The Generalized Arrhenius law in out of equilibrium systems

Serena Bradde\textsuperscript{1} and Giulio Biroli\textsuperscript{2}

\textsuperscript{1} Computational Biology program, Memorial Sloan-Kettering Cancer Center, New York, USA and Physics of Biological Systems, Institut Pasteur, 75724, Paris Cedex 15, France

\textsuperscript{2} Institut de Physique Théorique (IPhT) CEA Saclay and CNRS URA 2306, 91191 Gif Sur Yvette, France

In this work we provide a comprehensive analysis of the activation problem out of equilibrium. We generalize the Arrhenius law for systems driven by non conservative time independent forces, subjected to retarded friction and non-Markovian noise. The role of the energy function is now played by the out of equilibrium potential $\phi = -\lim_{T\to 0} T \log P_s$, with $P_s$ being the steady state probability distribution and $T$ the strength of the noise. We unveil the relationship between the generalized Arrhenius law and a time-reversal transformation discussed in the context of fluctuations theorems out of equilibrium. Moreover, we characterize the noise-activated trajectories by obtaining their explicit expressions and identifying their irreversible nature. Finally, we discuss a real biological application that illustrates our results.

PACS numbers: 05. 40. -a, 05. 70. -a, 05. 10. Gg

The Arrhenius law originally discovered for the rate of chemical reactions \cite{1} is one of the most important principles governing the behavior of systems characterized by several energy scales. The diffusivity of vacancies in crystals \cite{2}, the viscosity of strong super-cooled liquids \cite{3}, the rate of protein folding \cite{4} are just few examples showing the broad range of physical situations in which the Arrhenius law holds. This happens whenever the dynamics of the system at hand is dominated by a process requiring an activation energy $\Delta E$ much larger than the temperature. In this case, the time-scale characterizing the dynamical behavior is given by

$$\tau \simeq \tau_0 \exp \left( \frac{\Delta E}{T} \right),$$

thus leading to the same behavior for several observables related to the relaxation time (rates, diffusion coefficients, etc.). In this equation $\tau_0$ is, roughly speaking, the high-temperature time-scale \cite{3}.

The usual domain of applicability of the Arrhenius law is restricted to systems at equilibrium (or in a metastable equilibrium, as in the case of nucleation processes). In this case the Arrhenius law can be derived in several ways. The original proof is due to Kramers \cite{6}, who obtained it for a particle undergoing Langevin dynamics with white noise. A natural question is to what extent this can be generalized to non-equilibrium systems, which are nowadays witnessing a growing interest, e.g. gently shaken granular media, sheared fluids at low temperature, active matter, single molecule experiments, biochemical networks. These systems are out of equilibrium because are driven by forces that do not derive from a potential and are subjected to noise that does not necessarily correspond to a thermal bath at equilibrium.

If the force field is time-independent then a stationary probability distribution is generally reached at long time, but it is not given by the Gibbs-Boltzmann distribution $\exp(-E/T)$ since for non-conservative force fields there is no energy function to start with. In consequence, the usual intuition behind the Arrhenius law, which is based on a stochastic dynamics in an energy landscape characterized by a barrier $\Delta E$, becomes meaningless. Yet, experiments, such the ones of \cite{7} on shaken granular media, suggest that in the low noise regime a generalization of this law does exist. Despite some results obtained in specific contexts \cite{6,8,9,10,11,12,13,14}, a general and comprehensive analysis addressing this issue is still lacking. Here we fill this gap: for driven systems subjected to non-thermal and in general multiplicative non-Markovian noise we derive the generalization of the Arrhenius law and identify the corresponding rare noise-activated paths followed during the dynamics. Our theory applies to force fields characterized by more than one stable attractor and in the small noise limit, meaning that the timescale to ”jump” from one attractor to the other is much larger than the characteristic timescales to vibrate around them. As we shall explain, our results can be understood physically in terms of a generalization of time-reversal symmetry discovered in the context of fluctuations theorems \cite{14,15}. In order to illustrate our results we provide an explicit example inspired by bacteria evolution in the intestine caused by antibiotic administration \cite{17}.

