Non-equilibrium Statistical Mechanics and Evolution

Yile YING

June 7, 2018

Contents

This report summarises what I have learned in the last six weeks of Phys440 about non-equilibrium statistical mechanics (NESM) and its application in solving questions about evolution. A lot of developments in both non-equilibrium statistical mechanics and its application in evolution occurred in the last two, three decades. They are both very exciting, and sometimes "mind-blowing" to me. I will first introduce what are the major theories in non-equilibrium statistical mechanics; then motivate you (and myself) to enter the interdisciplinary area of NESM and evolution; in the end I will share some examples that I have explored in this interdisciplinary area. [1]

Please note that it is highly likely that I misunderstood some of the theories and wrote down wrong explanations or interpretations about them here, as everything in this report is purely based on searching and reading online in the last six weeks, and I have only done one statistical physics course so far. Please let me know if there is any mistake. It would be great to hear corrections and feedback!

1 What is Non-equilibrium Statistical Mechanics?

1.1 Equilibrium vs non-equilibrium

1.1.1 Equilibrium statistical mechanics

Statistical mechanics emerged in the nineteenth century from the works of Maxwell, Boltzmann, Gibbs etc. It started with descriptions of equilibrium systems where all observables of the systems are time independent and no current/flux (e.g. energy, entropy, matter) enters or leaves the systems. Such systems can be well quantified by a variety of state functions, such as the temperature, pressure, free energy, and partition function. By knowing only a small set of the state functions, one can easily specify the statistical or thermodynamic properties of the system.

However, equilibrium is usually just an approximation on the coarse grained level of a system - if we look closely enough, everything is still moving around. Similarly, an equilibrium process is based on the idealisation of being "quasi-static", which means the process is slow enough to allow the systems remain in equilibrium all the time. Besides, numerous natural systems are constantly out of equilibrium and cannot fit these two assumptions in any manner.

Therefore, it is quite often that we cannot describe a system or a process accurately enough by simply using the equilibrium statistical mechanics. Instead, we need to explicitly consider the dynamics of this system. Also, one question in equilibrium statistical mechanics really bothers me a lot, which is: how exactly the irreversibility described by the second law of thermodynamics arises from all the time reversal fundamental physics laws. To solve these issues, we need the non-equilibrium statistical mechanics.

1.1.2 Non-equilibrium statistical mechanics

Maybe the easiest way to define a non-equilibrium system is to say that it is not an equilibrium system. The observable of a non-equilibrium system is time dependent, or there is current/flux entering or leaving the system even if the system is in a steady state. Therefore, the system's state depends on the details of the dynamics and the coupling with the environment, it can exhibit various behaviours including relaxation, ageing, meta-stable states, and steady states [1].

It is hard to study non-equilibrium systems, both conceptually and technically, because we need a deep understanding of the dynamics, and we cannot just use a small set of parameters to specify the properties of the system anymore. Even in a steady state, the statistical or thermodynamical properties of a system are still difficult to describe, because "many of the basic state functions, including the temperature and entropy, are undefined for non-equilibrium states", and the distribution function of a non-equilibrium system is "fractal and non-analytic" [1].

Therefore, non-equilibrium statistical mechanics is still a work in progress, not a systematic or well-understood physical theory, especially compared to its equilibrium counterpart. Even in the

[1] Two papers provide good reviews related to this interdisciplinary area: *Life is Physics: Evolution as a Collective Phenomenon Far From Equilibrium* by Nigel Goldenfeld and Carl Woese, and *Origins of Life: a Problem for Physics, a Key Issues Review* by Sara Walker. In fact, many examples in this report can also be found in those two papers, but may be described in different ways or with different emphasis.

classical region, many questions are unsolved. For example, the turbulence is sometimes regarded as "the last great unsolved problem of classical physics".

Most of the early works in non-equilibrium statistical mechanics focus on the region near equilibrium, and quite a few theories are only applicable to specific situations. Only in the last two, three decades, there started to appear theorems such as the Fluctuation Theorem that are valid for general non-equilibrium situations. However, rigorous mathematical derivations for general situations are still rare, and there are many debates going on about the existing theorems in this area.

1.2 Main theories/approaches in NESM

As mentioned before, in equilibrium statistical mechanics, we use state variables and functions (e.g temperature, pressure, free energy) to define the system's states. On the contrast, in non-equilibrium realm, we study fluctuations (dynamics), which are determined by different control parameters (e.g. volume, force applied, temperature). In practice, the equations of motion are often written down from phenomenological considerations, and the approach is likely to be quite specific for each system, especially for small systems.

As the tool-set of equilibrium statistical mechanics is already well-established and relatively complete, the main goal in non-equilibrium analysis is to link the system's non-equilibrium behaviour to its equilibrium states and get access to the whole machinery of the equilibrium statistical mechanics.

There are two main approaches:

- Idealisation and approximation. For systems near equilibrium, or exhibiting some equilibrium-like states (local equilibrium, steady states), one can approximate the dynamics in a linear fashion, and write down the Hamiltonian of the systems with the equilibrium part plus the perturbation parts.

- Integrating/averaging over possible non-equilibrium trajectories to recover the equilibrium properties. This is mainly used in the Fluctuation Theorems, which is valid for general situations out of equilibrium including the far from equilibrium region.

More about these two approaches along with the main theorems in non-equilibrium statistical dynamics will be discussed below.

Besides, another branch in non-equilibrium statistical mechanics is the one focusing on the emergent states or collective phenomena of matter, which uses different but related approaches to ones mentioned before, such as non-equilibrium phase transitions. Those approaches are as important as others, I will discuss them in the end of this section.

1.2.1 Near equilibrium

As we may imagine, when a system deviate just a little bit from its equilibrium state, i.e. generated by small perturbations, it may relax fast back to equilibrium, and indeed, it was shown that a state near equilibrium decays exponentially towards equilibrium [2]. Besides, in near equilibrium region, the response (often time dependent) to the perturbations i.e. the change in the properties of the system that couples to the external field, can be expressed in a linear fashion to the perturbation (the external force/field), because we can ignore the higher order terms in the Taylor expansion of the expression of the dynamics. Therefore, the near equilibrium region is also often called the "linear response region", and can be understood by using the linear approximations and the tools from equilibrium thermodynamics.

