Korotayev, pp. 77–88. Moscow: Institut Afriki RAN. *In Russian* (Ёрл Т. К. Гавайские острова (800–1824 гг.). *Цивилизационные модели политогенеза* / Ред. Д. М. Бондаренко, А. В. Коротаев, с. 77–88. М.: Ин-т Африки РАН).

Engels F. 1884 [1978]. *The Origin of the Family, Private Property and the State.* Moscow: Izdatelstvo inostrannoi literatury. *In Russian* (Энгельс Ф. *Происхождение семьи, частной собственности и государства.* М.: Издательство иностранной литературы).

Fedorov G. V., and Polevoy L. L. 1984. ‘Kingdoms’ of Burebista and Detsebalus: Unions of Tribes or State? *Vestnik drevney istorii* 7: 58–80. *In Russian* (Федоров Г. В., Полевой Л. Л. «Царства» Буребисты и Децебала: союзы племен или государства? *Вестник древней истории* 7: 58–80).

Fenton U. N. 1978. Iroquois in History. *North American Indians* / Ed. by Yu. P. Averkieva, pp. 109–156. Moscow: Progress. *In Russian* (Фентон У. Н. Ирокезы в истории. *Североамериканские индейцы* / Ред. Ю. П. Аверкиева, с. 109–156. М.: Прогресс).

Filip Ya. 1961. *Celtic Civilization and its Heritage.* Prague: Izdatel'stvo Chekhoslovatskoy Akademii Nauk. *In Russian* (Филип Я. *Кельтская цивилизация и ее наследие*. Прага: Изд-во Чехословацкой Академии наук).

Fried M. H. 1967. *The Evolution of Political Society. An Essay in Political Anthropology*. New York, NY: Random House.

Frolov E. 1986. Birth of the Greek Policy. *Formation and Development of Early Class Societies. City and state* / Ed. by G. L. Kurbatov, E. D. Frolov, and I. Ya. Froyanov, pp. 8–99. Leningrad: Izdatelstvo Leningradskogo universiteta. *In Russian* (Фролов Э. Д. Рождение греческого полиса. *Становление и развитие раннеклассовых обществ. Город и государство* / Ред. Г. Л. Курбатов, Э. Д. Фролов, И. Я. Фроянов, с. 8–99. Л.: Издательство Ленинградского университета).

Gluckman M. 1987 [1940]. The Kingdom of the Zulu of South Africa. *African Political Systems* / Ed. by M. Fortes, and E. E. Evans-Pritchard, pp. 25–55. London – New York, NY: International African Institute.

Gluskina L. M. 1983. Problems of Crisis of the Polis. *Antique Greece. Problems of Development of Polis*. Vol. 2. Crisis of Polis / Ed. by E. S. Golubtsov, pp. 5–42. Moscow: Nauka. *In Russian* (Глускина Л. М. Проблемы кризиса полиса. *Античная Греция. Проблемы развития полиса.* Т. 2. *Кризис полиса* / Ред. Е. С. Голубцова, с. 5–42. М.: Наука).

Godiner E. S. 1991. Political Anthropology about the Origin of the State. *Ethnological Science abroad: Problems, Search, Solutions* / Ed. by S. Ya. Kozlov, and P. I. Puchkov, pp. 51–78. Moscow: Nauka. *In Russian* (Годинер Э. С. Политическая антропология о происхождении государства. *Этнологическая наука за рубежом: проблемы, поиски, решения* / Ред. С. Я. Козлов, П. И. Пучков, с. 51–78. М.: Наука).

Grinin A. L., and Grinin L. 2015. The Cybernetic Revolution and Historical Process. *Social Evolution & History* 14(1): 125–184.

Grinin A. L., and Grinin L. E. 2016. The Cybernetic Revolution and the Change of Technological Mode. *Vostochno-Evropeyskiy Nauchnyi Vestnik* 6(2): 5–15. *In Russian* (Гринин А. Л., Гринин Л. Е. Кибернетическая революция и смена технологического уклада. *Восточно-Европейский научный вестник* 6(2): 5–15).

Grinin L. E. 1997. Formations and Civilizations. Ch. 3. Problems and Aspects of Sociological Analysis of Historical Societies. Section 4. Distributive Relations. *Filosofiya i obschestvo* 4: 5–62. *In Russian* (Гринин Л. Е. Формации и цивилизации. Гл. 3. Проблемы и аспекты социологического анализа исторических обществ. Пар. 4. Распределительные отношения. *Философия и общество* 4: 5–62).

Grinin L. E. 1998a. Formations and Civilizations. Introduction into the Theory of Historical Process. *Filosofiya i obschestvo* 1: 7–54. *In Russian* (Гринин Л. Е. Формации и цивилизации. Введение в теорию исторического процесса. *Философия и общество* 1: 7–54).

Grinin L. E. 1998b. Formations and Civilizations. The Notion of Civilization within the Framework of the Theory of the Historical Process. *Filosofiya i obschestvo* 2: 5–89. *In Russian* (Гринин Л. Е. Формации и цивилизации. Понятие цивилизации в рамках теории исторического процесса. *Философия и общество* 2: 5–89).

Grinin L. E. 1998c. Formations and Civilizations. The Notion of Civilization within the Framework of the Theory of the Historical Process. *Filosofiya i obschestvo* 3: 5–54. *In Russian* (Гринин Л. Е. Формации и цивилизации. Понятие цивилизации в рамках теории исторического процесса. *Философия и общество* 3: 5–54).

Grinin L. E. 2001–2006. *State Formation as a Component of the Process of Transition from Primitiveness to Civilization (General Context of Social Evolution at the Early State Formation*). [The book was published in the journal *Filosofiya i obschestvo* from 2001 to 2006]. *In Russian* (Гринин Л. Е. *Генезис государства как составная часть процесса перехода от первобытности к цивилизации (общий контекст социальной эволюции при образовании раннего государства).* [Книга печаталась в журнале *Философия и общество* с 2001 по 2006 г.]

Grinin L. E. 2003a. *Productive Forces and Historical Process*. 2nd ed. Volgograd: 'Uchitel'. *In Russian* (Гринин Л. Е. *Производительные силы и исторический процесс.* 2-е изд. Волгоград: Учитель).

Grinin L. E. 2003b. The Early State and its Analogues. *Social Evolution & History* 1: 131–176.

Grinin L. E. 2004a. Democracy and Early State. *Social Evolution & History* 3(2): 93–149.

Grinin L. E. 2004b. The Early State and Its Analogues: A Comparative Analysis. *The Early State, Its Alternatives and Analogues* / Ed. by L. E. Grinin, R. L. Carneiro, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 88–136. Volgograd: Uchitel.

Grinin L. E. 2006a. Periodization of History: A Theoretic-Mathematical Analysis. *History & Mathematics: Analyzing and Modeling Global Development* / Ed. by L. E. Grinin, V. de Munck, and A. V. Korotayev, pp. 10–38. Moscow: KomKniga/ URSS.

Grinin L. E. 2006b. Periodization of History: A Theoretic-Mathematical Analysis. *History and Mathematics: Problems of Periodization of Historical Macroprocesses* / Ed. by L. E. Grinin, A. V. Korotayev, and S. Yu. Malkov, pp. 53–79. Moscow: Komkniga/URSS. *In Russian* (Гринин Л. Е. Периодизация истории: теоретико-математический анализ. *История и Математика*: *проблемы периодизации исторических макропроцессов* / Ред. Л. Е. Гринин, А. В. Коротаев, С. Ю. Малков, с. 53–79. М.: КомКнига/URSS).

Grinin L. E. 2007a. *State and Historical Process*: in 3 books. Moscow: URSS. *In Russian* (Гринин Л. Е. *Государство и исторический процесс*: в 3-х кн. М.: URSS).

Grinin L. E. 2007b. Industrial Revolutions as the Most Important Thresholds of History. *Man and Nature: Confrontation and Harmony* / Ed. by E. S. Kulpin, pp. 191–221. Moscow: IATs-Energiya. *In Russian* (Гринин Л. Е. Производственные революции как важнейшие рубежи истории. *Человек и природа: противостояние и гармония* / Ред. Э. С. Кульпин, с. 191–221. М.: ИАЦ-Энергия).

Grinin L. E. 2010. *The State and Historical Process. Evolution of Statehood: From the Early State to the Mature State.* 2nd ed. Moscow: Librocom. *In Russian* (Гринин Л. Е. *Государство и исторический процесс: Эволюция государственности: От раннего государства к зрелому.* Изд. 2-е, испр. Москва: Книжный Дом «Либроком»).

Grinin L. E. 2011. *The Evolution of Statehood. From Early State to Global Society.* Saarbrücken: Lambert Academic Publishing.

Grinin L. E. 2012. *Macrohistory and Globalization.* Volgograd: Uchitel.

Grinin L. E. 2016a. The New World Order and Philosophy. *Between Past Orthodoxies and the Future of Globalization. Contemporary Philosophical Problems* / Ed. by A. N. Chumakov, and W. C. Gay, pp. 143–156. URL: http://www.brill.com/products/book/between-past-orthodoxies-and-future-globalization.

Grinin L. E. 2016b. Evolution of World Order. *Evolution and Big History: Dimensions, Trends, and Forecasts* / Ed. by L. E. Grinin, and A. V. Korotayev, pp. 75–95. Volgograd: 'Uchitel' Publishing House.

Grinin L. E. 2017a. On Systemic Integration in the World System since the Bronze Age. *Social Evolution & History* 16(2): 76–111.

Grinin L. E. 2017b. The Processes of Systemic Integration in the World System. *Journal of Globalization Studies* 8(1): 97–118.

Grinin L. E., Carneiro R., Bondarenko D. M., Kradin N. N., and Korotayev A. V. (Eds.) 2004. *The Early State, Its Alternatives and Analogues*. Volgograd: Uchitel.

Grinin L. E., and Grinin A. L. 2015. *From Biface to Nanorobots. The World is on the Way to the Epoch of Self-Regulating Systems (History of Technologies and the Description of their Future).* Moscow: Uchitel. *In Russian* (Гринин Л. Е., Гринин А. Л. *От рубил до нанороботов. Мир на пути к эпохе самоуправляемых систем (история технологий и описание их будущего).* М.: Моск. ред. Изд-ва «Учитель»).

Grinin L. E., and Grinin A. L. 2016. *The Cybernetic Revolution and the Forthcoming Epoch of Self-Regulating Systems.* Moscow: Moscow branch of Uchitel Publishing House.

Grinin L. E., Bondarenko D. M., Kradin N. N., and Korotayev A. V. (Eds.) 2006. *Early State, Its Alternatives and Analogues.* Volgograd: Uchitel. *In Russian* (Гринин Л. Е., Бондаренко Д. М., Крадин Н. Н., Коротаев А. В. (Ред.) *Раннее государство, его альтернативы и аналоги*. Волгоград: Учитель).

Grinin L. E., and Korotayev A. V. 2006. Political Development of the World System: A Formal Quantitative Analysis. *History and Mathematics*. *Historical Dynamics and Development of Complex Societies* / Ed. by P. Turchin, L. Grinin, A. Korotaev, and V. de Munck, pp. 115–153. Moscow: KomKniga/URSS.

Grinin L. E., and Korotayev A. V. 2007. Political Development of the World System: Formal and Quantitative Analysis. *History and Mathematics: Macrohistorical Dynamics of Society and State* / Ed. by S. Yu. Malkov, L. E. Grinin, and A. V. Korotayev, pp. 49–101. Moscow: Komkniga/URSS. *In Russian* (Гринин Л. Е., Коротаев А. В. Политическое развитие Мир-Системы: формальный и количественный анализ. *История и Математика*: *Макроисторическая динамика общества и государства* / Отв. ред. С. Ю. Малков, Л. Е. Гринин, А. В. Коротаев, с. 49–101. М.: КомКнига/URSS).

Grinin L. E., and Korotayev A. V. 2008. *Social Macroevolution. Genesis and Transformations of the World System*. Moscow: LKI/URSS.

Grinin L. E., and Korotayev A. V. 2009a. The Epoch of the Initial Politogenesis. *Social Evolution & History* 8(1): 52–91.

Grinin L. E., and Korotayev A. V. 2009b. Social Macroevolution: Growth of the World System Integrity and a System of Phase Transitions. *World Futures* 65(7): 477–506.

Grinin L. E., and Korotayev A. V. 2012. Emergence of Chiefdoms and States: A Spectrum of Opinions. *Social Evolution & History* 11(2): 191–204.

Grinin L. E., and Korotayev A. V. 2013. The Origins of Globalization. *Globalization. Yesterday, Today, and Tomorrow* / Ed. by J. Sheffield, A. Korotayev, and L. Grinin, pp. 1–29. Litchfield Park, AZ: Emergent Publications.

Grinin L. E., and Korotayev A. V. 2014. Origins of Globalization in the Framework of the Afroeurasian World-System History. *Journal of Globalization Studies* 5(1): 32–64.

Grinin L. E., and Korotayev A. V. 2016. *Near East, India, and China in Globalization Processes*. Moscow: Moscow branch of Uchitel Publishing House. *In Russian* (Гринин Л. Е., Коротаев А. В. *Ближний Восток, Индия и Китай в глобализационных процессах*. М.: Моск. ред. изд-ва «Учитель»).

Grinin L. E., and Korotayev A. V. 2018. Origins of Globalization in the Framework of the Afroeurasian World-System History. *Comparing Globalizations, World-Systems Evolution and Global Futures* / Ed. by T. D. Hall, pp. 37–70. Springer. URL: https://doi.org/10.1007/978-3-319-68219-8_3.

Grinin L. E., Grinin A. L., and Korotayev A. 2017a. Forthcoming Kondratieff wave, Cybernetic Revolution, and Global Ageing. *Technological Forecasting & Social Change* 115: 52–68.

Grinin L., Grinin A., and Korotayev A. 2017b. The MANBRIC-Technologies in the Forthcoming Technological Revolution. *Industry 4.0: Entrepreneurship and Structural Change in the New Digital Landscape* / Ed. by T. Devezas, J. Leitão, and A. Sarygulov, pp. 243–261. Heidelberg – New York – Dordrecht – London: Springer.

Grinin L. E., Ilyin I. V., and Andreev A. I. 2016. World Order in the Past, Present, and Future. *Social Evolution and History* 15(1): 58–84.

Gumilev L. N. 1993. *Xiongnu*. St. Petersburg: Taim-aut-Kompass. *In Russian* (Гумилев Л. Н. *Хунну*. СПб.: Тайм-Аут-Компасс).

Gurevich A. Ya. 1970. *The Question of Genesis of Feudalism in Western Europe*. M.: Vysshaya shkola. *In Russian* (Гуревич А. Я. *Проблемы генезиса феодализма в Западной Европе*. М.: Высшая школа).

Gurevich A. Ya. 1999. Ancient Germans. Vikings. *Chosen works* / Ed. by A. Ya. Gurevich. Vol. 1. Moscow – St. Petersburg: Universitetskaya kniga. *In Russian* (Гуревич А. Я. *Проблемы генезиса феодализма в Западной Европе*. М.: Высшая школа).

Gurevich A. Ya. 2005. *Campaigns of Vikings*. Moscow: KDU. *In Russian* (Гуревич А. Я. *Походы викингов*. М.: КДУ).

Hellemans A., and Bunch B. 1988. *The Timetables of Science*. New York, NY: Simon and Schuster.

Ibn Khaldūn, `Abd al-Rahman. 1958. *The Muqaddimah: An Introduction to History*. New York, NY: Pantheon Books (Bollingen Series, 43).

Ibn Khaldūn, `Abd al-Rahmān. 2004. *Muqaddimat ibn Khaldūn.* al-Qāhirah: Dar al-Fajr li-l-Turāth.[37]

Ignatenko A. A. 1980. *Ibn-Khaldun*. Moscow: Mysl'. *In Russian* (Игнатенко А. А. *Ибн-Хальдун.* М.: Мысль).

[37] مقدمة ابن خلدون / تأليف عبد الرحمن ابن خلدون. **القاهرة: دار الفجر في التراث، 2004**.

'Inān, Muḥammad 'Abd Allāh. 1933. *Ibn Khaldūn: ḥayātu-hu wa-turāthu-hu al-fikriyy*. al-Qāhirah: Dār al-Kutub al-Misriyyah.[38]

Irons W. 2002. Cultural Capital, Raids for Cattle and Military Superiority of Traditional Cattle-Farmers. *Nomadic Alternative of Social Evolution* / Ed. by N. N. Kradin, D. M. Bondarenko, pp. 99–108. Moscow: TsTsRI RAN. *In Russian* (Айонс У. Культурный капитал, набеги за скотом и военное превосходство традиционных скотоводов. *Кочевая альтернатива социальной эволюции* / Ред. Н. Н. Крадин, Д. М. Бондаренко, с. 99–108. М.: ЦЦРИ РАН).

Irons W. 2004. Cultural Capital, Livestock Raiding, and the Military Advantage of Traditional Pastoralists. *The Early State, Its Alternatives and Analogues* / Ed. by L. E. Grinin, R. L. Carneiro, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 466–475. Volgograd: Uchitel.

Its R. F., Yakovlev A. G. 1967. To the Question of Social and Economic System of the Lyanshan Group of Nationality. *A Community and Social Organization of the East and the Southeast Asian Peoples* / Ed. by R. F. Its, pp. 64–106. Leningrad: Nauka. *In Russian* (Итс Р. Ф., Яковлев А. Г. К вопросу о социально-экономическом строе ляньшанской группы народности. *Община и социальная организация у народов Восточной и Юго-Восточной Азии* / Ред. Р. Ф. Итс, с. 64–106. Л.: Наука).

Kardini F. 1987. *Sources of Medieval Knights*. Moscow: Progress.

Kautsky K. 1931. *Materialistic Understanding of History*. Vol. 2. Moscow: GIZ. *In Russian* (Каутский К. *Материалистическое понимание истории*. Т. 2. М.: ГИЗ).

Kazankov A. A. 2002. *Aggression in Archaic Societies (By the example of Hunters-Gatherers of Semideserts*). Moscow: Institut Afriki RAN. *In Russian* (Казанков А. А. *Агрессия в архаических обществах (на примере охотников-собирателей полупустынь)*. М.: Ин-т Африки РАН).

Khazanov A. M. 1975. *Social History of Scythians*. Moscow: Nauka. *In Russian* (Хазанов А. М. *Социальная история скифов.* М.: Наука).

Khazanov A. M. 2002. Nomads of the Euroasian Steppes in Historical Retrospective. *Nomadic Alternative of Social Evolution* / Ed. by N. N. Kradin, and D. M. Bondarenko, pp. 37–58. Moscow: Institut Afriki RAN. *In Russian* (Хазанов А. М. Кочевники евразийских степей в исторической ретроспективе. *Кочевая альтернатива социальной эволюции* / Ред. Н. Н. Крадин, Д. М. Бондаренко, с. 37–58. М.: Институт Африки РАН).

Khazanov A. M. 2006. Nomads of the Euroasian Steppes in Historical Retrospective. *The Early State, its Alternatives and Analogues* / Ed. by L. E. Grinin, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 468–489. Volgograd: Uchitel. *In Russian* (Хазанов А. М. Кочевники евразийских степей в исторической ретроспективе. *Раннее государство, его альтернативы и аналоги* / Ред. Л. Е. Гринин, Д. М. Бондаренко, Н. Н. Крадин, А. В. Коротаев, с. 468–489. Волгоград: Учитель).

Khrapachevsky R. N. 2005. Military Power of Genghis Khan. Moscow: AST. *In Russian* (Храпачевский Р. Н. *Военная держава Чингисхана*. М.: АСТ).

[38] ابن خلدون : حياته و تراثه الفكري / تأليف محمد عبد الله عنان. القاهرة: دار الكتب المصرية، 1933.

Khudyakov Yu. S. 1991. *Armament of Central Asian Nomads in the Early and Developed Middle Ages*. Novosibirsk: Nauka. *In Russian* (Худяков Ю. С. *Вооружение центрально-азиатских кочевников в эпоху раннего и развитого средневековья.* Новосибирск: Наука).

Korotayev A. V. 1991. Some Economic Prerequisites of Class Formation and Politogenesis. *Archaic Society: Key Problems of Sociology of Development* / Ed. by A. V. Korotayev, and V. V. Chubarov, Vol. 1, pp. 136–191. Moscow: Institut istorii SSSR AN SSSR.

Korotayev A. V. 1997a. Genesis of Tribal Organization in the North-East Yemen Highlands. *Oriental Studies in the 20th Century: Achievements and Prospects*, pp. 194–198. Moscow – Budapest: Russian Orienalists' Association. ICANAS.

Korotayev A. V. 1997b. *Sabean Studies. Some General Trends and Factors of Evolution of Sabean Civilization*. Moscow: Vostochnaya literatura. *In Russian* (Коротаев А. В. *Сабейские этюды. Некоторые общие тенденции и факторы эволюции сабейской цивилизации*. М.: Восточная литература).

Korotayev A. V. 1998. Genesis and Development of Tribal Organization in the Northeast Yemeni Highlands. *Sociobiology of Ritual and Group Identity: A Homology of Animal and Human Behaviour. Concepts of Humans and Behaviour Patterns in the Cultures of the East and the West: Interdisciplinary Approach* / Ed. by M. Butovskaya, A. Korotayev, and O. Khristoforova, pp. 117–118. Moscow: Russian State University for Humanities.

Korotayev A. V. 2000a. From State to Chiefdom? From Chiefdom to Tribe? (Some General Tendencies of Evolution of the South Arabian Socio-Political Systems for the Last Three Thousand Years). *Early Forms of Social Organization. Genesis, Functioning, Historical Dynamics* / Ed. by V. A. Popov, pp. 224–302. St. Petersburg: Vostochnaya literatura. *In Russian* (Коротаев А. В. От государства к вождеству? От вождества к племени? (Некоторые общие тенденции эволюции южноаравийских социально-политических систем за последние три тысячи лет). *Ранние формы социальной организации. Генезис, функционирование, историческая динамика* / Ред. В. А. Попов, с. 224–302. СПб.: Восточная литература).

Korotayev A. V. 2000b. Tribe as a Form of Socio-Political Organization of Complex Non-Primitive Societies (Based on Materials of Northeast Yemen). *Alternative Ways to Civilization* / Ed. by N. N. Kradin, A. V. Korotayev, D. M. Bondarenko, and V. A. Lynsha, pp. 265–291. Moscow: Logos. *In Russian* (Коротаев А. В. Племя как форма социально-политической организации сложных непервобытных обществ (в основном по материалам Северо-восточного Йемена). *Альтернативные пути к цивилизации* / Ред. Н. Н. Крадин, А. В. Коротаев, Д. М. Бондаренко, В. А. Лынша, с. 265–291. М.: Логос).

Korotayev A. V. 2003. Penetration of Nomads to the Arabian South and Formation of Tribal Organization among the North-East Yemen Agricultural Population. *Nomadic Pathways in Social Evolution* / Ed. by N. N. Kradin, D. M. Bondarenko, and Th. J. Barfield, pp. 114–122. Moscow: Institute for African Studies.

Korotayev A. V. 2004. The Chiefdom: Precursor of the Tribe? *The Early State, Its Alternatives and Analogues* / Ed. by L. E. Grinin, R. L. Carneiro, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 300–325. Volgograd: Uchitel.

Korotayev A. 2005. A Compact Macromodel of World System Evolution. *Journal of World-Systems Research* 11(1): 79–93.

Korotayev A. V. 2006a. Periodization of History of the World System and Mathematical Macromodels of Socio-historical Processes. *History and Mathematics: Problems of Periodization of Historical Macroprocesses* / Ed. by L. E. Grinin, A. V. Korotayev, and S. Yu. Malkov, pp. 116–167. Moscow: Komkniga/URSS. *In Russian* (Коротаев А. В. Периодизация истории Мир-Системы и математические макромодели социально-исторических процессов. *История и Математика: Проблемы периодизации исторических макропроцессов* / Ред. Л. Е. Гринин, А. В. Коротаев, С. Ю. Малков, с. 116–167. М.: КомКнига/URSS).

Korotayev A. V. 2006b. *Social History of Yemen of the 10th Century BCE – 20th century CE*. Moscow: Komkniga/URSS. *In Russian* (Коротаев А. В. *Социальная история Йемена X в. до н. э. – XX в. н. э.* М.: КомКнига/URSS).

Korotayev A. 2006c. The World System Urbanization Dynamics: A Quantitative Analysis. *History & Mathematics: Historical Dynamics and Development of Complex Societies* / Ed. by P. Turchin, L. Grinin, A. Korotayev, and V. C. de Munck, pp. 44–62. Moscow: KomKniga/URSS.

Korotayev A. V. 2006d. The World System History Periodization and Mathematical Models of Socio-historical Processes. *History & Mathematics: Analyzing And Modeling Global Development* / Ed. by L. Grinin, V. de Munck, and A. Korotayev, pp. 39–98. Moscow: KomKniga/URSS.

Korotayev A. V. 2006e. *The Long-Term Political-Demographic Dynamics of Egypt: Cycles and Trends.* Moscow: Vostochnaya literatura. *In Russian* (Коротаев А. В. *Долгосрочная политико-демографическая динамика Египта: циклы и тенденции.* М.: Вост. лит-ра).

Korotayev A. 2007a. Compact Mathematical Models of World System Development, and How They can Help us to Clarify our Understanding of Globalization Processes. *Globalization as Evolutionary Process: Modeling Global Change* / Ed. by G. Modelski, T. Devezas, and W. R. Thompson, pp. 133–160. London: Routledge.

Korotayev A. V. 2007b. Macrodynamics of the World System Urbanization: A Quantitative Analysis. *History and Mathematics: Macrohistorical Dynamics of the Society and State* / Ed. by S. Yu. Malkov, L. E. Grinin, and A. V. Korotayev, pp. 21–39. Moscow: Komkniga/URSS. *In Russian* (Коротаев А. В. Макродинамика урбанизации Мир-Системы: количественный анализ. *История и Математика: Макроисторическая динамика общества и государства* / Ред. С. Ю. Малков, Л. Е. Гринин, А. В. Коротаев, с. 21–39. М.: КомКнига/URSS).

Korotayev A. V. 2007c. Sociological Theory of Abd al-Rahman ibn Khaldun (1332–1406) as a Methodological Basis for the Mathematical Modeling of Political-Demographic Dynamics of Medieval Egypt. *38. ICANAS. Abstracts* / Ed. by Z. Dilek, p. 310. Ankara: ICANAS.

Korotayev A. V. 2007d. Sociological Theory of Abd al-Rahman ibn Khaldun as a Methodological Basis for the Mathematical Modeling of Long-Term Political-Demographic Dynamics. *Mathematical Modelling of Historical Processes* / Ed. by G. G. Malinetsky, pp. 31–34. Moscow: Institut Prikladnoy matematiki im. M. V. Keldysha. *In Russian* (Коротаев А. В. Социологическая теория Абд ар-Рахмана ибн Халдуна как методологическая основа математического моделирования долгосрочной политико-демографической динамики. *Математическое моделирование исторических процессов* / Ред. Г. Г. Малинецкий, с. 31–34. М.: Институт прикладной математики им. М. В. Келдыша).

Korotayev A. 2008. Globalization and Mathematical Modeling of Global Development. *Hierarchy and Power in the History of Civilizations: Political Aspects of Modernity* / Ed. by L. E. Grinin, D. D. Beliaev, and A. V. Korotayev, pp. 225–240. Moscow: LIBROCOM/URSS.

Korotayev A. 2009. Compact Mathematical Models of the World System Development and Their Applicability to the Development of Local Solutions in Third World Countries. *Systemic Development: Local Solutions in a Global Environment* / Ed. by J. Sheffield, pp. 103−116. Litchfield Park, AZ: ISCE Publishing.

Korotayev A. 2012. Globalization and Mathematical Modeling of Global Development. *Globalistics and Globalization Studies* 1: 148–158.

Korotayev A. 2013. Globalization and Mathematical Modeling of Global Evolution. *Evolution: Development within Big History, Evolutionary and World-System Paradigms* / Ed. by L. E. Grinin, and A. V. Korotayev, pp. 69–83. Volgograd: Uchitel.

Korotayev A. V., and Grinin L. E. 2006. Urbanization and Political Development of the World System: A Comparative Quantitative Analysis. *History and Mathematics*. *Historical Dynamics and Development of Complex Societies* / Ed. P. Turchin, L. Grinin, A. Korotayev, and V. de Munk, pp. 115–153. Moscow: URSS.

Korotayev A. V., and Grinin L. E. 2007. Urbanization and Political Development of the World System: A Comparative Quantitative Analysis. *History and Mathematics: Macrohistorical Dynamics of the Society and State* / Ed. by S. Yu. Malkov, L. E. Grinin, and A. V. Korotayev, pp. 102–141. Moscow: Komkniga/URSS. *In Russian* (Коротаев А. В., Гринин Л. Е. Урбанизация и политическое развитие Мир-Системы: сравнительный количественный анализ. *История и Математика*: *Макроисторическая динамика общества и государства* / Ред. С. Ю. Малков, Л. Е. Гринин, А. В. Коротаев, с. 102–141. М.: КомКнига/URSS).

Korotayev A. V., and Grinin L. E. 2012. Global Urbanization and Political Development of the World System. *Globalistics and Globalization Studies* / Ed. by L. Grinin, I. Ilyin, and A. Korotayev, pp. 28–78. Moscow – Volgograd: Moscow University – Uchitel.

Korotayev A. V., and Grinin L. E. 2013. Urbanization and Political Development of the World System. *ENTELEQUIA revista interdisciplinar* 15 (especial 2013): 197–254.

Korotayev A. V., Komarova N. L., and Khaltourina D. A. 2007. *Laws of History. Cecular Cycles and Millennial Trends. Demography. Economy. Wars.* Moscow: Komkniga/URSS. *In Russian* (Коротаев А. В., Комарова Н. Л., Халтурина Д. А.

Законы истории. Вековые циклы и тысячелетние тренды. Демография. Экономика. Войны. М.: КомКнига/URSS).

Korotayev A. V., and Malkov A. S. 2016. A Compact Mathematical Model of the World System Economic and Demographic Growth, 1 CE – 1973 CE. *International Journal of Mathematical Models and Methods in Applied Sciences* 10: 200–209.

Korotayev A. V., Malkov A. S., and Khaltourina D. A. 2005. *Laws of History: Mathematical Modeling of Historical Macroprocesses. Demography, Economy, Wars*. Moscow: URSS. *In Russian* (Коротаев А. В., Малков А. С., Халтурина Д. А. *Законы истории. Математическое моделирование развития Мир-Системы. Демография. Экономика. Культура.* М.: КомКнига/URSS).

Korotayev A., Malkov A., and Khaltourina D. 2006a. *Introduction to Social Macrodynamics: Compact Macromodels of the World System Growth.* Moscow: KomKniga/URSS.

Korotayev A., Malkov A., and Khaltourina D. 2006b. *Introduction to Social Macrodynamics: Secular Cycles and Millennial Trends.* Moscow: KomKniga/URSS.

Korotayev A. V., Malkov A. S., and Khaltourina D. A. 2007. *Laws of History. Mathematical Modeling of Development of the World System. Demography. Economy. Culture*. Moscow: Komkniga/URSS. *In Russian* (Коротаев А. В., Малков А. С., Халтурина Д. А. *Законы истории. Математическое моделирование развития Мир-Системы. Демография. Экономика. Культура.* М.: КомКнига/URSS).

Korotayev A., and Khaltourina D. 2006. *Introduction to Social Macrodynamics: Secular Cycles and Millennial Trends in Africa.* Moscow: KomKniga/URSS.

Korotayev A., and Malkov S. 2012. Mathematical Models of the World-System Development. *Routledge Handbook of World-Systems Analysis* / Ed. by S. Babones, and C. Chase-Dunn, pp. 158–161. London: Routledge.

Korotayev A., and Zinkina J. 2017. Systemic Boundary Issues in the Light of Mathematical Modeling of World-System Evolution. *Journal of Globalization Studies* 8(1): 78–96.

Korsunsky A. R., and Günter R. 1984. *Decline and Fall of the Western Roman Empire and the Origin of Barbarian Kingdoms (up to the Middle of the 6th Century*). Moscow: MGU. *In Russian* (Корсунский А. Р., Гюнтер Р. *Упадок и гибель Западной Римской империи и возникновение варварских королевств (до середины VI в.).* М.: Изд-во МГУ).

Kottak C. Ph. 1980. *The Past in the Present; History, Ecology and Cultural Variation in Highland Madagascar*. Ann Arbor, MI: University of Michigan Press.

Kradin N. N. 1992. *Nomadic Societies.* Vladivostok: Dalnauka. *In Russian* (Крадин Н. Н. *Кочевые общества.* Владивосток: Дальнаука).

Kradin N. N. 2001a. *Empire of Xiongnu*. 2nd ed. Vladivostok: Dalnauka. *In Russian* (Крадин Н. Н. *Империя хунну*. 2-е изд. Владивосток: Дальнаука).

Kradin N. N. 2001b. Nomads in the World Historical Process. *Filosofiya i obschestvo* 2: 108–137. *In Russian* (Крадин Н. Н. Кочевники в мировом историческом процессе. *Философия и общество* 2: 108–137).

Kradin N. N., and Bondarenko D. M. (Eds.) 2002. *Nomadic Alternative of Social Evolution*. Moscow: Institut Afriki RAN. *In Russian* (Крадин Н. Н., Бондаренко Д. М. (Ред.). *Кочевая альтернатива социальной эволюции.* М.: Институт Африки РАН).

Kradin N. N., Bondarenko D. M., and Barfield T. J. (Eds.) 2003. *Nomadic Pathways in Social Evolution*. Moscow: Center for Civilizational and Regional Studies RAS.

Kremer M. 1993. Population Growth and Technological Change: One Million B.C. to 1990. *The Quarterly Journal of Economics* 108: 681–716.

Kubbel' L. E. 1988. Emergence of Private Property, Classes and State. *History of Primitive Society. The Epoch of Class Formation* / Ed. by Yu. V. Bromley, pp. 140–269. Moscow: Nauka. *In Russian* (Куббель Л. Е. Возникновение частной собственности, классов и государства. *История первобытного общества. Эпоха классообразования* / Ред. Ю. В. Бромлей, с. 140–269. М.: Наука).

Lamberg-Karlovsky K., and Sablov J. 1992. *Ancient Civilizations. Middle East and Mesoamerica*. Moscow: Nauka. *In Russian* (Ламберг-Карловски К., Саблов Дж. *Древние цивилизации. Ближний Восток и Мезоамерика*. М.: Наука).

Lesser A. 1968. War and the State. *War: The Anthropology of Armed Conflict and Aggression* / Ed. by M. Fried, M. Harris, and R. Murphy, pp. 92–96. Garden City, NY: The Natural History Press.

Lewis H. S. 1981. Warfare and the Origin of the State: Another Formulation. *The Study of the State* / Ed. by H. J. M. Claessen, and P. Skalník, pp. 201–221. The Hague: Mouton.

Lozny L. 1995. The Transition to Statehood in Central Europe. *Alternative Pathways to Early State* / Ed. by N. N. Kradin, and V. A. Lynsha, pp. 84–92. Vladivostok: Dal'nauka.

L'vova E. S. 1984. *Ethnography of Africa*. Moscow: MGU. *In Russian* (Львова Э. С. *Этнография Африки*. М.: МГУ).

L'vova E. S. 1995. Formation of Exploitation and Early Statehood among the Peoples in the Congo Basin. *Alternative Ways to the Early Statehood* / Ed. by N. N. Kradin, and V. A. Lynsha, pp. 151–164. Vladivostok: Dal'nauka. *In Russian* (Львова Э. С. Складывание эксплуатации и ранней государственности у народов бассейна Конго. *Альтернативные пути к ранней государственности* / Ред. Н. Н. Крадин, В. А. Лынша, с. 151–164. Владивосток: Дальнаука).

Mahdi M. 1937. *Ibn Khaldūn's Philosophy of History. A Study in the Philosophical Foundation of the Science of Culture.* London: Allen & Unwin.

Maquet J. 1974. *Civilizations of Africa to the South of Sahara. History, Technical Skills, Arts, Societies*. Moscow: Nauka. *In Russian* (Маке Ж. *Цивилизации Африки южнее Сахары. История, технические навыки, искусства, общества*. М.: Наука).

Marx K., and Engels F. 2004 [1845–1846]. *German Ideology*. New York: International Publishers.

Masson V. M. 1976. *Economy and Social System of Ancient Societies*. Leningrad: Nauka. *In Russian* (Массон В. М. *Экономика и социальный строй древних обществ.* Л.: Наука).

Masson V. M. 1980. Early Agricultural Societies and Formation of Settlements of the City Type. *Early Farmers*. *Ethnographic Essays* / Ed. by N. A. Butinov, and A. M. Reshetov, pp. 178–185. Leningrad: Nauka. *In Russian* (Массон В. М. Раннеземледельческие общества и формирование поселений городского типа. *Ранние земледельцы. Этнографические очерки* / Ред. Н. А. Бутинов, А. М. Решетов, с. 178–185. Л.: Наука).

Masson V. M. (Ed.) 1986. *Ancient Civilizations of the East*. Tashkent: FAN. *In Russian* (Массон В. М. (Ред.) *Древние цивилизации Востока.* Ташкент: ФАН).

Masson V. M. 1989. *First Civilizations*. Leningrad: Nauka. *In Russian* (Массон В. М. 1989. *Первые цивилизации*. Л.: Наука).

McNeill W. H. 1993. *Plagues and Peoples.* 2nd ed. New York, NY: Monticello.

McNeill W. H. 1995. The Changing Shape of World History. *History and Theory* 35: 8–26.

Miller J. 1984. *Kings and Kinsmen. Early Mbundu States in Angola*. Moscow: Nauka. *In Russian* (Миллер Дж. *Короли и сородичи. Ранние государства мбунду в Анголе*. М.: Наука).

Mongait A. L. 1974. *Archeology of Western Europe. Bronze and Iron Ages*. Moscow: Nauka. *In Russian* (Монгайт А. Л. *Археология Западной Европы. Бронзовый и железные века*. М.: Наука).

Morgan L. H. 1877. Ancient Society; or, Researches in the Lines of Human Progress from Savagery, Through Barbarism to Civilization. H. Holt.

Nefedov S. A. 2001. Method of Demographic Cycles. *Ural'skiy istoricheskiy vestnik* 7: 93–107. *In Russian* (Нефедов С. А. Метод демографических циклов. *Уральский Исторический Вестник* 7: 93–107).

Nefedov S. A. 2002a. Experience of Modeling of Demographic Cycle. *Informatsionniy bulleten' assotsiatsii 'Istoriya I kompyuter'* 29: 131–142. *In Russian* (Нефедов С. А. Опыт моделирования демографического цикла. *Информационный бюллетень ассоциации «История и компьютер»* 29: 131–142.)

Nefedov S. A. 2002b. About the Theory of Demographic Cycles. *Ekonomicheskaya istoriya* 8: 116–121. *In Russian* (Нефедов С. А. О теории демографических циклов. *Экономическая история* 8: 116–121).

Nefedov S. A. 2003. Theory of Demographic Cycles and Social Evolution of Ancient and Medieval Societies of the East. *Vostok* 3: 5–22. *In Russian* (Нефедов С. А. Теория демографических циклов и социальная эволюция древних и средневековых обществ Востока. *Восток* 3: 5–22).

Nefedov S. A. 2005. *Demographic and Structural Analysis of Social and Economic History of Russia. The End of the 15th – the Beginning of the 20th Century*. Yekaterinburg: UGGU Publishing House. *In Russian* (Нефедов С. А. 2005. *Демографи-*

чески-структурный анализ социально-экономической истории России. Конец XV – начало XX века. Екатеринбург: Изд-во УГГУ).

Nefedov S. A. 2007. *Concept of Demographic Cycles*. Yekaterinburg: EGGU. *In Russian* (Нефедов С. А. *Концепция демографических циклов*. Екатеринбург: ЕГГУ).

Neusykhin A. I. 1968. Prefeudal Period as a Transitional Stage of Development from the Tribal System to Early Feudal (Based on Materials of History of Western Europe of the Early Middle Ages). *Problems of History of Pre-Capitalistic Societies* / Ed. by L. V. Danilov, pp. 569–617. Moscow: Nauka. *In Russian* (Неусыхин А. И. Дофеодальный период как переходная стадия развития от родоплеменного строя к раннефеодальному (на материалах истории Западной Европы раннего средневековья). *Проблемы истории докапиталистических обществ* / Ред. Л. В. Данилова, с. 569–617. М.: Наука).