We start our analysis by focusing on a simplified model defined by the Langevin equation:

$$\gamma \frac{dx}{dt} = F + \eta$$

(1)

where $F$ is the force field, $\gamma$ is the viscosity, $\eta$ a Gaussian multiplicative noise, possibly corresponding to thermal fluctuations. We assume that $\eta$ has zero mean and variance $\langle \eta(t)\eta(t') \rangle = 2T\gamma \delta(t-t')\delta_{\alpha,\beta}$ where $H(x)$ is a generic positive function of $x$ and $T$ is the strength of the noise fluctuations (it only coincides with temperature for an equilibrated bath). The force field $F$ is non-conservative therefore the system is kept out of equilibrium: it dissipates heat with the reservoir and does (or receives) work because of the force field. More general
systems containing an inertial term, retarded friction and non-Markovian noise will be considered afterward. In order to study a simple but instructive case of activated non-equilibrium dynamics we focus on a force field characterized by three fixed points corresponding to $F = 0$ (two stable, $S_1, S_2$, and one unstable $U$). This is the counterpart of the usual process, a jump through an energy barrier, studied to derive the “equilibrium” Arrhenius law. Our derivation is based on the Martin-Siggia-Rose field theory [18], whose action using Ito convention is:

$$S = -\int_{t_i}^{t_f} dt \dot{x} \cdot \left( \frac{dx}{dt} - F(x) - TH(x) \dot{x} \right)$$

where $\dot{x}$ is the response field and we use time units such that $\gamma = 1$.

In the low noise limit, $T \to 0$, the dynamics can be described in the following way. The system first flows following the force field toward one of the stable points, whether this is $S_1$ or $S_2$ depends on the starting point, and then fluctuates locally around it. On much larger times, rare configurations of the noise eventually induce far away excursions allowing the system to escape from one basin of attraction to the other. The characterization of these dynamical processes can be obtained by computing the probability $P(S_1 \to S_2; t)$ that a system starting in $S_1$ at time $t_i$ is in $S_2$ at a very large time $t_f = t + t_i$. This is given by sum over all paths $[x(t)]$ that connect these two points in a time $t$. Each trajectory is weighted by the exponential of its corresponding action. In the low noise limit it is straightforward to check that the sum over paths is dominated by the saddle point contributions because one can pull out a $1/T$ factor in front of $S$ by redefining $\dot{x} \to \dot{x}/T$. The functional integral is thus approximated by the sum of the paths that verify:

$$\frac{\delta S}{\delta x_\alpha} = 0 : \frac{dx_\alpha}{dt} = -\sum_\beta \left[ \frac{\partial F_\beta}{\partial x_\alpha} \dot{x}_\beta + T \dot{x}_\beta \frac{\partial H(x)}{\partial x_\alpha} \right]$$

$$\frac{\delta S}{\delta \dot{x}_\alpha} = 0 : \frac{dx_\alpha}{dt} - F_\alpha = 2TH(x) \dot{x}_\alpha$$

(2)

In the equilibrium case, the solution of the saddle point equations can be constructed in terms of the downhill trajectory, that corresponds to the path going from the unstable point towards one of the stable ones, and the uphill trajectory that corresponds to a path going in the opposite direction. The extremal paths on which one has to sum to obtain $P(S_1 \to S_2; t)$ are all possible combinations of the uphill and the downhill ones concatenated in such a way to verify the boundary conditions $x(t_i) = x_{S_1}$ and $x(t_f) = x_{S_2}$. The detailed computation, which is analogous to the so called dilute instanton gas, is presented in [19]. In order to extend it to the non-equilibrium case we look for the generalization of the downhill and the uphill solutions. The former is immediately found: it corresponds to $\dot{x} = F$ and $\dot{x} = 0$ and leads to a null action. This means that the corresponding weight is of the order of one, as expected physically, since the downhill trajectory is a typical one and does not need to be activated by the noise. The main problem is to find the uphill trajectory going, say, from $S_1$ to $U$. In equilibrium conditions, corresponding to $F = -\nabla E$ and $H = 1$, the uphill trajectory is the time-reversed of the downhill one and reads $\dot{x} = \nabla E$ and $\dot{x} = x/T = \nabla E/T$. The corresponding action $S = -(E(U) - E(S_1))/T = -\Delta E/T$ leads to an Arrhenius weight, $e^{-\Delta E/T}$, for the uphill solution.

