Generalized statistical mechanics for superstatistical systems

Christian Beck
School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

Mesoscopic systems in a slowly fluctuating environment are often well described by superstatistical models. We develop a generalized statistical mechanics formalism for superstatistical systems, by mapping the superstatistical complex system onto a system of ordinary statistical mechanics with modified energy levels. We also briefly review recent examples of applications of the superstatistics concept for three very different subject areas, namely train delay statistics, turbulent tracer dynamics, and cancer survival statistics.

I. INTRODUCTION

Nonlinear dynamical processes often create a fluctuating environment for a given mesoscopic system [1]. This leads to mixing of the mesoscopic dynamics and that of the environment. If there is sufficient time scale separation, then very often superstatistical models yield a good effective description. The superstatistics concept has established itself as a powerful tool to describe quite general classes of complex systems [1–11, 22–26]. The basic idea is to characterize the complex system under consideration by a superposition of several statistics on different time scales, for example one corresponding to ordinary statistical mechanics (on a mesoscopic level modelled by a Langevin equation) and the other one corresponding to a slowly varying inverse temperature field \( \beta(\vec{x}, t) \) or some other relevant parameter.

There may be either spatial or temporal variations of the environment. The environment is represented by a suitable parameter entering the stochastic differential equation describing the mesoscopic system. The superstatistics concept can be applied in quite a general way, and a couple of interesting applications for a variety of complex systems have been pointed out recently [12–21, 27, 28]. Essential for this approach is the existence of sufficient time scale separation so that the system has enough time to relax to a local equilibrium state and stay within it for some time.

The stationary distributions of superstatistical systems, obtained by averaging over all \( \beta \), typically exhibit non-Gaussian behavior with fat tails, which can be a power law, or a stretched exponential, or other functional forms as well [4]. In general, the superstatistical parameter \( \beta \) need not to be an inverse temperature but can be an effective parameter in a stochastic differential equation, a volatility in finance, or just a local variance parameter extracted from some experimental time series. There are interesting applications in hydrodynamic turbulence [2, 20, 29, 30], for defect turbulence [12], for cosmic rays [13] and other scattering processes in high energy physics [31, 32], share price fluctuations [15, 27, 33, 34], solar flares [14], wind velocity fluctuations [17, 18], hydro-climatic fluctuations [19], the statistics of train departure delays [38] and survival statistics of cancer patients [39]. Maximum entropy principles can be generalized in a suitable way to yield the relevant probability distributions that characterize the various important universality classes in superstatistics [5, 40–43].

In this paper we shall develop a new theoretical approach to superstatistics, by formally mapping the superstatistical system onto a system of ordinary statistical mechanics where the energy levels are modified in a suitable way. This approach yields a new interesting theoretical tool to further develop the generalized statistical mechanics of superstatistical complex systems, and is described in detail in section 3. We also briefly review some recent examples of applications of superstatistical techniques. Our three examples, all from very different subject areas, are train delays on the British railway network, velocity signals in hydrodynamic turbulence, and the survival statistics of cancer patients.

II. REMINDER: WHAT IS SUPERSTATISTICS?

The concept is best illustrated by starting with a particular example of superstatistics, in fact the one that was considered first in [10, 11]. Consider the following well-known formula:

\[
\int_0^\infty d\beta f(\beta)e^{-\beta E} = \frac{1}{(1 + (q - 1)\beta_0 E)^{(q-1)/2}}
\]

where

\[
f(\beta) = \frac{1}{\Gamma\left(\frac{1}{q-1}\right)} \left(\frac{1}{(q-1)\beta_0}\right)^{\frac{q-1}{q-1}} \beta^{-\frac{1}{q-1}} \exp\left\{-\frac{\beta}{(q-1)\beta_0}\right\}
\]

(2)
is the $\chi^2$ (or $\Gamma$) probability distribution and $\beta_0$ and $q$ are parameters ($q > 1$).