Oosten J. 1996. Ideology and the Development of European Kingdoms. *Ideology and the Formation of Early States* / Ed. by H. J. M. Claessen, and J. G. Oosten, pp. 225–247. Leiden: Brill.

Oppenheim A. L. 1990. *Ancient Mesopotamia*. Moscow: Nauka. *In Russian* (Оппенхейм А. Л. *Древняя Месопотамия*. М.: Наука).

Orlova A. S., and L'vova E. S. 1978. *Historical Pages of the Great Savanna*. Moscow: Nauka. *In Russian* (Орлова А. С., Львова Э. С. *Страницы истории великой саванны*. М.: Наука).

Pershits A. I., and Khazanov A. M. (Eds.) 1978. *Primitive Periphery of Class Societies Prior to Great Geographical Discoveries (Problems of Historical Contacts)*. Moscow: Nauka. *In Russian* (Першиц А. И., Хазанов А. М. (Ред.) *Первобытная периферия классовых обществ до начала Великих географических открытий (проблемы исторических контактов)*. М.: Наука).

Pietkiewicz K. 2006. Grand Duchy of Lithuania. *The Early State, its Alternatives and Analogues* / Ed. by L. E. Grinin, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 304–334. Volgograd: Uchitel. *In Russian* (Петкевич К. Великое княжество Литовское. *Раннее государство, его альтернативы и аналоги* / Ред. Л. Е. Гринин, Д. М. Бондаренко, Н. Н. Крадин, А. В. Коротаев, с. 304–334. Волгоград: Учитель).

Podlazov A. V. 2000. *Theoretical Demography as the Basis of Mathematical History*. Moscow: IPM RAN. *In Russian* (Подлазов А. В. *Теоретическая демография как основа математической истории*. М.: ИПМ РАН).

Podlazov A. V. 2001. *The Basic Equation of Theoretical Demography and the Model of Global Demographic Transition.* Moscow: IPM RAN. *In Russian* (Подлазов А. В. *Основное уравнение теоретической демографии и модель глобального демографического перехода*. М.: ИПМ РАН).

Podlazov A. V. 2002. Theoretical Demography. Models of Population Growth and Global Demographic Transition. *New in Synergetics. A Look into the Third Millennium* / Ed. by G. G. Malinetsky, and S. P. Kurdyumov, pp. 324–345. Moscow: Nauka. *In Russian* (Подлазов А. В. Теоретическая демография. Модели роста народонаселения и глобального демографического перехода. *Новое в синергетике. Взгляд в третье тысячелетие* / Ред. Г. Г. Малинецкий, С. П. Курдюмов, с. 324–345. М.: Наука).

Potekhin I. I. 1954. South Africa. *Peoples of Africa* / Ed. by D. A. Olderogge, and I. I. Potekhin, pp. 524–616. Moscow: Izdatel'stvo Akademii nauk. *In Russian* (Потехин И. И. Южная Африка. *Народы Африки* / Ред. Д. А. Ольдерогге, И. И. Потехин, с. 524–616. М.: Изд-во Академии наук).

Preobrazhenskiy P. F. 2005. *Ethnology*. Moscow: Editorial URSS. *In Russian* (Преображенский П. Ф. *Курс этнологии*. М.: Едиториал УРСС).

Ratzel F. 1902. *Ethnography*: in 2 vols. St. Petersburg.: 'Prosveschenie'. *In Russian* (Ратцель Ф. *Народоведение*: в 2 т. СПб.: Просвещение).

Ritter E. A. 1968. *Shaka Zulu. The Rise of the Zulu Empire*. Moscow: Nauka. *In Russian* (Риттер Э. А. 1968. *Чака Зулу. Возвышение зулусской империи*. М.: Наука).

Roberts N. 1998. *The Holocene: An Environmental History.* Oxford: Blackwell.

Rybakov B. A. 1966a. Kievan Rus'. History of the USSR since the Ancient Times up to Now. Vol. 1. *Primitive Communal System. Ancient States of Transcaucasia and Central Asia. Ancient Rus* / Ed. by S. A. Pletneva, and B. A. Rybakov, pp. 476–572. Moscow: Nauka. *In Russian* (Рыбаков Б. А. Киевская Русь. *История СССР с древнейших времен до наших дней.* Т. 1. *Первобытнообщинный строй. Древнейшие государства Закавказья и Средней Азии. Древняя Русь* / Ред. С. А. Плетнева, Б. А. Рыбаков, с. 476–572. М.: Наука).

Rybakov B. A. 1966b. Rus' during an Era of 'The Song of Igor's Campaign'. Isolation of Independent Russian Principalities in the12th – the Beginning of the 13th Century. *History of the USSR since the Ancient Times up to Now.* Vol. 1. *Primitive-Communal System. Ancient States of Transcaucasia and Central Asia. Ancient Rus* / Ed. by S. A. Pletneva, and B. A. Rybakov, pp. 573–639. Moscow: Nauka. *In Russian* (Рыбаков Б. А. Русь в эпоху «Слова о полку Игореве». Обособление самостоятельных русских княжеств в XII – начале XIII в. *История СССР с древнейших времен до наших дней.* Т. 1. *Первобытнообщинный строй. Древнейшие государства Закавказья и Средней Азии. Древняя Русь* / Ред. С. А. Плетнева, Б. А. Рыбаков, с. 573–639. М.: Наука).

Sannikov S. V. 2002. The Structural-Comparative Typology of the Old German Societies (According to Written Sources). *Works of the 40th International Scientific Student's Conference 'Student and Scientific and Technical Progress'* / Ed. by S. G. Pyatkov, pp. 59–66. Novosibirsk: NGU. *In Russian* (Санников С. В. Структурно-компаративная типология древнегерманских обществ (по данным письменных источников). *Труды 40-й Международной научной студенческой конференции «Студент и научно-технический прогресс»* / Ред. С. Г. Пятков, с. 59–66. Новосибирск: НГУ).

Sannikov S. V. 2003. The Development of the Early Forms of the Royalty among the German People: Features of Politogenesis. *History and Sociology of State* / Ed. by A. P. Derevyanko, pp. 36–54. Novosibirsk: NGU. *In Russian* (Санников С. В. Развитие ранних форм королевской власти у германских народов: особенности политогенеза. *История и социология государства* / Ред. А. П. Деревянко, с. 36–54. Новосибирск: НГУ).

Sannikov S. V. 2005. The Religion in the Service of the Royalty in pre-Christian Germany and Scandinavia. *Vestnik NGU. Seriya Istoriya. Filologiya* 4(1): 33–42. *In Russian* (Санников С. В. Религия на службе королевской власти в дохристианской Германии и Скандинавии. *Вестник НГУ. Серия История. Филология* 4(1): 33–42).

Shchukin M. B. 2005. *Gothic Way (Goths, Rome and the Culture of the Chernyakhov)*. St. Petersburg: Philological faculty of St. Petersburg State University. *In Russian* (Щукин М. Б. *Готский путь (готы, Рим и черняховская культура)*. СПб.: Филологический факультет СПбГУ).

Service E. R. 1975. *The Origins of the State and Civilization. The Process of Cultural Evolution*. New York, NY: Norton.

Shkunaev S. V. 1989. The Community and Society of the Western Celts. Moscow: Nauka. *In Russian* (Шкунаев С. В. 1989. *Община и общество западных кельтов.* М.: Наука).

SIE. 1966. Maroboduus. *Soviet Historical Encyclopedia* / Ed. by E. M. Zhukov, Vol. 9, p. 123. Moscow: Sovetskaya entsiklopedia. *In Russian* (СИЭ. Маробод. *Советская историческая энциклопедия* / Ред. Е. М. Жуков, т. 9, с. 123. М.: Советская энциклопедия).

Taagepera R. 1976. Crisis around 2005 A.D.? A Technology-population Interaction Model. *General Systems* 21: 137–138.

Taagepera R. 1979. People, Skills, and Resources: An Interaction Model for World Population Growth. *Technological forecasting and social change* 13: 13–30.

Taagepera R. 1997. Expansion and Contraction Patterns of Large Polities: Context for Russia. *International Studies Quarterly* 41: 475–504.

Toynbee A. J. 1991. *A Study of History*. Moscow: Progress. *In Russian* (Тойнби А. Дж. *Постижение истории*. М.: Прогресс).

Tsirel S. V. 2004. On the Possible Reasons for the Hyperexponential Growth of the Earth Population**.** *Mathematical Modeling of Social and Economic Dynamics* / Ed. by M. G. Dmitriev, and A. P. Petrov, pp. 367–369. Moscow: Russian State Social University.

Turchin P. 2003. *Historical Dynamics: Why States Rise and Fall*. Princeton, NJ: Princeton University Press.

Turchin P. 2005. *War and Peace and War: Life Cycles of Imperial Nations*. New York, NY: Pi Press.

Turchin P. V. 2007. *Historical Dynamics. On the Way to Theoretical History.* Moscow: Komkniga/URSS. *In Russian* (Турчин П. В. *Историческая динамика. На пути к теоретической истории.* М.: КомКнига/URSS).

Turchin P. 2010. Warfare and the Evolution of Social Complexity: A Multilevel-selection Approach. *Structure and Dynamics* 4(3): 1–37.

Turchin P. 2015. *Ultrasociety: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth*. Chaplin, CT: Beresta Books.

Wright D. C. 2001. *The History of China.* Westport, CT: Greenwood.

Zinkina J., Ilyin I., and Korotayev A. 2017. The Early Stages of Globalization Evolution: Networks of Diffusion and Exchange of Domesticates, Technologies, and Luxury Goods. *Social Evolution & History* 16(1): 69–85.

Zinkina J., Korotayev A., and Andreev A. 2016. Circumscription Theory of the Origins of the State: A Cross-Cultural Re-Analysis. *Cliodynamics* 7(2): 187–203.

Zinkina J., Malkov A., and Korotayev A. 2014. A Mathematical Model of Technological, Economic, Demographic and Social Interaction between the Center and Periphery of the World System. *Socio-economic and Technological Innovations: Mechanisms and Institutions*. / Ed. by K. Mandal, N. Asheulova, and S. G. Kirdina, pp. 135–147. New Delhi: Narosa Publishing House.

10
Energy, Kondratieff Waves, Lead Economies, and Their Evolutionary Implications

William R. Thompson

Abstract

One way to look at the evolution of technological innovation is to develop ways to convert various types of matter into successively greater amounts of energy to fill sails, to spin cotton or to drive automobiles and air conditioners. One approach to interpreting Kondratieff waves (K-waves), associated with the leadership long cycle research program, emphasizes the role of intermittent but clustered technological innovations primarily pioneered by a lead economy, with various significant impacts on world politics. This approach is further distinguished by asserting that the K-wave pattern is discernible back to the 10th century and the economic breakthrough of Sung Dynasty China. While K-wave behavior has many widespread manifestations, the question raised in this essay is whether explanatory power is improved by giving a greater role to energy and energy transitions in the K-wave process(es). Eight specific implications are traced, ranging from the interaction of technological innovations and energy to cosmological interpretations. Our argument reflects a general theory of the evolution of complexity in all processes predicated on energy consumption. All 'natural entities', spanning physical, biological, and cultural phenomena, extract energy for survival, maintenance and reproductive purposes or, alternatively, put for resisting entropy.

Keywords: *Kondratieff waves, long cycle, energy, energy transitions, technological innovation, lead economy, evolutionary processes.*

One sign of a 'progressive' research program is whether its key assumptions are occasionally re-examined and revised as seems appropriate. The leadership long cycle program focuses on questions of informal governance in world politics but, unlike most other similar programs, emphasizes the role of technological innovation in lead economies, leading sectors, and Kondratieff waves. There is no need to jettison this emphasis. The lead economy-long wave should remain crucial to the program's explanatory infrastructure.

However, a case can be made for further elaborating how and why lead economies, technological innovation, and leading sectors are important and can best be interpreted. Elsewhere, I have argued (Thompson 2010) that technological innovation should not be divorced historically from interactions within a larger context of demographic changes, climate and disease factors. That is, technological innovation and the innovators are embedded in a larger socioeconomic fabric to which it and they respond. New technology is not an isolated, stand-alone driver. I would like to elaborate this type of argument further by incorporating energy considerations more explicitly as well. Rather than only emphasizing technological innovation in leader sectors *per se*, I suggest we consider the interactions of energy demands/consumption and technological innovation. It is not an either/or situation. Changes in energy sources need to be integrated with technological innovation and the technological innovation that is most important needs to be interpreted in terms of its significance for energy consumption (Goldstone 2002; Allen 2006; Griffin 2010). If we make this adjustment in core perspective, some things do not change. For instance, the indicators of technological innovation on which the research program has relied in the past do not need to change. But how they are viewed may require some adjustment. Moreover, there are also some interesting implications for speculating about future systemic leadership transitions.

In this paper, eight implications are highlighted. First, it is possible to argue that innovations in energy convertors or fuels are fundamental to the clusters of economic innovation that have been critical to long wave processes. This argument does not mean that the clusters of innovation are exclusively about energy factors but that energy considerations are closely linked to successive waves of innovation. A second implication pertains to the question of how far back in time one can trace K-waves. The leadership long cycle program finds evidence for K-waves activity back to the 10^{th} century in the form of technological innovations in Sung Dynasty China. But it is clear that the evidence is stronger after the late 18^{th} century British industrial revolution than before. One good reason is the two energy transitions that took place between the late 18^{th} and early 20^{th} centuries. The consequences of combining clustered technological innovation with energy transitions led to economic changes that are even more pronounced than in earlier centuries.

Third, one of the main foci of the leadership long cycle research program are long waves of economic growth which come in pairs or 'twin peaks' of clustered growth spikes. Energy considerations suggest reasons for these paired clusters of growth to be uneven in impact. The first peak should be less revolutionary in impact than the second because the first innovation wave must work within the prevailing economic landscape but the second wave has the advantage of building on the first wave's innovation set.

Since the leadership long cycle research program has focused primarily on the advent of technological innovation, adding energy considerations to the mix encourages an expansion of the focus to encompass resource acquisition and transportation activities as a fourth implication. Another implication of giving more attention to energy is the distinction between relative decline in production and export shares and achieving steady states in energy consumption. The steady state focus, in which periods of non-expansion of energy consumption predominate, may be more useful than focusing on, and debating, relative decline questions. This observation leads to a sixth implication in underlining the role of lead economies in leading the way through periods of energy transition and the development of reliance on new fuels. Steady states in energy consumption suggest that the gains from energy conversion processes have been maximized. New types of energy sources are needed to expand energy consumption. The next lead economy is likely to lead the way to the new types of energy sources.

Interpreting these processes in terms of energy acquisition and consumption makes it possible to link systemic leadership to ancient processes of development which helps to generalize the nature of the activities being examined. Further help in this regard is provided by a cosmological argument that energy consumption is the common denominator of the evolution of all natural processes. These last two implications reinforce the centrality of the processes focused upon and should help make the leadership long cycle research program seem less unorthodox overall – even while it proceeds from assumptions that are not widely accepted by scholars of world politics.

Before elaborating these implications further, it is probably helpful to first outline the leadership long cycle's perspective on Kondratieff waves (K-waves) and the role they play in the research program. Since they are an integral component to the overall theory, some of the associated components need to be introduced as well.

The K-Wave and the Leadership Long Cycle Argument[1]

Forty to sixty year Kondratieff waves (K-waves) are said to be driven by a host of different factors, including prices, technological innovation, energy transitions, demographic change, war, generational changes, investment, profits, and wages.[2] No doubt, there is something to be said for all of these claims in that the rhythms of long-term economic growth tend to encompass multiple phenomena. We simply have failed to sort out all of the interrelationships adequately. Yet it is difficult to proceed from the assumption that everything

---

[1] This overview section draws to some extent on a portion of Thompson (2007b).

[2] See, *e.g.*, Goldstein (1988) and Thompson (1990) for a discussion of the multiple Kondratieff wave interpretations. More recent variations include Devezas and Corredine (2001), and Rennstich (2008).

is related to everything else (even if it is). The leadership long cycle argument makes a number of assertions about how hierarchy is established in global politics – one of which privileges clustered, radical innovations in technology as the principal driving force of the K-wave.[3] This first set of assertions revolves around leading sectors which are industries built on radical innovations which have some potential of revolutionizing the way the economy is structured. Long-term growth is discontinuous and dependent on spurts in the development of these radical innovations. Radical innovations, in turn, generate new technology and industries characterized by high growth rates and alter the way old industries (characterized by slow growth rates) perform or, alternatively lead to their disappearance through Schumpeter's 'creative destruction' processes. Rapid growth on the part of the aggregate economy depends, of course, on the new, high growth sectors outperforming and more than offsetting the drag of the older, slow growth sectors.

It should be noted that these radical innovations are not simply a matter of the appearance of new hardware (Modelski 2001). Actors must learn how to cope with the implications of new technology and this takes time. Eventually, however, the effects become more routine as the new developments are assimilated, albeit unevenly around the world. Just how long this combination of hardware and perception process requires to work itself out, no doubt, is somewhat variable but probably approximates a generation. One generation is first exposed to the new technology and the following generation increasingly regards it as a routine way of doing things.

Initially, these leading sector trajectories were viewed as long waves or undulations of accelerated and slow economic growth. We have moved away from that conceptualization and now embrace the notion of a sequence of S-shaped growth curves. New sectors are introduced, grow quickly at first and then level off. Long-term economic growth is still subject to sequences of fast and slow growth but the underlying mechanism is the iterative introduction of new industries to replace old ones. Each new cluster of radical technological changes possesses an S-shaped trajectory that gradually flattens as its activities are perceived to become routine or even obsolete.

The introduction of radical innovations is monopolized by a single lead economy situated at or near the top of a global technological gradient on which the world's economies are organized hierarchically. At the bottom of the gradient, subsistence activities predominate. At the top, pioneering innovations for a time produce efficiency, productivity, and monopoly profits. The very frontiers of technology are extended with each radical inno-

---

[3] It shares the Schumpeterian emphasis on clustered, radical innovations with the Sussex school (Freeman and Louca 2001; Freeman and Perez 1988; Freeman and Soete 1997; Perez 2002) and Boswell and Chase-Dunn (2000).

vation in the ways in which commodities are produced. Technological innovation, imitation, and highly uneven diffusion makes movement up and down the gradient conceivable, but not necessarily all that likely. But as some other economies catch up eventually in harnessing the new technologies, the lead economy loses its lead.

Lead economies experience at least two waves of innovation in a process referred to as the 'twin peaks' phenomenon. The first wave (ascent) pushes a new economy to the top of the technological gradient. This highly destabilizing outcome encourages increased conflict and global warfare fought primarily among the states with economies situated near the top of the gradient. Thanks in part to the surpluses gained in the ascent wave and the consequent ability to organize a winning coalition, the lead economy's victory in the ensuing conflict is made more probable. Its resources are applied to funding capabilities of global reach (naval power later supplemented by air and space power) and coalitions of land and sea powers to defeat the most threatening adversaries.

The innovation lead in the first wave, intensive mobilization during the intensive conflict, and global war victory all combined to facilitate the lead economy's development of a second wave of clustered innovations. Most allies and rivals that participated in the global warfare emerge exhausted. The exception is the lead economy that actually profits from the conflict and extend its predominance as the premiere commercial-industrial and power with global reach. After the global war has ended, the coalition leadership in the global war has increasingly segued into something resembling systemic leadership. Yet, it is also in this immediate postwar era that other advanced economies narrow the gap with the economic leader's position. If the leader's first wave is one of ascent, the second wave of the pair is thus a catch up wave. As the system leader's capability foundation experiences relative decline after a few decades, so too does its opportunity to lead systemically.

Two other distinctive assumptions of this Kondratieff wave interpretation are that: 1) the perspective is evolutionary and 2) the K-wave pattern began to emerge faintly as early as the tenth century Sung China. No one argues that Kondratieff waves have been with us throughout recorded history. At some point, though, the long economic fluctuations with 40–60 year periodicity emerged. Only gradually were such processes likely to assume a shape that became easier to identify. In this case, the argument is that the first appearance of a paired K-wave pattern in economic innovations is found in the 10^{th} century in Sung China which is sometimes credited with developing the first economy with modern, industrialized features. Most importantly, the expansion of maritime trade in the South China Sea and the Indian Ocean, as well as the revived use of the Silk Roads on land, facilitated the transmission of long term, paired growth impulses to the other end of Eurasia via Venetian and Genoese intermediaries. Many of the economic innovations that later characterized western commercial and industrialized successes can be traced back to Chinese practices (Modelski and Thompson 1996; Hobson 2004). It is possible, therefore, to analyze nine twin-peaked processes or eighteen K-waves encompassing some one thousand years between 930 and 1973 (Modelski and Thompson 1996). Obviously, the claim that there have been as many as nineteen K-waves, counting the one that still seems to be in progress, is a major departure from K-wave convention. But there is no insistence that each set was as fully manifested as more recent ones. The K-wave process emerged only gradually and became most evident only in the past few centuries – a subject to which we will return.

No assumption is made that either technological change or capitalism suddenly emerged after the British industrial revolution. Both were amply exhibited for thousands of years, especially in activities involving long-distance commerce. But it was necessary to break free of economies dominated by relatively slow-moving agricultural dynamics fixated on interactions between climate, resource endowments, and population size. Early Chinese industrialism and commerce took a step in that direction. The process was aided and abetted subsequently by trading state behavior conducted by small Italian city-states and Portugal after the early Chinese experiment had failed.[4] Dutch, British, and U.S. innovations in commerce and industrialization of the past three to four centuries have contributed further to the increasing strength of long-term technological change rhythms.

Table 1 lists the lead economy history. Two Chinese (Northern and Southern Sung), two Italian (Genoa and Venice), a Portuguese, a Dutch, two British, and at least one U.S. set of paired innovation spurts are claimed. The radical innovations initially were largely focused on the development of the Chinese ‘national’ economy but not exclusively because trade's significance rose in the Southern Sung era. Thereafter, the emphasis shifted to commercial innovations through the 14th K-wave and industrial innovation courtesy of the British Industrial Revolution. The intention of the table is not to capture comprehensively everything that changed in each iteration but to draw attention to some of the more illustrative and profound changes around which each K-wave was focused.

[4] Part of China's problem was its distinctive threat environment and long struggle with nomadic attacks. The Mongols were able to defeat Sung China in part by using some of its technological innovations against the Chinese.

Table 1. Leading sector timing and indicators, from the 15th to 21st centuries

Lead Economy	Leading Sector Indicators	Start-up Phase	High Growth Phase
Portugal	Guinea Gold	1430–1460	1460–1494
	Indian Pepper	1494–1516	1516–1540
Netherlands	Baltic and Atlantic Trade	1540–1560	1560–1580
	Eastern Trade	1580–1609	1609–1640
Britain I	Amerasian Trade (especially sugar)	1640–1660	1660–1688
	Amerasian Trade	1688–1713	1713–1740
Britain II	Cotton, Iron	1740–1763	1763–1792
	Railroads, Steam	1792–1815	1815–1850
United States I	Steel, Chemicals, Electronics	1850–1873	1873–1914
	Motor Vehicles, Aviation, Electronics	1914–1945	1945–1973
United States II?	Information Industries	1973–2000	2000–2030
	?	2030–2050	2050–2080

One outcome of this pattern of economic leadership, seemingly new to the last millennium, is the development of a global system increasingly focused on the operations and management of long distance or inter-regional trade. This global system, initially Eurasian in scope and eventually planetary wide, functioned simultaneously with the more delimited foci of various regional systems. At the head of the global system (but not necessarily any of the various regional systems) is the lead economy that surges ahead of its competitors and rivals in an ascent K-wave only to find itself in an intensive bout of global warfare of generation length. Interestingly, while periods of conflict are found in the earlier paired sets of K-waves, successive rounds of global warfare only emerged halfway through the millennium in the 1490s. Western Europe was both multipolar and characterized by repeated and unsuccessful attempts, unlike most other regions, to acquire regional hegemony that were seen as being renewed in the 1490s.[5] As a consequence, global wars have combined and fused attempts of continental powers to assume the European hegemony with disputes over leadership at the global level. This process presumably ended in 1945 but could be transplanted to East Asia in the 21st century.

Ultimately, the K-wave process does not establish the world hegemony for the state possessing the lead economy. Rather, it propels the lead economy into the status of being the leading political-military-economic actor of a global system focused on long distance transactions, in marked distinction to regional power structures and attempts at territorial expansion in the home region. There is no need to equate the systemic leadership of Portugal in the 16th century with that of the United States in the second half of the 20th century beyond the minimal standard that both states established themselves as the leaders in global (*i.e.* interregional) economic innovation in their respective eras. The U.S. lead in 1945 was much greater in scope than the lead established by Portugal in 1517. Therefore, it is hardly surprising that the United States had a much stronger impact on shaping the postwar institutions of world order (as symbolized by the Bretton Woods package of the Generalized Agreement on Trade and Tariffs [GATT], the International Monetary Fund [IMF], and the United Nations [UN]) than did Portugal in the early 16th century.[6] In neither case was the system leader hegemonic. In both cases, the system leader had variable opportunities to shape the rules governing global system transactions.

The tripartite systemic leadership platform – leading sector growth (growth rates of leading sectors in the lead economy), leading sector share concentration (the lead economy's share of leading sector production among global powers), and global reach capability concentration (naval capability share) – is interrelated reciprocally. Leading sector growth leads to leading sector share concentration and global reach capabilities. Higher levels of global reach capabilities facilitate leading sector growth and leading sector share concentration. Yet, leading sector growth and share concentration also lead to military mobilization on land as well as at sea.

We have shown empirically (Reuveny and Thompson 2001, 2004) that the system leader's leading sector growth has been a systematic driver of the system leader's aggregate or national economic growth. Both of these variables,

[5] Rome, of course, had successfully unified most of what later became Western Europe but this early success was more a product of Mediterranean politics than it was 'European' politics. That is, Rome conquered the Mediterranean world and, in the process, peripheralized much of Europe to its empire centered in Italy.

[6] Portugal essentially created a protection racket regime in the western end of the Indian Ocean in which traders paid taxes to the Portuguese to be allowed to trade. The Portuguese could aspire to little more since their technological edge resided in ocean-going ships with cannon, as opposed to commodities that could be exchanged for Asian goods.

in turn, affect world economic growth positively while world economic growth influences the system leader's leading sector and national growth negatively. In some respects, then, the system leader is negatively affected by its own success. Its innovations contribute to world economic growth but as other economies improve their technological development, the advantageous position of the system leader is reduced.[7]

Eight Implications

Given the perspective outlined above, what might the increasing role of energy issues offer? Examining energy flows more closely should have payoffs for studying long economic waves or, more accurately, successive S-shaped technological growth trajectories. This is the first implication of incorporating energy into the leadership long cycle perspective. Smil (1994; see also Marchetti 1977), for instance, observes a close correspondence among Mensch's (1979) innovation cluster peaks, Schumpeter's peaks and troughs, and the introduction of new prime movers and fuels. Outlined in Table 2, Smil notes that each Kondratieff upswing was strongly influenced by the introduction of either new engines and new fuels, or both. The timing of these same early adoptions match the peaks of Mensch's (1979) innovation clusters (*i.e.* 1828 *vs.* 1830, 1880 *vs.* 1882, and 1937 *vs.* 1945) and the timing of Schumpeterian long wave trough centerpoints (1827 *vs.* 1828 and 1830, 1883 *vs.* 1882 and 1880, 1937 *vs.* 1945 and 1937). The midpoints of the Schumpeterian upswings are also roughly the midpoints of the prime mover/fuel eras. Smil regards this particular correspondence as more support for Mensch's argument that economic depressions stimulate new innovation waves.[8]

Finally, Smil also notes that a large number of the leading corporations in each prime mover era specialize in producing the new prime movers and associated fuels. Thus, the correspondences observed in Table 2 are hardly mysterious. Corporate activity provides the agency that links technological innovation and economic contraction and expansion. It is interesting to note, moreover, that Table 2 implicitly addresses the earlier implication about varied

[7] There are a number of other generalizations that can be made and that have been validated empirically about how the systemic leadership foundation influences other systemic processes, ranging from protectionism to the North-South cleavage (see Rasler and Thompson 1994; Reuveny and Thompson 2004; Thompson and Reuveny 2010).

[8] See as well Freeman's Sussex School emphasis (*e.g.*, Freeman and Perez 1988) on the key ingredients that will drive successive long waves. Most have an energy basis. At the same time, it should be noted that there is no standardization of K-wave periodicity as yet. Authors put forward approximations that sometimes overlap and sometimes do not. For instance, the Schumpeterian peaks in 1800 and 1856 in Table 2 do not exactly correspond to the relevant leadership long cycle high growth phases of 1763–1792 and 1815–1850. The 1911 and 1962 Schumpeterian peaks, though, do correspond with the 1873–1914 and 1945–1973 phases.

beats of the paired Kondratieffs. Focusing on the first column, the 1775–1830 period emphasized stationary steam engines while the 1830–1882 period stressed mobile steam engines, as found in trains and ships. The 1882–1945 period introduced internal combustion engines and steam turbines while the 1945–1990 period ushered in gas turbines. Note that engine power is substantially greater in the second period as compared to the first period when we look at these four eras as two sets of paired upswings.[9]

Table 2. Energy shifts and economic long waves, 1775–1990

Mensch Innovative Clusters Peaks	Schumpeterian Troughs	New Prime Movers and Fuels	Schumpeterian Peaks
		Stationary Steam Engines 1775–1830 (coal)	1800
1828	1827		
		Mobile Steam Engines 1830–1882 (coal)	1856
1880	1883		
		Steam Turbines and Internal Combustion Engines 1882–1945 (coal and crude oil)	1911
1937	1937		
		Gas Turbines 1945–1990 (coal, crude oil, and natural gas)	1962
	1990		

Source: Columns 1, 2, and 4 are based on Smil (1994: 240) who, in turn, drew on Mensch (1979) and Schumpeter (1939) for the peak and trough dates.

Nakicenovic (1991) considers these shifts (see Table 3) as substitution waves, with new technologies initially emerging in one era and becoming dominant in the next only to be supplanted by something else in a subsequent period. Precisely what comes next remains unclear. Natural gas sources of energy seem the most likely candidate at present but some mix of different sources will no doubt prevail. Which ones (or which mix) are selected, will depend

[9] Of course, each successive era also represents an expansion of engine power over the preceding era as well.

ultimately on changes in technology that make these alternative sources more reliable, safer, and less expensive.

The Smil and Nakicenovic tables, however, are suggestive about the role of energy transitions in the K-wave process. An energy transition is ongoing but not all that well advanced. It may take place later in the century and we think the hydrocarbon era is coming to an end but what will replace it remains vague. Substitution is ongoing slowly. No new fuels or engines (unless computers are seen as engines of a different kind) are yet evident either. If these generalizations are accurate, several possibilities are conceivable. If energy shifts have become a necessary part of the Kondratieff wave and have stalled for various reasons, does that portend parallel distortions to the shape of the current K-wave? The Sussex school (see, *e.g.*, Freeman and Perez 1988) argues that economic depressions result when there are delays in moving from one phase to the next due to the need to overcome resistance or obstacles to the next cluster of innovations. The current, protracted energy transition ultimately may come to be seen as such a delay.

Table 3. Clusters of pervasive technologies and substitution waves

Period	1750–1820	1800–1870	1850–1940	1920–2000	1980–2060
Dominant Systems	Water power, sails, turnpikes, iron castings, textiles	Coal, canals, iron, steam power, mechanical equipment	Railways, steam ships, heavy industry, steel, dyestuff, telegraph	Electric power, oil, cars, radio, TV, durables, petrochemicals	Gas, nuclear, aircraft, telecomm., information, photo-electronics
Emerging Systems	Mechanical equipment, coal, stationary steam, canals	Steel, city gas, indigo, telegraph, railways	Electricity, cars, trucks, radio, roads, oil, telephone, petrochemicals	Nuclear power, computers, gas, telecommunication, aircraft	Biotech., artificial intelligence, space industry and transport

Source: Based on Nakicenovic (1991: 486).

Alternatively, it may be that two energy transitions (first to coal and then to petroleum) were part of the K-wave history with fairly profound implications but that did not mean that energy shifts, at least in terms of fuels and engines, have become absolutely necessary to substitutions in clustered technology. Information technology, widely presumed to underlay contemporary technological changes, represents a different type of energy shift that may prove to be as difficult to assess while it is still ongoing as the shifts to coal and petroleum no doubt were.

The second implication follows from the first one. We discern 19 K-waves going back to the 10th century and Sung China. Roughly, most of the first two-thirds of this process was caught up in making use of wind for long-distance oceanic voyages which were carried out by relatively small states located on the fringe of Europe (Genoa, Venice, Portugal, the Netherlands, Britain). The voyages were profitable but harnessing wind was hardly new. The real innovations were focused on ship building (Venice, the Netherlands), improving navigation skills, or finding new routes (the Netherlands) to the Spice Islands. As impressive, profitable, and revolutionary for their time as these Asian and American trade connections were, they still seem to suffer in comparison with the revolutionary implications of new ways to manufacture products that were developed in the second half of the 18th century. One obvious explanation for this disjuncture is that an energy transition began in the late 18th century that substantially reinforced the impact of the Kondratieff process. From an evolutionary perspective, constant relationships are unlikely. Instead, they evolve over time, with some growing stronger and others becoming weaker. In this case, major energy transitions in the late 18th through early 20th centuries served to intensify the effects and consequences of clustered technological innovations.[10] The technological frontier was extended even more radically than in the past.

Another implication of giving more emphasis to the energy-technological innovation nexus is the nature of the twin peak phenomena. System leaders have tended to experience leads in innovation in sequential bursts of two upsurges, depicted in Table 4, that are separated by periods of global warfare. Hitherto, we have treated these paired innovation upsurges as equal. But in the context of interactions with energy, it takes time to transform the nature of energy conversion practices. As a consequence, the first burst in innovation tends to work within the prevailing economic landscape. The innovations may be radical but they are less likely to transform the economy to the full extent imaginable. The second one has the benefit of the earlier surge's changes and should be more revolutionary in its implications for how economic production is accomplished. Hence, the anticipated beat should not be 1–1 but, perhaps, something more like 1–1.5–2, with the second wave having a greater impact than the first. This differential beat rhythm is not a fact – merely a hypothesis taken and generalized from Griffin (2010: 123) who ar-

---

[10] A number of efforts to model K-waves based on aggregate data have been made without a great deal of success. Part of the problem is relying on the aggregate data but another part may be that the K-wave activity simply becomes more regular and therefore empirically discernible as we move toward the current period (see, *e.g.*, Korotayev and Tsirel 2010).

gues for a slow start for the British industrial revolution given the organic environment in which it began. She notes that the initial innovations relied on organic resources (horses, charcoal, and water) and then came to depend increasingly on inorganic resources (coal extracted from under the soil) with greater productivity as a result in a second surge. It may be that this differential beat is more discernible in more recent innovation surges. Nonetheless, the logic might well fit earlier growth surges too. Consider the Portuguese first growth surge based on West African pepper, slaves, and silver. Only in the second wave did the Portuguese enter the Indian Ocean. Or, the first Dutch growth surge was focused on its traditional Baltic trade. It is the second wave that is linked to the Dutch penetrating the Indian Ocean and the Spice Islands.[11] The initial 18th century British lead was predicated on its transportation of Asian products while the second wave was more focused on American production (*e.g.*, sugar and tobacco). It does not seem unwarranted to regard the first surge in the set to be more constrained by the environment in which the innovations occur in comparison to the second surge which can build on the first.

Table 4. The twin peak timing of leading sector growth surges and global war

First High Growth Surge	Global War	Second High Growth Surge
Portugal		
1460–1494	1494–1516	1516–1540
Netherlands		
1560–1580	1580–1609	1609–1640
Britain		
1660–1688	1688–1713	1713–1740
1763–1792	1792–1815	1815–1850
United States		
1873–1914	1914–1945	1945–1973

Incorporating energy obviously expands the focus on what lead economies need to do. This fourth implication is sketched in Fig. 1. Energy must have a source that can be tapped in some systematic matter.[12] Extraction and transportation from the source to production sites, therefore, becomes an important set of routines for the system leader either directly or indirectly. The focus on production sites (and commercial entrepots) is long standing and has been manifested in

[11] However, there are also strong incentives to re-examine Dutch energy utilization of peat and windmills. De Vries and van de Woude (1997) make a good case for calling the 17th century Netherlands the first modern economy.

[12] Keohane (1984: 32) argues that hegemons must control raw materials in addition to capital, markets, and competitive advantages in production. Once I thought a definitional emphasis on resource control was wrong but as long as the raw materials are focused on energy sources, I would now agree.

looking at sequences of pioneering and monopolizing leading sectors for periods of time. More storage and transportation of goods to their respective markets is the next step, followed by consumption, market share considerations, and waste associated with consumption.

Energy Source
↓
Extraction
↓
Transportation/Storage
↓
Commerce/Manufacture
↓
Transportation/Storage
↓
Consumption-Recycling/Destruction
↓
Waste/Environmental Degradation

Fig. 1. Energy flows

The leadership long cycle research program has focused primarily on the middle of this energy flow process, although the stress on naval power underlines the need for coercive protection of the two transportation links in the flow. Moreover, naval navigation hardware (compasses, rudders, and so forth) have also been standard foci (Devezas and Modelski 2008). But, fortunately, Bunker and Ciccantell (2005, 2007) have already analyzed the extraction-transportation and manufacture-transportation links. They do not look at what is manufactured; rather, they stress obtaining raw materials and building a transportation infrastructure. What is needed then is a synthesis of their model, perhaps subject to modifycations, into the leadership long cycle perspective. Waste is not exclusively a function of lead economy manufacture and consumption but it is likely to be a major, if not the major, source of problems associated with waste and, its corollary, environmental pollution. Were we to combine the production and consumption efforts of the lead economy and its main rivals, a lion's share of the generation of global wastes can be attributed to a small number of elite economies. Certainly, the lead economy is also a leader in waste and pollution production. Waste disposal and environmental degradation, thus, also become grist for the extended analytical mill.[13]

[13] Dealing with environmental degradations could well become a leading sector of the 21st century. See as well Chase-Dunn and Hall's (1997) iteration model and subsequent revisions that include environmental degradation as a function of economic productivity.

A fifth implication of giving more emphasis to energy is that some of the uncertainties of assessing relative decline may be eliminated. There are at least two problems that are affected. One is that it is remarkably difficult for most observers to distinguish between absolute and relative decline. Seeing no or little absolute decline, the popular reaction is what decline? Per capita income, for instance, falls in absolute decline phases but it is likely to improve in periods of relative decline. Without a clear impact on the quality of life, the notion of relative decline seems highly abstract. Relative decline is also difficult to gauge and even more difficult to assess in terms of its meaning. System leaders can enter into relative decline almost from the onset of their periods of predominance. Even so, any initial relative decline is apt to move very slowly and only pick up speed much later as competitors manage to catch up and perhaps, surpass the former leader. When other states and economies do transit past the incumbent leader, the relative decline becomes obvious. Before the point of transition, it is more nebulous even when many indicators point in the same direction.

The second problem lies with interpreting relative decline once it is recognized. How much decline makes a significant difference? If a system leader's lead diminishes by 10 %, is that huge, modest, or minor? Of course, that assessment must be contingent on the size of the gap between a leader and its followers. The greater is the size of the gap, the more room there is for relatively insignificant decline. But we have no practice in working out a metric that tells us when relative decline has reached significant proportions and when it has yet to pass some threshold mark. That has been especially the problem with interpreting U.S. relative decline. Its initial lead was quite commanding. Its rate of decline has been slow. It continues to possess a number of advantages over its rivals. Thus, it is not surprising that observers disagree contemporaneously about whether any decline has occurred.

One of the advantages of inputting more energy into the technological innovation box is that there is less emphasis on decline and more stress on attaining a steady state phase. Ascending economies tend to increase their consumption of energy. But at some point their increasing consumption levels off due to a combination of greater energy efficiency practices and reaching a point of optimal production given the types of energy sources that are available. The attainment of the phases of steady state energy consumption is quite clear in the British and U.S. cases.

Fig. 2 charts British consumption per capita as reported in Humphrey and Stanislaw (1979).[14]

[14] Humphrey and Stanislaw focus on mineral fuels and hydro-power and normalize their series in terms of 1800 = 100.