The naive guess for the generalization to the non-equilibrium case where one substitutes $-\nabla E$ with $F$, leading to $\dot{x} = -F$ and $\dot{x} = x/T$, does not work. In order to find the solution of this conundrum, we have to introduce the zero noise non-equilibrium potential $\phi = \lim_{T \to 0} -T \log P_s(x)$, with $P_s$ being the stationary probability distribution. In mathematical language $\phi$ is the large deviation function determining the probability of rare events in the zero noise limit. Plugging $P_s \simeq e^{-\phi/T}$ into the Fokker-Planck equation one finds that $\phi$ verifies the equation

$$\nabla \phi \cdot (F + H(x) \nabla \phi) + \mathcal{O}(T) = 0$$

(3)

Since $\phi$ somehow plays the role of the energy function, a natural generalization of the uphill solution can be found replacing $E$ with $\phi$ in the expression for $\dot{x}$, i.e. $\dot{x} = \nabla \phi/T$. Given the second equation of (2), this choice leads to:

$$\dot{x} = F + 2H(x) \nabla \phi \quad \dot{x} = \nabla \phi/T$$

(4)

which indeed provides a solution also for first equation of (2), as it can be checked by direct substitution and by using the derivative of eq. (3). The action associated to the uphill solution reads:

$$S = -\frac{1}{T} \int dt \nabla \phi \cdot \dot{x} + \nabla \phi \cdot (F + H(x) \nabla \phi)$$

(5)

Noticing that the last term between parenthesis is zero because of (3), we are left with the integral of a total derivative and, therefore, $S = -\Delta \phi/T$. Hence, we indeed obtain the generalized Arrhenius weight, $e^{-\Delta \phi/T}$, for the uphill solution. By combining together uphill and downhill solutions, as done in [19], one finds that

$$P(S_1 \to S_2; t) = P_s(x_{S_2}) \left( 1 - e^{-\frac{\Delta \phi}{T}} \right)$$

$$\tau(T) \simeq \tau_0 \exp \left( \frac{\Delta \phi}{T} \right)$$

where $\tau_0$ contains the sub-leading contributions to the generalized Arrhenius law coming from the Gaussian fluctuations around the instantons. On the basis of the results of [9] we expect that an explicit computation should lead to the result $\tau_0^{-1} = |\lambda_0|/(2\pi)|\prod_i(\omega_i^0/|\omega_i^0|)^{1/2}$
where $\lambda_0$ is the real part of the lowest eigenvalue of $\partial F_\alpha/\partial x_\beta$ evaluated in $U$ and the $\omega^U_i$s and the $\omega^S_i$s are the eigenvalues of the Hessian of $\phi$ evaluated in $U$ and $S_1$ respectively.

Note that in general the solution of a system of equations like (3) with fixed initial and final conditions is not straightforward at all, and can be only obtained numerically by the shooting method (one searches by trial and error the initial position and velocity that leads to the correct solution). Thus, being able to provide the explicit solution and, in top of that, finding that the action is the integral of a total derivative are quite unexpected simplifications. In equilibrium, they are due to the existence of the time-reversal symmetry [16]. Remarkably, a generalization of the time-reversal transformation, which was introduced in the context of fluctuations theorems out of equilibrium [12, 14–16], provides a physical explanation for the non-equilibrium case too. This transformation gives a relationship between the probability of a dynamical path and its time reversal in the so called adjoint dynamics corresponding to the Langevin evolution in a renormalized force field $F_A = -F - 2H(x)\nabla \phi$:

$$P([x(t)]; S_1 \rightarrow U) = P_A([x(-t)]; U \rightarrow S_1) \ e^{-\frac{\Delta \phi}{T}}. \quad (6)$$