We see that averaged ordinary Boltzmann factors $e^{-\beta E}$ with $\chi^2$-distributed $\beta$ yield an effective Boltzmann factor of $q$-exponential form, given by the right-hand side of eq. (1). The physical interpretation is that nonequilibrium systems with temperature fluctuations give rise to an effective description in terms of more general Boltzmann factors. In [10, 11] the $\chi^2$-distribution was advocated for $f(\beta)$, because at that time the aim was to better understand $q$-statistics [44] from a dynamical point of view. General $f(\beta)$ were then suggested in [1]. In that paper also the name ‘superstatistics’ was created. This name was simply an abbreviation for the fact that there is a superposition of two (or several) statistics. In no way this name wants to indicate that this type of statistics is ‘superior’ to others.

One can also construct dynamical realizations of superstatistics in terms of Langevin equations with parameters that fluctuate on large time scales [11]. These local Langevin equations describe the mesoscopic system under consideration. The situation is sketched in Fig. 1. The simplest example would be locally a linear Langevin equation

$$\dot{v} = -\gamma v + \sigma L(t)$$

with slowly fluctuating parameters $\gamma, \sigma$. Here $L(t)$ denotes Gaussian white noise. This describes the velocity $v$ of a Brownian particle that moves through spatial 'cells' with different local $\beta := \gamma/(2\sigma^2)$ in each cell (a nonequilibrium situation). If some probability distribution $f(\beta)$ of the inverse temperature $\beta$ for the various cells is given, then the conditional probability given some fixed $\beta$ in a given cell is Gaussian, $p(v|\beta) \sim e^{-\frac{1}{2}\beta v^2}$, the joint probability is $p(v, \beta) = f(\beta)p(v|\beta)$ and the marginal probability is $p(v) = \int_0^\infty f(\beta)p(v|\beta)d\beta$. Integration over $\beta$ yields effectively Boltzmann factors that are more general than Gaussian distributions, which depend on the specific properties of $f(\beta)$. If there are only finitely many cells, then the integral is understood to approximate the average over a large number of cells.

The principal idea of superstatistics is to generalize this example to much broader systems. For example, $\beta$ need not be an inverse temperature but can in principle be any intensive parameter. Most importantly, one can generalize to general probability densities $f(\beta)$ and general Hamiltonians. In all cases one obtains a superposition of two different statistics: that of $\beta$ and that of ordinary statistical mechanics. Superstatistics hence describes complex nonequilibrium systems with spatio-temporal fluctuations of an intensive parameter on a large scale. The effective Boltzmann factors $B(E)$ for such systems are given by

$$B(E) = \int_0^\infty f(\beta)e^{-\beta E}d\beta.$$

A lot of research has been done in this direction in recent years. If there is locally Gaussian behaviour, then the theory of superstatistics is clearly related to the theory of Gaussian scale mixtures. More generally one can prove a superstatistical generalization of fluctuation theorems [32], develop a variational principle for the large-energy
asymptotics of general superstatistics \[4\], proceed to generalize entropies for general superstatistics \[5, 40, 43\], let the \(q\)-values in eq. (1) fluctuate as well \[7\], and prove superstatistical versions of a Central Limit Theorem \[8\]. There are also relations with fractional reaction equations \[47\], random matrix theory \[16, 28, 35\], networks \[36\], and path integrals \[6\]. Very useful for practical applications is a superstatistical approach to time series analysis \[2, 18, 27\]. Applications have been pointed out for 3d hydrodynamic turbulence \[2, 20, 29, 30\], wind velocity fluctuations \[17, 18\], finance and economics \[21, 33, 34, 48\], blinking quantum dots \[49\], cosmic ray statistics \[13\] and quite generally scattering processes in particle physics \[31, 32\]. The concept has also been useful to analyze hydroclimatic fluctuations \[19\] as well as the statistics of train delays on the British railway network \[38\]. There are also medical applications \[39\].