For example, the diffusion equation we studied in Phys 305 is actually a result of the linear response theory. The linear approximation is adopted in the Fourier's law:

$$J = -K\nabla T \tag{1}$$

where J is the heat flux and K is the thermal conductivity. In far from equilibrium region, there are actually higher order terms (e.g. T^2, T^3), but they are negligible when the system is close to the equilibrium.

Some systems, though they are globally far from equilibrium states, at the same time are locally near thermodynamic equilibrium, so that the linear response theory is also applicable to these systems.

The scale of local equilibrium should be larger than the microscopic scale, but still small compared to the macroscopic scale (where observations are made). For example, the temperature deviation across the whole system can be large, but in a characteristic scale the temperature gradient can be small enough that we can "cut" the system into pieces for the linearization approximation to be valid, but also large enough that a macroscopic statistical description of each piece is also meaningful.

There are numerous equations describing linear response region and they were mostly derived during the 19th and early 20th centuries, e.g. Langevin Equation (Brownian Motion), the Kramers-Kronig relations, the fluctuation-Dissipation theorem and the Onsager Reciprocal relations. Here I will briefly state the main ideas of last three relations.

Both the Kramers-Kronig relations and the Fluctuation-Dissipation Theorem link the response to the energy dissipation of the system.

The Kramers-Kronig relations separate the response into two parts, the real part in the equations is the "reactive" part of the response, and the imaginary part is the "dissipative" part of the response, which describes how a system dissipates energy because of the external forces. The Kramers-Kronig relations imply that we can determine how a system reacts to the perturbations by observing its dissipative response, and vice versa.

Furthermore, the fluctuation-dissipation theorem relates thermal fluctuation in equilibrium to the dissipation response of the system. It assumes that the dissipation response to a small applied force is equivalent to the behaviour in a spontaneous fluctuation which is a reverse process for that response. For example, the Brownian Motion is the reverse equivalent process of drag. The object experiences drag (e.g. fluid resistance) and dissipates kinetic energy, turning it into heat. This process is equivalent to the situation when the object starts with zero energy in a fluid, and then hit by the molecules in the fluid, i.e. the heat energy in the fluid is converted into kinetic energy of the object. Therefore, the fluctuation-dissipation theorem allows us to study the linear response relaxation of a system from a prepared non-equilibrium state by understating its statistical fluctuation properties in equilibrium. [3]

The Onsager Reciprocal relations show that certain relations between different flows (responses) and forces in the linear response region are equivalent, and these relations are called "reciprocal relations". For example, in Phys 305, we have learned that the temperature difference between two systems in contact causes heat flow, and pressure difference causes matter (particle) flow. However, temperature difference can also cause matter flow (e.g. convection), and pressure difference can also cause heat flow. Moreover, the heat flow per unit of pressure difference and the matter flow per unit of temperature difference are equal. Therefore, the relations between temperature/pressure and heat/matter flow are reciprocal relations. The Onsager Reciprocal relations allow us to analyse complex dynamics in the near equilibrium region. [4, 5]

There are Some other famous equations describing the linear response region, e.g the Kubo formula, but they mainly deal with quantum systems so they are not of main interest here.

Moreover, fluid dynamics as the main branch of the non-equilibrium statistical mechanics can be applied to describe many systems (not necessary fluids) that are in local equilibrium.

1.2.2 Far from equilibrium

A system is far from equilibrium when the linear approximation is no longer valid and the higher order terms have to be taken into account. In such cases, the study becomes much harder since the equations of motion become much more complicated and it is hard to utilise the tools of equilibrium statistical mechanics. However, vast number of real-life phenomena are actually far from equilibrium, so the study of far from equilibrium statistical mechanics is highly interdisciplinary, engaged with many biophysicists, biochemists, geochemists, information theorists, even economists or social scientists.

Most statements that are valid in far from equilibrium region can be categorised into the Fluctuation Theorem [6]. It was firstly established by Denis Evans, E.G.D. Cohen, Gary Morriss and Debra Searles in the early 1990s, and then extended and generalised by Christopher Jarzynski and Gavin Crooks around 2000. Those theorems usually connect non-equilibrium behaviour of the system with equilibrium properties by methods such as "path integral". By applying "path integral", i.e, averaging the change over all possible paths from one state to another, one can recover the equilibrium state and utilise the tools of equilibrium statistical mechanics, e.g. the Boltzmann distribution.

The Fluctuation Theorem is closely related to the second law of thermodynamic, which states that, in macroscopic statistical level, it is impossible for the entropy of the system to decrease. However, in microscopical level, it is possible to have negative entropy change (entropy fluctuation), and the Fluctuation Theorem gives the exact mathematical expression for the probability if entropy will flow in the same or opposite direction to that dictated by the second law of thermodynamics. Therefore, it "provides a fresh look to our understanding of old questions such as the origin of irreversibility and the second law in statistical mechanics" [7]

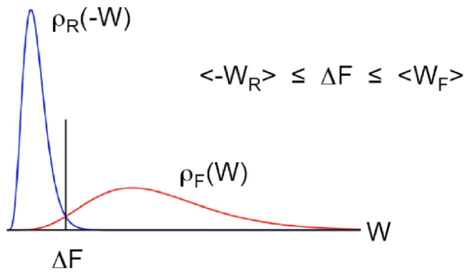

Figure 1: Work distributions of the forward and reverse processes [8].

In the beginning of 2000s Christopher Jarzynski introduced the nonequilibrium work relations, which can be seen as consequences of the The Fluctuation Theorems [9]. The work relations are mainly associated with the Helmholtz free energy differences ($\Delta F = F_B - F_A$) and the work done (W) during the transient from an equilibrium state A at temperature T to a general state B (the final free energy is defined when the system relaxes to the equilibrium after the work is done). In details, the work relations are[2] [8, 10]:

- The Jarzynski equality:
$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}, \tag{2}$$
where $\beta = (k_B T)^{-1}$, and the angular bracket denote an average over the possible paths (trajectories) from A to B. This equation provides means to predict Helmholtz free energy differences between two equilibrium states by the ensemble average of the non-equilibrium work. This equality also implies $\langle W \rangle \geq \Delta F$, which is related to the second law.