In this expression $P_A$ denotes the probability of paths evolving with the adjoint dynamics [20]. In order to identify the uphill trajectory one has to maximize the LHS. Instead of solving this difficult problem, one can solve the much easier one consisting in maximizing the RHS associated to the downhill path in the adjoint dynamics. Since the exponential term is path-independent (the boundary conditions are fixed), the most probable path for $T \rightarrow 0$ is just the one that follows the force field $F_A = -F - 2H(x)\nabla \phi$ going from the unstable point $U$ to the stable one $S_1$ [21]. Because of the identity (3), the time-reversal of this path is the one maximizing the LHS and, indeed, corresponds to the uphill solution. In conclusion, the generalization of the time reversal transformation provides the physical principle behind the generalization of the Arrhenius law and an explanation for the unexpected simplifications we discovered.

An important difference with the equilibrium case is that the large deviation function $\phi$ is not known a priori. Thus, it may seem that the previous results are of little use in practice. However, the explicit integration of eqs. in (4) between $S_1$ (or $S_2$) and $U$ and the evaluation of the corresponding action allows one to explicitly obtain $\Delta \phi$. Actually, this procedure is more general and allows one to completely reconstruct $\phi(x)$ up to a normalization constant. In order to do it, one can partition the phase space in basins of attraction associated to each stable point. The non-equilibrium potential for a point belonging to the basin of, say, $S_1$ is given by $\phi(x_{S_1})$ plus the action evaluated on the trajectory that connects $S_1$ to this point. This can be understood by noticing that under the time-reversal mapping the action $S$ transforms into the action for the downhill solution, which is null, plus $\Delta \phi/T$ between the final and the initial points. By comparing the values of $\Delta \phi$ for points along the separatrix one can completely reconstruct $\phi(x)$ up to a normalization constant. Mathematically, this is related to an underlying Hamilton-Jacobi theory that applies to eqs. like the ones in [2] [15] [22].

In order to show an explicit example of our theoretical framework, we focus on a biological-inspired example related to the study of the time evolution induced by an antibiotic treatment of bacterial microbes living in the human intestine [17]. This dynamical system was modeled in terms of an over-damped Langevin equation characterized by white noise and the non-conservative force field $F(x,y) = (\frac{x}{x+y} - cx, \frac{y}{x+y} - \psi xy - y)$. The variables $x$ and $y$ describe the concentrations of the bacteria which are respectively sensitive and tolerant to the antibiotic. The constant $c, \psi$ and $\epsilon$ are related to mortality rate, fitness and interactions between bacteria populations. This system exhibits a region of bistability for a particular choice of the parameters with two stable points $S_1 = (0, 1)$ and $S_2 = (1/\epsilon, 0)$, and an unstable one $U = (c\epsilon^{-1}, \frac{\psi+\epsilon(1-c\epsilon)}{\epsilon})$. In order to illustrate our theory, we compute the non-equilibrium potential $\phi$ by evaluating the action for all the extremal trajectories starting from stable points (Fig.1). It is double-well shaped with two minima associated to $S_1$ and $S_2$ and one maximum in correspondence to $U$. Then we characterize the optimal paths that interpolate between the two stable points, see Fig 2. The most striking difference with the equilibrium case is that uphill and downhill trajectories pass through different regions of the phase space. The reason for this can be understood by recalling that out-of-equilibrium systems substantiate a steady state probability current, $J = (F - T \nabla)P_x$. In the zero noise limit $J$ equals $(F + \nabla \phi)P_x$ and is orthogonal everywhere to the gradi-
and the system by the force field: $W_{\text{in}}$ and $W_{\text{out}}$ concerns the work done along the direction given by the probability current and downhill trajectories can be respectively decomposed in two perpendicular contributions: $\dot{x} = \gamma(t - t') \hat{J}_\alpha$ and $\hat{J}_\beta$ is a symmetric positive-definite operator associated to non-Markovian noise [25]. Physically, eq. (7) corresponds to a system which undergoes Newtonian dynamics in a non-potential force field and is coupled to an out of equilibrium thermal bath (only in the case $H = 1$ and for $\gamma_{\alpha,\beta}(t - t') = \gamma_{\alpha,\beta}$ the bath is at equilibrium). The main trick we used to analyze this case consists in reducing the problem to the one we already solved. This is done by showing that the general stochastic equation above can be rewritten in terms of overdamped Langevin equations characterized by white noise in an extended configuration space where new extra variables are introduced: the inertial term is handled by introducing the momentum, $\mathbf{p} = m \dot{x}$; whereas retarded friction and non-Markovian noise can be represented as the result of integrating out a bath of harmonic oscillators linearly coupled to the system and evolving by an over-damped Langevin equation characterized by white noise [26]. More details and the explicit calculations can be found in the SI text [27].