III. MAPPING SUPERSTATISTICS ONTO CONVENTIONAL STATISTICAL MECHANICS

Consider a system of ordinary statistical mechanics with energy levels \(E_i\) of microstate \(i\). We are looking at a canonical ensemble and \textit{a priori} the inverse temperature \(\beta\) is fixed. Now look at identical copies of the system but with different temperatures \(\beta_j\) in each spatial cell \(j\), at a given snapshot of time. This is a nonequilibrium situation.

Let \(\beta_0 = \int_0^\infty f(\beta) \beta d\beta\) be the average inverse temperature. We may formally consider a super-Hamiltonian describing the entire system which in the different spatial cells has effective energy levels \(\tilde{E}_i\), by writing

\[
\beta_0 \tilde{E}_i = \beta_j E_i.
\]

(5)

Apparently this means the super-Hamiltonian has energy levels \(\tilde{E}_i(j)\) given by

\[
\tilde{E}_i(j) = \frac{\beta_j}{\beta_0} E_i.
\]

(6)

in cell \(j\).

Since ordinary statistical mechanics is valid for \textit{arbitrary} energy levels, in particular also for the \(\tilde{E}_i(j)\), we may now do ordinary statistical mechanics for the super-Hamiltonian and introduce the partition function \(Z(\beta_0)\) of the entire system as

\[
Z(\beta_0) = \sum_{j,i} e^{-\beta_0 \tilde{E}_i(j)} = \sum_{j,i} e^{-\beta_j E_i} \approx \int_0^\infty f(\beta) e^{-\beta E_i} d\beta,
\]

(7)

(8)

(9)

where in the last step the sum over \(j\) is approximated by an integral. Since \(Z(\beta_0) = \sum_{j,i} e^{-\beta_0 \tilde{E}_i(j)}\) is an ordinary partition function (though with exotic, locally modified, energy levels), it is now possible to do ordinary statistical mechanics for this superstatistical nonequilibrium system, with all the known formulas.

We regard the free energy of the superstatistical system as a function of the mean inverse temperature \(\beta_0\) and define it as

\[
F(\beta_0) = -\frac{1}{\beta_0} \log Z(\beta_0).
\]

(10)

In the statistical mechanics formalism it is often convenient to work with the function \(\Psi(\beta_0) := \beta_0 F(\beta_0)\). Defining \(kT_0 = \frac{1}{\beta_0}\) one has

\[
F = U - T_0 S \quad \Psi = \beta_0 U - S
\]

(11)

(12)

But one has to be careful here what the meaning of the symbols \(U\) and \(S\) is: \(U\) is now the mean energy of the energy levels \(\tilde{E}_i\), rather than \(E_i\), and indeed this means that \(U\) is a global mean energy corresponding to the entire superstatistical system consisting of many cells. One has

\[
U = \sum_{j,i} p_i^{(j)} \tilde{E}_i(j)
\]

(13)
\[ S = -k \sum_{j,i} p_i^{(j)} \log p_i^{(j)} \]
\[ = -k \sum_{j,i} \frac{1}{Z(\beta_0)} e^{-\beta_0 \tilde{E}_i^{(j)}} (-\log Z(\beta_0) - \beta_0 \tilde{E}_i^{(j)}) \]
\[ = \beta_0 U + \log Z(\beta_0) \]

In this way we have formally mapped the superstatistical nonequilibrium system onto an (exotic) equilibrium system of ordinary statistical mechanics with average inverse temperature \( \beta_0 \) and a new type of Hamiltonian, corresponding to the energy levels \( \tilde{E}_i^{(j)} \). We should remark that the above idea of mapping superstatistics onto the statistical mechanics of an exotic Hamiltonian is completely new and different from previous attempts of developing a generalized statistical mechanics for superstatistical systems \[\text{[5, 27, 40, 41]}\].

### IV. POSSIBLE SUPERSTATISTICAL DISTRIBUTIONS \( f(\beta) \)

The distribution \( f(\beta) \) is determined by the dynamical large-scale structure of the complex system under consideration. There have been attempts to derive the specific form of \( f(\beta) \) relevant for a given complex system with given constraints from a generalized maximum entropy principle. We don’t elucidate this further here but refer to \[\text{[43]}\] and references therein for further details. Actually, what we want to do here is to proceed to practical applications. The relevant question is what type of \( f(\beta) \) are typically seen for experimental data as generated by a generic complex system.