- The Crook's Fluctuation Theorem:
$$\frac{\rho_F(W)}{\rho_R(-W)} = e^{-\beta(W - \Delta F)}, \tag{3}$$
where $\rho_F(W)$ and $\rho_R(-W)$ are the forward and reverse probability, respectively. The Crook's Fluctuation Theorem gives another way to determine the free energy difference, and the Jarzynski equation can actually be derived from it. It also implies that the work distributions of the forward and reverse processes cross each other at $W = \Delta F$, as shown in Figure 1.

- The third one "goes further, giving a prescription for constructing the entire equilibrium distribution" [8]:
$$\langle \delta(x - x_t) e^{-\beta W_t} \rangle = \frac{1}{Z_A} e^{-\beta H(x, \lambda_t)}, \tag{4}$$
where W_t is the work done up to time t for a trajectory x_t, Z_A is the partition function in the initial state, and $H(x, \lambda_t)$ is the Hamiltonian of the system at time t depends on the state x and the work parameter λ_t (related to W_t).

Note that the above approaches are not only useful for non-equilibrium transient state (initially prepared in an equilibrium state and later driven out of equilibrium but soon return to a new equilibrium once the driven force stops) or non-equilibrium ageing state (initially prepared in a non-equilibrium state and later put in contact with a thermal reservoir but fails to reach equilibrium

[2] The lecture notes prepared by E. Boksenbojm, B. Wynants and C. Jarzynski provide extensive derivations and explanations about the work relations. `https://arxiv.org/abs/1002.1230`

in observable or laboratory time scales), but also useful in analysing non-equilibrium steady states, which are systems driven by external forces but their properties do not change with time [7]. The steady states can also be analysed by linear response theories, depending on the conditions.

A drawback of the fluctuation theorems and the work relations is that one needs to get a sufficient number of sampling of the trajectories in order to perform the integral/averaging, which is sometimes not easy and limits the complexity of the systems that can be modelled.

1.2.3 Emergent states / collective phenomena

As we mentioned before, one of the main themes in statistical mechanics is to understand macroscopic phenomena from microscopic properties, and those macroscopic phenomena are often emergent states or collective phenomena, which one can not derive by simply adding up the individual components together. The theories about emergency are scattered in all regions of statistical mechanics, including equilibrium, near equilibrium or far from equilibrium regions, and some of them have deep connections with the theories we mentioned above. They are mainly studied in condensed matter physics for questions such as many-body problems, and its development in active matter is especially closely related to our topic.

One of the main tools to study emergent states or collective phenomena is through phase transition, where the macroscopic state or behaviour of a system changes (often discontinuously) because of the changes in the states and interaction among its microscopic components under external fields, e.g. a paramagnetic material undergoes symmetry breaking (becomes ferromagnetic) as the temperature decreases. In order to describe a phase transition, one needs to specify the order parameters of the system, e.g. the net magnetisation in a ferromagnetic system, and identify the "control parameter" that leads to phase transitions, e.g. the temperature. Another typical example is water turns to ice (which is also a process of symmetry breaking) as the temperature decreases or pressure increases; here the density is the order parameter, and temperature/pressure are the "control parameters" determined by external fields.

Phase transition has been extensively studied in equilibrium statistical mechanics, as one tries to minimise the free energy of a system at each state. However, the phase transitions in non-equilibrium regime still lacks a generalised theory. Some of the concepts in equilibrium can be applied to non-equilibrium such as "universality", but we still need new physics to emerge to fully understand the non-equilibrium regime. Even for those concepts applicable to non-equilibrium systems, we need extended or fresh models, e.g. we have the "directed percolation" model instead of the Ising model as one of the most important universality class in non-equilibrium. Furthermore, the discussion of mean-field theory, bifurcations, critical phenomena etc. are also useful for non-equilibrium systems with adaptations.

The powerful thing about phase transitions is that we do not need to know the exact details of the states and interactions at the microscopic levels to predict collective properties of the system, e.g. we do not need to know the exact forces of each magnets exerted on every other ones, or the exact interactions between the water molecules to predict the phase transition, and in other words, the behaviour of the collective is more of interest than the behaviour of individuals.

2 Why do we Use NESM to Study Evolution?

A trivial answer to this question is that, because all, or at least nearly all, the biological systems are operating at non-equilibrium states rather than equilibrium, and non-equilibrium statistical mechanics is maybe the most robust method to study this type of systems. However, one may wonder that: "Haven't we already understood evolution well enough, with the Darwinian theory proposed 200 years ago and with the modern development in genetics?", or "Can we really learn new things by explicitly studying the non-equilibrium properties of evolution using statistical mechanics which is physics not biology?". My answer to the first question is "no", our knowledge of evolution are actually far from complete; while my answer to the second question is "yes", and I believe we can gain new insights into both evolution and physics.

2.1 We do not know evolution well enough yet

There are two questions in evolution: 1) how life appears in the first place (the origin of life) 2) how life evolves afterwards. We do not have satisfactory answers for both questions.

2.1.1 the origin of life

When trying to answer the first question, one may also need to consider what is the exact definition of life, in order search for the origin of life. There is a long time debate on this definition, the different opinions on it can be grouped into the two main views on the origin of life: genetics first vs metabolism first. The proponents of genetics first approach believe that, the genetic inheritance is the essential unique and defining feature or life, and we should find out how this mechanism came to existence to understand the origin of life. On the opposite, the metabolism first view suggests that, the early life may not have any forms of genes at all but be just a group of molecules collectively reproducing together, like auto-catalytic sets, which can extract energy from the surroundings to persist and evolve.

However, both theories have their flaws. For the genetic view, the biggest challenge may be that to synthesis RNAs from prebiotic condition is proven to be hard, let alone DNAs; though scientists have found other possible genetic substrates such as TNA, but the transition mechanism from TNA to RNA is unclear. For the metabolism view, it is hard to define heredity of the auto-catalytic set, and the evolvability of the set is under doubt. Also early life probably had a very different biochemistry, so biochemistry based only on current samples cannot explain evolution. The paper by Sara I. Walker provides a good review of the limitations of both views [11].

The definition of life is certainly useful, as it can give us the criteria for evaluating different theories of the origin of life, but we should also look at all distinctive features together to understand life and its evolution, even if those features are not unique or defining, e.g. self-replication, irreversibility, energy extraction, the ability to predict future. Maybe, it is the collective property of these features that makes life unique to other systems.

Moreover, we can use physics to study each individual feature, e.g. we can model both the genetics and metabolism aspects of the origin of life by using statistical mechanics, and ask what physical properties are common to and distinctive of living things. As biological systems are fundamentally one type of organisation of matter, they must obey the basic laws of physics.