In conclusion, we provided a general and comprehensive analysis of the problem of activation out of equilibrium. We showed that the Arrhenius law holds for a very large class of out of equilibrium systems provided that the energy function is replaced by the non-equilibrium potential $\phi$. The most important difference with the equilibrium case is that the noise-activated paths are no longer related by a simple time reversal transformation: uphill and downhill paths visit different regions of the configuration space. We characterized these trajectories by obtaining their explicit expression, by identifying their irreversible nature and by unveiling how they are related to a generalization of the time-reversal transformation. We illustrated our results in an explicit example borrowed from biology where equilibrium is not even a well defined concept. We envision many possible interesting applications in different fields ranging from physics to economics, where equilibrium is a limiting assumption.

We thank J.-P. Bouchaud, J. Kurchan, M. Marsili and A. Silva for helpful discussions. GB acknowledges support from ERC grant NPRG-GLASS. SB acknowledges CEA-Saclay for hospitality.

Let us finally discuss the most general physical case we are able to treat, which is characterized by the following system of stochastic equations:

$$m \ddot{x}_\alpha + \sum_\beta \int_{-\infty}^{t} dt' \gamma_{\alpha,\beta}(t - t') \dot{x}_\beta = F_\alpha + \eta_\alpha$$

$$\langle \eta_\alpha(t) \eta_\beta(t') \rangle = 2T \delta_{\alpha,\beta}(t - t') H(x)$$

(7)

where $\gamma_{\alpha,\beta}(t - t')$ is a function related to retarded friction, symmetric with respect to the interchange of $\alpha, \beta$ and equal to zero for $t < t'$ and $\hat{\gamma}_{\alpha,\beta}(t - t')$ is a symmetric positive-definite operator associated to non-Markovian noise [25]. Physically, eq. (7) corresponds to a system which undergoes Newtonian dynamics in a non-potential force field and is coupled to an out of equilibrium thermal bath (only in the case $H = 1$ and for $\gamma_{\alpha,\beta}(t - t') + \gamma_{\alpha,\beta}(t' - t) = 2\gamma_{\alpha,\beta}$ the bath is at equilibrium). The main trick we used to analyze this case consists in reducing the problem to the one we already solved. This is done by showing that the general stochastic equation above can be rewritten in terms of overdamped Langevin equations characterized by white noise in an extended configuration space where new extra variables are introduced: the inertial term is handled by introducing the momentum, $\mathbf{p} = m \dot{x}$; whereas retarded friction and non-Markovian noise can be represented as the result of integrating out a bath of harmonic oscillators linearly coupled to the system and evolving by an overdamped Langevin equation characterized by white noise [26]. More details and the explicit calculations can be found in the SI text [27].

In conclusion, we provided a general and comprehensive analysis of the problem of activation out of equilibrium. We showed that the Arrhenius law holds for a very large class of out of equilibrium systems provided that the energy function is replaced by the non-equilibrium potential $\phi$. The most important difference with the equilibrium case is that the noise-activated paths are no longer related by a simple time reversal transformation: uphill and downhill paths visit different regions of the configuration space. We characterized these trajectories by obtaining their explicit expression, by identifying their irreversible nature and by unveiling how they are related to a generalization of the time-reversal transformation. We illustrated our results in an explicit example borrowed from biology where equilibrium is not even a well defined concept. We envision many possible interesting applications in different fields ranging from physics to economics, where equilibrium is a limiting assumption.