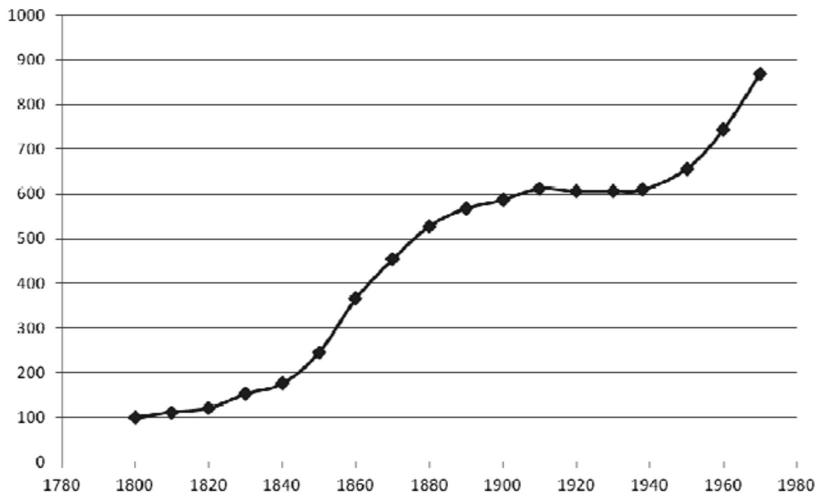


Fig. 2. British energy consumption per capita, 1800–1970

Not shown in Fig. 2 are estimates for the 18th century that suggest that energy consumption roughly doubled between 1700 and 1800 (47 to 100 on the index). Between 1800 and 1900, the increase in consumption per capita was nearly five-fold (100 in 1800 to 587 in 1900). The series peaked around 1910 and then went flat through World War II before beginning to ascend once again. The more contemporary (post-World War II) ascent, however, is associated with changes in fuel sources in a second energy transition. The flattening in the first half of the 20th century (and de-accelerating in the latter 19th century) presumably reflects the waning years of coal dependence as the principal fuel source, along with declining manufacturing activity.

Fig. 3 plots the U.S. energy consumption per capita in million BTUs.[15] Between 1950 and 1975, there was a 47 % increase (227 in 1950 to 333 in 1975). The series peaks in 1980 at 344 and stays flat through 2005, before declining in 2009. This last decline presumably reflects the global financial meltdown and losses in economic production and is thus likely to be temporary. Yet, overall, the series appears to have flattened from the 1970s on. As in the British case, there are multiple factors at work, including declining manufacturing demands and increased efficiency, but the combination of the two figures suggests that the flattening in Fig. 3 probably also reflects the waning years

[15] The data are taken from the U.S. Energy Information Administration's *Annual Energy Review, 2008* – see table 1.5 (Energy Consumption, Expenditures, and Emissions Indicators, Selected Years, 1949–2009), URL: http://www.eia.doe.gove/aer/pdf/pages/sec1_13pdf.

of the petroleum energy regime and the attainment of a steady state status in terms of energy consumption.[16]

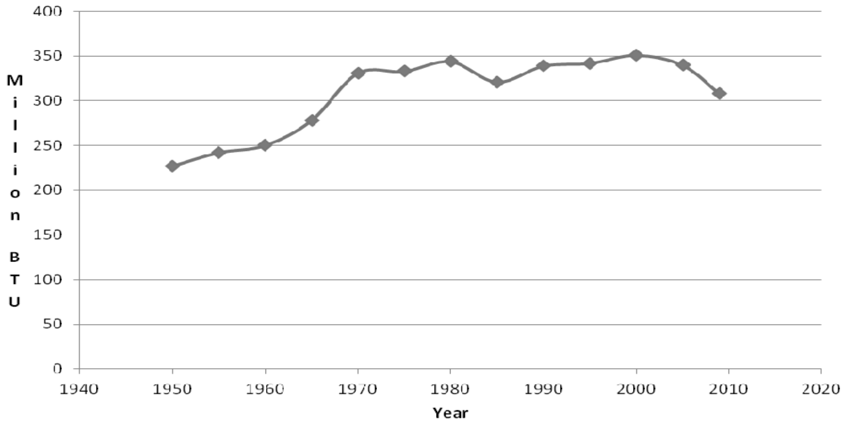


Fig. 3. The U.S. energy consumption per capita, 1950–2009

In this vein, LePoire (2009: 215) suggests that a transition to the Chinese leadership is a long way off. The Chinese energy consumption is very large but on a per capita basis is only about 10 % of the U.S. usage. That would imply that any plot of Chinese per capita consumption would show a positive trend perhaps for a number of years into the first half of the 21st century, other things being equal, but still not catching up to the leader. The other interesting facet of the Chinese consumption is that has been heavily dependent on coal and will probably continue to be reliant on coal through at least 2050.

From these observations, one might infer that the U.S. relative decline may easily be exaggerated, as are concerns about a transition to Chinese leadership in the near future. The real question from an energy perspective is which economy or economies will lead the way in replacing petroleum, especially in terms of automobile propulsion. Since we are in the very early stages of that movement, it is probably much too soon to tell – but it hints at what we might pay most attention.

The sixth implication is that leadership and energy transitions appear to have become increasingly intertwined. It makes sense that if lead economies are the vanguard of new and increased energy supply and consumption, they would also be an important agent in ushering in new eras of energy use. This tendency did not emerge full-blown with the advent of lead economies. Only

[16] A related issue is the quite significant extent to which the U.S. trade deficits are expanded by petroleum imports.

the last two lead economies, Britain and the United States, have been involved so far in the transitions depicted in Fig. 4.

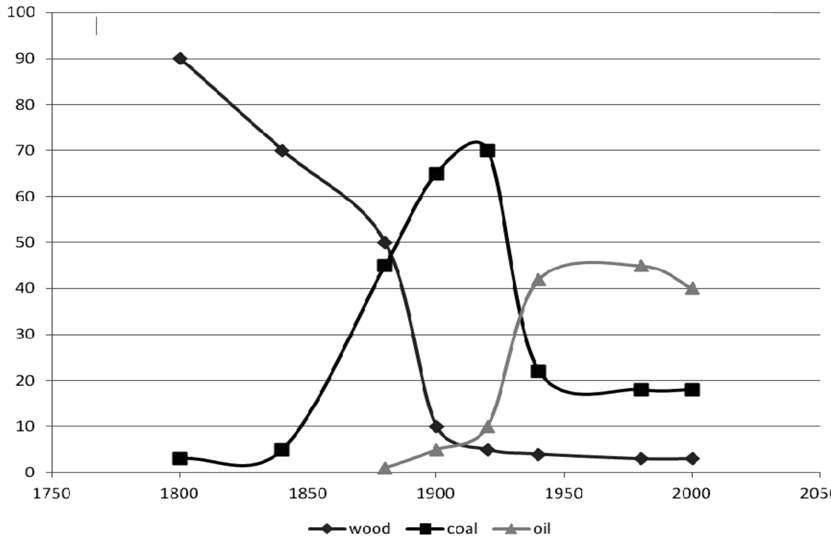


Fig. 4. Energy transitions in the United States

Britain led the shift to coal and competed intensely with the United States for control of petroleum reserves in the interwar years (Hugill 2011). By the beginning of World War II, the United States controlled some 50 % of the world's then known petroleum sources (Thompson 2007a).

It follows then that when we are speculating about leadership transition, it is not enough to simply look for innovation in a new wave of gadgets. We should also be looking for leadership toward a new era of energy use in which movement away from reliance on hydrocarbon sources is part of the pattern. In other words, the next lead economy will probably be the vanguard of employing alternative sources of energy – whether it be nuclear, solar, wind, natural gas, or some combination. It may also be that one reason for leadership transition is some inherent disadvantage in making the transition to the next era. Britain, for instance, was heavily committed to coal, did not possess large petroleum reserves at home, and was slow to make the switch to electricity. Given the pronounced U.S. reliance on petroleum, we may find that economies that are less dependent thanks to a lower level of development will encounter less inertia and resistance in the movement toward new energy sources.[17] Alternatively, the next lead economy is likely to need to have ample access to relatively inexpensive energy resources. The question may then hinge on the distribution of resource endowments.

Recognizing systemic leadership as a vanguard of new energy consumption practices creates opportunities to link contemporary processes to both ancient and cosmological processes. Early centers of ‘civilization’ developed similar resource acquisition networks and innovated novel ways to expand the supply of energy by building and maintaining irrigation canals and other ways to control water use. Sumer, the initial lead economy, is the example par excellence. What lead economies do is a more modern extension of older and even ancient political-economic practices and processes. We need to appreciate the continuity and to build on it analytically.

If a stronger connection to ancient developments is the seventh implication, an eighth is an intriguing link to a cosmological argument. Chaisson (2001: 120) contends that the ‘emergence, growth, and evolution of intricately complex structures’ is keyed to energy flows and governed by thermodynamic principles.

> Nature's many ordered systems can now be regarded as intricately complex structures evolving through a series of instabilities. In the neighborhood of a stable (equilibrium) regime, evolution is sluggish or nonexistent because small fluctuations are continually damped; destruction of structure is the typical behavior wherein disorder rules. By contrast, near a transition (energy) threshold, evolution accelerates and the final state depends on the probability of creating a fluctuation of a given type. Once this probability becomes appreciable, the system eventually reaches a unique though dynamic steady state, in which construction of structure wherein order rules is distinctly possible. Such states are thereafter starting points for further evolution to other states sometimes characterized by even greater order and complexity (Chaisson 2001: 78).

This argument (see also Adams 1975, 1982, 2010; Spier 2005, 2010) reflects a general theory of the evolution of complexity in all processes predicated on energy consumption. All ‘natural entities’, spanning physical, biological, and cultural phenomena, extract energy for survival, maintenance and reproductive purposes or, alternatively, put for resisting entropy. Greater complexity is achieved by tapping into greater quantities of matter and ener-

---

[17] One area worth more exploration are the implications of the system leader's dependence on weapons platforms developed in earlier global warfare but also reflecting a dependence on the prevailing energy regime. The commitment to the petroleum fueled twentieth century ships, tanks, and planes well into the 21st century would seem to be a good example.

gy. Table 5 offers a representative list of the 'free energy rate density' – an index of the amount of energy available per unit of mass – of various types of structures. All of these entities take energy from their environment to continue functioning. We are most familiar with our own participation in this fundamental process. Food allows us to live. Without food energy, we die. So it is with all other entities.[18]

Table 5. Some representative, estimated free energy rate densities

Structure	Average Densities
Galaxies	0,5
Stars	2
Planets	75
Plants	900
Animals	20,000
Human brains	150,000
Society	500,000

Note: The densities are expressed in erg units of energy per time per mass.
Source: Chaisson (2001: 139).

The attractiveness of this interpretation for our own purposes is that it provides a different way to view human efforts to improve their existence and quality of life. The basic process is one of energy acquisition and the expansion of how much energy is acquired. One way to look at the evolution of technological innovation, then, is the development of ways to convert various types of matter into successively greater amounts of energy to fill sails, to spin cotton or to drive automobiles and air conditioners. This process, over time, has moved along at different rates but is similar from the expansion of Sumer's resource acquisition network in the 4th millennium BCE to contemporary competitions to find ways to move automobiles by electricity or to convert solar energy into electricity. Political economies become successively more complex as energy densities are increased. But the process of acquiring and harnessing more and more sources of energy is not characterized by widespread innovation. It tends to occur first in one place and diffuse unevenly to other places that are in a position to emulate and, often, to improve on the initial innovations.[19]

[18] One interesting hypothesis is whether each successive lead economy is associated with significant improvements in the free energy rate density.

[19] There are certainly exceptions to this pattern. Agriculture, for instance, was invented independently in multiple places.

This basic pattern of pioneering innovations subject to uneven diffusion has structured long-term economic growth and is most clearly discernible in the Sung-Genoa-Venice-Portugal-Netherlands-Britain-United States succession in pioneering lead economies in the modern era of the last millennium. But it is not just successive clusters of innovation that is involved but also successive increases in the flow of energy acquired and energy density. The ability to convert sources of energy into successive advances in transportation and production is what long-term economic growth is all about.[20] Lead economies are thus principal agents in generating new drivers for economic development and growth. We should expect each successive leader to be associated with increased free energy rate densities. The leadership long cycle research program is organized very fundamentally around this insight. If the core process being examined also fits into a larger picture of parallel patterns in growth and development from the Big Bang on, so much the better. It reinforces the belief that the research program is on the right track. At the same time it also broadens and helps to justify lengthening the track on which the research program proceeds.

Technological innovation is about many things. The argument here is not that we scrap what has been said previously about the linkages between innovation and world politics. Rather, we need to broaden the nature of the inquiry by integrating energy considerations into the long cycle weave. The two perspectives are complementary because technological innovation and energy have been highly interdependent. Greater integration should enhance our understanding of both energy, the K-wave phenomenon, and processes of world politics.

References

Adams R. N. 1975. *Energy and Structure: A Theory of Social Power.* Austin, TX: University of Texas Press.

Adams R. N. 1982. *Paradoxical Harvest: Energy and Explanation in British History, 1870–1914.* Cambridge: Cambridge University Press.

Adams R. N. 2010. Energy, Complexity and Strategies of Evolution: As Illustrated by Maya Indians of Guatemala. *World Futures* 66: 470–503.

Allen R. C. 2006. *The British Industrial Revolution in Global Perspective: How Commerce Invented the Industrial Revolution and Modern Economic Growth.* URL: http://www. Nuffield.ox.ac.uk/users/allen/unpublished/econinvent-3.pdf.

[20] LePoire (2009: 217) offers an interesting frame on this problem by arguing for viewing history as a complex adaptive process in which succeeding phases of energy intensification over time have led to greater complexity. He thinks the succeeding phases are recognizable in five-fold expansions in energy intensity and dates them as follows: civilization (3000 BCE–400 CE), commerce (700–1720), industrialization (1720–1950), and knowledge-based (1950–?).

Boswell T., and Chase-Dunn C. 2000. *The Spiral of Capitalism and Socialism: Toward Global Democracy.* Boulder, CO: Lynne Rienner.

Bunker S. G., and Ciccantell P. S. 2005. *Globalization and the Race for Resources.* Baltimore, MD: Johns Hopkins University Press.

Bunker S. G., and Ciccantell P. S. 2007. *East Asia and the Global Economy: Japan's Ascent, with Implications for China's Future.* Baltimore, MD: Johns Hopkins University Press.

Chaisson E. J. 2001. *Cosmic Evolution: The Rise of Complexity in Nature.* Cambridge, MA: Harvard University Press.

Chase-Dunn Ch., and Hall Th. D. 1997. *Rise and Demise: Comparing World-Systems.* Boulder, CO: Westview Press.

Devezas T. C., and Corredine J. T. 2001. The Biological Determinants of Long-Wave Behavior in Socioeconomic Growth and Development. *Technological Forecasting and Social Change* 67: 1–57.

Devezas T. C., and Modelski G. 2008. The Portuguese as System-Builders: Technological Innovation in Early Globalization. *Globalization as Evolutionary Process: Modeling Global Change* / Ed. by G. Modelski, T. C. Devezas, and W. R. Thompson, pp. 30–57. London: Routledge.

Freeman Ch., and Louca F. 2001. *As Time Goes by: From the Industrial Revolution to the Information Revolution.* Oxford: Oxford University Press.

Freeman Ch., and Perez C. 1988. Structural Crises of Adjustment, Business Cycle, and Investment Behavior. *Technical Change and Economic Theory* / Ed. by G. Dosi, Ch. Freeman, R. Nelson, G. Silverberg, and L. Soete, pp. 38–66. London: Pinter.

Freeman Ch., and Soete L. 1997. *The Economics of Industrial Evolution.* 3rd ed. Cambridge, MA: MIT Press.

Goldstein J. 1988. *Long Cycles.* New Haven, CT: Yale University Press.

Goldstone J. A. 2002. Efflorescences and Economic Growth in World History: Rethinking the 'Rise of the West' and the British Industrial Revolution. *Journal of World History* 13(2): 323–389.

Griffin E. 2010. *A Short History of the British Industrial Revolution.* New York: Palgrave-Macmillan.

Hobson J. M. 2004. *The Eastern Origins of Western Civilization.* Cambridge: Cambridge University Press.

Hugill P. J. 2011. *The World Struggle for Oil: The Type II Transition Struggle between Britain and America.* Paper presented at the annual meeting of the International Studies Association. Montreal, Canada, March.

Humphrey W. S., and Stanislaw J. 1979. Economic Growth and Energy Consumption in the UK, 1700–1975. *Energy Policy* 7(1): 29–42.

Keohane R. O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy.* Princeton, NJ: Princeton University Press.

Korotayev A. V., and Tsirel S. V. 2010. A Spectral Analysis of World GDP Dynamics: Kondratieff Waves, Kuznets Swings, Juglar and Kitchin Cycles in Global Economic Development, and the 2008–2009 Economic Crisis. *Structure and Dynamics: eJournal of Anthropological and Related Sciences* 4(1): 1–55.

LePoire D. J. 2009. Exploration of Connections between Energy Use and Leadership Transitions. *Systemic Transitions: Past, Present, and Future* / Ed. by W. R. Thompson, pp. 205–222. New York: Palgrave-Macmillan.

Marchetti C. 1977. Primary Energy Substitution Models: On the Interaction between Energy and Society. *Technological Forecasting and Social Change* 10: 345–356.

Mensch G. 1979. *Stalemate in Technology.* Cambridge, MA: Ballinger.

Modelski G. 2001. What Causes K-Waves? *Technological Forecasting and Social Change* 68: 75–80.

Modelski G., and Thompson W. R. 1996. *Leading Sectors and World Powers: The Coevolution of Global Politics and Economics.* Columbia, SC: University of South Carolina Press.

Nakicenovic N. 1991. Diffusion of Pervasive Systems: A Case of Transport Infrastructure. *Diffusion of Technologies and Social Behavior* / Ed. by N. Nakicenovic, and A. Grubler, pp. 483–510. Berlin: Springer-Verlag.

Perez C. 2002. *Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages.* Cheltenham: Elgar.

Rasler K., and Thompson W. R. 1994. *The Great Powers and Global Struggle, 1490–1990*. Lexington, KY: University Press of Kentucky.

Rennstich J. K. 2008. *The Making of a Digital World: The Evolution of Technological Change and How It Shaped Our World.* New York: Palgrave Macmillan.

Reuveny R., and Thompson W. R. 2001. Leading Sectors, Lead Economies, and Their Impact on Economic Growth. *Review of International Political Economy* 8(4): 689–719.

Reuveny R., and Thompson W. R. 2004. *Growth, Trade and Systemic Leadership.* Ann Arbor, MI: University of Michigan Press.

Schumpeter J. A. 1939. *Business Cycles: A Theoretical and Statistical Analysis of the Capitalist Processes.* New York: McGraw-Hill.

Smil V. 1994. *Energy in World History.* Boulder, CO: Westview.

Spier F 2005. How Big History Works: Energy Flows and the Rise and Demise of Complexity. *Social Evolution & History* 4(1): 87–135.

Spier F. 2010. *Big History and the Future of Humanity.* London: Wiley-Blackwell.

Thompson W. R. 1990. Long Waves, Technological Innovation, and Relative Decline. *International Organization* 44: 201–33.

Thompson W. R. 2007a. Global War and the Foundations of U.S. Systemic Leadership. *America. War and Power, 1775–2000* / Ed. by J. Fuller, and L. Sondhaus, pp. 146–168. New York: Routledge.

Thompson W. R. 2007b. The Kondratieff Wave as Global Social Process. *World System History, UNESCO Encyclopedia of Life Support Systems* / Ed. by G. Modelski, and R. A. Denemark. Oxford: EOLSS Publishers. URL: http://www.eolss.net.

Thompson W. R. 2010. Synthesizing Secular, Demographic-Structural, Climate, and Systemic Leadership: Moving toward Explaining Domestic and World Politics in the Last Millennium. *Cliodynamics: The Journal of Theoretical and Mathematical History* 1(1): 26–57.

Thompson W. R., and Reuveny R. 2010. *Limits to Globalization and North-South Divergence.* London: Routledge.

Vries J. de, and van der Woude A. 1997. *The First Modern Economy: Success, Failure and Persistence of the Dutch Economy, 1500–1815.* Cambridge: Cambridge University Press.

11
Technological Dimension of Big History and the Cybernetic Revolution*

Leonid E. Grinin and Anton L. Grinin

Abstract

The present paper analyzes the evolution of technology from the beginning of the human history. We introduce a new paradigm to analyze the causes and trends of the global evolution. We also describe the direction of technological transformations, discuss and explain the present and forthcoming technological changes.

Our analysis of technological evolution mainly focuses on the second half of the 20th century. We present a detailed analysis of the latest technological revolution which we denote as 'Cybernetic', and give some forecasts about its development up to the end of the 21st century. It is shown that the development of various self-regulating systems will be the main trend of this revolution. We argue that the technological transition of the final phase of the Cybernetic Revolution will start in medicine, which is to be the keystone of technological convergence forming the system of MANBRIC-technologies (based on medicine, additive, nano, bio, robotic, IT and cognitive technologies). Today we are at the threshold of post-human revolution, the era of an intensive impact on the human body. The authors consider the directions of this revolution such as considerable life extension, organ replacement, BCIs, robotics, genome editing, etc. It is very important to understand the mechanisms of technological development and to measure the possible risks arising from them.

Keywords: *production revolutions, technological evolution, self-regulating systems, MANBRIC-technologies.*

Introduction. Between Human and Post-Human Revolutions

For centuries, technological changes were among the most fundamental drivers of social development, providing demographic growth and cultural progress. There is much information about the latest technological achievements appearing every day. However, for most of the human history the matter was diffe-

* The research is supported by the Russian Foundation for Basic Research (Project No. 17-02-00521-ОГН).

rent. For centuries and even millennia transformations have been undistinguished (Anuchin 1923; Lurie *et al.* 1939; Semyonov 1968; Chernousov *et al.* 2005; Belkind *et al*. 1956; Boas 1911; Kosven 1953; Kremkova 1936; Osipov 1959; Virginsky and Khoteenkov 1993; Sheypak 2009).

That is why it is very important to observe technological development comprehensively, where it becomes especially evident that, according to Fernand Braudel (1985), 'in reality, everything rested upon the very broad back of material life; when material life expanded, everything moved ahead'.

Technologies have been playing a significant role in the history of humankind from the very origin of *Homo sapiens*. Numerous facts show that already after 50,000 BP technologies were developed in various fields: from hunting and cooking to primitive painting. Achievements in such fields as agriculture, building, transportation, and many other human achievements could not have emerged without certain technologies. Thus, one can argue that *technologies play a very important role in Big History.* They played a special role in collective learning which is defined as the sixth threshold of increasing complexity. This *Homo Sapiens*' achievement which happened at the beginning of the Upper Paleolithic was probably one of the most important events in human history, and sometimes is termed as the *Human revolution* (*e.g.*, Shea 2006).[1] Today we are at the threshold of another important transition which is often called 'post-human revolution', which could bring quite radical changes to society and even transform the human biological nature.

1. TECHNOLOGICAL DIMENSION OF BIG HISTORY

Three Production Revolutions – Three Big History Thresholds

The whole historical progress can be divided into four big periods which we denote as four **production principles**[2]:

1. Hunter-Gatherer;[3]
2. Craft-Agrarian;
3. Trade-Industrial;
4. Scientific-Cybernetic.

Each production principle starts with great technological breakthrough which we denote as Production revolution. There were three such revolutions:

1) the Agrarian or Neolithic Revolution (12,000–10,000 – 5,500–3,000 BP);

2) the Industrial Revolution (the last third of the 15th – the first third of the 19th centuries);

3) the newest Cybernetic one (1950– the 2060/2070s).

---

[1] Sometimes we denote it as the Upper Paleolithic Revolution.

[2] See Grinin 2006a, 2006b, 2007a, 2007b, 2012; Grinin L. and Grinin A. 2013; Grinin A. and Grinin L. 2015.

[3] It lasted till the 12th mil. BP.

In respect of Big History these revolutions are tightly related to the main Big History thresholds (the Agricultural Threshold; Modern Revolution Threshold; and the Cybernetic Revolution is related to the Future Ninth Threshold[4]).

Production revolutions are technological breakthroughs which change the whole structure of society and the way of life. Each production revolution has its own cycle consisting of three phases: two innovative phases and between them – a modernization phase (see Tables 1, 2; Fig. 1).

At *the initial innovative phase* a new productive sector emerges. *The modernization phase* is a long period of distribution and development of innovations. It is a period of progressive innovations when the conditions gradually emerge for the final innovative breakthrough. At *the final innovative phase* new innovations dramatically spread and improve for the new production principle, which, at this time, attains full strength.

The Agrarian Revolution *was a great breakthrough from hunter-gatherer production principle to farming.* Its **initial phase** was a transition from hunting and gathering to primitive hoe agriculture and animal husbandry (that took place around 12,000–9,000 BP). The **final phase** was a transition to intensive agriculture (with large-scale irrigation and plowing) which started around 5,500 years ago (for more details see Grinin 2007a; Grinin A. and Grinin L. 2015; Grinin L. and Grinin A. 2016). These changes are also presented in Table 1.

Table 1. The phases of the Agrarian Revolution

Phases	Type	Name	Dates	Changes
Initial	Innovative	Manual agriculture	12,000–9,000 BP	Transition to primitive manual (hoe) agriculture and cattle-breeding
Middle	Modernization	Diffusion of agriculture	9,000–5,500 BP	Emergence of new domesticated plants and animals, development of complex agriculture, emergence of a complete set of agricultural instruments
Final	Innovative	Irrigated and plow agriculture	5,500–3,500 BP	Transition to irrigative or plow agriculture without irrigation

The Industrial Revolution *was a great breakthrough from craft-agrarian production principle to machine industry, marked by intentional search for and use of scientific and technological innovations in the production process.*

Its **initial phase** started in the last third of the 15th and 16th centuries with the development of shipping, technology and mechanization based on the watermill

[4] A number of Big History researches connect this threshold with the so-called singularity.

as well as with a 'more organic' division of labor. The **final phase** was the well-known breakthrough of the 18th and 19th centuries with the introduction of various machines and steam energy (for more details about Industrial Revolution see Grinin 2007b; Grinin A. and Grinin L. 2015; Grinin L. and Grinin A. 2016; Grinin and Korotayev 2015). These changes are presented in Table 2.

Table 2. The phases of the Industrial Revolution

Phases	Type	Name of the phase	Dates	Changes
Initial	Innovative	Manufacturing	the last third of the 15th – 16th centuries	Development of shipping, technology and mechanization on the basis of water engine, development of manufacture based on the division of labor and mechanization
Middle	Modernization	Diffusion of industrial enterprises	the 17th – early 18th centuries	Formation of complex industrial sector and capitalist economy, increase in mechanization and division of labor
Final	Innovative	Machinery	1730–the 1830s	Formation of sectors with the machine cycle of production using steam energy

The Cybernetic Revolution *is a great breakthrough from industrial production to the production and services based on self-regulating systems.*

Its **initial phase** dates back to the 1950–1990s. The breakthroughs occurred in the spheres of automation, energy production, synthetic materials production, space technologies, exploration of space and sea, agriculture, and especially in the development of electronic control facilities, communication and information. We assume that the **final phase** will begin in the nearest decades, *i.e.*, in the 2030s or a bit later, and will last until the 2070s.

We denote the initial phase of the Cybernetic Revolution as **a scientific-information** one, and the final one – as **a phase of self-regulating systems**. Today we are in its modernization phase which will probably last until the 2030s. This intermediate phase is a period of rapid distribution and improvement of the innovations made at the previous phase (*e.g.*, computers, Internet, cell phone, *etc.*). The technological and social conditions are also prepared for the future breakthrough. *We suppose that the final phase of the Cybernetic Revolution will lead to the emergence of various self-regulating systems* (see below).

The scheme of the Cybernetic Revolution is presented in Fig. 1.

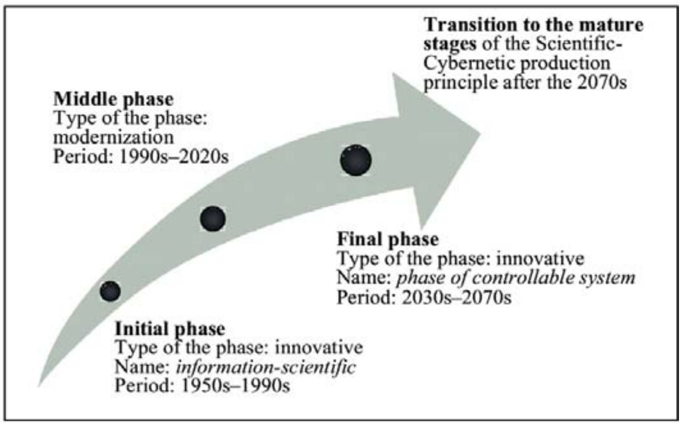


Fig. 1. The phases of the Cybernetic Revolution

Each phase of Big History is accompanied by the emergence of new evolutionary mechanisms. In particular, certain preconditions and preadaptations can be already detected within its previous phase. The same refers to the development of productive forces. Within previous production principle there appear some prerequisites of technologies, which flourish during the next production principle. For example, many mechanisms, engines and machines emerged within the Craft-Agrarian production principle, especially during its last centuries (the 12^{th} – 14^{th} centuries). But for Production revolution to start, there should occur some technological changes. Thus, the Industrial Production Revolution began at the end of the 15^{th} century and lasted until 1830.

Cybernetic Revolution, Self-Regulation and Artificial Intelligence in Terms of Big History

The theory of production revolutions proceeds from the assumption that the essence of these revolutions can be clearly observed only during their final phase. *The most important thing about the final phase of Cybernetic revolution will be a wide use of* ***self-regulation in different technological and bio-socio-technological systems***. The analysis of such systems can be based on cybernetics which is a transdisciplinary approach for exploring complex regulatory systems via the processes of receiving, transformation and transfer of information (see, *e.g.*, Wiener 1948; Beer 1959; Von Foerster and Zopf 1962; Umpleby and Dent 1999).

The most important characteristics and trends of Cybernetic Revolution are the following:

1. The increasing amount of information and growing complexity;
2. Consistent development of the system's abilities to the regulation and self-regulation;

3. Mass use of artificial materials with new properties;

4. Application and control of systems and processes of various nature including living material and new levels of organization of matter (including different nanoparticles as building blocks);

5. Miniaturization and microtization as a trend of the constantly decreasing mechanisms, electronic devices, implants, *etc.*;

6. Ubiquitous resource and energy saving;

7. Individualization as one of the most important technological trends;

8. Implementation of smart technologies and a trend towards humanization of their functions (use of the human language, voice, movements, *etc.*);

9. Control over human behavior and activity to eliminate the negative influence of the so-called human factor.

Some of these trends coincide wholly or partially with the perceptions of artificial intelligence and its future development (though this concept is quite vague and difficult to define). But other trends cannot be included into the concept of artificial intelligence (for more details see below).

Self-regulation can be defined as a system's ability to preserve stability and basic parameters within changing environment. Self-regulation as a broad concept incorporates various aspects of maintaining stable state of a system at all phases of Big History and especially at the biological and social ones. Self-regulation is of great importance for Big History since it is one of the most developed levels of growing complexity (see Grinin 2016).

Self-regulation has already revealed at the early phases of Big History, in fact, with the emergence of the first systems (*e.g.*, the first stars). The emergence of life is tightly connected with self-regulating systems.

In the course of chemical evolution chemical substances gradually became more complex until some of them got the ability for self-regulation. For example, lipids, which are able to change their form when interacting with water, while retaining its chemical structure. One of the most important features of living organisms is an existence of a code molecule. RNA is considered as the first self-replicating molecule.

The further formation of complex systems, such as DNA, proteins, enzymes, *etc.*, required the creation of a complex system of regulation. The more complicated the system became, the more complicated was its regulation. In order to overcome the entropy, systems sought to isolate themselves from direct contacts with the environment, forming protective (insulating) shells. Presumably that is how the first coacervates were formed, and later – the cells. A cell became the main self-regulating living system due to which many organisms were formed in the process of evolution. Biological systems demonstrate the complexity growing up to the level of self-regulation within evolution. Due to collective learning human society has also developed into a complex self-regulating system. Within the next decades the technological complexity is supposed to rapidly increase thus promoting the ability for self-regulation.

At present there are already many self-regulating systems around us, for example, self-driving cars, the artificial Earth satellites, pilotless planes, navigators laying the route for a driver. One more good example is life-supporting systems (such as medical ventilation apparatus or artificial heart). They can regulate a number of parameters, choose the most suitable mode and detect critical situations. The genetic engineering is also worth mentioning since it is used for the creation or changing biological and physiological self-regulating systems.

We suppose that during the final phase of the Cybernetic Revolution different *developmental trends should produce a cluster of technological innovations*. The medical sphere has unique opportunities to combine the abovementioned technologies into a single complex. In our opinion, *the general driver of this cluster will be medicine, which can connect additive technology, nanotechnology, biotechnology, robotics, information and cognitive technologies. We denote this technological cluster as a MANBRIC-complex (an acronym for the included technologies).*[5]

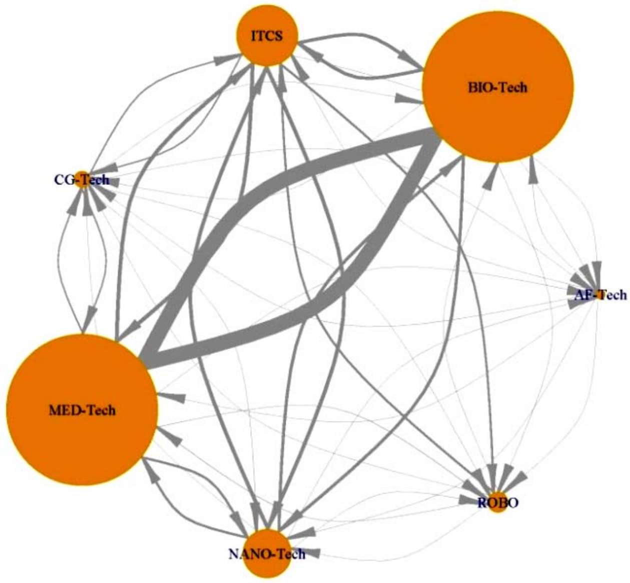


Fig. 2. The relationship between citation frequency in scientific publications and the technologies forming MANBRIC, according to the Web of Science, 2010–2015

[5] Namely: **M**edicine, **A**dditive, **N**ano, **B**io, **R**obotic, **I**nformation, and **C**ognitive technologies. For the convenience of pronunciation the technologies are listed not in order of priority.

Fig. 2 shows the citation frequency of MANBRIC-technologies in scientific publications and relations between the technologies forming the complex. The thickness of line demonstrates the intensity of interactions while the direction of arrows shows the sphere of application of technologies.

As one can see from this figure, medicine and biotechnologies are most closely related. There is also distinguished a separate direction of biomedicine (Pankhurst *et al.* 2003; Gupta A. and Gupta M. 2005).

The development of MANBRIC-complex can be examined, for example, via the analysis of patent applications in medicine, pharmaceuticals, and biotechnologies which also demonstrate *convergening growth rates* (Grinin L., Grinin A., and Korotayev 2016).[6]

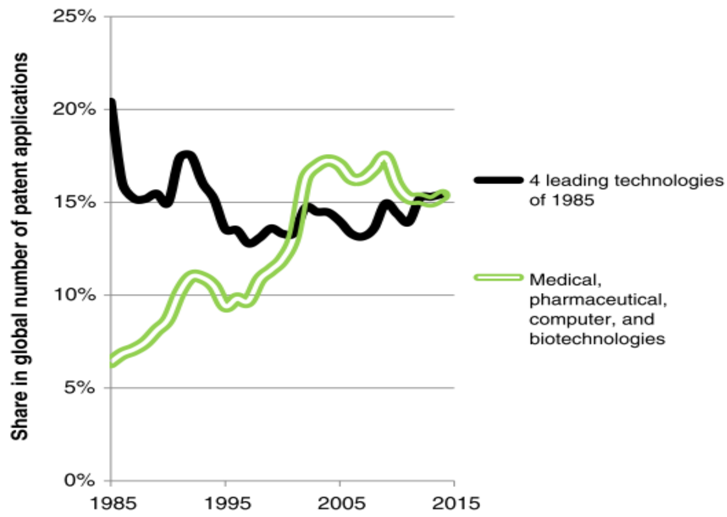


Fig. 3. Dynamics of the global combined share of four technologies with the highest share of patent applications in 1985 (electrical machinery, measurement, machine tools, and other special machines) in comparison with the dynamics of the global combined share of patent applications in four top categories (medical, pharmaceutical, computer, and biotechnologies), 1985–2014

Source: WIPO IP Statistics Data Center 2016.

The important question is in what sphere will the final phase of the Cybernetic Revolution start? First of all, one should remember that the 'breakthrough'

[6] See also Appendix to Ch. 9 in Grinin L. and Grinin A. 2015 at URL: https://www.socionauki.ru/book/files/ot_rubil_do_nano/online_version/9_chapter_appendix/266p.php.

sphere is usually quite narrow as it happened during the Industrial Revolution (when the breakthrough occurred in a specific field – cotton industry). In a similar way, given the general vector of scientific achievements and taking into account that a future breakthrough area should be commercially attractive, we think that the final phase (the one of self-regulating systems) of the Cybernetic revolution will begin in one of the recent branches of medicine. It probably has already formed (such as biomedicine or nanomedicine) or it can form as a result of the uptake of other innovative technologies into medicine.

It is important that in the nearest decades not only the developed but also developing countries will face the problems of population ageing, shortage of labor resources and the necessity to support a growing number of elderly people. The progress in medicine can contribute to the extension of working age (as well as to the general increase of the average life expectancy) of elderly people and to more active involvement of disabled people into labor activities. *Thus, elderly people and people with disabilities could more and more subsist for themselves.*

At present medicine is closely related to biotechnologies (see Fig. 2) through pharmaceuticals, gene technologies, new materials, *etc*. The distinctive feature of modern medical science is its 'bio-related trends' – a wide use of approaches based on the methods of molecular and cell biology.

At present medicine is highly computerized especially in the field of diagnostics, various automatic control systems have been developed; for example, for the control of breathing, nutrient supply to specific organs, blood pressure, control over the functioning of some internal organs, *etc.*

Medicine (supported by both government and private funding) has been a major influence on GDP.

Taking into consideration the anticipated faster growth rates of GDP in the developing countries and a rapid formation of the middle class there, one can suppose that in general, spending on health care will increase significantly. For example, in Germany a number of health care personnel constitute 22 % of the total number of employed people while the share of automobile industry is only 2.3 % (Nefiodow L. and Nefiodow S. 2014).

We have no opportunity to describe the whole range of MANBRIC-technologies with the equal attention. So in this paper we will focus on the most important spheres.

2. FUTURE TECHNOLOGIES

Big History, Technologies, and Rules of Evolution

When considering the issue of future technologies in terms of Big History, one should emphasize that the growing technological complexity is connected with some other aspects of Big History. Elsewhere we formulated a number of evolutionary rules, which can be applied for the analysis of different Big History directions (*e.g.*, Grinin, Markov, and Korotayev 2008, 2009; Grinin 2016; Grinin 2014; Grinin, Markov, and Korotayev 2017).

Among these evolutionary rules we single out three rules which are of particular importance for the development of technologies.

Rule 1. **Evolution occurs only in a small part of a system**
According to data obtained from Planck observatory, the Universe is composed of 5 % of ordinary (baryonic) matter, 24 % of dark matter, and 70 % of dark energy. Thus, the most bulk of our Universe is occupied by dark matter and energy which can hardly evolve. In living organisms, for example, an estimated percentage of the non-coding DNA reaches 98 %. In social evolution, according to some sources, the number of innovators in a society is about 3–5 %. The same is in evolution of technology, for example, only a small number of startup projects appear successful.

Rule 2. **Evolutionary block assemblage**
In evolution, there emerge some basic and more complex components which assemble in various combinations. In this sense, evolution is similar to construction, where ready-made units are used to build new creations. Thus, in cosmic evolution atoms are universal components for the formation of molecules while chemical evolution in space started with the emergence of a sufficient variety of atoms. On Earth the atoms and non-organic molecules launched the geological development, and later – the emergence of organic molecules, and eventually life. In biological evolution block assemblage principle of formation can be observed, for example, at the level of cells, tissues, and organs. Many molecules, for example, of DNA, also consist of peculiar semantic units. Similarly in social evolution religion or legal systems are often borrowed by other countries. All technologies are made according to this principle. For instance, a modern vehicle is a result of numerous technological achievements: from wheel, alloys, mechanical systems, plastic, fabrics to the onboard computer.

Rule 3. **The increasing diversity**
Variety is as a universal evolutionary trend. Thus, within cosmic evolution there appeared a growing number of chemical elements and molecules, as well as stars and planets. In biological evolution the number of species has been continually increasing for a long time. However, the growth cannot be constantly since evolution always balances around the optimum. Thus, it is not surprising that there are periods of reverse development and reduction of the diversity (*e.g.*, during mass extinctions). In social evolution, there is a growing diversity of political forms, cultures and religions.