It worth noting that, so far, we only have one sample of the origin of life, which is the life on the earth, and we may possess significant anthropic bias. It is possible that life may have appeared from a different origin or a different sequence of adopting the distinctive features on some other planets. Only if we uncover the mechanisms governing the transition from non-living to biological properties, that we will not be surprised if one day we find some life forms totally different from what we have already known.

2.1.2 the evolution of life

For the second question, the dominant evolutionary view since the last century has been the "Modern Synthesis", also called "Neo-Darwinism". It is a combination of Mendelian genetics and Darwins theory ("natural selection"), which explains evolution through the mechanism of heritable mutation and survival of the fittest.

The limitation of this theory is its very simplified evolution dynamics, where only very basic mechanisms, such as single point gene mutation and sexual recombination, are taken into consideration.

The genetics in "Morden Synthesis" is fairly linear and vertical, and does not account for gene interactions (e.g. epistasis) or mobile genetic elements such as horizontal gene transfer, transposons (jumping genes) and epigenetic (gene expression) inheritance. However, these non-vertical elements are important. For example, we humans, much complex than the simple worm *C. elegans*, evolved to have only one third of its genes [12, 13]. This illustrates that the interactions and combinations of genes affect the complexity of living systems more than plain number of genes. Also, the horizontal gene transfer is actually the main mechanism in the first billion years of evolution [14].

Moreover, the coupling between evolution and ecology is not clearly addressed, which leads to a fixed fitness landscape where the nature selection is based on static environment conditions. Thus, it can not accurately describe phenomena such as co-evolution (e.g. cyanobacteria and phages co-evolving through horizontal gen transfer), or evolutionary arm-race (e.g. cancer cells evolution in response to treatment; bacteria's resistance to antibiotics).

Besides, recent findings show that both mutations and horizontal gene transfer may not be completely random as described by Modern Synthesis. The genomes are actually built in a way that mutation is more likely to be useful than detrimental, especially under environmental stress [15, 16, 17].

Biologists have started to notice these shortcomings of Modern Synthesis recently. The Extended Evolutionary Synthesis as one of the new theories in evolution (not widely accepted yet), have adopted some advanced dynamics such as multilevel selection and epigenetic inheritance [18]. Physics is likely to speed up the understanding of these unexplained evolutionary phenomena by adopting theories or models from non-equilibrium statistical mechanics as we will discuss next.

2.2 We can understand evolution better from physics

2.2.1 Evolution is also a physics question

A major difference in physics from biology is that, in physics, the notion of life is absent: life or non-life, they are both some organisations of matter. Basically, we need to answer the question that, how and why "matter self-organises into hierarchies that are capable of generating feedback loops which connect multiple levels of organisation and are evolvable" [19]. Therefore, evolution is also a physics question, and the irreversible (and somewhat inevitable) emergent and continuous evolution of life can seek its explanation in the regime of non-equilibrium statistical mechanics.

We should be able to use physics terms to explain what features are striking or special about life and how they emerge or change through time. Non-living organisms can sometimes also have some life-like properties, e.g. self-replication and energy extraction, which is essentially all about strongly interacting collective properties of matter. One may even speculate that the evolution of life is similar to the evolution of technology, society or even the universe. Therefore, physics as a subject always searching for universal and fundamental laws, is promising and suitable to provide general principles for life and its evolution.

2.2.2 The physics approach and views are valuable for biology

The twentieth-century biology was dominated by reductionism, where biologists attempted to catalogue everything, from ecosystem, taxonomy, to cells, molecules and genes, and then tried to complete the puzzle of life [19]. However, life is not just a jigsaw puzzle that one can complete by finding and putting together all the pieces, rather, it is more like a dynamic hierarchy of multidimensional networks. For example, we literally know the roundworm *C. elegans* inside out - its entire anatomy. Even every neuron with the connections is known, and its full genome is mapped - but we still have no satisfying model to predict its behaviour [20].

Therefore, an interdisciplinary approach that embraces collective phenomena from statistical physics can be very useful here. The non-equilibrium statistical mechanics of living systems would naturally reflect emergent properties and allowable dynamical symmetries, as it goes beyond the behaviour of individual components to see how they self-organise into collective states.

Moreover, in biology, inductive reasoning is widely used, while in physics scientists more often use deductive reasoning (at least it seems to me like that). The limitation of the first approach is the risk of missing possibilities, especially in terms of evolution as we only have one sample so far. On the contrary, the deductive reasoning cares more about what is "impossible", and physics tends to think "which is not forbidden is mandatory" [19]. This type of deductive reasoning based effort can often direct or predict new findings. One example in biology is the groundbreaking discovery of the topoisomerases molecules, the enzymes helping to overwind or underwind DNA [21, 22]. It was its theoretically prediction led to their eventual discovery.

In addition, a physics-based approach to biology may be somewhat like the way a child learns, that can study living systems without much expectations heavily based on what we have already known about biology, and later using the known knowledge or new experiments to confirm the findings.

But of course, physicists will also have the bias from the physics background, and may oversimplify things. Therefore, one should be bold but also down to earth, and be aware that non-equilibrium statistical mechanics is not the only way to study evolution, though certainly an important one.

And anyways, I think we already have too many arbitrary divisions between different branches of science, thus we certainly should not limit ourselves anymore especially when we are doing some interdisciplinary things already.

2.3 We can also understand physics better

"Just as we find features of the atom, its stability, for instance, which are not reducible to mechanics, we may find features of the living cell which are not reducible to atomic physics, but whose appearance stands in a complementary relationship to those of atomic physics. – Max Delbrück, A Physicist Looks at Biology, 1949

Most scientists do not believe in new physics in biology, because it seems that technical limitations is the main reason preventing us from, for example, mapping/simulating all the neurons or metabolism pathways or evolutionary process. Nevertheless, we are still not able to exclude the possibility of finding undiscovered fundamental physics laws governing biology or evolution, as there are still so many unsolved problems without obvious direction of solutions, and we may have not tried hard enough yet.

Even if we will not discover any surprising new laws of physics, it is for sure that we will extend the frontier of physics by digging into the massive sample set of non-equilibrium systems in biology.