We thank J.-P. Bouchaud, J. Kurchan, M. Marsili and A. Silva for helpful discussions. GB acknowledges support from ERC grant NPRG-GLASS. SB acknowledges CEA-Saclay for hospitality.

[1] R. D. Levine. Molecular Reaction Dynamics (Cambridge University Press, 2005).
[2] R. J. Borg and G. J. Dienes, *An introduction to solid state diffusion* (Academic Press, 1988).
[3] L. Berthier and G. Biroli, Rev. Mod. Phys., 83, 587 (2011).
[4] M. L. Scalley and D. Baker, PNAS, 94, 10636 (1997).
[5] For simplicity of notations we henceforth set $k_B = 1$.
[6] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys., 62, 251 (1990).
[7] G. D’Anna and G. Gremaud, Nature, 413, 407 (2001).
[8] A. J. Bray and A. J. McKane, Phys. Rev. Lett., 62, 493 (1989).
[9] D. Borgis and M. Moreau, Physica A, 163, 877 (1990).
[10] C. Gardiner, *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences* (Springer Series in Synergetics, 2004).
[11] J. Prost, J.-F. Joanny, and J. M. R. Parrondo, Phys. Rev. Lett., 103, 090601 (2009).
[12] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, J. Stat. Mech., 2006, P08001 (2006).
[13] J. Wang, L. Xu, and E. Wang, PNAS, 105, 12271 (2008).
[14] T. Hatano and S. Sasa, Phys. Rev. Lett., 86, 3463 (2001).
[15] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, Phys. Rev. Lett., 87, 040601 (2001).
[16] J. Kurchan, *Lecture Notes of the Les Houches Summer School*, edited by T. Dauxois, S. Ruffo, and L. Cugliandolo, Lecture Notes of the Les Houches Summer School. (Oxford University Press, 2010) Chap. 2.
[17] V. Bucci, S. Bradde, G. Biroli, and J. Xavier, Plos. Comp. Bio. in press (2012).
[18] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford University Press, 2002).
[19] B. Caroli, C. Caroli, and B. Roulet, *Journal of Statistical Physics*, 26, 83 (1981).
[20] This relationship is more general and valid also at finite temperature if one replaces $\phi$ with $-T\log P_s$. Moreover, the initial and final point do not need to be stationary points.
[21] Stable (unstable) points with the standard dynamics remain stable (unstable) points for the adjoint dynamics. This can be proved by showing that the Hessian matrices at the critical points for the standard dynamics and its adjoint have the same eigenvalues.
[22] R. Graham, *Stochastic Nonlinear Systems in Physics, Chemistry, and Biology*, edited by L. Arnold and R. Lefever (Springer, 1981) p. 209.
[23] A simple and instructive case, suggested to us by J.-P. Bouchaud, is provided by a force field $F = -\nabla V + \nabla \times A$ such that rotational and irrotational contributions are orthogonal for any value of $x$. This property means that the rotational part makes the system flows along the iso-potential surfaces. In this case one can obtain explicitly several results discussed in the text. First, it is easy to check that the steady state distribution is given by the Boltzmann-like distribution $e^{-V/T}$ despite the fact that the system is out of equilibrium. Second, the steady state probability current reads $J/P_s = \nabla \times A$. In this case the decomposition of downhill and uphill trajectories discussed previously is very natural and follows directly from the definition of $F$.
[24] U. Seifert, Phys. Rev. Lett., 95, 040602 (2005).
[25] The other property verified by $\gamma_{\alpha,\beta}(t-t')$ is that $\gamma_{\alpha,\beta}(t-t') + \gamma_{\alpha,\beta}(t'-t)$ is a symmetric positive-definite operator, see the SI text.
[26] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, USA, 2011).
[27] See SI text.