In the evolution of technologies the growing diversity is very impressive. From 1980 to 2014 the number of patent applications has increased ten folds.[7]

The rules given above are used as examples to demonstrate that technological development we observe nowadays is not unique. Under similar conditions and preconditions, evolution in different systems may follow the similar patterns.

Thus, based on the evolutionary rules and the theory of production principles and other aspects we give some forecasts of upcoming technological revolution.

On Some Future Medical Technologies

Constant health monitoring as a self-regulating supersystem. During the final phase of the Cybernetic Revolution a very important direction of self-regulation can develop as different health monitoring systems for early diagnosis and preventing diseases. The key compounds of such devices are biosensors and similar tiny devices. One can easily imagine that in the future they will be able to become an integral part of a human life, providing a constant scanner of an organism or a certain organ and transmitting the information to a medical center in case of potential or real threats. On the whole, medicine will develop towards increasing individualization and personification through the selection for individual therapy while the use of mass drugs and standard therapeutic technologies will be reduced.

Economy, optimization of resource consumption, and miniaturization. The achievements in medicine will make a significant contribution to *the optimization of resource consumption*, for example, due to the targeted drug delivery and minimization of interference with the organism. Hospital treatment will be less used as the operations will be more targeted, and the rehabilitation period will be minimal. More people will be treated at home since the development of remote treatment is rather probable when doctors control the indices of a patient online and can make the necessary prescriptions remotely. It could sharply decrease a cost of medical treatment which now is exorbitant one for a great number of people. Saving money (as well as resources) is one of the most important directions for the economy.

Medicine develops in the direction of growing **miniaturization.** There is a trend of constantly decreasing size of instruments to micro and nanoscale (Peercy 2000). For example, repairing heart tissue destroyed by a heart attack usually requires invasive open-heart surgery. But now researchers have developed a technique that allows using a small needle to inject a repair patch, without opening up the chest cavity (Montgomery *et al.* 2017).

---

[7] About the dynamics of patent application see Appendix to Grinin L. and Grinin A. 2015, 274ff. URL: https://www.socionauki.ru/book/files/ot_rubil_do_nano/online_version/9_chapter_appendix/274p.php.

The perspective direction in medicine is slowing down the ageing process. It is highly probable that human medicine will significantly increase life expectancy. Already nowadays in some countries the average life expectancy is more than 80 years. We suppose that increase of life expectancy will occur as a result of a breakthrough in medical technologies in the 2030s – 2050s. In the 2050s the average life expectancy will increase by 15 years or even more.

It is quite possible that genetic methods will significantly increase life expectancy. In this respect, the study of telomeres, which were found to play an important role in cell division, seems to be promising (Slagboom, Droog, and Boomsma 1994).[8]

Transplantation. Another important branch of medicine is regeneration and transplantation of organs and tissues of a human body. At present medicine achieved great results in organ transplantation, (*e.g.*, heart, lungs, liver, pancreas, and kidneys). However, human donor organs are scarce, and people who donate donor organs without special agreement are brought to criminal responsibility all over the world.

Medicine and biotechnology will provide an opportunity to design different artificial organs, such as skin, retina, trachea, vessels, heart, ear, eye, limbs, liver, the lungs, pancreas, bladder, ovaries. Many of them are already designed today. Even new organs or combinations are possible. There is already an opportunity of tissue engineering. In laboratories scientists cultivate new cells to replace injured bone or cartilage. For example, recently, the soft artificial heart was created from silicone using a 3D-printing, lost-wax casting technique; it weighs 390 grams and has a volume of 679 cm^3. This artificial heart has a right and a left ventricle, just like a real human heart, though they are not separated by a septum but by an additional chamber. This chamber is in- and deflated by pressurized air and is required to pump fluid from the blood chambers, thus replacing the muscle contraction of the human heart (Cohrs, Petrou, and Loepfe *et al.* 2017).

This technology has the potential to develop cell therapy and methods of tissue regeneration.

One can expect that opportunity to 'deceive' the immune suppression will be one of the main breakthroughs in the field of regeneration and transplantion of organs and tissues.

Changing human reproductive capabilities is an especially important field of medicine. The number of incurable diseases causing infertility decreases. Nevertheless, the only opportunity for some patients is to use *in vitro* ferti-

---

[8] In 2009, Elizabeth H. Blackburn, Carol W. Greider and Jack Szostak were awarded the Nobel Prize for the discovery of how chromosomes are protected by telomeres and the enzyme telomerase from terminal underreplication.

lization. Besides, due to the development of medicine there increases a number of women who want to have children after their reproductive age is over (*e.g.*, at present, it is possible to grow an embryo outside the woman's body).

The Future of Biotechnology

One can suppose that at the very first stage of Cybernetic Revolution biotechnology, as an independent direction, will play a less important role than medicine. It will be rather an important component of medical technologies, providing breakthroughs in treatment of diseases or monitoring of organism functions. Genetic engineering will play important role in different spheres of biotechnologies (see below).

Gene modification. On the basis of the genetic data the most appropriate treatment will be adapted for individual patients, and if it is necessary the defective genes will be corrected. Presumably, first gene therapy will manifest itself in sport medicine as enormous investments are made in it and the best minds are engaged in this field.

When choosing the appearance of a future child (color of eyes, skin, *etc.*) gene therapy can be used. In future it might be possible that babies will be born almost by order, these will be 'the perfect babies' (Fukuyama 2002).[9] In other words, parents will choose desirable features of a child before his/her birth.

Achievements in self-regulation. The level of controllability will increase considerably within a number of important systems connected with biotechnologies. Thus, probably, while transforming an organism, scientists will insert not a separate useful gene (Simon, Priefer, and Pühler 1983), but a whole set of necessary genes which will operate depending on environmental conditions. Such characteristics will be extremely important in the case of climate changes which are quite probable. It will become possible to choose the most optimal varieties of seeds for a unique combination of weather conditions and territory. Consequently, huge databases of such plant varieties and variations will be created. It is quite possible that in the future the whole process of getting a transgenic plant will take place without human participation, thus, it will become self-regulating.

It is possible to assume that by the end of the final phase of the Cybernetic revolution the agricultural biotechnologies will be already developed to a degree that the modified products will be able to response even to the smallest fluctuations of local conditions. In other words, it will be possible for farmers to select individual fodder and drugs by means of programs and to order them

[9] It is difficult to say how 'perfect' they will be and what kind of problems will appear as a result of these technologies. *E.g.*, the possibility to predict the baby's gender resulted in gender imbalance in China. As a result, there are a disproportionate number of boys. Thus, we agree with Francis Fukuyama, who believes that the future achievements of the 'biotechnology revolution' should be accepted with great prudence (Fukuyama 2002).

via the Internet. Even an individual will be able to invent a houseplant hybrid suitable for the interior and to order its production and delivery.

The same refers to domestic animals: it will be possible to breed animals with peculiar characteristics within separate breeds of animals (or even by the individual order). It is probable that the selection of animals on the basis of genetic engineering will also develop in the direction of decreasing human participation.

Creation of new materials. In the 1940–1970s, one of the main directions was the development of industrial production of already known substances (*e.g.*, vitamins) or their analogues; however, during the same period there appeared the elements which did not exist in natural environment (*e.g.*, Humalog, which is a widely applied synthetic analogue of human insulin [Woollett 2012]). This sequence reminds the history of development of chemistry: at first people learned to produce the known substances, and then the artificial materials.

Due to biotechnologies many new materials are produced, for example, bioplastics. The main advantage of this material is that unlike ordinary plastic it is biodegradable. Thus, the main goal of bioplastics production is preserving environment, reducing the production of goods from non-renewable resources and cutting the discharging of carbon dioxide into the atmosphere. This is an important step to the creation of self-cleaning ecological systems in the future and also to the preservation of the environment.

The increase and cheapening of food production is a global challenge for the humankind taking into account that the population number will continue to increase for several more decades (first of all in the poor and poorest countries, especially in Africa), perhaps, reaching nine or more billion people (see UN Population … 2012). Biotechnologies can make a huge contribution to the solution of the problem.

Solution of Urban and Some Environmental Problems

Biotechnologies are successfully used for cleaning up oil spills, in wastewater treatment, *etc.* According to the Organization for Economic Cooperation and Development (OECD), the potential market for bioremediation, that is, the use of living microorganisms to degrade the environmental contaminants (including plants for soil purification), amounts tens of billions of dollars. Thus, important changes will certainly take place in the employment of biotechnologies for the solution of environmental problems. Here it is possible to assume that biotechnologies will be intruded first of all into the urban ecology. It is necessary to consider that in the coming decades the urban population will increase by 40–50 % (see, *e.g.*, NIC 2012). Among the problems which can be potentially solved by means of the development of biotechnologies, there are those related to water cleaning, recycling of waste, liquidation of stray animals (it will be promoted by introducing genes for sterility or something of that nature). Already today the micro-organisms for water cleaning are applied; with their help we also get bio-gas from waste recovery. But in the future these and similar problems will be solved by the development of self-regulating systems that will make it possible to solve a number of technical and scientific problems.

But the problem of ecological self-regulating systems, naturally, is not limited by the cities; it has to be extended to the cleaning of reservoirs and other ecosystems. The creation of ecological self-regulating systems will considerably reduce expenses and free huge territories occupied by waste deposits, as well as allow breeding fish in self-cleaning reservoirs.

One can assume that an important direction will be the creation of self-regulating environmental systems in resort and recreational territories which will provide the best conditions for rest and business.

Breakthrough in the Sphere of Resource Saving

Resource and energy saving is one of the main tasks and outcomes of introduction of biotechnology. The basic opportunities with respect to resources saving are connected with an opportunity to influence the genetic organization of living beings which at present serves the basis for the agricultural ('green') biotechnology which has already become a part of the initial phase of the Cybernetic Revolution. The breakthrough in this area is connected with *totipotency*, that is an ability of plants to form a full-fledged organism from a single cell. With the necessary gene transfer, one can make, for example, a variety of potato resistant to the Colorado beetle, or reduce the susceptibility to drought, cold and other stresses (Grinin *et al.* 2010). New agricultural technologies are of great importance for the developing countries. For example, genetically modified and pest resistant varieties of cotton plant and corn demand much smaller usage of insecticides which is more cost-effective and eco-friendly. The individualization is also noticeable in the animal genetic engineering which develops more slowly, but even now and in prospect it has an enormous value for agriculture and medicine (by means of genetic engineering it is possible to increase milk production, to improve quality of wool, *etc.*).

Biotechnology can help to solve many global issues, for example, to cheapen the production of medicines and foodstuffs including producing and making them in ecologically sound ways that can also keep or make the environment pristine, thereby considerably expanding their production. The solution to the food problem will proceed in different ways, in particular due to the mass production of food protein whose shortage is sharply perceived in many societies (at present the feed protein for animals is generally produced in this way). Even now there are results based on the production of food proteins or, for example, imitation meat. But so far such a production is too expensive. A gram of laboratorial meat costs US$ 1,000 dollars (Zagorski 2012) but this is part of the usual process from the laboratory to mass cheap production.

Artificial Intelligence, Robots, Nanotechnologies, Additive and Cognitive Technologies

Self-regulating systems and Artificial Intelligence. Within Big History the period after a new threshold is sometimes anticipated as a period of rapid development and even predominance of artificial intelligence. We agree that the future investigations in the fields of Big History and evolution are closely associated with the development of artificial intelligence. Thus, it turns to be important to define similarities and differences between self-regulating systems and artificial intelligence.

On the one hand, the notion of 'self-regulating systems' correlates rather closely with artificial intelligence which has become a subject of intensive research in the recent decades (see*, e.g.*, Poole, Mackworth, and Goebel 1998; Russell *et al.* 2003; Hutter 2005; Luger 2005; Neapolitan and Jiang 2012; Keller and Heiko 2014; Hengstler, Enkel, and Duelli 2016). 'Intelligent' machine is often defined as the one that takes actions that maximize its chance of success at some goal (*e.g.*, Russell *et al.* 2003). Of course, such a machine can be also considered as a self-regulating system. The notion of artificial intelligence is usually connected with machines, IT-technologies, robots, and sometimes equated with technical intelligence (Zhang *et al.* 2016).

On the other hand, *the notion of 'self-regulating systems' is wider than the notion 'artificial intelligence'* (AI) since the former includes various self-regulating systems that can function independently, but can hardly be regarded as Artificial Intelligence. For example, biotechnological systems designed to neutralize pollution, or the ones connected with human physiology (*e.g.*, artificial immunity on the basis of artificial antibodies, or systems based on the use of various other proteins or viruses, or genetic engineering technologies that are able to control certain physiological processes and so on). In addition we expect the emergence of self-regulating systems of mixed nature – for example, biochemiotechnical. One should also note that they can function within more complex systems, like a human organism. As examples of such self-regulating systems one may mention artificial organs grown in laboratories and incorporating a number of biosensors and other technical elements. Thus, any AI can be regarded as a self-regulating system, but not all self-regulating systems can be associated with AI.

Robots and self-driving cars in the future. The opportunities of using robots are undoubtedly vast. In particular, only these devices can help to solve the problem of care of growing numbers of elderly people and to some extent the associated problem of labor shortage. In general, there is no doubt that robots will play a significant role in the transition to self-regulating systems. We assume that in the 2020s certain although not revolutionary achievements in this area will occur, in the 2030–2040s we will witness a much more significant rise in robotics, but an explosive development of robots will happen a bit later in the 2050–2060s. By this time it is also possible to expect the creation of really 'smart' robots. We believe that in the next two decades robotics will develop rapidly in the service sector. At present there are many publications on how robots may replace humans in many fields. We agree that the changes in this sphere will be enormous yet; they will take several decades to occur.

In future robotic servants may replace household chores as well as perform some complicated tasks, for example, they will be more and more involved into investigation of space bodies and other tasks that can be dangerous for humans (military, rescue, space activities, *etc.*). Hardly all of them will be anthropomorphous; their design will be most likely defined by functions. However, universal robots are also likely to emerge.

Robots will play a very important role in medicine, for example in surgery and in the sphere of social nursing care. The number and variety of surgical robots grow every year. According to some forecasts, surgical robotics market will grow up to US$ 28.8 billion by 2020.[10] Surgical robotic systems are a combination of equipment, accessories, software, and services, which help doctors to perform minimally-invasive surgeries including gynecological, cardiac, neurological, and orthopedic. Robotic systems allow surgeons to automate the surgical procedure, improve efficacy and precision during the procedure, and minimizing post-surgical complications.

Robotic systems will continue to be used extensively in transportation, in particular they will also be used in the development of self-driving vehicles. The latter might be especially important. Taking into account the above-described 'meaning' of the Cybernetic Revolution (as a revolution of self-regulating systems), the breakthrough will most probably occur in the direction of autonomous transport. Vehicles and other transport systems will become self-driving and will use the electric vehicle technologies. Even today, there are attempts of realizing this opportunity. A vivid example here is Tesla's self-driving cars. But other groups of companies also announced their self-driving cars. For example, 'Mercedes-Benz' has presented the concept of driverless car (della Cava 2015). *Google* works to create such a car by 2020 (see Muoio 2015), but it already tests the Toyota self-driving car in California (and arranges joint projects with *Ford*). Just as in 1997 the computer defeated the world chess champion, recently self-driving car has beaten the racing driver at speeds over 200 kilometers per hour. Some researchers even work to understand how to make self-driving cars become capable of making moral and ethical decisions just like humans do. Any decision that involves risk of harm to a human or even an animal is considered to be an ethical decision. It also includes quite rare situations when a collision is unavoidable, but a decision can be made as to

[10] URL: https://www.alliedmarketresearch.com/surgical-robotics-market?surgical-robotics-market.

which obstacle to collide with. Researchers believe that by algorithms it is possible to make self-driving car decide whether to use a sophisticated algorithm or a simple rule such as 'always stay in the lane' (Sütfeld, Gast *et al.* 2017).

However, the development of such systems as self-regulating systems is an important forerunner of the forthcoming start of the final phase of the Cybernetic Revolution (in the 2030s). The self-driving electric vehicles with a new accumulator together with roads allowing free recharge can become a powerful source of technological development during the final phase of the Cybernetic Revolution.

The Future of Nanotechnology

Nanotechnologies have become one of the most popular technologies of modern times and they have many prospects.[11] First of all, different nanocoatings will rapidly develop. Nanocoatings are used in different fields: industry, aircraft building, and electronics. The components of nanoelectronics, photonics, neuroelectronic interfaces and nanoelectromechanical systems are also promising. They will allow further micronization of devices. We believe that self-regulating technologies (*e.g.*, self-cleaning coatings which regulate the temperature) will gain a widespread use. Similar technologies have already been created.

For example, recently in the University of Central Florida a flexible anti-reflection film on smartphones and tablets was made. It makes the screen bright and sharp as well as scratch resistant and self-cleaning. The film contains tiny uniform dimples, each about 100 nanometers in diameter (about one one-thousandth of the width of a human hair) (Guanjun Tan *et al.* 2017).

In future, the nanotechnologies will provide excellent opportunities for the self-assemblage of nanoelements and nanodevices. It will become possible to make a transition to controlled self-assemblage of nanosystems, creation of three-dimensional networks, nanorobots, *etc*. One may also speak about the use of molecular devices, atomic design. There are rather attractive prospects in the development of nanomechanics, nanomachinery, and nanorobotics. Long ago there started to develop the idea of creation of computers that process and store information through not special condition of environment (magnetic, electric, and optical) but through nanotechnologies, for example, via silicone (the main material in the production of semi-conductor devices) chips replaced by carbon nanotubes. In this case a bit of information can be written in the form of a cluster, for example, of 100 atoms. This would reduce their size several-fold and at the same time increase quick response. Quantum technologies (*e.g.*, the creation of quantum computer) will be one of the most important technological breakthroughs in this context. It is very difficult to predict which path the de-

---

[11] See Appendix to Grinin L. and Grinin A. 2015, p. 288ff. URL: https://www.socionauki.ru/book/files/ot_rubil_do_nano/online_version/10_chapter_appendix/288p.php.

velopment of information technologies will follow, but one can assume that as a result of the completion of the Cybernetic revolution, the information capacity will increase by an order of magnitude.

The Future of 3D-printers

The opportunities provided by 3D-printers are great: from building to cooking, from a house workshop to museums, from medicine to children toys, from training models to design. At present, 3D-printing is used in aircraft construction and rocket engineering to produce individual elements, for example, support stand for an aircraft engine (see, *e.g.*, Turichin 2015). And just because they are used in such high-technology spheres their development needs considerable investments.

In fact, 3D-printers constitute a universal house workshop or a universal production or, construction factory. And in the future they will acquire new functions and incorporate new subsystems.

Additive 3D-printing (*i.e.*, merging (fusing) of materials and creation of certain objects) is a very promising direction. Thus, in future 3D-printers will help to produce any material needed, even the biological one. Great opportunities are especially associated with the opportunities to grow human organs and tissues, including through the usage of patient's own tissues. Soon it will suffice to have a sketch and to make (to 'print', 'fuse') any detail at home or in a 3D-printing center. It will also be possible to organize a small single-piece production. Engineers could also develop simple 3D-food printers which can print, for example, candies or pizza.

Undoubtedly, the development of additive technologies will be connected with other directions of MANBRIC-complex, for example, with robotics (additive technologies will be used to create robots, and at the same time the robots themselves will use additive technologies in their activities).[12]

Cognitive technologies. Neural interfaces or brain-computer interfaces (BCI). A brain-computer interface (BCI) is a direct communication pathway between brain and an external device. This technology implements the interaction between brain and computer systems that can be realized via electrode contact with head skin or via electrodes implanted into brain. The implementation of neural interfaces is already wide-spread, for example, in artificial visual systems or bionics. The most notable device is the cochlear implant, which has been implanted in more than 220,000 people worldwide.

[12] *E.g.*, the Stormram 4, as the robot is named, is made from 3D-printed plastic and is driven by air pressure. This robot can be used in an MRI scanner. Carrying out a biopsy (removing a piece of tissue) during a breast cancer scan in an MRI significantly increases accuracy. The Stormram 4 is a stimulus for the entire diagnostic phase of breast cancer; the accurate needle control, effectively real-time MRI scanning and a single, thin-needle biopsy enable quicker and more accurate diagnoses to be made (University of Twente 2017).

In three or four decades, disabled people will get another chance in life. BCIs may improve rehabilitation for people with strokes, head trauma, and other disorders. At present there already exist devices which allow paralyzed people to speak, write and even work at the computer as, for example, in the case of the famous astrophysicist, Stephen Hawking.

Those who can afford it and want to increase their abilities will be able to replace their body parts by bionic ones.

Also in three or four decades, small scalp electrodes will make remote brain control possible. So people will be able to turn TV on only by thinking about it. In the future neural interfaces can be applied not only in medicine, but also in daily pursuits, for example to control condition of a driver's or an operator's brain and in case of falling asleep to awake him automatically. In general the achievements in cognitive science are already in use and their application will increase even more in the areas which move towards self-regulating systems – from medicine to robotics, from cybernetics to problems of artificial intelligence, and, of course, for the military purposes. However, serious technical and social difficulties can hamper the development of this direction (see below). After surpassing these constraints, the development of neural interfaces will promptly reach a new level.

Some ideas on other future technologies. Smart devices. Everyday technologies become more self-regulating, complicated, and more intelligent. Their names speak for themselves. The word 'smart' is used as a prefix for many devices. Today smartphones have become ubiquitous, while smartwatches are becoming popular, people watch smart TV, and in schools they use smart boards. Here are just a few examples of smart things: smart kettle, smart swimsuit, smart stroller, smart cup, smart rope, smart T-shirt which tracks your posture, smart cane with GPS for elderly people, smart bottle which automatically tracks hydration and temperature, smart highway with nanocoating which changes its color according to the weather and warns drivers of potential risk. There is also developed a concept of a smart city with smart traffic signs and traffic light signals, as well as smart cars. This will also allow time and resource saving. Exoskeleton will allow people to perform hard work with fewer efforts. We assume that this trend will continue and thus, in three or four decades the majority of everyday devices will be smart. An absolute majority of them will be connected to the smartphone and Internet. One can predict that we will live in smart homes with smart kitchens, while a smart climate control system will maintain the required temperature 24 hours a day.

Mobile phone as an integrating device. The key feature of the future technologies is that most of them will be integrated with mobile phone or similar devices. A mobile phone will be a universal control panel and analytical center. It will collect all data from smart technological devices, for example: how many meters one walks, how many calories one consumes, how many hours one sleeps, how much money one spends, how many hours one plays basketball and how many points one scores. The mobile phone will become a powerful means of control not only over an individual, but over pets and children. For example, a smart bracelet that monitors a child's clean hands and signals if the child takes, for example, unwashed fruit. At any time parents can check the purity of their child's hands using a special application in smartphone.

Thus, the important future technological trend is the development of virtual reality through different devices especially mobile phone or in another form of such integrative device (as it is known there are different ideas on this future forms, *e.g.*, glasses). It is quite probable that such devices will be able to adopt the functions and become a new type of sensory organ and source of information for people. Thus, special glasses will allow connecting vision and hearing with the high resolution virtual reality devices. In the future, virtual reality may be not only seen but also felt. A small band with a device on the arm is already designed which will enable users touch the object in virtual reality.

Conclusion. Will the Development of the Cybernetic Revolution proceed in the Direction of Cyborgization?

There is no doubt that future development within Big History and evolutionary paradigms is connected with the development of intelligence and transformation of intellectual creatures. As to the direction and speed of this transformation there are many points of view, including those (which we do not share) that AI will be able to unite billions of people's minds into a new system (Kurzweil 2000) or that humans will soon become immortal (see below). On the other hand, the development of medicine and self-regulating systems, which will constitute the nucleus of changes in the final phase of the Cybernetic revolution, will undoubtedly lead to the increasing interference in human body. In this context, we would like to conclude the paper by the reflections about the ways this interference in human body can change the human biological nature and transform a human into a cyborg. A very popular word 'cyborg' (short for 'cybernetic organism') derives from the word 'cybernetic'. Cyborg is defined as a theoretical or fictional being with both organic and biomechatronic parts.[13]

The term 'cyborg' often applied to an organism that has restored function or enhanced abilities due to the integration of some artificial component or technology that relies on some sort of feedback. It is obviously that many achievements in medicine will impel our civilization to the state in which more and more humans can become partial cyborgs. Thus, we are following the path of development of self-regulating systems of a new type which will be constituted by the elements of different origin: biological and artificial. All that we

[13] The term was coined in 1960 by Manfred Clynes and Nathan S. Kline (Halacy 1965).

have written about artificial organs and tissues will contribute to the breakthrough in the field of both production of absolutely new materials which will expand the implementation of non-biological elements in the human body. Thus, the Cybernetic Revolution is closely connected with the process that can be designated as cyborgization. We should be aware of the fact that this actually means not only the formation of a new direction in medicine, but also the moving towards the *cyborgization* of a human being. Of course, this can cause a certain and quite reasonable anxiety. On the other hand, expanding the opportunities for not just a long but also an active life is hardly possible without significant support for the sensory organs and other parts of the body which weaken as a result of ageing and other reasons. Finally, glasses or contact lenses, artificial teeth, tooth fillings, bones, aerophones, artificial blood vessels, mitral valves, *etc.* allow hundreds of millions of people to live and work and these people still remain humans. The same is true with respect to more complex systems and functions. Thus, people with disabilities can benefit from the development of medicine and cyborgization as they will be able to significantly compensate their drawbacks. However, we suppose that the idea that some day the human body will be fully replaced by non-biological material and only the brain or the organs which support the senses will remain are just fantasy. This will never come true (the well-known ideas about such future for humankind are presented by Kurzweil [1999]). People who propose such solutions, for example, to replace supposedly less lasting and comfortable biological material by the technological inventions (such as replacement of haematocytes by billions of nanorobots, *etc.*) in their forecasts try to use the outdated logic that was widespread several decades ago in science fiction or scary stories: the replacement of biological organisms with technical ones. The modern logic of scientific and technological progress including the latest achievements in bioengineering shows the shift towards the synthesis of biological forms and technical solutions into a unified system. Still there are numerous obstacles here. Let us take, for example, the above-mentioned possible opportunities for brain control which may be hampered by the immune rejection in the first turn. Second, many nanostructures, for example, nanopipes, which had been predicted a bright future appeared very toxic for human body (Kotov *et al.* 2009). Third, the implantation of external devices leads to traumatization of the whole organism despite all serious attempts to reduce this impact. Another problem is the different electric conductance of biological material and of a technical device, though there is certain progress in the solution of this problem (Abidian and Martin 2009). But even if we solve these problems we will still need some powerful software capable to handle brain signals.

Technical achievements can hardly replace the biological mechanisms which have been selected for many millions of years. On the contrary, we should follow the path of 'repair', improvement, the development of self-regulation and support of biological mechanisms via some technical solutions. The human brain is very tightly connected with the body and sensory organs, most of its functions are based on the control of the body that does not imply its full-fledged work outside its biological foundation. The opportunities of science and medicine to replace worn organs will increase but the biological foundations of a human will always exist and must prevail. If one can help the human body by different means including methods of activization of immune system, opportunities of genetics, the methods of blocking or decelerating the process of ageing, *etc.*, it is much more reasonable to preserve the human biological foundation. In any case, in the nearest decades in the process of cyborgization quite radical breakthroughs are possible, but nevertheless the process of cyborgization will not go too far. Thus, we believe that in the next hundred years the human lifestyle and biological nature will experience crucial changes which can become a turning point in the transition to the post-human society. However, these changes, no matter how profound they are, will be very far from the images drawn by modern wishful-thinking scientists.

References

Abidian M. R., and Martin D. C. 2009. Multifunctional Nanobiomaterials for Neural Interfaces. *Advanced Functional Materials* 19(4): 573–585. DOI: 10.1002/adfm.200801473.

Anuchin D. N. 1923. *The Discovery of Fire and Methods to Get It.* Moscow – Petrograd: Gosudarstvennoye izdatel'stvo. *In Russian* (Анучин Д. Н. *Открытие огня и способы его добывания.* М. – Петроград: Государственное издательство).

Beer S. 1959. *Cybernetics and Management*. London: English Universities Press. DOI: 10.2307/3007308.

Belkind L. D., Konfederatov I. Ya., and Shneiberg Ya. A. 1956. *The History of Technology*. A textbook for energetic and electro-technical higher education establishments. Moscow – Leningrad: Gosenergoizdat. *In Russian* (Белькинд Л. Д., Конфедератов И. Я., Шнейберг Я. А. *История техники*: учеб. для энергетических и электротехнических высших учебных заведений и факультетов. М. – Л.: Госэнергоиздат).

Boas F. 1911. *The Mind of Primitive Man.* A course of lectures delivered before the Lowell Institute, Boston, Mass., and the National University of Mexico, 1910–1911. New York: The Macmillan Company.

Braudel F. 1985. *La dynamique du capitalisme.* Paris: Les Editions Arthaud.

della Cava M. 2015. *Captain Kirk, Your Mercedes F 015 Awaits*. URL: http://www.usatoday.com/story/tech/2015/01/06/mercedes-f-015-luxury-in-motion-debuts-at-consumer-electronics-show/21359239.

von Foerster H., and Zopf G. W. 1962. Principles of Self-Organization. *Self-Organization* / Ed. by M. C. Yovits, and S. Cameron, pp. 31–50. New York: Pergamon Press.

Chernousov P. I., Mapelman V. M., and Golubev O. V. 2005. *The Iron Metallurgy in the History of Civilizations*. Moscow: MISiS. *In Russian* (Черноусов П. И., Мапельман В. М., Голубев О. В. *Металлургия железа в истории цивилизаций*. М.: МИСиС).

Cohrs N. H., Petrou A., Loepfe M., *et al.* 2017. A Soft Total Artificial Heart-First Concept Evaluation on a Hybrid Mock Circulation. *Artificial Organs*. DOI: 10.1111/aor.12956

Fukuyama F. 2002. *Our Post-human Future: Consequences of the Biotechnology Revolution*. New York: Farrar, Straus, and Giroux.

Grinin A. L. 2016. Self-Regulation as a Global Evolutionary Mega-Trend. *Evolution: Evolution and Big History: Dimensions, Trends, and Forecasts* / Ed. by L. E. Grinin, and A. V. Korotayev, pp. 139–164. Volgograd: Uchitel.

Grinin A. L., and Grinin L. E. 2015. The Cybernetic Revolution and Historical Process. *Social Evolution & History* 1: 125–184.

Grinin A. L., Kholodova V. P., and Kuznetsov V. V. 2010. A Comparative Analysis of Physiological Mechanisms of Salt-Endurance of Different Varieties of Mustard. *Vestnik Rossiyskogo Universiteta Druzhby Narodov* 1: 27–38. *In Russian* (Гринин А. Л., Холодова В. П., Кузнецов В. В. Сравнительный анализ физиологических механизмов солеустойчивсти различных сортов горчицы. *Вестник Российского университета дружбы народов* 1: 27–38).

Grinin L. E. 2006a. Periodization of History: A Theoretic-Mathematical Analysis. *History and Mathematics: Analyzing and Modeling Global Development* / Ed. by L. E. Grinin, V. de Munck, and A. Korotayev, pp. 10–38. Moscow: KomKniga.

Grinin L. E. 2006b. *The Productive Forces and Historical Process*. 3rd ed. Moscow: KomKniga. *In Russian* (Гринин Л. Е. *Производительные силы и исторический процесс*. 3-е изд. М.: КомКнига/URSS).

Grinin L. E. 2007a. Production Revolutions and the Periodization of History. *Herald of the Russian Academy of Sciences* 77(2): 150–156.

Grinin L. E. 2007b. Production Revolutions and Periodization of History: A Comparative and Theoretical-Mathematical Approach. *Social Evolution & History* 6(2): 75–120.

Grinin L. E. 2012. Kondratieff Waves, Technological Modes and Theory of Production Revolutions. *Kondratieff Waves: Aspects and Prospects: Yearbook* / Ed. by A. A. Akaev, R. S. Greenberg, L. E. Grinin, A. V. Korotayev, and S. Yu. Malkov, pp. 222–262. Volgograd: Uchitel. *In Russian* (Гринин Л. Е. Кондратьевские волны, технологические уклады и теория производственных революций. *Кондратьевские волны: аспекты и перспективы: ежегодник* / Отв. ред. А. А. Акаев, Р. С. Гринберг, Л. Е. Гринин, А. В. Коротаев, С. Ю. Малков, с. 222–262. Волгоград: Учитель).

Grinin L. E. 2014. The Star-Galaxy Era of Big History in the Light of Universal Evolutionary Principles. *Teaching and Researching Big History: Exploring a New Scholarly Field* / Ed. by L. E. Grinin, D. Baker, E. Quaedackers, and A. V. Korotayev, pp. 163–187. Volgograd: 'Uchitel' Publishing House.

Grinin L. E., and Grinin A. L. 2013. Macroevolution of Technology. *Evolution: Development within Big History, Evolutionary and World-System Paradigms* / Ed. by L. E. Grinin, and A. V. Korotayev, pp. 143–178. Volgograd: 'Uchitel' Publishing House.

Grinin L., and Grinin A. 2016. *The Cybernetic Revolution and the Forthcoming Epoch of Self-Regulating Systems.* Moscow: Moscow branch of 'Uchitel' Publishing House.

Grinin L. E., Grinin A. L., and Korotayev A. V. 2016. Forthcoming Kondratieff Wave, Cybernetic Revolution, and Global Ageing. *Technological Forecasting and Social Change.* DOI: 10.1016/j.techfore.2016.09.017.

Grinin L. E., and Korotayev A. V. 2015. *Great Divergence and Great Convergence: A Global Perspective*. Springer.

Grinin L. E., Korotayev A. V., and Markov A. V. 2017. Biological and Social Phases of Big History: Evolutionary Principles and Mechanisms. *From Big Bang to Galactic Civilizations. A Big History Anthology.* Volume III. *The Ways that Big History Works: Cosmos, Life, Society and Our Future* / Ed. by B. Rodrigue, L. Grinin, and A. Korotayev, pp. 158–199. Primus Books.

Grinin L. E., Markov A. V., and Korotayev A. V. 2008. Macroevolution in Animated Nature. Moscow: LKI/URSS. *In Russian* (Гринин Л. Е., Марков А. В., Коротаев А. В. *Макроэволюция в живой природе и обществе.* М.: ЛКИ/URSS).

Grinin L. E., Markov A. V., and Korotayev A. V. 2009. Aromorphoses in Biological and Social Evolution: Some General Rules for Biological and Social Forms of Macroevolution. *Social Evolution & History* 8(2): 6–50.

Gupta A. K., and Gupta M. 2005. Synthesis and Surface Engineering of Iron Oxide Nanoparticles for Biomedical Applications. *Biomaterials* 26(18): 3995–4021. DOI: https://doi.org/10.1016/j.biomaterials.2004.10.012.

Halacy D. S. Jr. 1965. *Cyborg: Evolution of the Superman.* Harper.

Hengstler M., Enkel E., and Duelli S. 2016. Applied Artificial Intelligence and Trust – The Case of Autonomous Vehicles and Medical Assistance Devices. *Technological Forecasting and Social Change.* Elsevier 105: 105–120.

Hutter M. 2005. *Universal Artificial Intelligence, Machine Learning.* DOI: 10.1145/1358628.1358961.

Keller J., and Heiko A. 2014. The Influence of Information and Communication Technology (ICT) on Future Foresight Processes – Results from a Delphi Survey. *Technological Forecasting and Social Change. Elsevier* 85: 81–92.

Kosven M. O. 1953. *Essays on the History of Primitive Culture*. Moscow. *In Russian* (Косвен М. О. *Очерки истории первобытной культуры*. М.).

Kotov N. A., Winter J. O., Clements I. P. *et al.* 2009. Nanomaterials for Neural Interfaces. *Advanced Materials* 21(40): 3970–4004. URL: http://dx.doi.org/10.1002/adma.200801984.

Kremkova V. M. 1936. The 10th – 12th Centuries. From the History of Agriculture in England and France. *History of Agriculture* / Ed. by N. Y. Vavilov, and V. P. Williams, pp. 125–163. Moscow: Academy of Sciences Press. *In Russian* (Кремкова В. М. Века X–XII. Из истории сельскохозяйственной техники Англии и Франции. *История агрикультуры* / Ред. Н. И. Вавилов, В. Р. Вильямс и др., с. 125–162. М.: Изд-во Академии наук).

Kurzweil R. 1999. *Age of Spiritual Machines*. New York, NY: Penguin Books.

Kurzweil R. 2000. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence.* N.Y: Penguin Groups.

Luger G. F. 2005. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving.* Pearson education.

Lurie I., Lyapunova K., Mathieu M., Piotrovsky B., and Flitner M. 1939. *Essays on History of Technology in the Ancient East.* Moscow; Leningrad: USSR Academy of Sciences Publishing. *In Russian* (Лурье И., Ляпунова К., Матье М., Пиотровский Б., Флитнер М. *Очерки по истории техники древнего Востока.* М. – Л.: Изд-во АН СССР).

Montgomery M., Ahadian S., Locke Huyer D. *et al.* 2017. Flexible Shape-memory Scaffold for Minimally Invasive Delivery of Functional Tissues. *Nature Materials.* DOI: 10.1038/nmat4956.

Muoio D. 2015. 10 Companies Making a Bold Bet that They'll have Self-driving Cars on the Road by 2020. *Business Insider*, October 12. URL: http://www.businessinsider.com/google-apple-tesla-race-to-develop-self-driving-cars-by-2020-2015-10.

Neapolitan R. E., and Jiang X. 2012. Contemporary Artificial Intelligence**.** *Chapman & Hall/CRC*.

Nefiodow L., and Nefiodow S. 2014. *The Sixth Kondratieff. The New Long Wave of the World Economy*. Rhein-Sieg-Verlag: Sankt Augustin.

NIC – National Intelligence Council. 2012. *Global Trends 2030: Alternative Worlds*. URL: www.dni.gov/nic/globaltrends.

Osipov G. V. 1959. *Technology and Social Progress*. Moscow: AN USSR Press. *In Russian* (Осипов Г. В. *Техника и общественный прогресс*. М.: Издательство АН СССР).

Pankhurst Q. A. *et al.* 2003. Applications of Magnetic Nanoparticles in Biomedicine. *Journal of Physics D: Applied Physics* 36(13): R167. URL: http://stacks.iop.org/0022-3727/36/i=13/a=201.

Peercy P. S. 2000. The Drive to Miniaturization. *Nature* 406(6799): 1023–1026.

Poole D. L., Mackworth A., and Goebel R. G. 1998. Computational Intelligence and Knowledge. *Computational Intelligence: A Logical Approach* (Ci): 1–22. URL: https://www.cs.ubc.ca/~poole/ci.html.

Russell S. J. *et al.* 2003. *Artificial Intelligence: A Modern Approach.* Pearson Education Limited.

Semyonov S. A. 1968. *The Development of Technology during the Stone Age*. Leningrad: Nauka. *In Russian* (Семенов С. А. *Развитие техники в каменном веке*. Л.: Наука).

Shea J. J. 2006. *The Human Revolution Rethought, Evolutionary Anthropology: Issues, News, and Reviews.* Edinburgh: Edinburgh University Press. DOI: 10.1002/evan.20085.

Sheypak A. A. 2009. *History of Science and Technology. Materials and Technologies*. Part 1. Moscow: MGIU. *In Russian* (Шейпак А. А. *История науки и техники. Материалы и технологии*. Ч. 1. М.: МГИУ).

Simon R., Priefer U., and Pühler A. 1983. A Broad Host Range Mobilization System for in Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Nature Biotechnology* 1(9): 784–791. DOI: 10.1038/nbt1183-784.

Slagboom P. E., Droog S., and Boomsma D. I. 1994. Genetic Determination of Telomere Size in Humans: A Twin Study of Three Age Groups. *American Journal of Human Genetics* 55: 876–882.

Guanjun Tan, Jiun-Haw Lee, Yi-Hsin Lan, **Mao-Kuo Wei, Lung-Han Peng, I-Chun Cheng, Shin-Tson Wu. 2017.** Broadband Antireflection Film with Moth-eye-like Structure for Flexible Display Applications. *Optica* 4(7): 678. DOI: 10.1364/ OPTICA.4.000678.

Sütfeld L. R., Gast R., König P., and Pipa G. 2017. Using Virtual Reality to Assess Ethical Decisions in Road Traffic Scenarios: Applicability of Value-of-Life-Based Models and Influences of Time Pressure. *Frontiers in Behavioral Neuroscience* 11. DOI: 10.3389/fnbeh.2017.00122.