Physics can be much enriched by studying the non-equilibrium phenomena abundant in life but very rare in pure physical systems. For example, inspired by biology observations, a new branch in condensed matter physics emerged in 1995, called "active matter", which studies matter made of "agents" that can use energy to move or apply force on other objects.

Besides, studying biological or ecological phenomena that are also present in purely physical processes can give us insights into non-equilibrium statistical mechanics by revealing principles of self-organisation and so on. For example, when a group of scientists were studying how the flow of liquid in a pipe is changing from being laminar to turbulent, they drew analogies from a predator-pray co-evolving ecosystem and identified a certain type of phase transition in fluids shown by experiments [23].

Indeed, much of the current interest in non-equilibrium statistical mechanics actually stems from progress in biology, e.g. it is when reconciling the details of biomolecular processes that scientists started to question the thermodynamics of small systems. "And the fact that a living system in equilibrium is oxymoronic (or dead) has necessitated the development of new frameworks for understanding otherwise familiar statistical physics" [5].

3 Examples of Using NESM to Study Evolution

3.1 Examples of using the extended second law of thermodynamics

One of the mysteries about life is its irreversibility, that we never saw anything "ungrow", e.g. a big tree never "reverse grows" back to a tiny sprout. The second law of thermodynamics is a very tempting tool to make use of here, as "it is very good at permitting or forbidding things when we do not know much about the system" [24]. As we discussed in Section 1.2.2, the Fluctuation Theorems in NESM provided extended versions of the second law in non-equilibrium thermodynamics, and Jeremy England, a physicists in MIT, started to use them in 2013 to determine, under what physical conditions where no life exists, life or life-like physical properties will emerge and persist. During the past five years, his lab has derived theoretical minimal cost for self-replication, precision and self-repair, and the thermodynamic probabilities of adaptation. Besides, computer simulations were designed to test and further explore the emergence of life-like features. The work related to self-replication and adaptation is presented below.

3.1.1 Self-replication

He firstly started with the Crooks Fluctuation Theorem as shown in Equation 3, as it implies that, the more energy/entropy a process dissipates, the less likely for the reverse process to happen. He extended the Crooks equation into the following form[3]:

$$\langle W \rangle_{x \to x'} - \langle \Delta E \rangle_{x \to x'} + T\Delta S \geq k_B T \ln\left(\frac{\rho_{x \to x'}}{\rho_{x' \to x}}\right). \tag{5}$$

This equation explicitly shows that the irreversibly of a process puts a lower bound on the amount of heat and entropy produced during the forward process, which is an addition to the notion that only the process with non-negative energy/entropy production is permitted.

If we assume a self-replicator has both the forward and reverse processes, with the forward replication rate (the growth rate) at g, and backward rate δ (we can think of $1/\delta$ as the durability). The above relation can be expressed as:

$$\langle \Delta Q \rangle_g \geq k_B T \ln\left(\frac{g}{\delta}\right) - T\Delta S, \tag{6}$$

where $\langle \Delta Q \rangle_g = \langle W \rangle_g - \langle \Delta E \rangle_g$. This inequality indicates the energy cost of self replication $\langle \Delta Q \rangle_g$, and sometimes this cost of making life is more dominated by the irreversibility, which is different from the cost of creating only a temporary disorder. Using the bacteria E. coli as an example, the theoretical minimum heat dissipation for a single replication calculated from Equation 6 is roughly one-sixth of what it actually does, which shows the bacterium respect the bound. The extra heat produced is normal because it has to do a lot more than simply maximising its replication rate to survive, e.g. adjusting its internal machinery to the fluctuating environment) [25].

We can rearrange the equation further to get the expression for the maximum net growth rate:

$$g_{max} - \delta = \frac{e^{\frac{\langle \Delta Q \rangle_g}{k_B T} + \frac{\Delta S}{k_B}} - 1}{\frac{1}{\delta}}. \tag{7}$$

It is obvious from this equation that, the more energy $\langle \Delta Q \rangle_g$ one dissipates to its environment, the faster it can replicate. The dissipated energy can be from either the energy (E) initially stored in the replicator (e.g. hydrolysis of carbohydrates or lipids), or the work done (W) by external forces in the environment (e.g. absorption of energy from light during photosynthesis). Thus, "the empirical biological fact that reproductive fitness is intimately linked to efficient metabolism now has a clear and simple basis in physics" [26].

We can also imply from Equation 7 that, a self-replicator can increase its net growth rate by decreasing its durability ($1/\delta$) and increasing the entropy, which is to say, one strategy for winning the competition among other self-replicators is to be more prone to spontaneous degradation and simpler in construction [26].

This sounds a bit counter-intuitive, as we would think that being durable (e.g. living longer) and complex is crucial for a specie in survival games. However, it is thought-provoking to me when I think about how viruses or bacteria not only survive but sometimes even thrive in the battle with antibiotics while having simple forms and short life time.

3.1.2 Dissipative adaptation

Previously, we discussed the probability for a forward process to happen and applied it on self-replication. Here, we will look at what affects the direction of evolution a system will choose under the prediction of extended second law of thermodynamics. This is described by two of the papers of England's lab [27, 28].

It again starts with the Crook's equation as shown in Equation 3. After mathematical rearranging and approximation (more complicated than before), we arrived at the following expression comparing the probability to go from the initial state **I** to state **II** or state **II** at a given time τ.

$$\begin{aligned}\ln\left[\frac{\rho(\mathbf{I} \to \mathbf{II}; \tau)}{\rho(\mathbf{I} \to \mathbf{III}; \tau)}\right] \simeq \; & \underbrace{-\ln\left[\frac{\langle \Omega_{BZ} \rangle_{\mathbf{II}}}{\langle \Omega_{BZ} \rangle_{\mathbf{III}}}\right]}_{\text{"distance" from equilibrium}} + \underbrace{\ln\left[\frac{\rho(\mathbf{II} \to \mathbf{I}; \tau)}{\rho(\mathbf{III} \to \mathbf{I}; \tau)}\right]}_{\text{reverse process (accessibility)}} \\ & + \underbrace{(\Psi_{\mathbf{I} \to \mathbf{II}} - \Psi_{\mathbf{I} \to \mathbf{III}})}_{\text{average dissipation}} - \underbrace{(\Phi_{I \to \mathbf{II}} - \Phi_{I \to \mathbf{III}})}_{\text{fluctuations of dissipation}}.\end{aligned} \tag{8}$$

[3] All the equations here from England's papers are rearranged for easier interpretation based on my own understanding, and adapted to be consistent with the symbols used in this article.