There seem to be three different superstatistics that are of utmost importance \[\text{[2]}\]. These are (a) \( \chi^2 \)-superstatistics (= Tsallis statistics), (b) inverse \( \chi^2 \)-superstatistics, and (c) lognormal superstatistics.

In case (a), \( f(\beta) \) is given by the Gamma distribution

\[ f(\beta) = \frac{1}{\Gamma(n/2)} \left( \frac{n}{2\beta_0} \right)^{n/2} \beta^{n/2-1} e^{-n\beta/2\beta_0}, \]

where again \( \beta_0 \) is the average of \( \beta \). This generates generalized Boltzmann factors \( B(E) \) that decay with a power law. \( n \) is a parameter characterizing the number of degrees of freedom.

In case (b), \( f(\beta) \) is given by

\[ f(\beta) = \frac{\beta_0}{\Gamma\left(\frac{n}{2}\right)} \left( \frac{n\beta_0}{2} \right)^{n/2} \beta^{-n/2-2} e^{-n\beta_0/2\beta}. \]

In this case the generalized Boltzmann factors \( B(E) \sim \int f(\beta) e^{-\beta E} \) decay as \( e^{-\beta\sqrt{E}} \) for large \( E \).

Finally, in case (c) \( f(\beta) \) is given by the lognormal distribution

\[ f(\beta) = \frac{1}{\sqrt{2\pi}s\beta} \exp \left\{ \frac{-(\ln \beta - \mu)^2}{2s^2} \right\}, \]

where \( \mu \) and \( s \) are suitable parameters. In the remaining sections, we briefly describe one example for each of these three different cases.

### V. TRAIN DEPARTURE DELAYS

Traffic delays on the British railway network are reasonably well described by \( \chi^2 \)-superstatistics. The probability density of observed train departure delays of length \( t \) has been analyzed in detail in \[\text{[38]}\]. Millions of departure times
were automatically stored and evaluated. The 0th-order theoretical model for the waiting time $t$ is a Poisson process which predicts that the waiting time distribution until the train finally departs is $P(t|\beta) = \beta e^{-\beta t}$, where $\beta$ is some parameter. But this does not agree with the actually observed data \[38\]. A much better fit is given by a $q$-exponential, see Fig. 2.

What may cause this power law that fits the data? The idea is that there are fluctuations in the parameter $\beta$ as well. These fluctuations describe large-scale temporal or spatial variations of the British rail network environment, which take place on a much larger time scale than the actual train departures. $\beta$-fluctuations are e.g. produced at the begin of the holiday season with lots of passengers, or if there are problems with the track or bad weather conditions. Also there can be extreme events such as derailments, industrial action, terror alerts, etc. The observed long-term distribution of train delays is then a mixture of exponential distributions where the parameter $\beta$ fluctuates:

$$p(t) = \int_0^\infty f(\beta)p(t|\beta)d\beta = \int_0^\infty f(\beta)\beta e^{-\beta t}.$$  \quad (22)

For a $\chi^2$-distributed $\beta$ with $n$ degrees of freedom one obtains

$$p(t) = C \cdot \left(1 + b(q - 1)t\right)^{-\frac{1}{q-1}},$$  \quad (23)

where $q = 1 + 2/(n + 2)$, $b = 2\beta_0/(2 - q)$ and $C$ is a normalization constant.

VI. TURBULENT FLOWS

Various aspects of hydrodynamic turbulence are quite well described by lognormal superstatistics. In this case the mesoscopic local dynamics corresponds to a single tracer particle that is advected by the turbulent flow. The environment of the tracer particle changes. For a while it will see regions of strong turbulent activity, then it will move
of given vorticity surrounding the test particle. Further details are described in [30].