Turichin G. A. 2015. *Additive Technologies in Modern Production.* Report at the 2nd International Seminar 'Basic Technologies of the First half of the 20th Century (Structural and Cyclical Analysis)'. Saint-Petersburg, 1–2 October. *In Russian* (Туричин Г. А. *Аддитивные технологии в современном производстве.* Доклад на втором международном семинаре «Базисные технологии первой половины XIX в. (структурно-циклический анализ)». Санкт-Петербург, 1–2 октября).

Umpleby S. A., and Dent E. B. 1999. The Origins and Purposes of Several Traditions in Systems Theory and Cybernetics. *Cybernetics and Systems* 30(2): 79–103. DOI: 10.1080/019697299125299.

Virginsky V. S. and Khoteenkov V. F. 1993. *The Essays on the History of Science and Technology from Ancient Times to the Mid-fifteenth Centuries*. Moscow: Prosveschenie. *In Russian* (Виргинский В. С., Хотеенков В. Ф. *Очерки истории науки и техники с древнейших времен до середины XV века*. М.: Просвещение).

UN Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat**. 2012.** *World Population Prospects: The 2010 Revision.* URL: http://esa.un.org/unpd/wpp/index.htm.

University of Twente. 2017. 3D-Printed Robot Aims to Fight Cancer. *ScienceDaily*. URL: www.sciencedaily.com/releases/2017/07/170703121134.htm. Accessed February 5, 2018.

Wiener N. 1948. *Cybernetics, or Control and Communication in the Animal and the Machine*. Scientific American. *JSTOR* 179(5): 14–19.

WIPO = World Intellectual Property Organization. **2016.** *World Intellectual Property Organization database*. URL: http://www.wipo.int/portal/en/index.html.

Woollett R. 2012. Innovation in Biotechnology: Current and Future States. *Clinical Pharmacology and Therapeutics* 91(1): 17–20. DOI: 10.1038/clpt.2011.219.

Zagorski I. 2012. Not by Meat Alone: Leather Jackets will be Grown in a Lab. *Vesti.ru*. September 20. URL: http://www.vesti.ru/doc.html?id=9120 84&cid=2161. *In Russian* (Загорский И. 2012. Не мясом единым: кожаные куртки будут выращивать в лаборатории. *Вести.ру*, 20 сентября).

Zhang Y. *et al.* 2016. Technology Roadmapping for Competitive Technical Intelligence. *Technological Forecasting and Social Change* 110: 175–186. DOI: 10.1016/j.techfore.2015.11.029.

12
Global Society as Singularity and Point of Transition to the New Phase of Social Evolution

Sergey V. Dobrolyubov

Abstract

In this article social evolution is considered as a process consisting of three phases: Adaptive, Structural and Cognitive, which are separated by two phase transitions or by two singularities – the Neolithic and the Global. The mechanism of social evolution at these phases is different and is based on different institutional means of cognition and competition. At the current Structural Phase, competition of individuals leads to inequality, and competition of societies leads to extension of societies. Social inequality and exploitation of the periphery become institutional tools for the development. The expansion of societies and evolutionary limitations of its growth lead to life cycles of societies. The maximum size of society increases in the process of evolution and tends to cover all humankind. The Global Society is a final point of structural evolution, and transition to it is singularity. It will be a metamorphosis of the society's nature. The mechanism of further social evolution at the Cognitive Phase will rely directly on individual's need for cognition and self-realization, and not on the special social institutions. Mathematical model of the primary transformations dynamics at structural phase is described by the equation $T(n) = -11214 + 1893\,n$, where $T(n)$ – is the moment of evolutionary transformation, and n – is the ordinal number of transformation. Global singularity is predicted by this model in AD 3930.

Keywords: *social evolution, phase of evolution, singularity, global society, sociogenesis.*

The notion of 'singularity' has different meanings. One is purely *mathematical*, in this case we denote discontinuity in a function where its value rushes to infinity and becomes indefinite. On the other hand, the notion of 'singularity' is used as a *metaphor* for the initial, final or trigger state of a process when its properties are also indefinite, *e.g.*, 'Big Bang', 'black hole', *etc*. In this case, we are more interested in the nature of event, and not at what moment it happened. I apply the notion of singularity to social evolution in this metaphoric sense, although the article proposes a mathematical model of social evolution and explains how this process will come to a singularity not in a mathematical limit's sense.

Scientific trend in social studies leads to a wide use of the notion of singularity in the math sense. Indeed, modern society faces acceleration of changes. The extrapolation of some current trends indicates that the rate of changes has mathematical singularity. There are a number of its interpretations: technological singularity (Kurzweil 2005), demographic singularity (Foerster *et al*. 1960; Kapitza 2006), evolutional singularity – a topic frequently touched on by Russian historians and evolutionists (Snooks 1996; Diakonoff 1999; Panov 2005; Nazaretyan 2016). Different estimates of singularity's moment give a similar result – the first half or middle of the 21st century (Nazaretyan 2016). For illustration, let us consider the interpretation of 'planetary evolution' by hyperbolic curve (see Fig. 1) known in Russian scientific community as 'Snooks-Panov vertical' (Panov 2005; Nazaretyan 2016). This is a chain of geological, biological, social and technological transformations presented as sequence of phase transitions (revolutions), which is described by equation

$$t_n = t^* - T / \alpha^n, \qquad \text{(Eq. 1)}$$

where t_n and n are the moment and the ordinal number of phase transitions (revolutions); α – coefficient of evolution acceleration, showing in what ratio the next interval between revolutions is shorter than the previous one; T – duration of the entire time interval; t^* – the point (moment) of the singularity at which period between phase transitions tends to zero. This singularity is interpreted either as a social catastrophe or as a transition of society into a new unknown quality (Nazaretyan 2016).

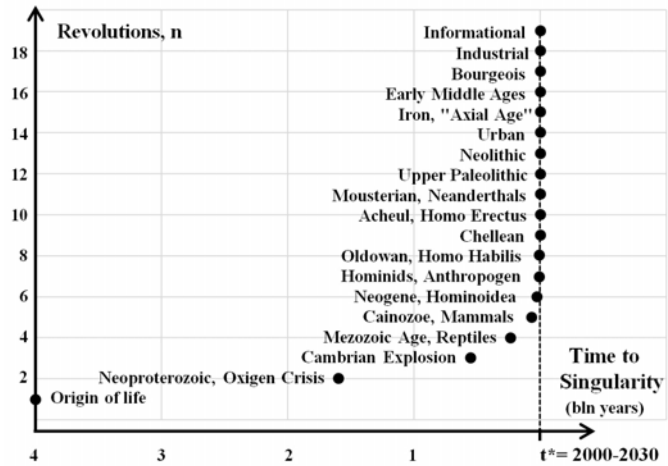


Fig. 1. Planetary evolution

Source: Panov 2005.

Yet, such an alarmist extrapolation is criticized. The Russian scientist Andrey Korotayev, who makes extensive use of mathematical methods in historical studies, argues (Korotayev 2009, 2015) that none of the real processes comes to a singularity. Processes tend to develop according to the S-shaped logistic curve, and we may better use the notion of a 'singularity zone' as a metaphor of a phase transition to a new state of process.

Mathematical interpretation of singularity requires a rigorous parameter, the same over the entire measurement range, for example, the *number* of people for demographic growth (Foerster *et al.* 1960) or the *number* of transistors per integrated circuit for technological growth (Moore 1965). The concepts of technological and planetary singularity do not satisfy this requirement of rigor, because they only technically measure time intervals, but in fact arbitrarily combine dissimilar transformations: the Cambrian explosion, the appearance of primates, the urban revolution, the appearance of electricity, the digital revolution, *etc*. Strict mathematical operations with non-strict values of events do not allow us strict formulation of mathematical singularity, neither biological, nor social, nor even technological. Besides, when calculating intervals this way, we exclude from consideration society itself and the nature of its transformation in a possible true singularity.

Recognizing the fact of technological changes acceleration, I will focus on the nature of social evolution and direction of social changes in attempt to understand *what* social novelty may arise in a singularity, rather than *when* it may happens. We need a better understanding of the comparative significance of evolutionary events (transformations), such as the Neolithic, Urban and Industrial revolution, which we are going to involve in mathematical speculation.

Singularities and Phases of Social Evolution

Let us first consider social events, which we could understand as singularities, revolutions, *etc*.

There is a question of whether we can use the notion of 'event' to characterize long and stepwise transformations such as emergence of *Homo sapience* or Neolithic transition? Yes, we can. It is conditional. In fact, there are no events in nature; they are only our concepts of changes. The 'moment of event' is an ideal notion, the same as geometric notion of a point. We always have to provide a definition of what we consider as an event at this time interval. For instance, the appearance of *Homo sapiens* and the beginning of its social evolution can be understood as an event only within a larger scale process, *e.g.* within the Big History perspective (Christian 2005). However this 'event' can be considered as a long process in the scale of the hominids' evolution.

Thus, singularity is a metaphor for the appearance of a process of a new nature or a transition from one process to another. We just should not mix the processes of different scales and maintain consistency between events and process of each scale. For example, the social evolution of humankind has only two singularities – the initial and the final. However, if we are able to distinguish between different subprocesses within this evolution, then we can understand phase transitions as singularities of these smaller scale subprocesses.

The *initial singularity* of social evolution is a 'moment' of a new phenomenon appearance – symbolic content of consciousness. Content development is a new type of evolution. One can localize separation of symbolic evolution from biological between 200,000 and 40,000 BC.

Human being, from this moment on, gets the ability to infinitely abstract and complicate ideas and accumulate knowledge. Human beings sequentially adapted more complex concepts to reality. Concepts have passed selection through practice and can be interpreted as 'memes' – gene analogs (Dawkins 1976). Evolution of symbolic content is a process of deeper and more complete understanding of reality and of human being place in it. To a certain extent, social evolution is identical to Cognition.

Social Evolution ≡ Cognition

Evolution of notions proceeds in all spheres of human beings activity – production and ideological, because all human interactions have symbolic component (Mead 1934). By the way, Marx' economic reductionism stems precisely from the fact that he ignores this symbolic component of all actions. For example, he relies on the idea that the main difference between man and apes is ability to produce (Marx 1987), rather than ability to create new symbolic meanings. Of course, ideas are tied to material practices, especially in social (collective) form such as social consciousness, but in this way we can only explain conservatism of idea-practice bundle, but not their development.

Tools and technologies have not evolved by themselves; they are artifacts of people's representations or traces of ideas' evolution in these areas. Social relations also have not evolved by themselves. Relations are reproduced by people in the process of actualization of their representations about how to act. In order to change the actions, a person must first ideate a new action. Human being in all aspects of rational activity operates with meanings; thus social evolution is evolution of meanings that are materialized in artifacts and social structures (Dobrolyubov 2012a).

Relation between material and ideological sphere of human activity is not cause-and-effect; this relation is correlative. Technologies and ideologies correlate through cognition, which is common to them. People conceptually cognize and practically master reality. They not only improve material technologies (tools, weapons, building construction, *etc*.), but they also change understanding of reality and, most importantly, change their attitude to reality and attitude to their own place and role in it, *i.e*., they change their values and evolutionarily elevate the status of a human being.

Human beings have sequentially displaced supernatural causality out of cognized phenomena; they complicate understanding of natural phenomena and representation of supernatural beings behind phenomena (spirits, totemic deities, gods, God). Along with that, they alter their own role in dealing with supernatural essence. The significance and value of a human being in his own understanding has increased in the course of evolution. Religious, moral and social concepts and, accordingly, social interactions became more humanistic and individualistic. The autonomy of human beings within society was gradually growing; means of social order maintenance were going through humanization; ways of coercion to labor gradually become less rigid (*Ibid.*).

However, within this process we can distinguish different *social mechanisms* that have pushed forward cognition and changes in technologies and ideologies.

The individual desire for cognition is an aspect of a broader contradictory need for immediate self-realization. However, in a social (not individual) form, cognition is mediated by institutionalized social interactions that contain individual and group competition and cooperation. Institutional way of their realization is an evolutionary mechanism of cognition and development.

One may recognize three types of such evolutional mechanism and three phases of social evolution – *Adaptive, Structural* and *Cognitive*. These phases are separated by two singularities – the Neolithic transition of gatherers/hunters bands to a settled society and the Global transition of multinations' social structure to a single society. Both phase transitions expel the previous mechanism of development and introduce the other. These mechanisms use different social means for competition-cooperation and different social means for cognition. In fact, these are three different evolutions at these phases, in which course different features evolve. Put it simply, one can say that adaptive skills evolved in the first phase, social structure is evolving in the contemporary phase, and after transition to a single global society only knowledge and technology will evolve.

Singularity:	Neolithic →	Global →	
Phase of Evolution:	Adaptive	Structural	Cognitive
Cognitive capacity:	Restrained	Regulated	Released
Cognition outcome:	Adaptation	Social development	Knowledge
Competition outcome:	Resettlement	Structure's expansion	Cognition

The Adaptive Phase of Social Evolution

The *Cognition* was not an explicit and conscious type of activity in gatherers/hunters bands; it was not demanded and not stimulated by social institutions. Moreover, there was an 'ideological' barrier in consciousness for change of practice. Beliefs, superstitions, taboos, rituals, *etc*. fixed in mind practices that have already existed. Such fixation happend due to the weak role of rationality and instinctive reliance on the proven solutions. The weakness of rationality was expressed in an ensoulment (animation) and mystification of all objects of nature. The human being put himself in a servant position in regards to supernatural powers, which, as he believed, were behind objects and had a legitimate arbitrariness and power over events and human destiny. Human being should rather dread and respect this power than to better understand and explain the nature of the phenomenon. Each insight and change in practice happened rarely and required the overcoming of the relevant 'ideology'. Changes were based on their immediate effect rather than on rational analyze of phenomena. Thus, cognition was open just in the direction of diversification of attainments and techniques, their adaptation to different natural niches.

Competition had a specific mode at the Adaptive Phase of evolution. *Competition of individuals* within the group had no evolutionary consequences for the group, *i.e.* did not lead to the development of the group structure, which human species inherited from the hominid pack. Only positions of individuals in the informal hierarchy can be changed. *Competition of groups* also did not have structural evolutionary consequences; it led only to adaptive variations in the size of a group, its predominant activity, its habitat, and so on.

Social evolution at this phase was similar tothe natural evolution; better adapted groups displaced less adapted groups from the habitat and, therefore, this led to migration and extension of species habitat but not to social structure development. Untilthere were no obstacles for resettlement, evolution of cognition and understanding of reality was restrained, and evolution of society's structure did not begin.

The Structural Phase of Social Evolution

The Neolithic Revolution occurred about 10,000 BC as transition of group to neighboring community. We may define it as a point of singularity of adaptive evolution that introduced a new form of society and a new mechanism of its evolution. It is evolution of social structure's dimension and complexity. The contemporary Structural Phase of social evolution has begun since this 'moment'. Of course, modern society and Neolithic settlements have essential differences but they have the same mechanism of evolutionary changes. It combines a peculiar social form of individual and group competition and a peculiar social mode of cognition. Competition and cooperation now lead to structural consequences. The competition of *individuals* results in social stratification and institualization of inequality, manifested in the ruling elites, stratums, classes. The cooperation of *individuals* leads to their consolidation into larger cohesive and solidary societies. In turn, the *competition of societies* results in their expansion (growth, merging, *etc.*).

Now different types of social structures (*e.g.*, political, economic, ideological) tend to expand. However, such expansion occurs in terms of evolutionary limitations of social structure's growth. For the integration of growing diversity in a wider social format the society requires more advanced (productive, effective) technologies and more advanced (universal, humanistic) ideologies. Society no longer restrains cognition but *regulates* it but still does not release individual cognition fully. Cognition and, accordingly, development are based not directly on human curiosity and initiative, *i.e.*, on natural need for self-realization, but on social mechanism of their regulation (promotion/limitations). This mechanism uses internal social inequality as a tool for the development through unequal exchange between society's members (*i.e.*, through exploitation of individuals). On the other hand, it uses external society's inequality for the development through unequal exchange between societies (*i.e.*, through exploitation of society's periphery or some societies by others). The competition of individuals and societies reproduces internal and external inequalities, but the successful exploitation of external inequality can mitigate negative consequences of the internal one.

At present, stratification of society has to occur for its development. Only elite can form demand for cognition and development of technologies. It is elite (ruler, royal court, nobility, state bureaucracy, priesthood) which demanded the development of weapons, monumental representation of cult and authority, elitist consumption, art, *etc.* Their objectives are strengthening of elite's internal status and the success of society in external competition.

The other side of this mechanism is the presence of low strata within society, exploitation of which allowed freeing up resources for thinkers, engineers, architects, artists, *etc*. More stratified societies evolved faster than low stratified, all other things being equal. Another consequence of this mechanism is the presence of inequality between societies. Society may exploit resources of other societies, what groups of gatherers/hunters cannot do. The most notable example is the classic Athens, which during its hegemony used resources of other poleis for weaponing, civil and cult building, development of art, theater, science, *etc*. Modern societies also use financial, economic and political hegemony, although in hidden forms of unequal exchange, for obtaining resources that allowed them to free significant part of population for fundamental science, technologies development, space exploration, *etc*. The *World-systems* analysis described the mechanism of such core-periphery interactions (Wallerstein 2004). Inequality in all its forms is a source and prerequisite for the development of society at the contemporary Structural Phase of social evolution.

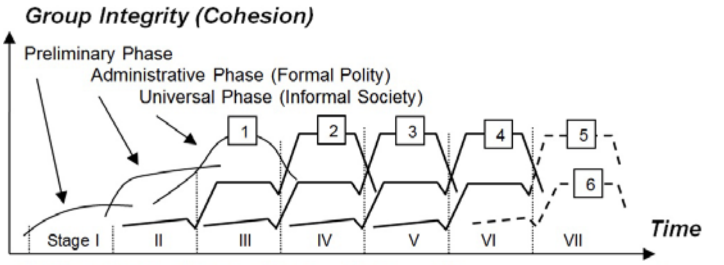
Life Cycle of Societies at the Structural Phase

The phenomenon of rise and fall of large social structures – *life cycle* of civilizations – has emerged at this phase, since there are evolutionary constraints for their structural growth. Emergence, extension, sophistication and final decay of societies become a form of their evolution. Groups also emerge and decay but their life cycle has no evolutionary consequences for their internal structure. A neighboring community becomes initial social format that may serve as starting point of extended life cycles of societies.

Consolidation of individuals within societies and competition of societies lead to expansion of the socio-political structure. At each step of expansion, the society has repeatedly undergone two key transformations and relevant phases – administrative and universal (Dobrolyubov 2009). At the *Administrative Phase*, one of the competing polities (states) subordinates the others and unites them (coercively or voluntarily) into a single political structure. The interaction and communication of individuals based on common procedures leads to standardization of practices and values. Mental reflexing of one's own similarity with others leads to the formation of a collective consciousness and self-awareness. At the next *Universal Phase* this social consciousness carries group solidarity and ensures informal cohesion of society. The entire cycle of rise and decay of society integrity also contains the *Preliminary Phase* of the beginning of societies' competition and the *Final Phase* of group cohesion dissolution (see Fig. 2). The transition from the Administrative to the Universal Phase is related with the transfer of the border of *we-they* perception from the collective identity of one format to the collective identity of a wider format. This transfer is accompanied by crises of traditional social identity, values and collective solidarity. It entails an aggravation of all kinds of group conflicts – social, ethnic, religious. After the crisis the society becomes a universal cohesive social subject that begins to compete in a broader social environment and repeats the cycle of expanding the formal political structure, and then consolidating the informal society in a wider format. The extension of social structure is accompanied by its complication.

Polity-Society may incrementally expand up to the maximal format, which is limited by the current evolutionary conditions – technological and ideological. Note that the ability of societies for political and especially military expansion increases faster than their ability to develop advanced values and ideologies that are associated with more stable and conservative culture, religion, traditions, *etc*. Therefore, even early states were able to expand the administrative structure through warfare up to large civilizational dimensions (*e.g.*, the Inca empire), but they were never able to universalize a society of this size.

When the growth of the social and political structure reaches an evolutionary limit, a social entity ceases further expansion and depletes the accessible periphery. This leads to a decrease in consumption and an increase in social tension. On the other hand, this entity attempts to integrate and universalize an over diverse social structure in a single society but does not fulfill it. As a result, political entity does not acquire broader collective identity and loses traditional collective identity and social solidarity of its core. For example, modern Western Europe had gone far in terms of informal social integration. This process is accompanied by the dissolution of the national collective identity as a main social identity of individuals. If someone attempts to turn back to nation-states, Europe risks not acquire the pan-European collectivity and at the same time lose national collectivity, and then lose any social basis for collective cohesion.



1000 B.C.	753 B.C.	509 B.C.	287 B.C.	27 B.C.	312	–	**Rome**
1250 B.C.	1000 B.C.	750 B.C.	500 B.C.	250 B.C.	–	–	**Athens**
500	750	1000	1250	1500	1750	2000	**Europe**
600	850	1100	1350	1600	1900		**Russia**

Rome Societies: 1 – Rome Town community; **2** – Rome Polis – Latium (Late Republic); **3** – Italic Nation-state (Early Empire); **4** – Rome Civilization-state (Late Empire).
Athens: 1 – Athens Town community; **2** – Athens Polis: **3** – Greek Territorial Society.
European Societies: 1 – Manors and Town communities; **2** – County, Duchies, Principalities and Town's Republics; **3** – Mono-ethnic Nation-states; **4** – Poly-ethnic Nation-states; **5** – European Civilization-state; **6** – Global Society-state.
Russia: 1 – Town communities; **2** – Independent Principalities, Town's Republics; **3** – Mono-ethnic Russian state; **4** – Poly-ethnic Russian Nation-state; **5** – Russian (Eurasian) Civilization-state.

Fig. 2. Civilizational cycle of society-state genesis (Dobrolyubov 2009, 2012b)

Any large political system, if it does not complete transition to a single universal society, with time loses system functioning and becomes an easy target of a less civilized but more solidary neighboring states, migrating barbarians, organized sects, radical movements. It can also become an easy victim of natural disasters, climate change, *etc*. The collapse of a large civilization leads to the emergence of many smaller social actors with a more primitive level of social development corresponded to invading societies. New entities inherit some technologies and ideologies, but they have their own cultural and social codes, and they begin their own cycle of expansion and development just with the better initial conditions that allow them to reach a wider size and greater social complexity in subsequent development.

The duration of the phase of maturation of collective solidarity in each of the formats is on average about 250 years (see Fig. 2). We can explicitly observe such phases of structural growth in a number of historical societies – Athens, Rome, Europe, Russia (Dobrolyubov 2009, 2012b). The average duration of the whole cycle of the society growth is: for growth up to complex chiefdoms – three phases or about 750 years, up to early states – about 1,000 years, up to territorial states – about 1,250 years, up to large civilizations – 1,500 years and more (Dobrolyubov 2012a).

The Macro-Evolutionary Diagram

All social structures eventually collapse, but the maximum size of society and the relevant level of social complexity increase in the course of evolution. Therefore, the social evolution at the Structural Phase can be presented as a macro-sequence of max social formats in the axes of technology and ideology complexity (see Fig. 3). This sequence begins with neighboring settlements and continues up to a global society. Certain societies recapitulate the path of growth from small to large formats and from lesser to greater social complexity, and finally collapse. In other words, we have to distinguish between the macro-evolution of the max structural formats and the meso-evolution through these formats of certain societies in their life cycles. Hereafter, I will use the concepts of macro and meso-evolutions in this specific sense.

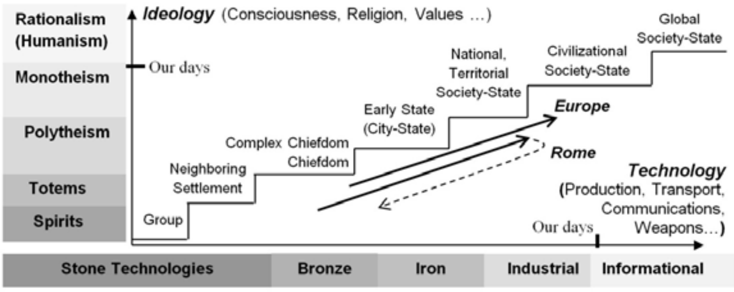


Fig. 3. Social formats in evolutionary ordinates

It should be noted that different classifications of societies between the settlement and the early state are being discussed. There are analogues and alternatives of the chiefdom, complex chiefdom and an early state (Kradin 2008; Grinin 2004, 2011; Grinin and Korotayev 2011). Besides, the role of tribal formations has not been fully clarified. Therefore, Grinin and Korotayev suggest using a more general classification: *medium-complex society – complex society – early state*. For our analysis, it is important that all alternative paths eventually merge in the state and that the same number of levels of complexity of analogs or alternatives exist along this way. The same is true of the early state. This is a broader concept than a city-state, but pristine states in the primary centers of civilization (Mesopotamia, Egypt) emerged precisely as city-states. Therefore, in the future, we will use the concept of city-state as characteristic of this social format.

Also note that the terms 'territorial', 'national' and 'civilizational' are used here in a specific sense. These social formats characterize internal social complexity that is going to be universalized in a single society. Territorial society unites a relatively homogenous and related (*e.g.*, mono-ethnic) environment, whereas a nation, as a rule, is a more complex multiethnic and multicultural entity, which, therefore, requires more 'abstract' and 'artificial' values and ideologies. In this sense, a mono-ethnic nation-state can be considered as a synonym for territorial society-state. In addition, national way of universalization is ultimately assimilation of cultures and languages (Romanization in Rome, Anglicization in Britain), whereas civilizational universalization supposes preservation of integrated cultures; that requires even more advanced values. Civilizational format is a supranational one. It characterizes multinational and often multi-confessional society. There can be both formal polity (state) and an informal universal society in these formats.

Macro-sequence of the max formats is objective and therefore linear. We can interpret the max formats and the relevant levels of material and ideological development as *evolutionary platforms* (Dobrolyubov 2012a). Nevertheless, each civilization recapitulates development up to its platform from lower levels and smaller formats, therefore societies with different evolutionary levels always coexist at the Structural Phase of Evolution.

Despite alternatives and analogues (Grinin 2004) in the past and multiple modernities (Eisenstadt 2000) in our days, all lines of developments will inevitably merge into a single global society, which will mean completion of Structural Phase of evolution.

One of the competing societies in its meso-evolution will finally make this global macro-transition. The further evolution of the global society will proceed in a stable format, as it took place at the Adaptive Phase in the stable format of the band. However, the global society will differ from bands and modern societies that it will not be able to have life cycles and will change the mechanism of further evolution and, therefore, its social nature.

It is obvious that the society requires more efficient technologies in the fields of production, communication, transport, weapon at each step toward wider format but it also needs a more sophisticated consciousness and more universal values which are often understood only as a result of changes in technologies or production, for example in the concept of social formations (Marx 1977) or in the concept of techno-humanitarian balance (Nazaretyan 2009). In fact, ideologies as well as technologies to the same extent are prerequisites for the integration of a wider social variety. For example, the transition from groups of gatherers/hunters to neighboring settlements and chiefdoms requires the development of religion from belief in spirits of objects to belief in totem deities that are the emblems of more universal supernatural powers representing larger social entities: clan, community, chiefdom. In its turn, the transition to the early state requires a more universal mythology, containing pantheon of gods. This allows integration of societies by collecting their sanctuaries in one center, for example the Acropolis in Athens or the Capitol in Rome. Moreover, gods have to acquire a human guise, or at least human behavioral traits, to facilitate formation of behavioral ideology, (*e.g*., an ideology of heroism, of citizenship, *etc.*). Informal integration of larger poly-ethnic and multicultural societies, such as the Roman Empire, requires even more universal ideology – monotheism that understands a human being more personally and allows unification through values more universal than a kinship, ethnic or any cultural affiliation.

In terms of fragmentation advanced ideology of large multi-ethnic states becomes superfluous for more primitive successors. For example, at the beginning of the new life cycle of European civilization (see Fig. 2), Christian humanism and aspiration to human perfection were unclaimed in medieval societies and were reduced to formal practices – abstinence, prohibitions, asceticism. Of course, the universality of monotheism was politically beneficial to the barbarian rulers who contributed to the spread of Christianity. Only the *Renaissance* has rediscovered humanism in Christianity and came to exaltation of human being, but then the *Enlightenment* liberated humanism from religious packing and introduced the secular and even anti-clerical ideology of individual freedoms and human rights.

However, at the beginning of *Modernity*, freedom is understood more as freedom of competition and, consequently, freedom of social stratification. The slogans of the French revolution – *Fraternité* and *Égalité* are Christian (and communist) and not at all liberal or bourgeois. Modernity's liberal ideology and values were aimed at formal status rather than the actual position of a human being in society. In our view, the integration of global diversity in a single society will require greater universality of values and more humanistic understanding of human being than liberalism provides as an ideology of formal rights and free competition in the market economy.

Modern societies are far from completing evolutionary macro-sequence (see Fig. 3). National societies should first undergo transformation to a society-state of civilizational format (Europe is trying to do it now) and only then transformation into a society-state of the global format (see Fig. 2). This path implies the crises of reformatting of existing societies. Thus, the social and political structure cannot have 'sustainable development' at this distance.

The forthcoming conflict has visibly shown its civilizational nature (Huntington 1996) when ideological opposition democracy – communism, which was historically accidental, has disappeared. The more the West acts as a solidary collective actor, the more other societies are self-aware at the same level of integrity and, therefore civilizational boundaries begin to show up where they did not matter before. In particular, the confrontation between Europe and Russia is growing as fast as Europe becomes a distinct social agent (Dobrolyubov 2012c).

The movement towards a global society contains contradictory also in terms of values; their development periodically is demanded in opposite directions. For example, the universality of the European consciousness aids to overcome national egoisms in the course of formation of a common European collective identity and collective agency. Achieving this goal, universalism makes Europe vulnerable to the influx of migrants who are foreign to this universalism and who bring rigor particularism to Europe (religious, cultural, and even clannish). Part of the problem is that European universalism exceeds the needs of local civilization integration; it is rather a cosmopolitan and globalist but not particular European. There is a contradiction here. If Europe does not acquire an explicit understanding of *We* as distinct from *They* and does not associate this collective identity with selfishness and even isolationism (which, incidentally, has American consciousness) that would be sufficient for leadership in civilizational competition, then Europe can simply disappear as a collective carrier of values. European consciousness is ready for global universalization, but does not have sufficient group solidarity and collective agency to lead such integration.

Nevertheless, catastrophic consequences of the collapse of large civilizations are mitigating evolutionarily, because their life cycles are not fully synchronized. At the same time several civilizations which are the carriers of common achievements, are involved in the World-System. For example, the Arabic civilization became an intermediary between Greco-Roman and modern European civilization; this allowed restoration in Europe some of its own ancient achievements a millennium later. The collapse of civilizations cuts off the peaks of its development – the most artificial and refined elitist practice, science, engineering, the most advanced social institutions, the system of elite education, civility and so on.

The development of civilizations through rise and decay does not allow us to reconcile the concept of stadial (or unilinear) evolution shared by the founders of evolutionism (Lewis Morgan, Friedrich Engels, Herbert Spencer and others) with concept of multilinear evolution offered by neo-evolutionists (Leslie White, Julian Steward, Marshall Sahlins and others). The actual process is both progressively stadial and cyclic. Only the transition to macro-observation allows us to ignore evolutionary 'failures' and distinguish mature forms of local civilizations from the historical flow and interpret them in different ways: as social formations (Marx 1977), as phases of historical process (Diakonoff 1999), as world's civilizations (Yakovets 1999), as production principles (Grinin 2007), as evolutionary platforms (Dobrolyubov 2012a).

Acceleration or Cyclic Recurrence?

One can agree with the statement of the acceleration of historical time. However, when some researchers describe this acceleration by the hyperbolic curve (see Fig. 1), they are, in fact, artificially 'hurry up' evolution by ascribing higher evolutionary importance to current transformations.

There is no objection, when scientists distinguish different evolutions (Grinin *et al.* 2011), successive phases of historical process (Diakonoff 1999), world civilizations (Yakovets 1999) or otherwise classify periods of development. However, in order to construct a model for evolution's acceleration based on durations of these periods, one should first prove that the used phase transitions that break up the process into periods have equal significance throughout the considered interval. If we begin to consider social or even planetary evolution in more and more specific technological transformations, we equate the significance of these transformations, *i.e.*, recognize Neolithic changes in society equal to changes caused by invention of the Internet.

For example, the historians (Diakonoff 1999; Yakovets 1999) consider, with minor differences, the following sequence of historical phases: Prehistoric, Neolithic, Early-class, Antique, Medieval, Pre-Industrial, Industrial, and Post-Industrial, which Alexander Panov used in his model (Panov 2005). However, the last four 'phases of world history' according to Diakonoff or 'world civilizations' according to Yakovets are historical phases of European civilization, *i.e*., they are locally Western, and not world ones. Of course, Western achievements are diffused in the World-system as well as the Greco-Roman, Arabic or Chinese achievements in the past, but the evolutionary issue is whether these advancements are irreversible? Civilizations have life cycles and they lost civilization many times in past history. So, there is a reasonable question: which Western social and technological advancements will remain in non-Western societies in case of collapse of the leading Western civilization? We do not know for sure.

In fact, different civilizations are undergoing similar structural transformations. If we take a closer look at the Greco-Roman civilization, we will find transformations and development phases, similar to European ones (see Fig. 4). Rome and Greece began their development from the 'dark ages', fragmented social entities with natural economies (it is an antique analogue of the Early Middle Ages in Europe). Then polity of city-state format appeared: poleis in Greece and Rome, and town republics, principalities, duchies in Europe. Then, the universal societies were formed in a city-state format, which was accompanied by a cultural explosion – the 'Axial Age' revolution (Jaspers 1953), manifested as the Classicism in Greece, the Hellenization in Rome, and the Renaissance in Europe. Finally, large universal societies have emerged with commodity production and market economy. Some historians define Roman society at this stage as 'proto-bourgeois' or 'capitalistic' (Semyonov 2003: 164; Vassiliev 2008).

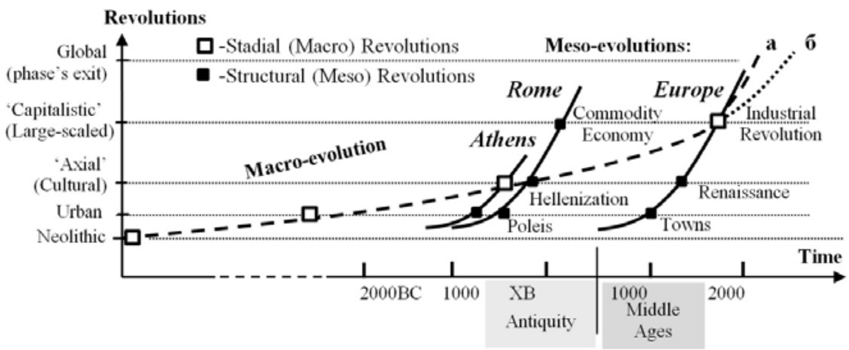


Fig. 4. Illustration of macro-evolution as a sequence of meso-evolutions

It seems that every large civilization follows the same path of structural development and, therefore, it has its own singularity, in fact – its own collapse. By applying the hyperbolic model to the events of antiquity Andrey Korotayev, as a sort of science joke, calculated the 'Korotayev-Archimedes singularity' in AD 115 (Korotayev 2015).

If we look at the historical process in terms of macro-evolution, we can state that the Urban Revolution occurred when towns first appeared in early civilizations. This statement is also true for other *stadial* revolutions. The following civilizations, including Greco-Roman and modern European, just repeat this way, starting with a lower level (complex chiefdoms), but at higher overall evolutionary level of the World-system. Each civilization has to recapitulate urban transition because towns and city-states are social formats that a growing society passes in the life cycle of its genesis. Other macro-evolutionary Revolutions are also related with the new social formats, which are also the stages of the genesis of every specific society. Each civilization recapitulates the urban, then cultural transformation, and then the transition to a large universal society with a commodity economy. Thus, we describe different meso-evolutions of specific civilizations by the same 'revolutions' as the macro-evolutionary process, but we should distinguish *Stadial Revolutions* and ordinary *Structural Revolutions* of specific societies.

Note that the axis of 'revolutions' in Fig. 4 is not fully stadial in aspects which are different from the structural one. Indeed, the transition to the universal society in a seemingly same structural format may have different institutional appearance. Roman society of imperial period acquired large-scale economy with largest regional market that demanded large-scale commodity production. This production required a large number of slaves, thus, their owners were predominant economic agents. However, we do not find industrial revolution in this ancient 'capitalism', though the Roman engineering and technological level, which was subsequently lost, was comparable with pre-bourgeois European (mechanisms, water actuators, mechanical mowers, steam turbine, *etc.*). The true European medieval invention is an *individual* economic agent, protected by the *institution* of private property, for which in fact the previous development cycle of Roman law and practices was required. Slaveholders demanded inventions to achieve mainly high-status and not economic objectives – for spectacular shows to impress a crowd, for monumental construction, weaponing, *etc*. Only an individual economic agent began to demand inventions for the sake of individual profit. As a result, the typical *structural* revolution of transition to the universal societies in Europe (in fact, 'national' and 'capitalistic') has acquired the features of a *stadial* revolution in all other aspects.

The universality of big societies gives them some similarities in social relations and consciousness. We can consider Roman society in a certain sense as a consumer society; this society gave individuals considerable autonomy, it had a large 'proletariat', it was cosmopolitan, *etc*. These signs are symptoms of *ancient modernity* that arose in the course of overcoming the traditional society. Later on, the society acquired *postmodern* signs of deconstruction, decadence, indifference and fatigue.

Linear stadial approaches, for example, Marxist ones, tend to exclude from consideration the entire chain of stadial transformations in each particular ancient society. They describe Roman civilization by the general stadial level, such as slavery, agrarian society, Antiquity, the 'Axial Age', *etc*. Although the historical phases of Roman society (monarchy, republic, and empire) have stadial differences in production and social intercourse. By the way, the World-system approach, unlike the Marxist one, notes this gradation (Grinin and Korotayev 2009).

Until recently Western societies repeated structural transitions that have already taken place in the past societies (see Fig. 4). Of course, modern society faced new phenomena – the digital revolution, Internet and social networks, genetic engineering, *etc*. However, we can place these 'phase transitions' on the sequence where the Neolithic, Urban and other stadial revolutions are located, only if the Western civilization completes the Structural Phase of evolution. Then, indeed, the meso-evolution of Western society will coincide with the macro-evolution of society 'in general' or of humanity as a whole (curve **a**, Fig. 4). However, in the event of Western civilization collapse and social primitivization, as was the case with all previous historical civilizations, future observers of the process will assign less importance to the transformations of Western society, as we do now with regards to Roman society. Future observers will smooth out the course of evolution (curve **b**, Fig. 4) and assign the averaged characteristics to the societies that existed before them.

Linear-stadial evolutionary approaches assign a single evolutionary level to the Greco-Roman civilization in order to artificially inflate stadial level of the

Early Middle Ages. In reality, ancient society was highly civilized and technically advanced. Its decay has led to societies' return to the 'dark ages', to the pre-state level of social organization, to the loss of culture, knowledge, technologies, *etc*. Of course, at this time feudalism began to form as unique system of legal relations. The feudalism in the course of its long evolution gave birth to modern society. However, one cannot attribute evolutionary perfection of the modern society to its medieval embryos.

The macro-evolution of social formats shown in Fig. 3 as a sequence of stadial revolutions is a slower process compared with meso-evolutions shown as civilizations' development cycles. Even the leading European (Western) civilization, after it becomes politically unified state-society (this process is delayed for a while), will need at least one more 250-year phase for political unification of the global society and for its values universalization (see Fig. 2). This is a long historical period, even if we assume that under present conditions the phase's duration might be reduced.

The very idea of infinite growth of evolution rate seems doubtful. For example, changes in biological evolution cannot go faster than life span of organisms (generation of organisms) – a kind of *biological 'quantum'* of evolutionary time. Evolution at a higher rate is physically impossible. Of course, organisms can change faster, but only in ontogenesis; and such changes are not evolutionary. The concept of social evolution also loses its meaning when we begin to consider the changes that occur faster than the *social 'quantum'* of evolutionary time. This is the period of existence of conservative carriers of social structures, institutions and values, such as a mature individual, generation of people, solidarity communities (societies), *etc*. At the structural phase of social evolution, such a quantum of time is the phase of sociogenesis, lasting for about 200–300 years (see Fig. 2).