The first term in Equation 8 is about how far the state **II** or **III** is from its corresponding equilibrium state, where Ω_{BZ} is the ratio between the distribution of the state and its corresponding Boltzmann distribution. It shows the tendency of the system to evolve towards thermal equilibrium, and the smaller the "distance" from equilibrium is, the more likely the state will be chosen, which is not surprising.

The second term in Equation 8 shows the reversibility of the transition, which is similar to what we have discussed in Section 3.1.1. Apart from regarding the reversibility as the inverse durability as in Section 3.1.1, we can also think of it as the accessibility of the states. Apparently, not all states are equally accessible on a finite time scale. A higher probability for the reverse process to happen can indicate a lower energy barrier (i.e higher accessibility) between the initial states **I** and the final states **II** or **III**, so that the corresponding forward process gets easier.

The third and the forth term are both about dissipation, with the third one for the average energy a forward process dissipates and the forth one for the fluctuation (reliability) of that dissipation. As we discussed in Section 3.1.1, energy dissipation is closely related to the metabolism rate and its ability to absorb energy. Here we can see that the system tends to evolve in the direction that it can reliably dissipate more energy (high Ψ and low Φ). Moreover, when the first two terms are fixed, i.e. when state **II** and **III** are equally far from equilibrium and accessible in a given time, the one with an exceptional history of absorption and dissipation will be favoured, and that is why it is referred to as "Dissipative Adaptation" in England's papers.

This prediction was later further explored by computer simulations of spring networks and toy chemical models [29, 30]. In both simulations, the evolution results of the networks were that, they either self-organised to extremely good at extracting and dissipating energy from the external fields, or they ended up being in a shape that could merely absorb or dissipate any energy, especially when the environment is challenging. This is a very interesting phenomenon, as it shows that a systems tends to be fine-tuned to the energy source in the environment, even without the mechanisms of self-replication and natural selection [31].

3.2 Examples of using phase transitions

From 1990s, scientists started to map evolution phenomena to phase transitions in a variety of contexts, where biological systems are sometimes "argued to be poised at the critical point between order and disorder (colloquially this is sometimes phrased as 'poised at the edge of chaos')" [11].

The main challenge in adopting phase transition analysis is to define the order parameters and then find the control parameters. The emergence of homochirality in life is one of the phenomena where phase transition approach finds its immediate application [32, 33, 34].

The homochirality of life refers to the facts that the amino acids (the basic compounds for proteins) in organisms are primarily left-handed and the carbohydrates in DNA and RNA are mostly right-handed, while both left and right chiral forms equally exist in abiotic environments. Therefore, we can treat the homochirality of matter as the order parameter, and the corresponding spontaneous symmetry breaking (from heterchiral to homochiral) happens when life forms. Thus we can build different models to figure out what the "control parameters" are, which indicates the condition for life to evolve, e.g. "autogenic rate of polymers" [33], "the fidelity of enzymatic reactions" [33].

Besides, phase transitions are also used to explain genotype selections, and the competition between gene mutation and recombination [19], where the high homologous recombination rate of genes (exchanged between two similar or identical strand of DNA) resembles high temperature, which leads to diverse genotypes (the order parameter), and low recombination rate leads to stable favourable genotypes where the mutated gene is more likely to be preserved [35, 36]. As we mentioned before, one of the major puzzles in evolution is how the gene transfer mechanism shifted from mainly being horizontal in the early era to mainly being vertical in most of organisms nowadays. Computer simulations with physics models or techniques developed in non-equilibrium statistical mechanics to analyse emergent properties such as phase transitions may be able to demystify this shift [37].

Note that, not all the phase transition analysis here uses the methods developed in the non-equilibrium regime.

3.3 Side examples (not explicitly using NESM)

3.3.1 Combine with information theories

Along with the development in computer science and artificial intelligence, there is also a trend of using information theory (which is related to thermodynamics) to better understand life, starting from the pioneers such as von Neumann and Turing.

It is identified that life has quite distinctive information architecture comparing to non-living systems. For example, the exceptional feature of separation of information storage (DNA) and processing (proteins and other machinery) was noticed by von Neumann and adopted in computer design. Some scientists believe that it is the way life controls and manages information (the algorithm part), rather than the chemical components (the hardware part), that distinguish life from non-life [38, 39]. The transition from non-life to life could be that when "information gains direct and context-dependent causal efficacy over the matter in which it is instantiated", i.e. the material substrate of life can be manipulate by an abstract and non-physical entity ("algorithmic information") that responds to its environment and its own condition within the hierarchy of biological systems, e.g. the gene expression management, or population control [40].

3.3.2 Combine with Geochemistry

One view on the origin of life is that life appeared as the irreversible non-equilibrium thermodynamics result of the geochemistry on Earth, where a huge amount of energy has been "trapped" and needed to be released [41, 42].

As the temperature and pressure on the early earth decreased, the rates of the abiotic reactions slowed down, while the drive towards equilibrium persisted, that was when the energy got "trapped". During this time, enzymatic reactions, one of the main distinctive properties of life, emerged as they could release the energy stress faster. Hence, abiotic reaction pathways were gradually replaced by enzymatic ones, so life, incorporating enzymatic reactions emerged as a planetary response to provide a relaxation channel for the thermodynamic stress on the primitive earth.

One of the evidence for this theory is that, life using chemical energy appeared earlier than that using light energy [43]. For example, the tube worms and various other creatures found in the extreme conditions of hot hydrothermal vents are proposed to be the living fossils similar to early life forms. Besides, it was shown that the Krebs cycle (one of the most important metabolic pathways) is possible under abiotic conditions at high temperature and pressure with pure mineral catalyst, but later the pure mineral catalysts were replaced by more efficient metalloenzymes (proteins with metal ions) as the condition cooled down [41].

The promising part of this theory is that we can make predictions on how organisms with different metabolic strategies may appear in different geochemical environment, which may give us a framework for evaluating possibility of life on the exoplanets.

However, we still need an extension of this theory to explain the emergence of heritable and evolvable structures, i.e. how life kept developing afterwards.

Note that, the transition from abiotic to living state in this theory has also been analysed by phase transitions [42].