One has β T by the time scale ratio \( \tau \) which agrees very well with experimentally measured data of the acceleration statistics (Fig. 3). In 'Lagrangian turbulence' one is interested in the statistics of velocity differences \( \vec{u}(t) := \vec{u}(t + \tau) - \vec{u}(t) \) of the particle on a small time scale \( \tau \). For \( \tau \to 0 \) this velocity difference becomes the local acceleration \( \vec{a}(t) = \vec{u}(t)/\tau \). A superstatistical Lagrangian model for 3-d velocity differences of the tracer particle has been developed in [30]. The mesoscopic dynamics is a superstatistical Langevin equation of the form

\[
\vec{a} = -\gamma \vec{u} + B \vec{n} \times \vec{u} + \sigma \vec{L}(t).
\]

Here \( \gamma \) and \( B \) are constants. The term proportional to \( B \) introduces some rotational movement of the particle, simulating vortices in the flow. The noise strength \( \sigma \) and the unit vector \( \vec{n} \) evolve stochastically on a large time scale \( T_\sigma \) and \( T_B \), respectively. \( T_\sigma \) is of the same order of magnitude as the Kolmogorov time scale \( \tau_\eta \). In this model the Reynolds number \( R_\lambda \) is basically given by the time scale ratio \( T_\sigma \gamma / T_L / \tau_\eta \sim R_\lambda \gg 1 \). The time scale \( T_B \gg \tau_\eta \) describes the average life time of a region of given vorticity surrounding the test particle. Further details are described in [30].

The parameter \( \beta \) is again defined to be \( \beta := 2\gamma/\sigma^2 \), but it does not have the meaning of a physical inverse temperature in the flow. Rather, one has \( \beta = 1/\nu^1/2 \langle \epsilon \rangle^{-1/2} \), where \( \nu \) is the kinematic viscosity and \( \langle \epsilon \rangle \) is the average energy dissipation. \( \epsilon \) is known to fluctuate in turbulent flows. Kolmogorov’s theory of 1961 suggests a lognormal distribution for \( \epsilon \), which leads us to lognormal superstatistics. For very small \( \tau \) the 1-d acceleration component of the particle is given by \( a_x = u_x/\tau \) and one gets out of the model the 1-point distribution

\[
p(a_x) = \frac{\tau}{2\pi s} \int_0^\infty d\beta \beta^{-1/2} \exp \left\{ \frac{-\log \beta}{2s^2} \right\} e^{-\frac{1}{2}\beta^2 \sigma^2}.
\]

which agrees very well with experimentally measured data of the acceleration statistics (Fig. 3).

The 3-d superstatistical model of [30] predicts correlations between the three acceleration components. An intrinsic property of the model is that the acceleration \( a_x \) in \( x \) direction is not statistically independent of the acceleration \( a_y \) in \( y \)-direction. We may study the ratio \( R := p(a_x, a_y)/(p(a_x)p(a_y)) \) of the joint probability \( p(a_x, a_y) \) to the product of 1-point probabilities \( p(a_x) \) and \( p(a_y) \). For independent acceleration components this ratio would always be given by \( R = 1 \), whereas the 3-d superstatistical model yields the prediction

\[
R = \frac{\int_0^\infty \beta f(\beta)e^{-\frac{1}{2}\beta^2 \sigma^2(a_x^2 + a_y^2)}d\beta}{\int_0^\infty \beta^1/2 f(\beta)e^{-\frac{1}{2}\beta^2 \sigma^2 a_x^2}d\beta \int_0^\infty \beta^1/2 f(\beta)e^{-\frac{1}{2}\beta^2 \sigma^2 a_y^2}d\beta}.
\]

The trivial result \( R = 1 \) is obtained only for \( f(\beta) = \delta(\beta - \beta_0) \), i.e. no fluctuations in the parameter \( \beta \) at all. Fig. 4 shows \( R := p(a_x, a_y)/(p(a_x)p(a_y)) \) as predicted by lognormal superstatistics: Experimental measurements of acceleration correlations yield very similar results to those predicted above [30, 51].
VII. SURVIVAL STATISTICS OF CANCER PATIENTS