An unjustified transfer of evolution 'arrow' from socio-structural changes to technological changes leads to too optimistic assessment of the evolution acceleration and of the 'moment' of its singularity with an error of at least several centuries, if not thousands of years. Technological singularities change neither the nature of society, nor the role of human beings in it. Singularity of social evolution is possible, but it will be a social rather than technological event; it will not be caused directly by technological changes, no matter how impressive they are, but will be caused by their organic link to social phenomena.

The Rate of Social Evolution

The question of social evolution rate remains. But how can we measure it?

If the growth of the maximum format of society takes place at the Structural phase of evolution, then it is logical to use this structural step as a measure of evolutionary progress of society. I once again remind here that we are talking about the format of informal *society*, and not about the format of a *state* structure that can run far ahead. Though history gives us information mainly about political structures (conquests, centralization of states, *etc*.), and not about informal society as a cohesive community with common collective identity. In fact, each case of a large state or empire formation testifies to the appearance of a cohesive core of a smaller format. For example, the Macedonian, Mongolian or Incas' expansion rather speaks about the existence of relatively narrow ethnic cohesive social core on which the rulers relied than about informal societies of civilization format emergence. Most empires remained formal and fragile political structures, and never became universal societies of such wide format. We should take this into account when determining the moment of the actual evolutionary transformation.

Besides one should use only the first cases of the structural formation to universal societies, which are true or pristine evolutionary transformations. The subsequent recapitulation of these transformations by other societies is not actually a macro evolutionary one, but is an ordinary structural transformation in their life cycle. The first cases of the primary formation of universal societies in each format are summarized in the table.

Table

First transition to universal society in format of	T	Year	Events that indicate this transition
Band	T_0	40,000 BC	Hunters/gatherers bands
Settlement	T_1	10,000 BC	Late Natufian Neolithic settlements in the Middle East 10,800–9,500 BC (Munro 2003; Barker 2009)
Chiefdom	T_2	8,000 BC	Walls and tower of Jericho 8,350–7,370 BC (Kenyon 1981)
Complex Chiefdom	T_3	5,000 BC	Urban revolution 5,000 BC (Childe 1950); Eridu at Ubaid period in Schumer 6,500–4,100 BC (Mallowan 1970)
City-State	T_4	3,000 BC	Political centralization of Egypt 3,000 BC
Territorial Society-State	T_5	1,500 BC	'World power' in Egypt – the New Kingdom (1,549–1,069 BC)
National Society-State	T_6	27 BC	Romanized Italic national core within Rome Empire, 27 BC.
Civilizational Society-State	T_7	–	Universal Society-State of Europe at Universal phase (2,000–2,250 AD)
Global Society-State	T_8	–	Global Universal Society-State

These data require explanations. Let us comment through some points.

Homo sapiens inherited the initial format of society – a *band* from a flock of hominids, and the first structural transformation was, in fact, a Neolithic transition to agriculture and permanent *settlements*. Therefore, we will not use the starting point of social evolution (T_0) as it is not a structural transition. Evolution at the Adaptive phase had a different mechanism and occurred more slowly than at the Structural phase. Of course, the rate of evolution at the Adaptive phase can be a separate subject of study.

Chiefdom and complex chiefdom were studied using the examples of societies of Polynesia, Oceania, America, *etc*. that are closer to our time. We know practically nothing about chiefdoms and complex chiefdoms (or medium-complex and complex societies) in the primary centers of civilization in Mesopotamia and the Middle East. The archeology provides us a predominantly urban line of their development. Nevertheless, we may relate development of their political center with a certain level of informal organization. The appearance of the first looking like towns settlements (Jericho) we can relate to the formation of the chiefdom (T_2) and we can relate the urban revolution in Mesopotamia, which occurred around 5,000 BC to the universal societies within complex chiefdoms (T_3).

Why can we state so? We know that maturation of universal and cohesive society in a certain format leads to the attempts of its administrative expansion. This fact is represented in the scheme of sociogenesis (see Fig. 2) as imposition of two phases – the universal phase of one society and the administrative phase of society of the following format. We can use it as the markers of successful completion of previous transition. When a cohesive collective identity is formed in complex chiefdom, it inevitably attempts to seize other chiefdoms' capitals and build political superstructure of the city-state format, which leads to political formation of urban policies, nomes, *etc*. Thus, we can use it as an indicator of completion of universal and cohesive society formation within complex chiefdom.

The same applies to the next transformation (T_4). The centralization of the territorial format of the state reveals that the society in the previous format of the city-state has already been universalized and has acquired common identity and collective cohesion. This indicator, perhaps, slightly shifted to the right along the time axis, but not more than for the 250-year phase of sociogenesis society's genesis that is not so significant for the evolutionary scheme. By the way, the phase of existence of social identity and cohesion of a certain format (200–300 years) proves to be a period of stability for excessively large states that exceed the size of their own informal society, after which they disintegrate, for example, Akkad (2,316–2,137 BC), Assyria (1,353–1,000 BC), *etc*. The periodic collapse of territorial states led to oscillation of the sociogenesis societies' genesis around the city-state format. It simply means that macro-evolution is getting stuck. The short periods (200–300 years) of large states' existence indicate that the societies of city-state format still remain the actual actors at this time.

Only the appearance of large multi-ethnic empires which have been stable during two or three phases of their genesis (500–600 years) and which have projected power to remote periphery for a long time, indicates that they have a cohesive ethnic core that can be understood as the universal society and collective identity of territorial format (T_5). The new Kingdom of Egypt (1,549–1,069 BC) first formed a single universal society and overcame internal competition of nomes, which persisted throughout the Old Kingdom (2,686–2,181 BC). It was the first stable 'world power' with unprecedented prosperity and stability (Shaw 2000).

The next evolutionary transformation (T_6) is the formation of a universal national society. The appearance of the first nations is usually associated with the modern Europe. In our opinion, the first nation was Romanized Italy as metropole of the Roman Empire during the Principate. It had the structural traits of the European nations and possessed sufficient degree of 'artificiality' and 'abstractness' of ideology, which included a kind of Roman nationalism based on the ideas of Roman exceptionality, superiority over 'others', contrasting themselves with barbarians, *etc*. This core was originally multi-ethnic and composed of the Latins, Italics, Etruscans, Gauls, *etc*., who were Romanized. After more than a century of national crisis, including the Gracchi reforms, allied and civil wars between parts of Italy, this society was universalized and acquired a single social identity and internal cohesion. After the political structuring of the Empire in 27 BC (T_6) this community became a nation and a collective metropole for the subordinate provinces. Italy as a part of the Empire had universal citizenship and retained republican institutions – the Senate, the court and the rights of the individual, self-government of cities, *etc*. The strength and cohesion of this national community allowed the Empire to exist for about five hundred years. Although it should be said that this national community failed to universalize the imperial society of the civilizational format.

European societies have retraced the structural path from complex chiefdoms to nations that have already been passed by others (see Fig. 2) but with a new technological and value quality. Europe is just preparing for the most important step of its structural genesis – transition to the universal society of civilizational format (T_7), *i.e.* to a single European state-society. Such a step has not yet been made by any of the societies. This evolutionary transformation may occur (of course, may not occur) during the current 250-year civilizational universal phase. Accordingly, the last structural evolutionary transformation (T_8) will be the transition to a single universal global society within the global state.

Now, we have the data points, and we can try to mathematically interpret the course of structural evolution. The best approximation shown in Fig. 5 is described by the linear equation

$$T(n) = -11214 + 1893\,n, \qquad \text{(Eq. 2)}$$

where $T(n)$ – is the moment of evolutionary transformation, n – is the ordinal number of transformation.

If we put the next civilizational transformations in the equation, we get that Europe should make transition to the universal society in $T_7 = 2037$. Of course, we can only use this as an illustration because the error is ±125 years, since the dates used are time intervals and not the points.

However, some interesting results can still be obtained from this interpretation. For example, the rate of *structural* evolution at the contemporary phase is constant and is around V = 1,893 *years per transition*, which obviously follows from the data used. Indeed, evolutionary transitions to the new social format take place approximately every two millennia.

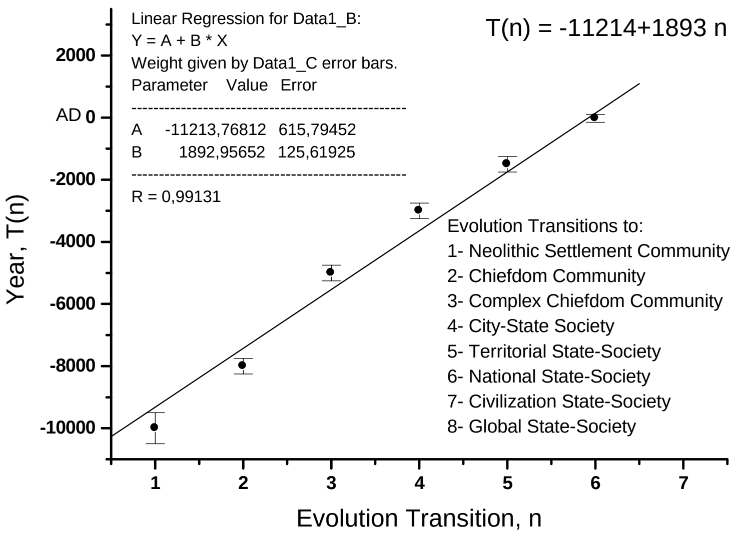


Fig. 5. Approximation of social evolution at the Structural Phase

Even if we shift the dates and use other examples of societies, for example, take the Old Kingdom of Egypt (2,686–2,181 BC) as the first transition to the territorial format society, or European nations as the primary formation of nations, it will not have significant impact on the rate of macroevolution. This is likely to affect the degree of correlation. Only the total number of transformations can significantly affect the rate of evolution.

The result obtained for a global transition T_8 = AD 3930 is strange only at first glance. In fact, it follows from the assumption that the global transition will be completed by another, rather than by modern Western, civilization. That assumption is based on the previous cases of societies' genesis. None of the civilizations of the past has made two macro-evolutionary transitions (*i.e.*, primary) in the same cycle of genesis, although each of them passed through a chain of meso-evolutionary transformations (in fact, secondary) from the chiefdom to the wide formats.

Numerous processes in modern society may have singularities in the mathematical sense (demographic, economic, and technological) that can destabilize social system. Nevertheless, until social evolution completes the Structural Phase, the society will have a simple response to the growth of any parameter in blow-up regime and even of a number of parameters – it simply collapses and rolls back along the evolutionary scheme (see Fig. 3) without transition in any new quality. That always happened in the course of past social evolution.

In the event of the collapse of Western civilization due to the new migration of peoples (today, it seems, these are migrants from non-Western countries with prevalence of intolerant, and clannish, authoritarian social consciousness and values), we may face a new genesis of a new civilization, which in full cycle can take around 1,750–2,000 thousand years (see Fig. 2). However, it is just an assumption. Western civilization may continue to lead the process of globalization and complete the Structural Phase of evolution, but even in this case it will take one or two phases of genesis – 250–500 years.

The Nature of the Global Singularity

The formation of a universal global society will become a true singularity related to the *metamorphosis* of society's nature. The conditions for a transition towards the global society and its new characteristics are a very interesting subject of a separate study. We only briefly denote them. The global society will have new features due to the fact that it will be *single* and will have *no periphery*, which will not allow using the present evolutionary mechanisms for further social development.

At the current Structural phase, the competition of societies, social and core-periphery inequality as well as the presence of the upper class are necessary elements for the development. Global societies' sustainable development is impossible with these institutions. It will have to overcome social and regional inequalities, conflict forms of competition, consumer orientation of consciousness, stop the depletion of the world's natural resources and enter into symbiosis with nature. The society will have to liberate a human being from wage labor, which is a form of forced labor, and give an individual opportunity of voluntary, *i.e.*, completely free activity.

Such society looks like a Marxist utopia only at the modern level of technological development and universality of value system. And if technologies are able to develop progressively, then the values can reach a new level only through the crisis of existing value system and existing rationality.

Meanwhile, evolution has enough time and can wait. Sooner or later the society will come to its new social nature. Human being's desire for self-affirmation and self-realization in cognition will move forward further evolution at the Cognitive Phase. The social changes will be more rational, whereas now they are unintended outcome of societies' competition. The responsibility for the continuation of evolution, which was previously transferred from biological selection to the competition of societies, will now be transferred directly to human rationality. However, evolution will remain an objective and inevitable process, since cognizable reality is objective and evolutionary purposes of reason, human or artificial, in reality are also objective.

Conclusion

The article proposes division of social evolution into three phases: adaptive, structural and cognitive. The period of the structural phase is about 10,000 BC – 4,000 AD. The dynamics of primary structural transformations at this phase is described by the equation $T(\mathrm{n}) = -11214 + 1893\,n$, where $T(n)$ – is the moment of evolutionary transformation, and n – is the ordinal number of transformation. The rate of structural evolution is V = 1,893 years per transition. Global singularity as the completion of the structural phase of social evolution and the transition to the universal global society-state is predicted by the model for 3930.

References

Barker G. 2009. *The Agricultural Revolution in Prehistory: Why did Foragers Become Farmers?* Oxford: Oxford University Press.

Childe G. 1950. The Urban Revolution. *Town Planning Review* 21: 3–17.

Christian D. 2005. Macrohistory: The Play of Scales. *Social Evolution and History* 4(1): 22–59.

Dawkins R. 1976. *The Selfish Gene.* Oxford: Oxford University Press.

Diakonoff I. M. 1999. *The Paths of History. From the Prehistoric Man to Modern Times*. Cambridge: Cambridge University Press.

Dobrolyubov S. V. 2009. Theory of Society's Genesis. *Social Evolution and History* 8(1): 221–255.

Dobrolyubov S. V. 2012a. Sociogenesis vs. Marx Evolutionary Determinism: The Anthropic Mechanism of Social Dynamics. *Social Evolution and History* 11(1): 82–123.

Dobrolyubov S. V. 2012b. Polis, National and Civilizational Cycles of Russian Society's Genesis. *Mir Rossii* 2: 116–143. *In Russian* (Добролюбов С. В. Полисный, национальный и цивилизационный циклы социогенеза российской общности. *Мир России* 2: 116–143).

Dobrolyubov S. V. 2012c. Collective Agent and Collective Subject: Some Basis for Conceptualizing. *Sotsiologicheskiy zhurnal* 2: 53–80. *In Russian* (Добролюбов С. В. Коллективный агент и коллективный субъект: некоторые основания концептуализации. *Социологический журнал* 2: 53–80).

Eisenstadt S. N. 2000. Multiple Modernities. *Daedalus* 129(1): 1–29.

Foerster H. von, Mora P. M., and Amiot L. W. 1960. Doomsday: Friday, 13 November, A.D. 2026. *Science* 132(3436): 1291–1295.

Grinin L. E. 2004. The Early State and Its Analogues: A Comparative Analysis. *The Early State, its Alternatives and Analogues* / Ed. by L. E. Grinin, R. L. Carneiro, D. M. Bondarenko, N. N. Kradin, and A. V. Korotayev, pp. 88–136. Volgograd: 'Uchitel' Publishing House .

Grinin L. E. 2007. Production Revolutions and Periodization of History: A Comparative and Theoretical-Mathematical Approach. *Social Evolution and History* 6(2): 75–120.

Grinin L. 2011. Complex Chiefdom: Precursor of the State or Its Analogue? *Social Evolution and History* 10(1): 234–275.

Grinin L. E., and Korotayev A. V. 2009. Social Macroevolution: Growth of World System Integrity and a System of Phase Transitions. *World Futures* 65(7): 477–506.

Grinin L. E., and Korotayev A. V. 2011. Chiefdoms and their Analogues: Alternatives of Social Evolution at the Societal Level of Medium Cultural Complexity. *Social Evolution and History* 10(1): 276–335.

Grinin L. E., Korotayev A. M., and Markov A. V. 2011. Biological and Social Phases of Big History: Similarities and Differences of Evolutionary Principles and Mechanisms. *Evolution: A Big History Perspective* / Ed. by L. E. Grinin, A. V. Korotayev, and B. Rodrigue, pp. 158–198. Volgograd: 'Uchitel' Publishing House.

Huntington S. P. 1996. *The Clash of Civilizations and the Remaking of World Order*. NY: Simon & Schuster.

Jaspers K. 1953. *The Origin and Goal of History.* New Haven, CT: Yale University Press.

Kapitza S. P. 2006. *Global Population Blow-up and After. The Demographic Revolution and Information Society*. A Report to the Club of Rome. Moscow, Hamburg.

Kenyon K. 1981. Excavations at Jericho. Vol. 3. *The Architecture and Stratigraphy of the Tell* / Ed. by T. A. Holland. London: British School of Archaeology in Jerusalem.

Korotayev A. V. 2009. New Technologies and Future Scenarios, or is the Singularity Near? *History and Synergetic* / Ed. by S. Yu. Malkov, , and A. V. Korotayev, pp. 183–191. Moscow: URSS. *In Russian* (Коротаев А. В. Новые технологии и сценарии будущего, или сингулярность уже рядом? *История и синергетика* / Ред. С. Ю. Малков и А. В. Коротаев, с. 183–191. М.: УРСС).

Korotayev A. V. 2015. *Singular Points of Megahistory*. Report at 2nd Symposium of 'Megahistory and Global Evolution', October 27–29. Moscow.

Kradin N. N. 2008. Early State Theory and the Evolution of Pastoral Nomads. *Social Evolution and History* 7(1): 107–130.

Kurzweil R. 2005. *The Singularity is Near.* NY: Viking Books.

Mallowan M. 1970. The Development of Cities from al- 'Ubaid to the End of Uruk 5, Chapter VIII / Ed. by I. E. S. Edwards, C. J. Gadd, and N. G. L. Hammond, *The Cambridge Ancient History*. 3 rd ed. Vol. 1, part 1: Prolegomena and Prehistory, pp. 327–462. Cambridge: Cambridge University Press.

Marx K. 1987 [1846]. *The German Ideology*. London: Lawrence and Wishart Ltd.

Marx K. 1977 [1859]. *A Contribution to the Critique* of *Political Economy*. Moscow: Progress Publishers.

Mead G. 1934. *Mind, Self, and Society*. Chicago, IL: University of Chicago Press.

Moore G. 1965. Cramming More Components onto Integrated Circuits. *Electronics Magazine* 38(8): 114–117.

Munro N. 2003. Small Game, the Younger Dryas, and the Transition to Agriculture in the Southern Levant. *Mitteilungen der Gesellschaft für Urgeschichte* 12: 47–71.

Nazaretyan A. P. 2009. Technology, Psychology and Catastrophes: On the Evolution of Non-Violence in Human History. *Social Evolution* & *History* 8(2): 102–132.

Nazaretyan A. P. 2016. *Non-Linear Futures: The 'Mysterious Singularity' in View of Mega-History. Between Past Orthodoxies and the Future of Globalization.* Leiden, Boston: Brill.

Panov A. D. 2005. Scaling Law of the Biological Evolution and the Hypothesis of the Self-Consistent Galaxy Origin of Life. *Advances in Space Research* 36: 220–225.

Semyoenov Y. I. 2003. *The Philosophy of History*. Moscow: Sovremennye tetradi. *In Russian* (Семенов Ю. И. 2003. *Философия истории*. М.: Современные тетради).

Shaw I. (Ed.) 2000. *The Oxford History of Ancient Egypt.* Oxford: Oxford University Press.

Snooks G. D. 1996. *The Dynamic Society. Exploring the Sources of Global Change.* London and New York: Routledge.

Vassiliev L. S. 2008. Factors of History. *Istoricheskaya psikhologiya i sotsiologiya istorii* 1: 169–195. *In Russian* (Васильев Л. С. Факторы истории. *Историческая психология и социология истории* 1: 169–195).

Wallerstein I. 2004. *World-Systems Analysis: An Introduction*. Durham, NC: Duke University Press.

Yakovets Y. V. 1999. *Cycles. Crisis. Forecasts*. Moscow: Nauka. *In Russian* (Яковец Ю. В. *Циклы. Кризисы. Прогнозы*. М.: Наука).

13
The Punctuated Equilibrium Macropattern of World System Urbanization and the Factors that Give Rise to that Macropattern

Antony Harper

Abstract

Change in complex systems, system evolution, is itself complex. This paper is about world system evolution as it is reflected in the pattern of urbanization over the last 5,000 years. It will be shown that the pattern of urbanization as part of the immensely complex world system exhibits non-linearity in that it is neither smooth nor continuous but rather is punctuated by periods of rapid change interspersed between periods of stasis. This pattern was first described in biological systems by Eldridge and Gould (1972) for speciation, and much of the pattern of urbanization reflects the characteristics of punctuated equilibrium first described by those two authors. Specifically, this paper will investigate the phenomenon of punctuated equilibrium reflected in both the macro-pattern of urbanization over historic time, i.e. the evidence for punctuated equilibrium as reflected by data on urbanization and on the level of state development, and possible mechanisms for such punctuated behavior including the general model of self-organized criticality as developed by Per Bak (1996), the role of hypercycle formation in punctuated equilibrium, the role of aromorphic processes, and the interaction between population, carrying capacity, and level of technology as represented by a very general math model.

Keywords: *world system, urbanization, punctuated equilibrium.*

Introduction

Science is a human endeavor that identifies patterns in nature and then attempts to explain and give predictive value to those patterns by using tested, hypothetical reasoning. History is rightly identified as a branch of science by a number of scholars, chief among them Peter Turchin (2003). This inclusion of history within the domain of science is due to its overlap with the social sciences, is also due to its domain subscribed by the domain of biology, a science, as *history* is the history of a particular species, our species, *Homo sapiens*, and finally is due to its being the last segment of a temporal record extending through the domain of paleontology to limits set only by what contemporary cosmology has discovered with that temporal record. All of these conditions are clearly under the scrutiny of a scientific lens, a lens that has been given sharp focus due to the quantitative efforts established by the Russian school of macrosocial dynamics led by Andrey Korotayev and Leonid Grinin and by the efforts of Peter Turchin's newly christened Cliodynamics. Even so, history is only recently coming under formal scientific scrutiny and has much to reveal. [Parenthetically, I would like to point out that the efforts of traditional historians, past and present, are in no way devalued by a more quantitative approach, but that a *scientific* history is not simply revelatory in its own right but also complimentary and supplementary to the contributions of traditional historians and other social scientists.] It should further be noted that underlying the notions of explanation and prediction is not just a fundamental understanding of phenomenon or pattern but also of mechanism, *i.e.* how events and processes occur. It is to these ends, to explanation, prediction, and the revelation of mechanism, that this paper is focused, and, more specifically, on the reality of (relatively) abrupt change as it manifests itself in the macro-pattern of urbanization over the last 5,000 years of world system history.

As the overarching model applied in this paper is that the human world is a world system, a part of which is the process of urbanization, a note should be made here of the fact that many world system scholars do not agree with the extension backward in time of the existence of the world system to the beginning of urbanization, historical urbanization, at approximately 3000 BCE. Chief among these critics is the doyen of world system analysis, Immanuel Wallerstein, but there are others. The position, however, is taken that the world system, albeit a loosely connected world system, does extend to that temporal horizon, and this position follows the lead of the late Andre Gunder Frank (Frank and Gills 2014) and others who suggest that the world system characteristics of flow of capital, core-periphery structure, and cycles of hegemony and revolution are all evident phenomena prior to 1500 CE. It is within the context of a world system that the research for this paper was done and done explicitly to address the punctuated pattern of urbanization over the last five thousand years.

Abrupt change in nature has been a subject of scientific study for some time. Phase changes in matter are an obvious example (see, *e.g*., Sole 2011 and Scheffer 2009) but more recently phase changes have also been identified and investigated in the study of mathematical chaos, in ecological literature, and in a variety of other fields. In 1972 a remarkable paper was coauthored by Niles Eldridge and the late Stephen J. Gould on the reality of punctuated change in the process of speciation. In this paper Eldridge and Gould (1972) showed that a literal interpretation of the evidence of the fossil record in light of the accepted theory of allopatric speciation suggested that the event of speciation was rapid, contra Darwin *et al.*, not stately as the conventional wisdom of gradualism would suggest. The event of speciation itself would occur in small, peripherally isolated populations of a larger biological species population. Lastly the actual event of speciation would consequently hardly ever be subject to preservation in the fossil record. Eldridge and Gould also identified a second reality, that of stasis in the fossil record, and they defined stasis as oscillation about some average value, the oscillations being a result of microevolutionary adjustments to relatively stable local environments. The paper went on to explicitly document examples of punctuation and stasis in both Pleistocene snails and Paleozoic brachiopods, specialty areas of research of both authors. Clearly, this paper 'Punctuated Equilibrium: An Alternative to Phyletic Gradualism' was a watershed event with respect to evolutionary understanding and forever changed the context within which evolutionary events would be understood. It was not without controversy.

Eldridge and Gould went further than simply identifying a pattern in the fossil record, they established the reality of biological mechanisms being the motive forces for evolutionary patterns in the record, that is that the fossil record would only reveal pattern in which biological mechanisms, both extant and extinct, created. Specifically in the case of punctuated equilibrium, allopatric speciation was proposed as the mechanism that would produce the stasis and punctuation evident in the fossil record. While it would be inappropriate here to delve into a detailed explanation of allopatric speciation, there are two aspects of that theory which should be brought to the attention of the reader, the geographic distribution of local populations within a given species population and within that distribution the potential for isolation and therefore (relatively) rapid evolution of small, peripheral local populations. These populations would not preserve well in the fossil record but would change rapidly and would therefore represent tipping points in the speciation process. To reiterate, small, (relatively) isolated populations in a species population are potentially subject to rapid change from which new species may arise, an event that is biologically explainable but is not well resolved in the fossil record. Consequently, the temporal distribution of fossils should (and does) reveal both stasis and punctuations. It is a position of this paper that this pattern of (relative) stasis and punctuation occurs by default of context over the course of Big History but is specifically evident in the macro-pattern of urbanization over the last five thousand years of world system history.

A further comment on mechanism is needed here. The late Per Bak (1996), in his research on self-organized criticality (SOC) established a general mechanism for the occurrence of punctuated evolution, and in fact in a brief conversation with Stephen Gould he asked if Gould had established a theoretical explanation for punctuated equilibrium, upon which Gould acerbically replied, 'Punctuated Equilibrium is a theory'. Sadly, Gould did not recognize the depth of Bak's question, that there had to be a mechanism of a mechanism, that punctuated change required context, constraint, and process in order to occur, and that understood from this broader perspective (the mechanism) would yield greater understanding of the phenomenon of punctuated equilibrium in general. What Bak (1996) describes and Flyberg *et al.* (1993) proposed by way of mechanism of punctuated equilibrium is simply this; given a set of species with a range of selective values, using a specific selection process, the population of species is upgraded in a way that describes all species surviving above a given threshold selection. This process mimics the pattern of punctuated equilibrium in that periods of stasis (above some minimum selective value) are punctuated by periods of avalanches of change. This biological and theoretical aside is mentioned here, as it will be suggested that the periods of punctuation in urban area evolution over time may have an analogous mechanism. The emphasis here is on 'may'.

There is one more evolutionary perspective that should be addressed here with respect to the parallel processes of biological and social evolution, that of the Russian school of evolutionary thought. Specifically, the concept of aromorphosis (Grinin, Markov, and Korotayev 2009) has relevance here. Aromorphosis is the evolutionary process by which the evolution of grades of organization occurs. Further, this process can occur relatively quickly and represents a qualitative jump in organization of either a biological or social system. Such events are usually followed by diversification in structure, although the diversification may occur sometime in the future, for example, in biological evolution the origin of the mammalian grade of organization occurred in the Triassic, however the adaptive radiation, that is the diversification of mammals, occurred in the Cenozoic, only after the extinction of (most of) the dinosaurs, which in turn swept the adaptive landscape relatively bare but available for occupation.

It is the position here that the concept of aromorphosis as it applies to social systems has significance with respect to the punctuated phases of urbanization investigated in this paper. After all, urbanization at its base depends on agricultural surplus, a surplus that allows the development of increased social structure and in turn relies ultimately on energy sources to be exploited more efficiently to support that social structure and its evolution. So, there is a positive feedback between the ability to produce food, the efficiency with which the food is produced, the release of human potential to invest in improved technology that will affect population growth and the social organization of that increased (and increasing) population, a fraction of which has been freed from, in fact divorced from, direct food production. Such changes in social structure, that is the evolution of social structure, have not occurred continuously, but exhibit punctuation, and both the details of these aromorphic effects and their mechanisms will be addressed later in this paper.

It should be re-emphasized at this point that while identifying and understanding mechanism is an important goal, the initial and primary intent of this paper is to show that the pattern of stasis and punctuation represents the mode of urban area evolution within the world system, and of course it has also been suggested that this pattern is characteristic of phenomenon within the larger context of Big History, as the concept of punctuated equilibrium certainly occurs well before the advent of cities or even of the genus *Homo*. However and again, the primary focus of this paper is to show that this punctuated pattern is very much a part of the pattern of urbanization over the last five thousand years.

In order to demonstrate the evidence for punctuation and stasis in the evolution of urbanization within a world system context, modified data for maximum urban area will be used, and in graphical presentation it will be shown that there are three periods of rapid change which are positioned between three periods of stasis. It will further be shown that the three periods of rapid change differ from the periods of (relative) stasis not only in rate of change but also exhibit other characteristics which distinguish them. It will also be shown that associated with the initiation of these periods of rapid change is a threshold ratio value characterizing the level of urbanization at the initiation point of rapid change. Beyond these indicators of change, it will be shown that the nature of the cue to begin periods of rapid change is the existence of a tipping point in the early stages of each of the three periods of punctuation, that is the periods of rapid change, although at present it can only be casually inferred that the tipping points involved are indirect rather than direct. An inference will then be drawn that the change in tempo and mode of urbanization involves an interplay between rural, or at least less urban, and (more) urban populations. Finally, the rudiments of a model mechanism for the alternation between relative stasis and punctuation will be presented, the one which it is hoped will be understood to be at the level of a modest proposal only.

It is important to note here that three papers were previously published, Grinin and Korotayev (2006), Korotayev (2006), and Korotayev and Grinin (2006), each also with quantitative data exhibiting punctuated change. If, for example, one views Diagrams 7, 9, and 10 of Korotayev (2006), Diagrams 1, 3, and 6 of Grinin and Korotayev (2006), or Diagrams 1, 3, 4, 6, 7, and 10 of Korotayev and Grinin (2006) as a representative sample, it can be seen that change over time of the dependent variable in each graph, for example, the number of developed states, world urban population, megaurbanization, particularly for cities with a population larger than 10,000, and territory controlled by developed states, is not monotonic, in other words not smooth and continuous, but is disjunct with periods of (relatively) rapid change punctuating periods of fluctuation about some mean, all indicative of a pattern of punctuated equilibrium. A more detailed discussion of these findings will be undertaken in the forthcoming paper.

Materials and Methods

The data used in this study came from three sources, Chandler's *Four Thousand Years of Urban Growth: An Historical Census*, *World Cities* by George Modelski, and the U.S. Census Bureau section on historical demography and are the same sources used in previous investigations (Harper 2010 and others). Both global and maximum urban area population data were natural log transformed to allow graphical comparison over orders of magnitude. The specific data used to demonstrate punctuation and stasis were the continuous accumulation (= summation) of differences between consecutive values of maximum urban area over the last five thousand years of world system history. All graphical representation was done using SciDAVis software, and all regressions were done on a TI-84+ calculator. Further, it is a fundamental assumption of this research that all urban areas exhibit a Pareto/Zipf distribution.

In order to show the distinction between periods of punctuated change and those of relative stasis, it will be necessary first to compare the overall trends in both total world system population and maximum urban area size. Once this has been done, punctuation and stasis will be demonstrated, and it will be shown that these periods quantitatively differ in a disjunct manner. In turn, it will be shown that the periods of quantitative disjunction reveal the presence of unique events associated with them, and in fact these change-motivating events that may be termed tipping points, as they meet a condition associated with tipping points, that of distinct changes in diversity (Lamberson and Page 2012).

Finally, whenever research attempts to penetrate the depths of past time, there is always the criticism that the forces at play then were different than those at play now, permanently obscuring clear knowledge of the past. In fact, beginning with James Hutton and Charles Lyell, this criticism is the basis for the countering notion of uniformitarianism. Historical urban scaling studies (Ortman *et al.* 2013, 2014) have shown that the scaled relationships determined for contemporary urban areas also hold for past urban areas. This note is mentioned to underpin the fact that the methods used are applicable over the depth of time studied here.

Results

In Fig. 1a to follow the natural log-transformed data for both maximum urban area and the total world system are plotted against time, and it can be seen that there is general agreement in shape between the two plots.

Specifically, an almost linear plot for total population (top curve) and a less linear but similar trend for the maximum urban area magnitude parallel each other over time up until the beginning of the 19^{th} century, point 48 on the x-axis, at which point a steep increase in slope is apparent, being due of course to the early stages of the last phase of the Industrial Revolution (Grinin and Korotayev 2015). At a macro-level these plots are effectively tandem. In further support of this tandemness, when the upper plot, lnT over time, is plotted against the lower plot,

$\ln C_{max}$ over time, (see Fig. 1b) and a linear regression is generated for the plot, the relationship shown is essentially linear with $R^2 = .9179$, the implication of this second graph being that there is a high correlation between the two plots in Fig. 1a. Since the slope of the linear regression in Fig. 1b is slightly greater than one, the rate of growth of urbanization, as characterized by the magnitude of population for maximum urban areas, with the assumption that urban areas of varying (slightly) greater than that of the world system population as a whole.

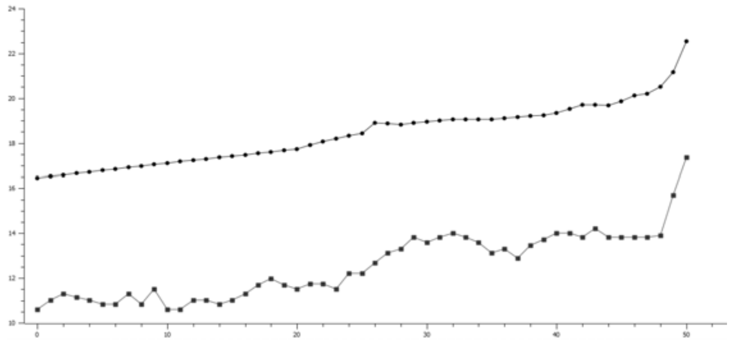


Fig. 1a. The x-axis represents time in centuries, and the y-axis represents the natural log-transformed values of both the total world system population, top graph, and the maximum urban area values over the last five thousand years

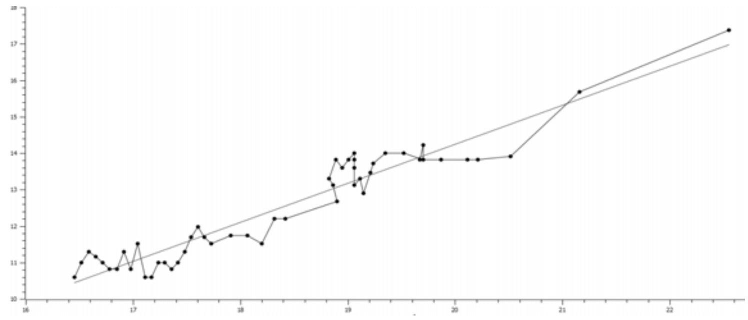


Fig. 1b. The x-axis represents lnT, where T = total population of the world system for a given century, and the y-axis represents $\ln C_{max}$, where C_{max} = population of the maximum urban area for a given century. The linear regression for this plot is: $\ln C_{max} = 1.0700 + 7.1594 \ln T$, $R^2 = .9179$

But there is also evidence for non-linear and non-monotonic behavior of the system as a whole in this second graph, as on closer inspection, the plot of maximum urban area values certainly does exhibit a general tendency of increase but does so with significant and abrupt changes in slope that are both positive and negative. The upper curve does reveal two events of abrupt positive change in slope, one at point 26, that is 400 BCE, for one century which is then followed by a lower slope than the data prior to 400 BCE, and the second at point 48, a population trend that we are currently in and the one aptly described by both Von Forester *et al*. (1960) and Korotayev *et al*. (2006a). However, the maximum urban area values show more frequent changes, three subsets of which are continuously positive, the first from 1600 BCE to 1200 BCE, the second from 700 BCE to 100 BCE, and the third, which we are, again, currently experiencing, from 1800 CE through to the present (and most probably beyond). Each of these periods of rapid, positive change punctuate periods in which maximum urban area values fluctuate about some mean. It will be shown that these fluctuations appear to have both maximum and minimum values and that each succeeding period of fluctuation does so at a higher average level of maximum urban area population. These periods of fluctuation are analogous to periods of stasis a la punctuated equilibrium in that until conditions occur for rapid increase in maximum urban area values, that is punctuation of the equilibrium, any significant increase in maximum urban area value is relatively shortly followed by a decrease in that value. In other words, the fluctuations in maximum urban area size do so around a mean.

These changes in pattern from ones of fluctuation about some mean and with maximum and minimum limits alternating with continuous and rapid increase are more clearly represented by inspecting the pattern of summed differences in the natural log values of maximum urban area populations. The individual differences in $\ln C_{max}$ from century to century are represented in Fig. 2a, and it can be seen that the values of these differences simply fluctuate about an average value which is above zero but show no indication of alternating periods of stasis and punctuation. Fig. 2b represents the plot of summed differences.

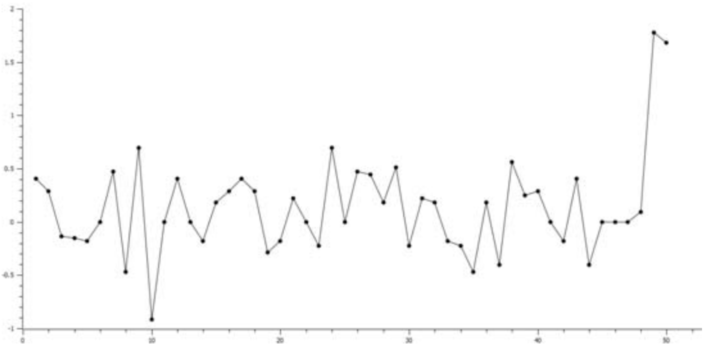


Fig. 2a. The x-axis represents time in centuries, and the y-axis represents the differences in maximum urban area values from consecutive centuries. Note that there is no distinction between periods of stasis and periods of punctuation

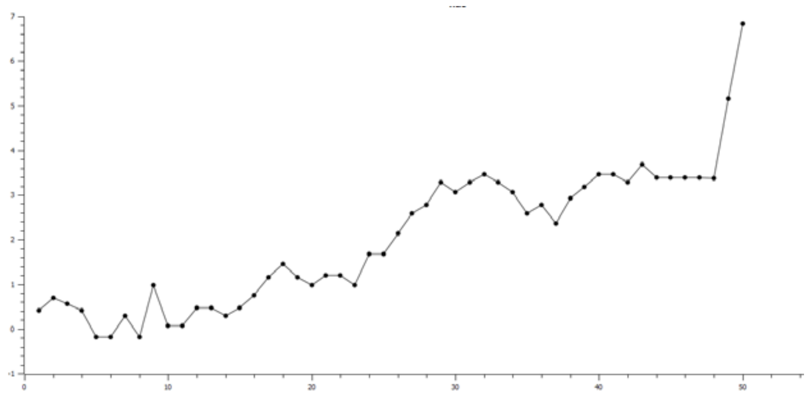


Fig. 2b. The x-axis represents time in centuries, and the y-axis represents the summation of the differences in maximum urban area values from consecutive centuries, *i.e.* $\sum\Delta\ln C_{max}$/century. The three periods of rapid change, from 1600 BCE to 1200 BCE, from 700 BCE to 100 BCE, and from 1800 CE through the present, are more prominently represented as are the alternate periods of stasis

If in turn, these three periods of rapid change and their preceding stasis are represented individually, the change from consecutive phases can be seen even more definitively. In Fig. 3 the period of stasis followed by rapid change are clearly displayed, and linear regressions of each are given. It should be noted that both the slopes and R^2 values are distinctly different for each. In the prior period of stasis the slope is essentially zero, and the R^2 value is extremely low, where as in the following period of punctuation the slope is .3020 and R^2 is .9853, that is that approximately 98 % of the variation in the distribution of points is represented by the regression, while the regression for the period of stasis represents almost none of the variation. Considering the next set of stasis and punctuation periods, *i.e.*, is from 1200 BCE to 700 BCE and from 700 BCE to 100 BCE, the same basic pattern is revealed. These data are represented in Fig. 4. Again, as in Fig. 3 both the slopes and R^2 values are characteristic of each of the consecutive periods of stasis followed by punctuation. Also, the Y-intercept value for this period of stasis is greater than the previous period, that is 2.3378 when compared to .3441, representing a higher plateau, a higher average value for the maximum urban area as a consequence of the previous period of punctuation. The final set of stasis and punctuation periods are represented in Fig. 5 below. As in the previous two sets of stasis followed by punctuation, both the slopes and values for variation are characteristic of their respective time periods, and, further, the Y-intercept value is yet again higher than the previous

Y-intercept value, as would be expected due to the period of punctuated growth separating the two periods of stasis.