4 Personal Comments

4.1 About non-equilibrium statistical mechanics

- One thing that confuses me a bit is the scope of the statistical mechanics as compared to thermodynamics, kinetics and statistical physics, which are all closely related. It seems to me that statistical mechanics emphasises more on the link between microscopic properties and macroscopic phenomena, while the kinetic theory is more about the rate of change, and thermodynamics cares more about the states (e.g. temperature, energy) of a system and transitions between them. In terms of studying biological evolution, we care about both the conditions for life to occur or evolve (covered by thermodynamics) as well as the rate of changes e.g. reproduction/mutation (covered by kinetics), we also need to understand the underlying dynamics from maybe the molecular level to explain evolution which is a macroscopic phenomena. Therefore, I chose "statistical mechanics" as the main term here. "Statistical physics" seems to be a common way to address statistical mechanics, but as we are discussing an interdisciplinary area, "statistical mechanics" sounds more appropriate to me than "statistical physics".

- Non-equilibrium statistical mechanics (NESM) is such a cooooool area! It seems that one can use it to study anything and it already has applications literally everywhere: condensed matter physics, quantum field theory, biology, neuroscience and deep learning, economics, sociology, etc!! I even feel like if I'm allowed to choose only one subject to study, it would be non-equilibrium statistical mechanics, because it's like a key to anything I'm interested!
- Non-equilibrium statistical mechanics is certainly not an easy subject to study, and I already experienced that during the last two months...As it is still an developing area, there is even no single textbook that can provide a good general idea in this field. Though there are already many textbooks or lecture notes about NESM, they are all VERY different from each other, which is very confusing and frustrating......
- (A joke:) The good thing of NESM is also that: it's still being developed, so it's still possible to put your name onto some equation you may discover one day haha :D ! Look at how many equations named after someone in this area......They are very confusing.....and they change through time, e.g. the BBGKY hierarchy, which is named after Bogoliubov, Born, Green, Kirkwood, and Yvon, who contributed to the theory one after another...

4.2 About the motivation for this interdisciplinary area

- More and more I feel that interdisciplinary areas are my favourites. At heart I think Im always looking for good stories that explain things I see in the world all around me, and I like those ones that can explain various aspect of my life at the same time. And I really don't want to limit myself to just one thing.
- energy, entropy, life, chaos, irreversibility.... these keywords are somehow just too attractive and magical to not be curious about.....maybe I have too much fantasy in my mind and should be more down to earth haha!
- The approach used in this area is quite nice: theory derivation + data analysis + computer simulation, and sometime cooperate with biology lab. Again, I like this kind of diverse approach.

4.3 About the physics theories in evolution

- I very like the work done by England's lab using the extended second law of thermodynamics, and actually it were their papers that led me into this area. Maybe the descriptions of those works I mentioned in this report sound quite speculative, but they are actually doing very solid work in my opinion. I like how brief and thought-provoking the equations are at the same time.
- The phase transitions tool is quite widely used as I noticed, but how it was applied was a bit "brutal" or "rough", or maybe a bit too simplified. It's not saying that phase transitions are useless, but I really need more time to understand it and its applications.
- The information theory of life looks quite abstract to me, maybe because my knowledge of information is very limited. Maybe I should learn some information theory, not just for understanding evolution, but mainly to upgrade myself to a more modern human being.
- The "geo-thermal stress theory" is the one amuses me the most and also upsets me sometimes. As essentially, according to my understanding, we (life) are just like the lightnings in the storm, or even just like...the "fart" of the earth.....on my god...it is a little bit hard to accept it is the truth that we are just helping the earth to release stress! But on the other hand, it is actually such an important mission, and I have more excuse now to eat more so that I can help to release more trapped energy for the earth!

4.4 About this directed individual study

- It was quite nice during the first several weeks when I explored different areas in biophysics, but then suddenly became quite frustrating when I started to look into the details of NESM in Evolution, as it has been quite hard to find nice general and introductory material for this topic, especially NESM, and almost no one near me is familiar with this topic.
- However, I'm truly feeling happy now and even a bit proud of myself after all the frustrations. It actually feels quite nice after I summarised all the stuff I've read to this report, because the report can serve as a good "dictionary" for me in the future if I want to go further in this area! Thanks Petrick for the encouragement and support!!
- A major skill I learned from this process is how to gain high level understandings in a short time by skipping a lot of technical or logical details. In the beginning I was very uncomfortable with skipping details, as it was not the way we used to learn most of the time, or I would say, usually teachers will skip those details for us and tell us the high level understandings directly. But now I began to feel comfortable to do this process myself, though still need more practice.

References

[1] James C. Reid, Sarah J. Brookes, Denis J. Evans, and Debra J. Searles. The dissipation function: Its relationship to entropy production, theorems for nonequilibrium systems and observations on its extrema. In *Beyond the Second Law*, Understanding Complex Systems, pages 31–47. Springer, Berlin, Heidelberg, 2014.

[2] Michael Cross. Equilibrium versus nonequilibrium, 2008. `http://www.cmp.caltech.edu/~mcc/BNU/Notes1_2.pdf`, date accessed: 26/05/2018.

[3] Wikipedia. Fluctuation-dissipation theorem, 2018. `https://en.wikipedia.org/wiki/Fluctuation-dissipation_theorem`, Date accessed: 26/05/2018.

[4] Lars Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37:405–426, Feb 1931.

[5] Abigail Klopper. Topics in non-equilibrium physics. *Nature Physics*, 11:103, 2015-02-03.

[6] Ralf Eichhorn, Hans Fogedby, Alberto Imparato, and Carlos Meja-Monasterio. Foundations and applications of non-equilibrium statistical mechanics. *Physica Scripta*, 86(5):050201, 2012.

[7] Ritort Felix. *Nonequilibrium Fluctuations in Small Systems: From Physics to Biology*, chapter 2, pages 31–123. Wiley-Blackwell, 2008.

[8] Eliran Boksenbojm, Bram Wynants, and Christopher Jarzynski. Nonequilibrium thermodynamics at the microscale: Work relations and the second law. *Physica A: Statistical Mechanics and its Applications*, 389(20):4406 – 4417, 2010.

[9] Reid, J. C., Sevick, E. M., and Evans, D. J. A unified description of two theorems in non-equilibrium statistical mechanics: The fluctuation theorem and the work relation. *Europhys. Lett.*, 72(5):726–732, 2005.