Data of the survival statistics of cancer patients can be well fitted using models based on inverse $\chi^2$-superstatistics. A superstatistical model of the progression of metastasis and the corresponding survival statistics of cancer patients has been developed in [39]. The final result that comes out of that model is the following prediction for the probability density function of survival time $t$ of a randomly chosen patient that is diagnosed with cancer at $t = 0$:

$$p(t) = \int_0^\infty t^{n-1} \lambda^n e^{-\lambda t} \frac{\lambda_0(n\lambda_0/2)^{n/2}}{\Gamma(n/2)} \lambda^{-n/2-2} e^{-\lambda_0 t/2} d\lambda$$  \hspace{1cm} (27)

This can also be written as

$$p(t) = \frac{(n\lambda_0)^{3n/4}}{\Gamma(n/2)\Gamma(n/2)} \left( \frac{t}{2} \right)^{3n/4} \left[ \frac{\sqrt{2n\lambda_0 t}}{n} K_{n/2+1} \left( \sqrt{2n\lambda_0 t} \right) ight]$$

$$- K_{n/2} \left( \sqrt{2n\lambda_0 t} \right),$$  \hspace{1cm} (28)

where $K_\nu(z)$ is the modified Bessel function. Note that this is inverse $\chi^2$-superstatistics. The role of the parameter $\beta$ is now played by the parameter $\lambda$, which in a sense describes how aggressively the cancer propagates. This parameter has fluctuations from patient to patient.

The above formula based on inverse $\chi^2$-superstatistics is in good agreement with real data of the survival statistics of breast cancer patients in the US. The superstatistical formula fits the observed distribution very well, both in a linear and logarithmic plot (see Fig.5).

When looking at the time scales in the above figures one should keep in mind that the data shown are survival distributions conditioned on the fact that death occurs from cancer. Many patients, in particular if they are diagnosed at an early stage and treated accordingly, will live a long healthy life and will die from something else than cancer. These cases are not included in the data shown above.

VIII. CONCLUSION AND OUTLOOK

In this paper we have dealt with mesoscopic and other complex systems that are embedded into a temporally changing or spatially fluctuating environment. If there is sufficient time scale separation, then a mixture of different statistics (a superstatistical description) is an appropriate method to describe these types of complex systems.
FIG. 5: Survival time statistics of breast cancer patients diagnosed with cancer at \( t = 0 \), both in a linear and double logarithmic plot. The solid line is the superstatistical model prediction \[39\].

Basically this means that one mixes ordinary statistical mechanics with another statistics of e.g. local temperature fluctuations.

In the first part of the paper we have pointed out how superstatistical complex systems can be mapped onto systems of ordinary statistical mechanics. The key point is that one deforms the effective energy levels in a suitable way and then applies the well-known techniques of ordinary statistical mechanics to this (exotic type of) super-Hamiltonian.

In the second part of the paper we summarized a few recent applications of the superstatistical approach to real-world problems, which covered quite a range of different subject areas. We studied train delay statistics, turbulent tracer dynamics, and survival statistics of cancer patients. Many other areas may benefit from a generalized statistical mechanics formalism for superstatistical systems as well.

In a year where there is a good chance to finally experimentally confirm the long-awaited Higgs particle, it might be appropriate to end this paper by mentioning that scattering processes in high energy physics can also be well-described by superstatistical models. The experimentally observed power laws of differential cross sections and energy spectra at very high energies have been modelled in terms of superstatistical generalized statistical mechanics \[13, 31, 32, 52\]. Superstatistical techniques have also been recently used to describe the space-time foam in string theory \[53\], and a generalized statistical mechanics model underlying chaotic types of vacuum fluctuations yields a Higgs mass prediction of 154 GeV \[54, 55\]. It seems there is a lot of scope for relevant contributions of generalized statistical mechanics in high energy physics and quantum field theory.
Acknowledgements

I am very grateful to Dr. Hugo Touchette for providing Fig. 1.