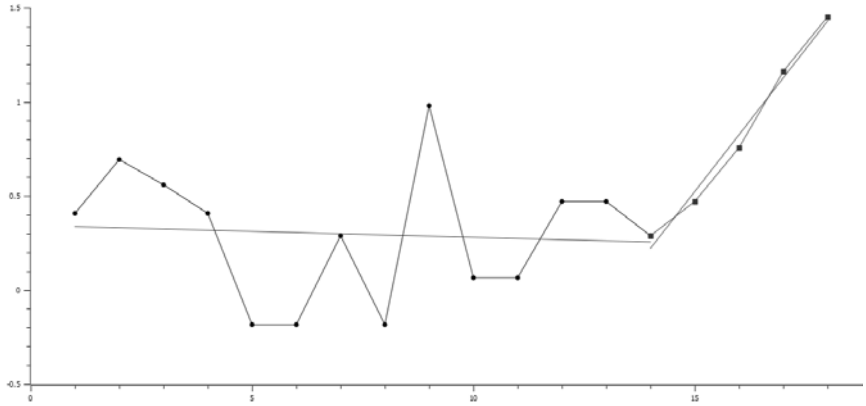


Fig. 3. Axes as in Fig. 2. The linear regressions for the periods, 2900 BCE to 1600 BCE and 1600 BCE to 1200 BCE are respectively: Y = –.0064X + .3441, R^2 = .0060, and Y = = .3040X – 4.0055, R^2 = .9853

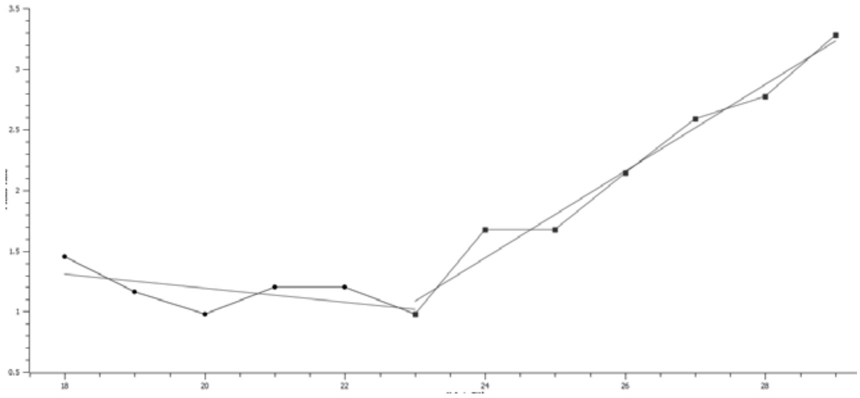


Fig. 4. Axes as in Fig. 2. The linear regressions for the periods 1200 BCE to 700 BCE and from 700 BCE to 100 BCE are respectively: Y = –.0573X + 2.3378, R2 = .3759, and Y = = .3579X – 7.1455, R2 = .9732

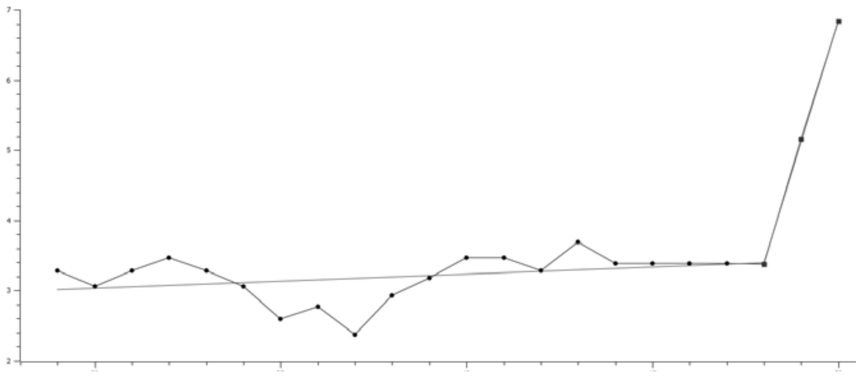


Fig. 5. Axes as in Fig. 2. The linear regressions for periods 100 BCE to 1800 CE and from 1800 CE to 2000 CE are respectively: Y = .0198X + 2.4411, R2 = .1302, and Y = 1.7301X – – 79.6479, R2 = .9998

As the previous results imply, there must be some change in the mode of urbanization and the relationship of urbanization to its reverse, this reverse process being labeled ruralization. One line of investigation which will be followed here will be to search for evidence of tipping points, specifically, and these tipping points first occurring from relative stasis to a punctuated mode of urbanization, and then, of course, whether the tips also exist from a mode of punctuation returning to relative stasis. A tipping point will be understood to be a response to continuous change of a given variable in a discontinuous way, and the evidence for such changes will be a disjunct distribution in diversity on either side of the tip as calculated by: $D = \sum 1/p^2$, where *p* represents the frequency of some class of urbanization.

With the previous description of the definition of D in mind, please consider Fig. 6, which represents the distribution of D over the last five thousand years. It is apparent that over the period of time represented that both the magnitude of D and the magnitude of changes in the magnitude of D decrease. Further, there appear to be a number of instances in which D changes precipitously, specifically from 1600 BCE to 1200 BCE almost continuously, from 700 BCE to 600 BCE, and from 1800 CE to 1900 CE. While the change in D from 800 CE to 900 CE appears to represent a significant change visually, it is clear that this peak does not represent a tipping point in which a stasis mode changes to one of punctuation, as it does not occur within the bounds of one of the three periods of punctuation identified previously.

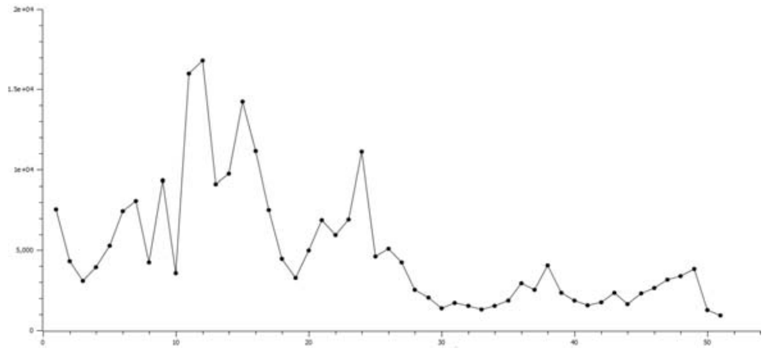


Fig. 6. The X-axis represents time, and the Y-axis represents D. D was computed over a range of values which were fractions of the maximum urban area population for a given century. The graph implies that both the magnitude of changes in D and the absolute values of D decrease with time

Fig. 7 immediately below represents the period of time from 1900 BCE to 1 CE which also includes the period of punctuation from 700 BCE to 100 BCE which marks the shift from the preceding period of stasis, 1200 BCE to 100 BCE, to the following period of stasis, 100 BCE to 1800 BCE. This graphical morphology is consistent with the occurrence of a tipping point. Further, the linear regressions for the three respective time periods have slopes of 1263.1728, –6502.259, and –772.6426. There are two other events meeting the criteria of tipping which are represented in Figs 8 and 9 to follow. Note that Fig. 10 which is a modified form of Fig. 9 is included for completeness, that is to show that the period of 400,000 years following the impact of the 14^{th} century, effectively a reset period where the maximum urban area size is approximately one million, emphasizes the change in the magnitude of D with respect to the succeeding tip period. In both Figs 9 and 10 the pattern of regression is as in Fig. 8 with a positive slope for the period preliminary to the tip period, the very negative slope of the tip period itself, and then a less negative slope for the period succeeding the tip period.

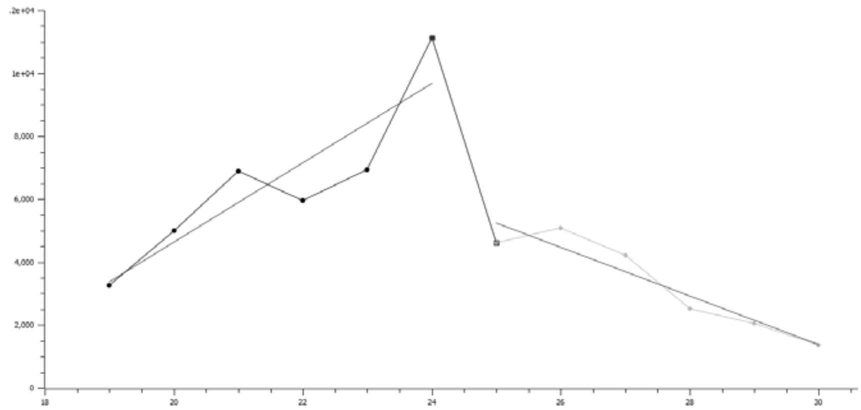


Fig. 7. The X-axis represents time from 1200 BCE to 100 BCE, and the Y-axis represents D. Note that the sequence of regressions from left to right in the graph is respectively: Y = = 1263.4141X – 20,633.4141, R^2 = .8033, Y = −6502.259X + + 167,174.305, R^2 = 1, and −772.6426X + 24,558.0734 R^2 = = .8963, a sequence which would be expected to span a tipping event

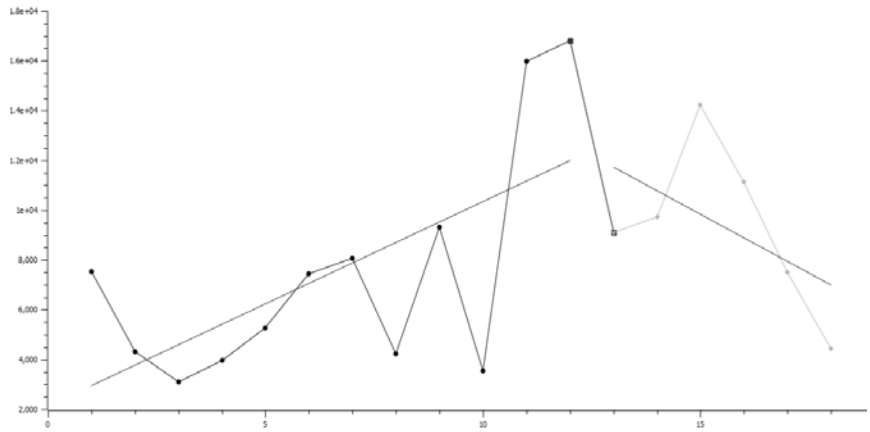


Fig. 8. Axes as in Fig. 7 but over the time span, 3000 BCE to 1200 BCE. Y = 820.5444X + 2134.3555 R^2 = .4088, −7700.452X + + 109,217.86, R^2 = 1, and −946.5634X + 24033.7208 R^2 = = .2869

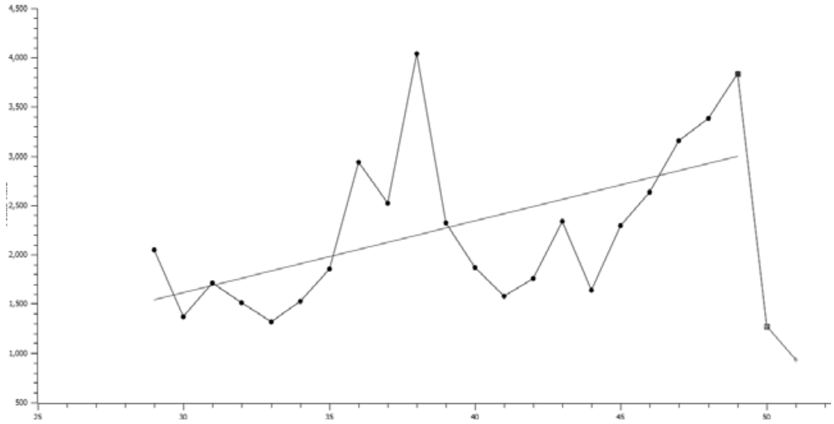


Fig. 9. Axes as in Fig. 8 but over the time span, 100 BCE to 2000 CE. The linear regressions for the periods, 100 CE to 1800 CE, 1800 CE to 1900 CE, and 1900 CE to 2000 CE are respectively: Y = 72.9020X – 574.5061 R^2 = .3163, Y = –2563.411X – – 129,446.947, R^2 = 1, and Y = –347.759X + 18,644.347, R^2 = 1

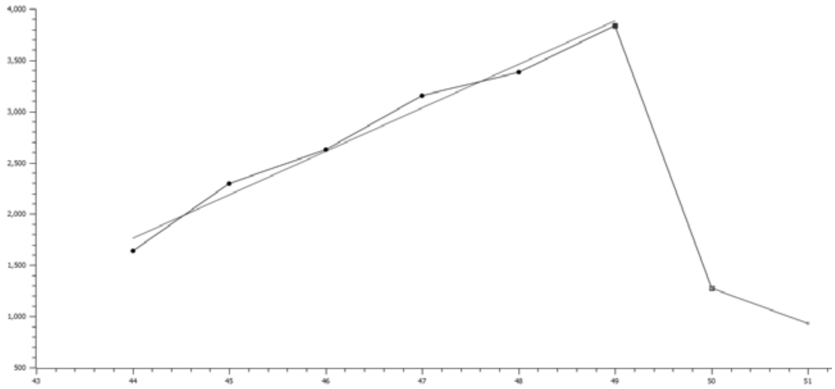


Fig. 10. Axes as in Fig. 7 but over the time span, 1400 CE to 2000 CE. The linear regressions for succeeding periods, 1400 CE to 1800 CE, 1800 CE to 1900 CE, and 1900 CE to 2000 CE are respectively: Y = 423.1254X – 16,852.8231 R^2 = .9844, Y = –2563.411X – 129,446.947 R^2 = 1, and Y = –347.759X + 18,644.347 R^2 = 1

One further point, in Fig. 11 to follow a graph of the ratio of maximum urban area magnitude to total world system population is given. Note that the four lowest points, all equal to or less than .0014, a value that appears to be a threshold point for system tipping. Three of these minimum values coincide with the beginning of each of the three periods of punctuation. The significance of the first value, the one not coincident with the beginning of a phase of punctuation, will be treated in the discussion.

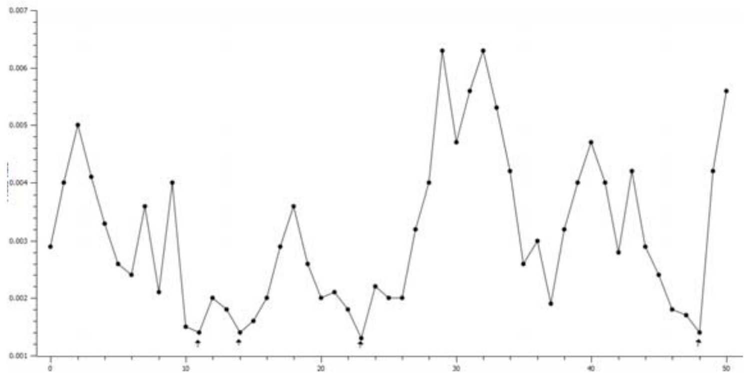


Fig. 11. The X-axis represents time in intervals of one hundred years and begins at 3000 BCE. The Y-axis represents the ration, C_{max}/T, where T = total population. Note the four arrows representing minimum values for this ratio. The timing of three of these low values coincides with the beginning of a phase of punctuation

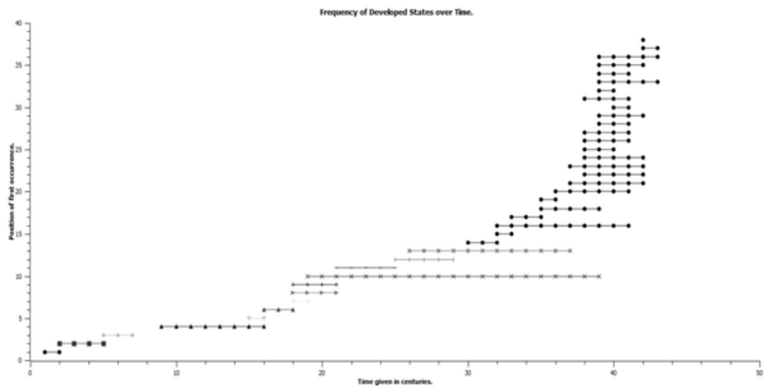


Fig. 12. The x-axis represents time in century increments beginning with Ur III, 2111 BCE and extending to 1900 CE. The y-axis represents the chronological order of first occurrence of the developed states represented in Table 1 of Grinin and Korotayev (2006)

Fig. 12 is introduced in concert with Figs 5, 9, and 10. Grinin and Korotayev (2006) suggest that developed states could not arise until a significant number of early states had formed and analogously the same holds true for the origin of mature states, *i.e.* that the existence of a significant number of developed states would be required prior to the advent of mature statehood. It can be seen that the frequency of developed states increases dramatically in the last seven centuries. This is illustrated more clearly in Figs 13a and 13b. In Fig. 13a there is a clear increase in the frequency of existent developed states from 1300 CE (note the arrow at 1300 CE) to 1800 CE with a significant decrease in the frequency of developed states in 1900 CE. If the accumulation of developed state-centuries is graphed against time as in Fig. 13b, the graph appears sigmoid and the 14th century, again noted by an arrow, is at the beginning of this accelerated phase. All three of these graphs, Figs 12, 13a, and 13b, indicate a marked increase in the frequency of extant developed states and at a time over which the Industrial Revolution occurred, a time leading to and overlapping with the origin of mature states.

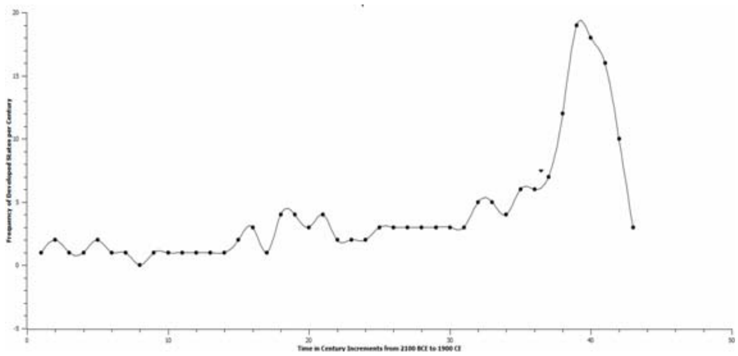


Fig. 13a. The occurrence of developed states over the last 4,100 years

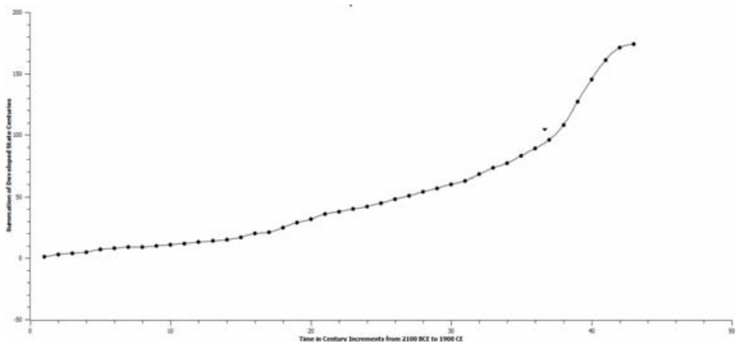


Fig. 13b. The accumulation of developed states centuries over the last 4,100 years

A Preliminary Math Model of Punctuated Growth

An intuitively constructed mathematical model of punctuated growth is presented here, a model that is only a step above a reasonably well organized verbal metaphor, but such mental constructs do have their utility, as I hope will be seen shortly. Over 50 years ago Richard Levins (1968) suggested that there are three conditions affecting the functional nature of math models: generality, precision, and reality, only two of which would be operational at any given time for a specific model. The model presented here is general and precise; however, its ability to reflect the reality of punctuated growth for any given system is limited, and in fact may well be non-existent. The precision of this model is in my opinion only significant in that explicit values of a variety of variable can be predicted, but as was mentioned, those values do not have reality. Even so, having a general model has its heuristic value. One more comment here, a caveat, this model was specifically constructed to investigate the interplay between population, carrying capacity, and technology and makes no distinction between the world system population as a whole and the ever growing segment of that population dwelling in urban environments.

The model will be described first, data generated by the model will be presented, and then an explanation of the significance of the model will follow in the Discussion. To follow are the four differential equations of the model with supplementary equations and definitions of all constants:

$$dN_1/dt = r_1N_1[K - (N_1 + N_2)],\ dN_2/dt = r_2N_2[K - (N_1 + N_2)],\ dK/dt = [T - (N_1 + N_2)]/K,\ dT/dt = T/[K - (N_1 + N_2)],$$

where N = population size, K = carrying capacity, $r_1 = r_{01} + aN_2$, $r_2 = r_{02} + bN_1$, T = level of technology, and a and b as well as r_{01} and r_{02} are fitted constants. Initial and constant values: $N_1 = 1.5$, $N_2 = 1$, $K = 5$, $T = 2$, $a = .016$, $b = .011$, $r_{01} = .015$, $r_{02} = .02$, $dt = .1$.

Even at this level of relative model simplicity, the math modelling software, STELLA, had to be used to generate the following graphs. The model includes two interacting populations that exhibit positive feedback. This portion of the model was presented in Harper (2016) and is meant to show that population interaction of this nature leads to the form of growth identified by Korotayev *et al.* (2006b) as hyperbolic. This effect is produced by the structure of the rates of growth of both populations, that is $r_1 = r_{01} + aN_2$ and $r_2 = r_{02} + bN_1$. These equations represent the condition that while each population has its base rate, either r_{01} or r_{02}, the growth rate of each is enhanced by and dependent upon the other population. This dependency is represented by the terms, aN_2 and bN_1, where the constants, a and b, are greater than zero in value and represent the contribution of each population to the other's growth rate. However, complex systems, and the world system is nothing if it is not complex, have their limits.

Symbolic representation of limited population growth was first developed by Verhulst, and the same form of that representation, $K - N$, is used here. Specifically, the hyperbolic nature of these first two equations is limited by the term, $K - (N_1 + N_2)$, implying that the two interacting populations occupy and exploit the same environment. Also however, the two populations can improve their environments, thus increasing the carrying capacity of the environment they both share. This environmental improvement is brought about by the use of appropriate technology. Consequently, that environment is subject to change, and in this model is represented by the third differential equation, $dK/dt = [T - (N_1 + N_2)]/K$, in which it can be seen that as the total population grows with respect to a given level of technology, the rate of change of the environmental carrying capacity decreases. As a result, the fourth equation in the model supplies a simple model of technological change in relationship to both population growth and change in carrying capacity. That equation is: $dT/dt = T/[K - (N_1 + N_2)]$, and its structure is such that as the carrying capacity of the total population is approached, technological improvement, here represented simply by the quantitative change of the variable, *T*, is increased. The results of a simulation with the constant and initial values given above, is shown in the following graph which is a result of the simulation running over 27 iterations:

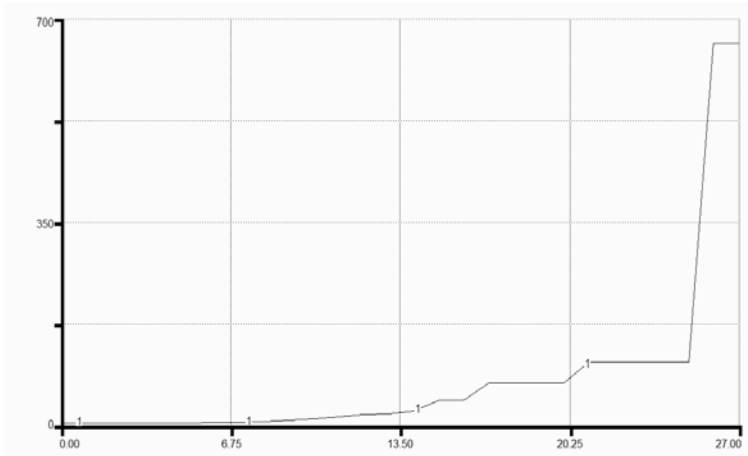


Fig. 14. The X-axis represents time in undefined units, and the Y-axis represents the magnitude of the total populations of both N_1 and N_2. The stepwise nature of the graph is clearly evident

It is evident that this graph, while representing population increase over time, does not represent continuous population growth but is characteristic of the structure of the graphs in Figs 2b, 3, 4, and 5. If the growths of both the carrying capacity and technology are represented over the same periods of time, the result appears as follows:

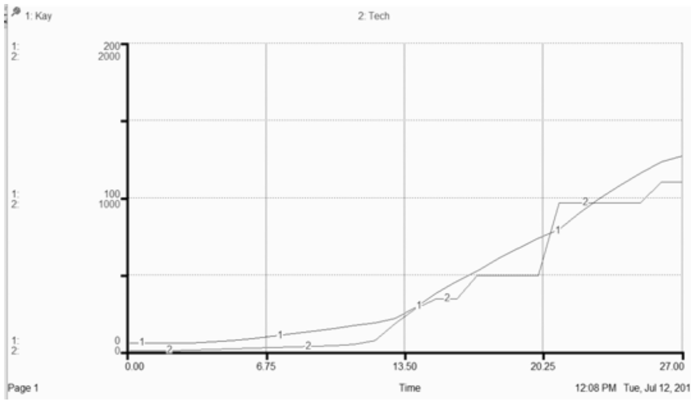


Fig. 15. The X-axis represents time in unspecified units, and the Y-axis represents the magnitudes of both carrying capacity, Kay, and technology, again in unspecified units

It can be seen that while the growth of technology is punctuated, the growth of the system carrying capacity is smooth(er) but with changing slopes indicating changing rates, but not of a punctuated nature. But if population size and the magnitude of technology are graphed as in Fig. 16 it can be seen that both seem to change in a similar fashion, although population change in this graph does not follow change in technology as suggested by Grinin and Korotayev (2006). This may be due to the resolution of the model, where *dt* = 1. The nature of these three graphs and their significance will be treated in the Discussion.

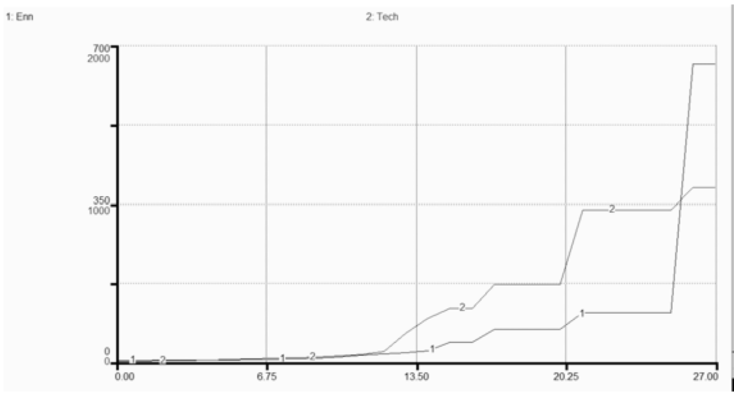


Fig. 16. The X-axis represents time in unspecified units, and the Y-axis represents the magnitudes of both technological developments and total population size, also in unspecified units. The designation, Enn, represents (urban) population size

The nature of these three graphs and their significance will be treated in the Discussion.

Discussion

Preliminary to the discussion of the data introduced here, an elaboration of the previously introduced phenomenon of aromorphosis is necessary, as this is, as punctuated equilibrium is, one more model borrowed from the biological sciences to elucidate the mechanism of given social science processes. Specifically, aromorphosis provides as model for change in level of organization of, in the case of this paper, a form of social order, that of urbanization. The substance of what follows is taken from Grinin, Markov and Korotayev (2009).

If one considers Figs 2b, 3, 4, and 5, the punctuated phases of change associated with the growth of urbanization are clearly represented, and in Fig. 2b specifically clear changes in the order of magnitude of maximum urban area can be seen, for example from populations of 100,000 to 1,000,000 during the period from 2100 BCE to 100 BCE, some 2,000 years but in which the actual changes in order of magnitude occur over much shorter times, approximately half that time, specifically from 1600 BCE to 1200 BCE and from 700 BCE to 100 BCE. It is these periods of relatively rapid change for which aromorphosis is being invoked.

Grinin, Markov, and Korotayev suggest rules for the occurrence of aromorphosis. Specifically, they suggest four necessary and sufficient conditions in order for aromorphosis to occur. First, the system about to undergo aromorphosis must exhibit enough diversity. This is required for the establishment of positive feedback between different segments of the system. [Note here that positive feedback underlays the rapidity with which an aromorphic change may occur.] In any evolutionary process for change to occur selection must occur, and for selection to occur there must be something to select. With regard to the actual process, block assemblage is one way in which an aromorphic jump may occur, however, there must be a diversity of blocks from which to choose. This is their second rule. The third rule suggests that there is a direct connection, a direct proportional relationship between diversity and innovation. The fourth rule requires that there be efficient selection mechanisms leading to intergroup competition.

There are a variety of consequences to be noted for aromorphic processes. First, since selection occurs, there will be loss; selection always incurs a cost. Second, aromorphic change permits new adaptive zones to be exploited. Just as the evolution of the mammalian level or organization opened a variety of new adaptive zones and in a sometimes delayed fashion, that is mammals originating in the Triassic but only beginning their radiation in the early Cenozoic, for example, shrews, primates, hominins, the consequences of social aromorphosis may be delayed. Grinin, Markov, and Korotayev give the example of iron metallurgy originating at a period of time well before its spread after the (so-called) Bronze Age Collapse. Third, aromorphosis enhances system resilience, since multiple functions can occur at the newly evolved level of organization that could not exist previously. Fourth, as just alluded to, multiple (new) functions imply increased complexity, and, fifth, with this newly acquired complexity new evolutionary changes can occur. Finally, aromorphosis permits previous limitations to the system to be exceeded. In fact, this is at the basis for the evolution of a new and higher level of organization to begin with.

The data of this study reveal two broad conclusions, that the summed pattern of change in maximum urban area population is similar to the pattern of punctuated equilibrium for speciation proposed by Eldridge and Gould and that the stasis phases are relatively different from the phases that represent punctuation. In fact, the pattern is more broadly speaking one of stasis punctuated by phase changes or phase transitions to further states of stasis. There are then at least three issues to be addressed, first, to confirm the punctuated nature of the evolution of maximum urban areas and its implication for the pattern of urbanization in general, second, to identify cues for the change from stasis to punctuation and the return to stasis, and, third, to suggest some form of mechanism for these changes.

The pattern of punctuated growth in urbanization has been alluded to by the following authors, Korotayev (2006), Grinin and Korotayev (2006), and Korotayev and Grinin (2006) and a mechanism mirroring punctuated growth has been proposed by Grinin, Markov, and Korotayev (2009). Regarding the dynamics of urbanization over time Korotayev has shown that while the overall pattern of urbanization is quadratic hyperbolic, because of the multiple agents operating within the world system over time, but largely mediated by feedback between population growth and the growth of technology intertwined with the accumulation of per capita surplus, the pattern of urbanization over time is not monotonic, that is smooth and continuous, but rather consists of periods of (relative) stasis punctuated by periods of more rapid growth. Review of Diagrams 7, 8, 9, 10, and 11 in Korotayev (2006) will confirm this. Also, in this paper Korotayev divides the periods of urbanization into sets A, one through three, and B, one through three, corresponding to periods of rapid growth, the A set, and cyclical change, the B set.

In tandem with this paper are two papers by Grinin and Korotayev (2006) and Korotayev and Grinin (2006) in which the specifics of political development both from the perspective of a quantitative analysis, the former reference, and a comparative analysis, the latter analysis, are given. In the first reference it is shown that the evolution of early, developed, and mature states corresponds to the periods of relative stasis, with the periods of punctuation being inserted in between each of these periods. In particular, the authors suggest that developed states are not possible until a significant number of early states exist, with the same holding for the existence of mature states, the implication that urbanization itself created the environment for further urbanization. It is interesting that this verbal model suggests a threshold or tipping point has to be breached before the next level of both urbanization and state evolution could occur, also supporting the position of this paper. However, as will be seen further on these tipping points can be exceeded in differing fashions depending on the context of the level of development (and complexity) of the world system. As will be seen shortly, this has particular pertinence with respect to the degree of fit to the linear regressions applied to both punctuated and especially stasis phases of urbanization.

As was noted in the initial portion of the Results section, in broad terms, the pattern of both population increase and maximum urban area increase is quite similar, and a very good linear fit of one set of log-transformed data against the other occurs with R^2 = .9179. However, as was also mentioned previously, on only slightly closer inspection, the pattern of urban area increase is quite different in detail. Specifically, there are three periods of stasis punctuated by three periods of rapid change, and these periods of rapid change are qualitatively different not only with respect to the rate of continuous change noted but also with respect to the variability of these periods of punctuation. These periods of punctuation are from 1600 BCE to 1200 BCE, from 700 BCE to 100 BCE, and from 1800 CE to 2000 CE, and linear regressions of the data of the three phases explain respectively 98.58 %, 97.32 %, and 99.98 % of the variability exhibited in each set of data. With respect to the three periods of stasis, from 2900 BCE to 1600 BCE, from 1200 BCE to 700 BCE, and from 100 BCE to 1800 CE, and their respective linear regressions explain respectively, .6 %, 37.59 %, and 13.02 % of the variability. Clearly, during periods of stasis there is much greater variability, a fact that seems incongruous with respect to the term, stasis. The point of course here is that these respective phases of the summed values of log-transformed values of maximum urban area magnitude are different both quantitatively and qualitatively. With regard then to the functional classification of state levels of evolution described by Grinin and Korotayev (2006), phases of stasis represent continuous adaptation of the world system at any given level of state organization, a process that will lead ultimately to movement to the next higher level of state organization. If these phase changes from punctuation to stasis and from stasis to punctuation are considered at a slightly finer scale, it can be inferred that interaction between differing segments of the total world system population are associated with these phase changes.

Regarding the surficial contradiction of the nature of the periods of stasis with respect to their variability, one has only to consider the slopes of their linear regressions to see that the variability represents oscillations about a mean for periods of stasis. Specifically, the slopes of the periods of stasis are all close to zero or horizontal, that is –.0064, –.0573, and .0198. This plateau-like nature of these periods has been noted before (Harper 2010; Korotayev 2006; Grinin and Korotayev 2006; Korotayev and Grinin 2006), and in slightly different form have been noted by Morris (2013). He specifically suggests that there have been plateaus with respect to maximum urban area magnitude of ten thousand, one hundred thousand, and one million, and twenty-five million, and while I would use slightly different values of forty-thousand, one hundred thousand, one million, and most recently something above thirty-five million, given the margin of error in the proxy data used, the notion of stabilizing values of maximum urban area magnitude existing over periods of time is in agreement with the data at hand.

It is interesting that the interaction between rural and urban populations seems to play a major role in the establishment of periods of stasis and punctuation in the macropattern of urbanization. During periods of stasis the average maximum size of urban areas remains static but gamma, basically the variable indicating the distribution of urban areas as defined by a Zipf-Pareto distribution, increases implying that greater urbanization is occurring at urban area sizes below that of maximum. In other words, as the total population of the world system increases, even though the top end of the distribution of urban areas remains relatively constant, the total urbanized population can potentially increase due to the expanded sizes of sub-maximal urban areas. If one calculates the size of sub-maximal urban area populations while holding the maximum urban area population constant and allowing gamma to increase, the sizes themselves increase, and this is in fact the pattern seen from the 14th century to the beginning of the 19th century. During this period of time maximum urban area size remained essentially constant at approximately one million or slightly more, and yet the total population of the world system increased from 350 million to 813 million and gamma increased from 1.3483 to 1.4112, a clear indication of increasing urban area size below maximum. Further, the increase in population was accommodated in a predictable way a la a Zipf-Pareto distribution. This effect is alluded to in both Diagrams 1 and 5 of Korotayev and Grinin (2006) in which the territory of developed states increased at a greater rate than did the urbanized population of the world system over essentially that same period of time. What is also implied here is that movement from the rural population to urban areas, more so than the birth rate endemic to urban areas simply because the death rate characteristic of urban areas is higher than that of rural areas and also very often is higher than urban birth rates, is responsible for this growth in urban population. Further, with regard to the level of evolved state, early, developed, or mature, the change from a lower level of state to a higher level requires that urbanization precedes the change. In fact, it is this increase in urbanization, especially at maximum size that contributes to the change from early to developed state and then from developed to mature state. However, phases of stasis may well involve increase in urbanization at lower levels of urban area size. Increase in urbanization does not always occur simply by dint of movement.

As previously implied this is unquestionably not the end of the process, as technology plays a significant role in the evolution of both urbanization and level of state development. Korotayev and Grinin (2006) make the point that technology leads the process of urbanization and improvements, new evolved levels of technology, must occur prior to any increase in urbanization. They use a particularly effective example of the advent and spread of iron technology, suggested by them to initially originate with the Hittites. However, on collapse of the Bronze Age (at the eastern end of the Mediterranean) and due to the attendant collapse of extensive trade routes at that time the production of bronze implements of war and peace, that is agriculture, was not possible. Instead, iron technology was adopted due to the relatively easy access of iron. This technology required time to spread, so, since the evolution of urbanization follows the evolution of technology, urbanization at this time was delayed as was the evolution of the developed state.

A further point needs to be made here and is dependent on the work of Per Bak (1996) and Flyberg *et al.* (1993). These researchers have shown using a simple computer model of evolution that evolving systems self-organize in a punctuated fashion up to a critical point, a point which Bak labeled as the state of self-organized criticality (SOC). Further, the research of Bak, Flyberg, and others has demonstrated that the stepwise changes are due to changes in maximum fitness, with variability in the system being due to the range of fitnesses, which when minimum fitnesses are eliminated push the system to the next higher level. While I am not in the position of fitting such model expectations, the data at hand do fit the pattern of punctuated change and similarity of appearance is suggestive, that is if it looks like a duck, quacks like a duck, flies like a duck, and so forth, it is probably a duck. Unquestionably however, there is potential fallibility in this reasoning, since the duck in question could also in reality be a goose. Even so, I am invoking the reasoning of Bak *et al*., as they present a model consistent with the data at hand in that as a consequence one would expect fluctuations above (or about) some minimum or median value of stasis until a cue occurs for the system to move to the next higher level, and this is what the data here on the summed values of natural log-transformed magnitudes of maximum urban area reveal. [Parenthetically, I would like to mention that this research assumes that the world system has not reached the ultimate state of self-organized criticality but quite clearly appears to be headed in a direction that will lead at least to a level of self-organized criticality limited by the degree of urbanization and state organization, as any self-organizing system would unless of course collapse of some magnitude ensues.]

In summary, the data presented in this paper, the data presented by Morris (2013), by Grinin and Korotayev (2006), Korotayev (2006), by Bak (1996) and by Flyberg *et al*. (1993) either support the notion of stasis in the macropattern of urbanization, that is the existence of some limiting mean about which oscillation occurs, or provide an underlying mechanism for the existence of such a mean. It should be noted that saltation from one stasis mean to the next requires an increase in either energy sources available or the technology to more efficiently exploit already available resources. This in turn supports the concept presented by Grinin and Korotayev (2006) that, as was mentioned previously, technological improvement precedes the development of urbanization to the next level. It should further be noted that the process being explained fits well with the concept of aromorphosis introduced at the beginning of this discussion (Grinin, Markov, and Korotayev 2009). That the narrative of human history should no longer be considered as a sequence of individual events per se, at least at the macro-scale, and that there is a clear evidence for punctuated change in human history at least as it is reflected in the evolution of urbanization and also for the existence of logico-deductive models to explain that mode of change should now be clear. What follows, even if somewhat repetitious, supports this notion.