[10] C. Jarzynskia. Nonequilibrium work relations: foundations and applications. *The European Physical Journal B*, 64(3):331–340, 2008-08.

[11] Sara Imari Walker. Origins of life: a problem for physics, a key issues review. *Reports on Progress in Physics*, 80(9):092601, 2017.

[12] Michele Clamp, Ben Fry, Mike Kamal, Xiaohui Xie, James Cuff, Michael F. Lin, Manolis Kellis, Kerstin Lindblad-Toh, and Eric S. Lander. Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences*, 104(49):19428–19433, 2007.

[13] Stephan Jay Gould. Humbled by the genome's mysteries, 2011. `https://www.nytimes.com/2001/02/19/opinion/humbled-by-the-genome-s-mysteries.html`, date accessed: 26/05/2018.

[14] Jordana Cepelewicz. Seeing emergent physics behind evolution, 2017. `https://www.quantamagazine.org/seeing-emergent-physics-behind-evolution-20170831/`, date accessed: 26/05/2018.

[15] Oliver J. Rando and Kevin J. Verstrepen. Timescales of genetic and epigenetic inheritance. *Cell*, 128(4):655–668, 2007-02-23.

[16] Gary M. Dunny. The peptide pheromone-inducible conjugation system of enterococcus faecalis plasmid pCF10: cellcell signalling, gene transfer, complexity and evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1483):1185–1193, 2007-07-29.

[17] Rodrigo S. Galhardo, P. J. Hastings, and Susan M. Rosenberg. Mutation as a stress response and the regulation of evolvability. *Critical Reviews in Biochemistry and Molecular Biology*, 42(5):399–435, 2007-01-01.

[18] WikiPedia. Extended evolutionary synthesis, 2018. `http://www.cmp.caltech.edu/~mcc/BNU/Notes1_2.pdf`, date accessed: 26/05/2018.

[19] Nigel Goldenfeld and Carl Woese. Life is physics: Evolution as a collective phenomenon far from equilibrium. *Annual Review of Condensed Matter Physics*, 2(1):375–399, 2011.

[20] John Rennie. How brain waves surf sound waves to process speech, 2018. `https://www.quantamagazine.org/neuroscience-critics-learn-how-brain-waves-link-to-speech-20180522/`, date accessed: 26/05/2018.

[21] J. C. Wang. Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine. *Quarterly Reviews of Biophysics*, 31(2):107–144, 1998-05.

[22] James C. Wang. A journey in the world of DNA rings and beyond. *Annual Review of Biochemistry*, 78:31–54, 2009.

[23] Hong-Yan Shih, Tsung-Lin Hsieh, and Nigel Goldenfeld. Ecological collapse and the emergence of travelling waves at the onset of shear turbulence. *Nature Physics*, 12(3):245–248, 2016-03.

[24] Jeremy England. No turning back: The nonequilibrium statistical thermodynamics of becoming (and remaining) life-like. *MIT Physics Colloquium*, 2017. `https://www.youtube.com/watch?v=10cVVHKCRWw`, Date accessed: 21/05/2018.

[25] Mark Buchanan. Dissipate to replicate. *Nature Physics*, 11:104, 2015-02-03.

[26] Jeremy L. England. Statistical physics of self-replication. *The Journal of Chemical Physics*, 139(12):121923, 2013-08-21.

[27] Jeremy L. England. Dissipative adaptation in driven self-assembly. *Nature Nanotechnology*, 10(11):919–923, 2015-11.

[28] Robert Marsland III and Jeremy England. Far-from-equilibrium distribution from near-steady-state work fluctuations. *Physical Review E*, 92(5), 2015-11-16.

[29] Jordan M. Horowitz and Jeremy L. England. Spontaneous fine-tuning to environment in many-species chemical reaction networks. *Proceedings of the National Academy of Sciences*, 114(29):7565–7570, 2017-07-18.

[30] Tal Kachman, Jeremy A. Owen, and Jeremy L. England. Self-organized resonance during search of a diverse chemical space. *Physical Review Letters*, 119(3):038001, 2017-07-21.

[31] Jeremy England. Why trees dont ungrow, 2017. https://aeon.co/essays/does-the-flow-of-heat-help-us-understand-the-origin-of-life, date accessed: 26/05/2018.

[32] Yukio Saito and Hiroyuki Hyuga. Complete homochirality induced by nonlinear autocatalysis and recycling. *Journal of the Physical Society of Japan*, 73(1):33–35, 2004-01-15.

[33] Marcelo Gleiser and Sara Imari Walker. An extended model for the evolution of prebiotic homochirality: A bottom-up approach to the origin of life. *Origins of Life and Evolution of Biospheres*, 38(4):293, 2008-08-01.

[34] P. G. H. Sandars. A toy model for the generation of homochirality during polymerization. *Origins of life and evolution of the biosphere*, 33(6):575–587, 2003-12-01.

[35] Kalin Vetsigian and Nigel Goldenfeld. Global divergence of microbial genome sequences mediated by propagating fronts. *Proceedings of the National Academy of Sciences*, 102(20):7332–7337, 2005-05-17.

[36] Richard A. Neher and Boris I. Shraiman. Competition between recombination and epistasis can cause a transition from allele to genotype selection. *Proceedings of the National Academy of Sciences*, 106(16):6866–6871, 2009-04-21.

[37] Nigel Goldenfeld and Carl Woese. Biology's next revolution. *Nature*, 445:369, 2007-01-24.

[38] George F. R. Ellis. Top-down causation and emergence: some comments on mechanisms. *Interface Focus*, 2(1):126–140, 2012-02-06.

[39] G. Auletta, G. F. R. Ellis, and L. Jaeger. Top-down causation by information control: from a philosophical problem to a scientific research programme. *Journal of The Royal Society Interface*, 5(27):1159–1172, 2008-10-06.

[40] Sara Imari Walker and Paul C. W. Davies. The algorithmic origins of life. *Journal of The Royal Society Interface*, 10(79):20120869, 2013-02-06.

[41] Everett L. Shock and Eric S. Boyd. Principles of geobiochemistry. *Elements*, 11(6):395, 2015.

[42] Morowitz Harold and Smith Eric. Energy flow and the organization of life. *Complexity*, 13(1):51–59, 2007.

[43] Decker K., Jungermann K., and Thauer R. K. Energy production in anaerobic organisms. *Angewandte Chemie International Edition in English*, 9(2):138–158, 1970.