If Fig. 1a is inspected and compared with Fig. 2b, there is clear agreement with respect to overall form, but Fig. 2b shows clearer distinction between periods of stasis and periods of punctuation. It is for this reason that the summed values of changes in the natural log-transformed maximum urban area size were used. Also, to reaffirm, a graph of the individual values of the urban area data would not on inspection reveal any significant difference between periods of stasis and those of punctuation. Given the step-wise appearance of Fig. 2b, what is implied by its form? First, it is apparent that history at the level of maximum urban area evolution, and therefore urban area evolution in general and based on the assumption of urban areas being distributed in a Pareto-like fashion, has not simply been 'one damned thing after another' (Toynbee 1976), but that there have been distinct periods or phases of history, each representing different modes of change. Again, the mechanism for these differing modes of change fit very clearly with the concept of social aromorphosis as developed by Grinin, Markov, and Korotayev (2009) and with the very general model of punctuated change in biological evolution as presented by Flyberg, Sneppen, and Bak (1993). Although, the combinations of stasis and punctuation do repeat, so, perhaps Toynbee should have said that history 'is just one stasis phase followed by one phase of punctuation after another'. On the one hand, periods of stasis represent fluctuations about some mean maximum urban area value of change or above some minimum threshold of the same a la Bak. On the other hand, phases or periods of punctuation represent movement of the world system as represented by maximum urban area change to new and higher levels of stasis. These two distinct phases have associated with them distinctly different ranges of variation as mentioned previously, stasis with exceptionally low values and punctuation with exceptionally high values. The implication of the former, I believe, suggests some form of system searching, while the latter suggests some relatively rapid change, perhaps very closely canalized or perhaps of an autocatalytic fashion that occurs so (relatively) quickly that the system itself is responding in an exponential or hyperbolic way. A brief discussion of the following time periods, 2900 BCE to 1200 BCE, 1200 BCE to 100 BCE, and 100 BCE to 2000 CE will follow. Each of these periods represents a combination of an initial period of stasis followed by a period of punctuation.

The period from 2900 BCE to 1200 BCE is the most problematic, as it exhibits several sequential fluctuations, only one of which is identified here as associated with a period of punctuation, the last and smallest fluctuation (see Fig. 3). The R^2 value for the stasis phase of this period, .0060, is very low, confirming that the range of data, that is the fluctuation exhibited by the data, is considerable. All other fluctuations during this period are associated, I believe, with the initial establishment of empire and how that initial establishment induced its own ecological overshoot as described in the book of the same name (Catton 1983) and as a result do not fit the criteria established for a phase of punctuation. (Parenthetically, aromorphosis must play a significant role directly in the initial establishment of empires, for example, the Akkadian and Ur III empires, but this appears not to be associated with a change in the level of urbanization, as urbanization should precede such territorial extension. However, such events in relation to energy acquisition are worthy of future research.) Note, for instance, that the period from 2100 BCE to 2000 BCE represents a decline in the Akkadian Empire and is also associated with very harsh climatic changes. The actual punctuation phase begins at 1600 BCE and follows a decrease in summed change from the previous century, in fact this point at which the punctuation begins represents the regressed mean value of the previous period of stasis as can be seen from the graph. The beginning of this phase is a point shared by both initial stasis and the following phase of punctuation. Interestingly, this pattern is consistent with the change from stasis to punctuation in the remaining two time periods as well. Of note here also is the fact that the ratio of maximum urban area value over the total population of the world system is at or less than .0014, a threshold minimum that holds for similar points in the other two periods. Why this particular minimum should hold is not clear.

In the second combined period of stasis and punctuation, from 1200 BCE to 100 BCE, a change from stasis to punctuation at 700 BCE is exhibited, and this occurs a century after the (approximate) beginning of the Axial Age as established by Karl Jaspers (1953). The stasis before this point in time has R^2 = .3759, implying that the range of fluctuation was considerably lower than that of the previous stasis phase, while that of the phase of punctuation to follow is R^2 = .9732, which is consistent with the R^2 values for the other two phases of punctuation. It is also interesting here that the stasis phase is shorter than the phase of punctuation, five hundred years as opposed to six hundred years, not inconsistent with the data of Bak but intuitively surprising in light of the length of the stasis phase of the previous period investigated. This period of lengthy punctuation reaches a regressed mean of one million for the size of the maximum urban area in the next stasis phase and spans the time when the first quasi-global empires were established, the Han Empire and the Roman Empire. Trade also became essentially global at that time with the establishment of the Silk Road and the water routes of the Mediterranean Sea, Indian Ocean, and points east into Indonesia and along the eastern coast of China. As will be seen, this level of complexity required a considerable period of search time before cuing the next phase of punctuation. In more quantitative terms, the system search achieved low ratios for $C_{max}/T \leq .0014$ twice during this period of 2100 years, only one of which is associated with the phase of punctuated growth that we are now in.

The third and final period to be considered extends from 100 BCE to 2000 CE, is the longest of the three periods, and is the one we are all currently living in. The phase of stasis associated with period of twenty-one hundred years is the longest of the three phases investigated, and involves three peaks, two troughs, and a four hundred year period of actual stasis with little fluctuation in maximum urban area magnitude before entering the final period of punctuation. This period of time spans the apogees of the Han and Rome empires, the European Dark Ages, two monumentally significant plagues, Justinian's and the Black Death, the European Renaissance, the Age of Discovery, the evolution of democracies, and the list goes on and on, and with a variance value of R^2 = .1302 is clearly a period of time when the distribution of points, of the change in maximum urban area magnitude, varies considerably about a linearly regressed mean. This is then a period of system's search before entering the current period of punctuated growth and represents an extended period of time over which new technologies affecting both the discovery of new energy sources, that is the fossil fuel triad of coal, oil, and gas, were not accessed in any substantial way until the middle of the 18th century CE. In turn, the three points of the current period of punctuation extending from 1800 CE to 2000 CE and into the current time have a variance value of R^2 = .9998. This is a period of time in which the end phase of the Industrial Revolution (Grinin and Korotayev 2015) propelled the world into the Industrial Age followed by the post-Industrial Age, a time when the size of the maximum urban area increased from 1.1 million in 1800 CE to the megacity of Tokyo, which currently is approximately hosting a population of thirty-five million, and a time in which the exponent, γ, a measure of urban distribution, decreases from γ = 1.4112 to 1.2099. The implication of this decrease is that there are now far more urban areas in which our population is housed than ever before, and consequently the last two hundred years have witnessed a massive movement of humanity from a largely rural existence to a largely urban one.

As mentioned previously, this interplay between rural and urban populations, one currently shifting the balance in favor of urban occupation to the extent of 70 % of the world system population by 2050 has of course been going on since the Neolithic Revolution. What the previous portion of this discussion suggests is that new levels of urbanization have been achieved in a step-wise fashion, and that, further, this occurs against a demographic backdrop of more of less continuous increase in world system population. The relationship between the population as a whole and the process of urbanization will be addressed next.

As can clearly be seen from Fig. 1, both the total population of the world system and the magnitude of maximum urban area show similar patterns of development over the last five thousand years, however, as has been already addressed, the detail of these patterns differs with the development of total population exhibiting a smoother pattern over all, while maximum urban area magnitude both increases and decreases against the background of more or less steady population change. Further, the exponent, γ, which characterizes the pattern of urbanization per time point and is predicated on the distribution of urban areas being Pareto-Zipf-like, also changes in a fluctuating pattern, in fact the mirror image of that of urban areas. It is particularly interesting to note that per stasis phase, the total population increases, while at the end point of each phase, the final value of maximum urban area magnitude is at the regression mean. In other words, while total population increases, and one would intuitively expect that maximum urban area magnitude would also, in the final point of that phase, one shared with the succeeding phase of punctuation, the maximum urban area value is exactly at the regressed mean. However, at the same time, the value of γ is maximum for the phase. What does this imply? It implies, I believe, that the increased population is housed in a greater abundance of lower level urban areas and that populations that were essentially not urbanized at the beginning of a given phase of stasis now are.

The event of stasis is then characterized by the following process: After a period of continuous maximum urban area increase and against a background of more or less continuous total population change, over the period of time associated with stasis, the increased total population is then housed in more urban areas of smaller scale. The process of stasis, of accommodating greater total population within a framework of both urban and rural environments, is done so also against a stable maximum urban area magnitude and exhibits fluctuations about this stable maximum until the final fluctuation achieves some tipping point at which maximum urban area magnitude increases for an extended period of time. Each upward movement of the system during stasis, that is each increase in maximum urban area magnitude with its attendant change in the frequency of lesser urban areas, can be thought of as an attempt to challenge the upper limit of the system set by the maximum urban area magnitude of that stasis period.

There are also minimum sizes with respect to the regressed mean of any given stasis phase which seem to be roughly the regressed mean of the previous stasis phase. However, when the ratio of maximum urban area magnitude to total world system population is calculated, there appears to be a threshold minimum value at and below which is initiated the next phase of punctuation. That value is .0014, and this represents another way of considering the context in which a phase of punctuation originates. There are four such points as represented in Fig. 11, the last three of which correspond to the initiation of phases of punctuation. The low value point does not represent such an initiation and is probably a consequence of external factors; climatic factors are the most obvious candidate. It is proposed, then, that some sort of cue or trigger exists in which rapid, continuous urbanization follows from one stasis level to the next and as previously mentioned is controlled by a tipping point.

Tipping points are points in a process in which the process itself changes rapidly. They may be thought of as phase changes, and there are broadly two types of tipping points, direct and indirect (Lamberson and Page 2012). Direct tips are those in which a continuous increase in a given data set produces at some point a rapid increase in that same data set, while an indirect tip is one in which continuous change in one type of input causes a rapid change in a different type of output. An example of the former is, and an example of the latter is a diabetic slipping into a coma due to low blood sugar. Also, tips can be characterized by changes in either entropy or diversity before and after the point of tipping. In the case of entropy, the Shannon-Weaver index is used, $H = -\sum p \ln p$, where p is the frequency of some entity, while changes in diversity can be monitored by the formula: $D = \sum 1/p^2$. The second formula is used here.

Figs 7, 8, and 9 represent graphs of the tips associated with the phases of punctuation from 1200 BCE to 100 BCE, from 3000 BCE to 1200 BCE, and from 100 BCE to 2000 CE. These graphs were listed out of chronological order, as the graph of change from 1200 BCE to 100 BCE was thought to be the most characteristic of tipping and was therefore presented first as a model against which the other two graphs could be compared. The tipping points themselves in each of these time periods fall between 700 BCE and 600 BCE, 1800 BCE and 1700 BCE, and between 1800 CE and 1900 CE respectively. In each case the tip occurs after the actual punctuation phase begins. Why should this be?

It is proposed here that the system continually fluctuates/oscillates around a stasis level the ceiling of which is continually tested. In this context then it is easy to understand that over time the testing of the ceiling of each stasis period will produce increasing sub-phases only the last of which will induce punctuation. As a consequence of this, a tip to initiate punctuation would be expected to occur after the actual punctuation phase had begun.

Any proposal of mechanism at the present point of this research must be put in the category of speculation, and so what follows is exactly that. However, it is hoped that by presenting such a proposal thought will be stimulated to consider ways of testing the proposed mechanism or perhaps even addressing the problem, that of what causes the occurrence of tipping points and their attendant phases of punctuation, from an entirely different context. Given this preamble, what follows is a mechanism consistent with the data used.

First, it should be noted that in general terms a mechanism already exists, that of Flyberg *et al*. (1993) in which an evolving system moves to a position above a minimum selective threshold and then proceeds to test the upper boundaries of the selective environment until a new and higher minimum is achieved. The explicit problem that I see here is how this general notion can be matched to the specifics of the urban-rural dance that occurs during stasis to ultimately reach a tipping point. During phases of stasis diversity exhibits a general trend of increase, and as part of the mechanism to generate a tipping point it is proposed that this is a necessary condition. However, while increased diversity is necessary, it is not sufficient, because diversity, and here diversity is meant as a proxy for specialization, does not imply linkage within the system. It simply refers to the differentiation of parts within the system. The linkages that are necessary really imply the establishment of positive feedback which will bring about the necessary urban growth to institute a phase of punctuation. In turn, positive feedback is a consequence of establishing hypercycles, the entity proposed by Eigen and Schuster (1977) to exceed the limitations of the reproductive (and information) paradox that he discovered and now bears his name. The question then is: What process establishes (a) hypercycle(s) which in turn create(s) the positive feedback to further create a punctuated phase of urban growth? (For a simple and lucid explanation of what hypercycles are all about, see Smith and Szathmary (1995), Chapter 4 The Evolution of Templates, Section 4.5 The Hypercycle.)

To answer this question the work of Padgett *et al.* in Padgett and Powell (2012) will be cited. Specifically this work is focused on how organizations and markets emerge from a milieu, an environment, of firms, skills, and exchanges. They suggest that such emergence is autocatalytic, the result of which is the development of one or more hypercycles and give a detailed explanation of emergence and probability of survival under a variety of conditions, for example, a depauperate environment, an enriched environment, restricted emergence to a single hypercycle, and emergence involving the creation of multiple hypercycles. The model to be used here is simply a hypercycle involving two entities, a rural population with its attendant characteristics and an urban population with its own adaptive characteristics. These two components exchange goods and services and in doing so create a hypercycle that sustains both. It should be noted here that this relationship has been noted earlier (Harper 2016) and in that model the emergence of hyperbolic growth is represented. The specific assumptions and points of the current model are as follows:

1. The world system is a self-organizing system.

2. As previously asserted, the model of self-organization is that of a hypercycle.

3. Hypercycles are sets of interconnected components which feed off each other and supply each other with the essential resources for existence.

4. Hypercycles have limits with respect to their level of complexity and rate of function.

5. During periods of stasis the world system hypercycle(s) organize in order to make the next 'jump', that is to enter a phase of punctuation.

6. In order to do this, diversity of components is maximized.

7. A punctuation phase then ensues with its attendant loss of diversity and stops at the next (proximal) limit.

8. Stasis begins again and diversity expands, albeit in a nonlinear and fluctuating fashion.

This verbal model can also be complimented by very general mathematics.

A preliminary math model of punctuated growth was also presented, and the graphical data generated by this math model shows that a system operating with the explicit structure and constraints of the model exhibits punctuated growth. What is the implication of this type of growth for the world system? I believe that punctuated growth is an inherent part of the process of world system growth as reflected by urbanization and in fact at a very coarse level of simulation is due to the interactions between population, carrying capacity, and technology; the model certainly implies this. Fig. 14 shows a pattern of population growth not unlike that of world system population growth shown in any number of publications by Korotayev and his colleagues, for example, Korotayev, Malkov, and Khaltourina (2006 a, 2006b, and 2006c) and is also similar to the patterns represented in Grinin and Korotayev (2006) and Korotayev and Grinin (2006). However, what is worth of a relatively simple model, a general model, of such system behavior?

There are several benefits of such a model; three will be given here. As at its present stage of development the model only reveals a general level of similarity with the data presented here and not explicitly matched values, its overall structure then is reminiscent of the general nature of such systems. The rapid change in population size is due to the hyperbolic nature of the growth of two interacting populations within the model. This is certainly in keeping with the pattern of world system population growth over the last five thousand years, the changes in urbanization level are, however, not continuous, also in keeping with world system behavior over the historical period, as are changes in technology. Of a more dubious nature is the pattern of carrying capacity change, which is not discontinuous but continuous with changing slopes associated with the punctuated changes in both population size and the level of technology. The model then suggests a mode of further research into the (more) explicit interactions between these three variables.

System size is also revelatory. The system represented by this math model is a small one, nothing like the level of complexity of the world system even though it exhibits some of the same behavior. Perhaps what is implied then is that the real system is either heavily buffered or that at different levels of organization processes at sublevels reach thresholds, tipping points, that cause response at higher levels. The process of social aromorphosis as detailed by Grinin, Markov, and Korotayev (2009) certainly would be consonant with such changes, and specifically associated with aromorphic changes would be changes in the level of evolution of the state.

Thirdly, the graphical system behavior suggests that thresholds are breached with respect to what the carrying capacity will allow, and the system as a whole responds very quickly due to the interaction of the two population systems. Considering hypercycle formation as a motivation for aromorphic change then suggests that in the actual data for world system urbanization change the timing of such hypercycle formation must occur prior to punctuated change, and it is during these periods that evidence for such hypercycle formation should be searched for. While the periods of stasis as represented by the model do not exhibit fluctuations, it should be noted that such periods do exist over significant segments of model time, and this length of time is scaled to the punctuated change to follow.

The model in both its verbal and mathematical forms proposed above is barely a shadow of its future self, and clearly there is a sense of vast incompleteness. First, the model barely exceeds the level of a narrative description. Second, mechanisms let alone mechanisms within mechanisms have yet to be defined. For example, consider the following two questions: How is diversity actually maximized? What establishes hypercycle limits and rate of function? Third, and for the moment finally, a specific form of the simple two component hypercycle of rural-urban interaction has yet to be established, as has the analogue of Eigen's paradox with respect to the change in level of urbanization that the world system has exhibited over the last five thousand years of its history. However, even with this sense of 'vast incompleteness' identified, it is hoped that the proposed model will further thought and research on the mechanisms of punctuation and stasis that have characterized world system self-organization and therefore evolution over historic time.

Summary

1. Change in complex systems is neither simple in occurrence nor in mechanism.

2. The phenomenon of punctuated equilibrium as first proposed and described by Eldridge and Gould has broader application than simply to the process of speciation and is common to a variety of Big History scenarios.

3. While it is understood that punctuated equilibrium falls within the domain of biology and is therefore a phenomenon best studied within the domain(s) of science, it is the position of this paper that history is a science.

4. The focus of this paper is the occurrence of punctuated equilibrium in the evolution of urbanization over the course of world system history, and world system analysis gives this paper its primary context.

5. The work of Per Bak and others is also used not only to add to the context of this study but also to provide a general mechanism for the process of punctuated equilibrium.

6. Natural log-transformed data for both maximum urban area magnitude and the magnitude of total world system population constitute the raw data for this study.

7. The summed data for maximum urban area magnitude was used, as it provided clearer evidence for the distinction between stasis and punctuation than did the unsummed data.

8. It is assumed that per increment of time urban areas have a Pareto/Zipf distribution.

9. There is clear similarity in pattern between the graphed data for maximum urban area and for total population over time. However, on close inspection, the data for maximum urban area reveals three phases of stasis separated by three phases of punctuation.

10. Punctuated patterns in world system change have been recognized previously with respect to urbanization, state evolution including state territory size, and the fact that state evolution has produced three different levels of organization, early, developed, and mature states which fit nicely with the pattern of punctuations and stasis investigated here.

11. Each type of phase is characterized by its own R^2 values. Those of phases of stasis are very low, implying great variation, and those of phases of punctuation are very high, implying that the process of punctuation is very directed.

12. The research of Leonid Grinin and Andrey Korotayev with respect to both the formal and comparative quantitative analysis change of urbanization over time and the work of Andrey Korotayev on urbanization dynamics have particular significance here.

13. A simple measure of diversity was used to identify the presence of tipping points from stasis to punctuation.

14. The existence of tipping points in the evolution of world system urbanization implies that continuous change in diversity is interrupted at key points as the system switches from mode of a stasis to a mode of punctuation.

15. Tipping points occur after the apparent initiation of punctuation due to the fact that during stasis the system fluctuates in urban area distribution as represented by γ, that is the system searches for the next higher level of urbanization.

16. Minimum values for the ratio, $C_{max}/T \leq .0014$, occur four times in world system history, three of which are associated with the apparent initiation of a phase of punctuation.

17. The plateau-like nature of maximum urban area evolution has also been noted by Morris.

18. World system history at least as revealed here by the data on maximum urban area magnitude is not just an occurrence of events, one after another, that is historically contingent events with little overall context, but is on the contrary a process with structure. That structure is a consequence of phases of stasis alternating with phases of punctuation.

19. A mechanism is proposed in which it is assumed that the world system is a self-organizing system and uses as its mode of self-organization the hypercycle. Hypercycle evolution is tied to the gain (= stasis) and loss (= punctuation) of the diversity of urban area sizes.

20. The process of social aromorphosis is also evident as a contributing mechanism for punctuated change.

21. A simple math model of punctuated change was developed to show the general characteristics of that change as a result of the interaction between population, carrying capacity, and technology.

Mathematical Appendix

Determination of p-values for $D = \sum p_m^{-2}$.

Given that the distributions of urban areas can be given by the formula, $F = \alpha C^{-\gamma}$, where F = frequency, $\alpha = C^{\gamma}$, and γ = a fitted constant determined numerically from the equation, $C_{max}^{\gamma} - C_{max} - - (\gamma - 1)T = 0$ (for details of the derivation see Harper 2010) and assuming that the actual sizes for urban areas follow a distribution of 2^m, where m = the index of urbanization, so that $C_{max} = = C_{max}2^0$, and urban areas having a frequency of 2 have m = 1, that is for the magnitude of $C_1 = [C_{max}/2]^{\gamma}$, and for m = 2, $C_2 = [C_{max}/2^2]^{\gamma}$, and so forth, then to find the total urban population for each m simply multiple the previous formula by 2^m, that is $C_{max}2^{m(1 - 1/\gamma)}$. It is these values that are used for m = 0, 1, 2, and 3 that are used to determine the appropriate p-value for the computation of $D = \sum p_m^{-2}$ in the following way: 1. Compute C_m for m = 0, 1, 2, and 3. 2. Sum all values. 3. Divide each C_m by the total to compute each p-value. 4. Use the formula above to compute D.

References

Bak P. 1996. *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus.

Catton W. R. 1983. *Overshoot: The Ecological Basis of Revolutionary Change*. Chicago: University of Illinois Press.

Chandler T. 1987. *Four Thousand Years of Urban Growth: An Historical Census*. Harvard: Edwin Mellon Press.

Eigen M., and Schuster P. 1977. The Hypercycle: A Principle of Natural Self-Organization. Part A: Emergence of the Hypercycle. *Die Naturwissenschaften* 64: 541–565. URL: http://jaguar.biologie.hu-berlin.de/~wolf ram/pages/seminar_theoretische_biologie_2007/literatur/schaber/Eigen1977Naturwissenschaften64.pdf.

Eldridge N., and Gould S. J. 1972. Punctuated Equilibrium: An Alternative to Phyletic Gradualism. *Models in Paleobiology* / Ed. by T. J. M. Schopf, pp. 82–115. San Francisco: Freeman Cooper.

Frank A. G., and Gills B. 2014. *World System: Five Hundred Years or Five Thousand*. New York: Routledge.

Flyberg H., Sneppen K., and Bak P. 1993. Mean Field Theory for a Simple Model of Evolution. *Physical Review Letters* 71: 4087.

Forester H. von, Mora P. M., and Amiot L. W. 1960. Doomsday: Friday, November 13, A.D. 2026. *Science* 132: 1291–1295.

Grinin L., and Korotayev A. 2006. Political Development of the World System: A Formal Quantitative Analysis. *History and Mathematics: Historical Dynamics and Development of Complex Societies* / Ed. by P. Turchin, L. E. Grinin, V. C. de Munck, and A. V. Korotayev, pp. 115–153. Moscow: KomKniga.

Grinin L., and Korotayev A. 2015. *Great Divergence and Great Convergence: A Global Perspective*. Springer.

Grinin L. E., Markov A. V., and Korotayev A. V. 2009. Aromorphosis in Biological and Social Evolution: Some General Rules for Biological and Social Forms of Macroevolution. *Social Evolution and History* 8(2): 6–50.

Harper T. 2010. The Trajectory of the World System over the Last 5000 Years of History. *History and Mathematics: Processes and Models of Global Dynamics* / Ed. by L. Grinin, P. Herrmann, A. Korotayev, and A. Tausch, pp. 13–63. Volgograd: Uchitel.

Harper T. 2016. Punctuated Equilibrium, a Pattern Common to Big History that also Characterizes the Pattern of Urbanization. *Social Evolution & History* 16(1): 86–127.

Jaspers K. 1953. *The Origin and Goal of History.* New York: Routledge and Kegan Paul.

Korotayev A. 2006. The World System Urbanization Dynamics: A Quantitative Analysis. *History and Mathematics. Historical Dynamics and Development of Complex Societies* / Ed. by P. Turchin, L. Grinin, A. Korotayev, and V. de Munck, pp. 44–62. Moscow: KomKniga/URSS.

Korotayev A., and Grinin L. 2006. Urbanization and Political Development of the World System: A Comparative Quantitative Analysis. *History and Mathematics. Historical Dynamics and Development of Complex Societies* / Ed. by P. Turchin, L. Grinin, A. Korotayev, and V. de Munck, pp. 115–153. Moscow: KomKniga/URSS.

Korotayev A., Malkov A., and Khaltourina D. 2006a. *Introduction to Social Macrodynamics: Compact Macromodels of the World System Growth.* Moscow: Editorial URSS.

Korotayev A., Malkov A., and Khaltourina D. 2006b. *Introduction to Social Macrodynamics: Secular Cycles and Millennial Trends.* Moscow: Editorial URSS.

Korotayev A., Malkov A., and Khaltourina D. 2006c. *Introduction to Social Macrodynamics: Secular Cycles and Millennial Trends in Africa.* Moscow: Editorial URSS.

Lamberson P. J., and Page S. 2012. *Tipping Points.* URL: http://www.santafe.edu/media/workingpapers/12-02-002.pdf.

Levins R. 1968. *Evolution in Changing Environments*. Princeton: Princeton University Press.

Modelski G. 2003. *World Cities: –3000 to 2000*. Washington, D.C.: FAROS 2000.

Morris I. 2013. *The Measure of Civilization: How Social Development Decides the Fate of Nations*. Princeton: Princeton University Press.

Ortman S. G., Cabaniss A. H. F., and Bettencourt L. M. A. 2013. *Urban Scaling in Pre-Hispanic Mexico.* SFI Working Papers. 2013-01-001. URL: http://www.santafe.edu/media/workingpapers/13-01-001.pdf.

Ortman S. G., Cabaniss A. H. F., Sturm J. O., and Bettencourt L. M. A. 2014. The Pre-History of Urban Scaling. *PLOS*. URL: http://journals.plos.org/plosone/article?id=10. 1371/journal.pone.0087902.

Padgett J. F., McMahan P., and Xing Zhong 2012. *The Emergence of Organizations and Markets*. Princeton: Princeton University Press.

Scheffer M. 2009. *Critical Transitions in Nature and Society*. Princeton: Princeton University Press.

Smith J. M., and Szathmary E. 1995. *Major Transitions in Evolution*. New York: Oxford University Press.

Sole R. V. 2011. *Phase Transitions*. Princeton: Princeton University Press.

Toynbee A. 1976. *The Toynbee-Ikeda Dialogues: Man Himself Must Choose*. New York: Harper and Row.

Turchin P. 2003. *Historical Dynamics: Why States Rise and Fall*. Princeton: Princeton University Press.

U.S. Census Bureau Historical Estimates of World Population. U.S. Census Bureau. URL: https://www.census.gov/population/international/data/worldpop/table_history.php.

Contributors

Valery A. Anisimov received his PhD from the St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences. He now works at A2iA SA. His research interests include patterns recognition, image analysis, artificial intelligence, molecular biology, macroevolution and abiogenesis. He is the co-author of several articles and one book.

David C. Baker, Macquarie University, Sydney, Australia, Researcher and Tutor. PhD in Big History at Macquarie University in Sydney under David Christian. He now teaches Big History at the universities of Amsterdam, Eindhoven, and Rotterdam alongside Fred Spier and Esther Quaedackers. His research interests include Universal Darwinism, collective learning, the rise of complexity, and demographic-structural theory. He is also engaged with the IBHA Publications Committee in setting up a Big History academic journal.

Eric J. Chaisson is Director of the Wright Center for Innovative Science Education at Tufts University, where he is also Research Professor of Physics and Astronomy and Research Professor of Education. He is also an Associate of the Harvard College Observatory and serves with the Faculty of Arts and Sciences at Harvard University. Trained initially in atomic physics, Chaisson obtained his doctorate in astrophysics from Harvard University. During his tenure as Associate Professor at the Harvard-Smithsonian Center for Astrophysics, his research focused largely on the radio astronomical study of interstellar gas clouds. This work won him fellowships from the National Academy of Sciences and the Sloan Foundation, as well as Harvard's B. J. Bok Prize for original contributions to astrophysics and Harvard's Smith-Weld Prize for literary merit. Previous to his current positions, he spent several years at the (Hubble) Space Telescope Science Institute at Johns Hopkins University. He has written nearly 200 publications, most of them in the professional journals, and has authored or coauthored 12 books. Chaisson's major research interests are currently twofold: His scientific research addresses an interdisciplinary, thermodynamic study of physical and biological phenomena, thereby seeking the origin, evolution, and unification of galaxies, stars, planets, and life in the Universe. His educational research engages experienced teachers and computer animators to create better methods, technological aids, and novel curricula to enthuse teachers and instruct students in all aspects of natural science. He teaches an annual undergraduate course at Harvard University on the subject of cosmic evolution, which combines both of these research and educational goals. Chaisson holds membership in many American and international scientific organizations, several honor societies, and dozens of academic, public, and federal advisory committees.

Sergey V. Dobrolyubov graduated from Moscow Technical Institute (MADI) and began his professional career as an engineer investigating the causes of failures of complex equipment. After unexpected collapse of the Soviet Union, he began to wonder about the reasons for stability and dynamics of social structures. Due to his engineering past, he seeks more substantive descriptions of social phenomena and vague social categories. He is an independent scholar whose research interests include social evolution, theory of society, sociology, values, Marxism. His main theme is the cyclical genesis of societies and civilizations. He is a frequent author of the leading Russian sociological journals.

Tom Gehrels is Professor of Planetary Sciences and an astronomer at the University of Arizona in Tucson (USA). During World War II, he was active in the Dutch Resistance and an organizer for the Special Operations Executive (SOE). He pioneered the first photometric system for asteroids in the 1950s, as well as a system of analyzing wavelength dependence for polarization of stars and planets in the 1960s. He was also the Principal Investigator for the Imaging Photopolarimeter experiment on the Pioneer 10 and Pioneer 11 first flybys of Jupiter and Saturn in the 1970s. Dr. Gehrels initiated the Space Science Series of textbooks, for which he was general editor. He also initiated the Spacewatch Program and, until 1997, was its Principal Investigator (PI) for electronic surveying; their goal was to obtain statistics about asteroids and comets, including near-Earth asteroids. He has participated in the discovery of over 4,000 asteroids, as well as a number of comets. Besides teaching at the University of Arizona, Dr. Gehrels also lectures at the Physical Research Laboratory in Ahmedabad (India), where he is a Lifetime Fellow. This is a United Nations program for graduate students from a variety of countries, from Uzbekistan to North Korea. Presently, his research focuses on universal evolution. He was a winner of the 2007 Harold Masursky Award for his outstanding service to planetary science.

Anton L. Grinin, PhD in Biological Sciences, Leading Research Fellow of Volgograd Centre for Social Research. His main research interests include biotechnologies, global technological transformations and forecasts. He is the co-author of the monographs *From Biface to Nanorobots: The World on the Way to the Epoch of Self-Regulating Systems* (2015; 'Uchitel' Publishing House, in Russian), *The Cybernetic Revolution and the Forthcoming Epoch of Self-Regulating Systems* (2016; 'Uchitel' Publishing House, in English) and a number of articles including 'Macroevolution of Technology' and 'Global Technological Transformations'.

Leonid E. Grinin is Senior Research Professor at the Institute for Oriental Studies of the Russian Academy of Sciences in Moscow and Senior Research Professor at the Laboratory for Destabilization Risk Monitoring at the National Research University Higher School of Economics. He is the Editor-in-Chief of the journal *Age of Globalization* (in Russian), as well as a co-editor of the international journals *Social Evolution & History* and the *Journal of Globalization Studies*. Dr. Grinin is the author of over 500 scholarly publications in Russian and English, including 31 monographs. These monographs include *Philosophy, Sociology, and the Theory of History* (2007; in Russian); *Productive Forces and Historical Process* (2006, in Russian); *State and Historical Process* (3 vols, 2009–2010, in Russian); *Social Macroevolution: World System Transformations* (2009; in Russian, with A. Korotayev); *Macroevolution in Biological and Social Systems* (2008; in Russian, with A. Markov and A. Korotayev); *Global Crisis in Retrospective: A Brief History of Upswings and Crises* (2010; in Russian, with A. Korotayev); *The Evolution of Statehood: From Early State to Global Society* (2011); *The Cycles of Development of Modern World System* (2011; in Russian, with A. Korotayev and S. Tsirel); *From Confucius to Comte: The Formation of the Theory, Methodology and Philosophy of History* (2012; in Russian);

Macrohistory and Globalization (2012); *Cycles, Crises, and Traps of the Modern World-System* (2012; in Russian, with A. Korotayev); *Great Divergence and Great Convergence* (2015; with A. Korotayev); *Economic Cycles, Crises, and the Global Periphery* (2016; with Arno Tausch and A. Korotayev); *Islamism, Arab Spring, and the Future of Democracy: World System and World Values Perspectives* (2019; with Andrey Korotayev and Arno Tausch).

Antony Harper, a researcher in the general area of applied mathematics to human historical processes, also understood as quantitative history and more appropriately, using Peter Turchin's term, as cliodynamics. He is currently associated with both New Trier College and Benedictine University. His primary research interests are modelling the pattern of urbanization over recorded history, *i.e.*, from approximately 3000 BCE to the present. More currently these interests include understanding the consequences of this urbanization pattern, applying the model of punctuated equilibrium, originally propounded by Eldridge and Gould, to the pattern of urbanization, understanding why complex systems in general should exhibit the PE pattern of growth, and within the constraints of punctuated equilibrium beginning to understand the context and mechanisms that bring about punctuated change, *i.e.*, quite frankly, understanding what causes historic tipping points.

Andrey V. Korotayev is Senior Research Professor at the International Laboratory for Political Demography and Social Macrodynamics at the Russian Presidential Academy of National Economy and Public Administration, Head of the Laboratory for Destabilization Risk Monitoring at the National Research University Higher School of Economics, Senior Research Professor of the Oriental Institute and Institute for African Studies, Russian Academy of Sciences, as well as a Professor at the Faculty of Global Studies of the Moscow State University. He is the author of over 600 scholarly publications, including such monographs as *Ancient Yemen* (1995), *World Religions and Social Evolution of the Old World Oikumene Civilizations: A Cross-Cultural Perspective* (2004), *Introduction to Social Macrodynamics: Compact Macromodels of the World System Growth* (2006), *Introduction to Social Macrodynamics: Secular Cycles and Millennial Trends* (2006), *Great Divergence and Great Convergence* (2015, with Leonid Grinin), and *Economic Cycles, Crises, and the Global Periphery* (2016, with Arno Tausch and Leonid Grinin). At present, together with Askar Akaev and Sergey Malkov, he coordinates the Russian Academy of Sciences Presidium Project 'Complex System Analysis and Mathematical Modeling of Global Dynamics'. He is a laureate of the Russian Science Support Foundation in 'The Best Economists of the Russian Academy of Sciences' Nomination (2006).

David LePoire has a PhD in Computer Science from DePaul University and BS in physics from CalTech. He has worked in environmental and energy areas for many governmental agencies over the past 25 years. Topics include uncertainty techniques, pathway analysis, particle detection tools, and physics-based modeling. He has also explored historical trends in energy, science, and environmental transitions. His research interests include complex adaptive systems, logistical transitions, the role of energy and environment in history, and the application of new technologies to solve current energy and environmental issues.

Edmundas Lekevičius works at the Center for Ecology and Environmental Studies, Vilnius University, Lithuania. He has been working in the field of ecology and theoretical biology since 1978, and now teaches ecology, adaptation theory and ecosystem evolution at Vilnius University. He is the author of GAT, the general adaptation theory (*Elements of General Adaptation Theory*. Vilnius: Mokslas, 1986) and the deductive version of the ecosystem evolution theory (*The Origin of Ecosystems by Means of Natural Selection.* Vilnius: Institute of Ecology, 2002). Here are some of his recently published articles (without co-authors): 'Ecosystem Evolution: Main Stages and Possible Mechanisms' (*Zhurnal Obshchei Biologii* [Moscow], 2003, 64(5): 371–388); 'The Russian Paradigm in Ecology and Evolutionary Biology: *pro et contra*' (*Acta Zoologica Lithuanica*, 2006, 16(1): 3–19); 'Vacant Niches in Nature, Ecology, and Evolutionary Theory: A Mini-Review' (*Ekologija* [Vilnius], 2009, 55(3–4): 165–174). E-mail: elekevicius@gmail.com.

Alexander V. Markov is Senior Research Fellow of the Institute for Paleontology of the Russian Academy of Sciences. He is the author of more than 140 scientific publications in zoology, paleontology, evolution theory, historical dynamics of biodiversity, and in other fields of evolutionary biology, including monographs: *Morphology, Systematics and Phylogeny of Sea Urchins of the Schizasteridae Family* (1994); *Quantitative Laws of Macroevolution: Experience of Systematic Approach Use for the Analysis of Supraspecific Taxons* (1998; with E. B. Naymark); *Macroevolution in Biological and Social Systems* (2008; with Leonid Grinin, Andrey Korotayev), *Hyperbolic Growth in Biological and Social Systems* (2009; with Andrey Korotayev), *Human Evolution* (two volumes, 2011). Dr. Markov is a member of the Editorial Board of *Journal of General Biology*, an author of numerous popular science publications, the founder and author of the research and education portal 'Problems of Evolution' (URL: http://www.evolbiol.ru).

William R. Thompson is Distinguished Professor and the Donald A. Rogers Professor of Political Science at Indiana University, Bloomington. He is also currently the editor-in-chief of the Oxford University Press's *Oxford Research Encyclopedia in Politics*. He is a former president of the International Studies Association (2005–2006), editor-in-chief of *International Studies Quarterly* (1994–1998 and 2009–2013), and has been on the editorial boards of numerous journals. He has published a number of articles in leading political journals and is the author, coauthor, or editor of such books as *The Arc of War: Origins, Escalation and Transformation* (2011, with Jack S. Levy) and *Transition Scenarios: China and the United States in the Twenty-First Century* (2013; with David P. Rapkin) (both published by University of Chicago Press), *How Rivalries End* (2013; University of Pennsylvania Press), *Asian Rivalries: Conflict, Escalation, and Limitations on Two-Level Games* (2011; Stanford University Press), *Globalization and Global History* (2006; Routledge), *Strategic Rivalry: Space, Position, and Conflict Escalation in World Politics* (2008; Cambridge University Press), and *Puzzles of the Democratic Peace: Theory, Geopolitics and the Transformation of World Politics* (2005; Palgrave-Macmillan). He has also received the World Society Foundation's Award of Excellence in World Society Research, the International Global Research Association and Moscow State University's V. I. Vernadsky Gold Medal of Honor (for contribution to global studies), and the International N. D. Kondratieff Foundation and Russian Academy of Natural Sciences' Silver Kondratieff Medal (for contribution to the social sciences).

History and Mathematics

Guidelines for Contributors

Preparation of manuscripts. Articles should generally be no longer than 100,000 symbols. The paper's abstract should not exceed 150 words. A separate sheet should give the author's brief CV (up to 250 words). A list of keywords should be supplied. Figures of good quality should be submitted as separate files.

Bibliographical references should be given in parentheses in standard author-date form in the body of the text: (Duffy, Morrison, and Macdonald 2002; Crumley 1987: 164–165; 1995: 4; 2001; Claessen 1985: 196–198, 201, 207; 2000, 2002; Chu *et al*. 2003: 29).

Examples of references in text:

'In a larger population there will be proportionally more people lucky or smart enough to come up with new ideas' (Kremer 1993: 685); according to Rothmaler (1976: 127–129).

A complete list of references cited, arranged alphabetically by author's surname, should be typed at the end of the article along the following lines:

A) Journal articles:

Schaffer W. M. 1985. Order and Chaos in Ecological Systems. *Ecology* 11(1): 103–109.

B) Edited volumes:

Frank A. G., and Gills B. K. (Eds.) 1993. *The World System: Five Hundred Years of Five Thousand?* London: Routledge.

C) Contributions to edited volumes:

Humphrey N. K. 1976. The Social Function of Intellect. *Growing Points in Ethology* / Ed. by P. P. G. Bateson, and R. A. Hinde, pp. 303–317. Cambridge: Cambridge University Press.

D) Monographs:

Berryman A. A. 1981. *Population Systems: A General Introduction.* New York, NY: Plenum Press.

E) Internet resource:

a) If the site does not mention the publication date:

Chase-Dunn C., and Hall T. D. 1999. *The Chesapeake World-System.* URL: http://www.wsarch.ucr.edu/archive/papers/c-d&hall/asa99b/asa99b.htm.

b) If possible, put the publication date:

U.S. Bureau of the Census. 2008. *World Population Information.* URL: http://www.census.gov/ipc/www/world.html. Date accessed: 24.02.2008.

Quotations. Single inverted commas should be used except for quotations within quotations, which should have double inverted commas. Quotations of more than 60 words should be set off from the text with an extra line of space above and below, and typed without inverted commas.

The articles are to be sent to the following email addresses:

Leonid Grinin lgrinin@mail.ru

Andrey Korotayev akorotayev@gmail.com

'History and Mathematics' homepage on our website: http://www.sociostudies.org/almanac/ham/