Solving noncomputable problems
using quantum-classical hybrid.

Michail Zak
Jet Propulsion Laboratory
California Institute of Technology

Abstract.

The challenge of this paper is to relate quantum-inspired dynamics represented by a self-supervised system, to solutions of noncomputable problems. In the self-supervised systems, the role of actuators is played by the probability produced by the corresponding Liouville equation. Following the Madelung equation that belongs to this class, non-Newtonian properties such as randomness, entanglement, and probability interference typical for quantum systems have been described in [1]. It has been demonstrated there, that such systems exist in the ***mathematical*** world: they are presented by ODE coupled with their Liouville equation, but they belong neither to Newtonian nor to quantum physics. The central point of this paper is the application of the self-supervised systems to finding global maximum of functions that is no-where differential, but everywhere continuous (such as Weierstrass functions)

1.Introduction.

The basic idea of this paper is to exploit a new kind of dynamical systems that would preserve superposition of random solutions, while allowing one to measure its state variables using classical methods. In other words, such a hybrid system would reinforce the advantages and minimize limitations of both quantum and classical aspects. The central point is the application of the self-supervised systems to finding global maximum of functions that is no-where differential, but everywhere continuous (such as Weierstrass functions). We will start this section with brief description of self-supervised systems that were introduced and discussed in our previous publications [1,2,3,4]

In order to illuminate specific features of self-supervised systems, we will start with control dynamics that described by a system of ODE:

$$\frac{d\mathbf{v}}{dt} = \mathbf{F}[\mathbf{v}, U] \qquad (1)$$

Here

$\mathbf{v} = v_1, v_2, \ldots v_n$ is the vector of state variables to be controlled,

$u = u_1, u_2, \ldots u_m$ is the control vector that represents ***external*** actuators.

Let us compare the control system Eq. (1) with the following system

$$\frac{d\mathbf{v}}{dt} = \mathbf{F}[\rho(\mathbf{v})] \qquad (2)$$

where the probability ρ is introduced via the Liouville equation corresponding to Eq. (2)

$$\frac{\partial \rho}{\partial t} + \nabla \bullet (\rho \mathbf{F}) = 0 \qquad (3)$$

It describes the continuity of the probability density flow originated by the error distribution

$$\rho_0 = \rho(t = 0) \qquad (4)$$

in the initial condition of ODE (3).

Comparison of Eqs.(1) and (2) shows that they have similar structure, and the role of the external actuator U in the control system (1) is played by the term $\rho(\mathbf{v})$ in the system (2). However the origins of these actuators are fundamentally different: the actuator U represents an external force, while the actuator $\rho(\mathbf{v})$ is an internal one. Indeed it is defined by Eq. (3) that, in turn, uniquely follows from Eq. (2). That is why the system (2),(3) can be called self- controlled, or self-supervised.

From the physical viewpoint, the feedback from the Liouville equation is a fundamental step in our approach: in Newtonian dynamics, the probability never explicitly enters the equation of motion. In addition to that, the Liouville equation generated by Eq. (2) is nonlinear with respect to the probability density ρ

$$\frac{\partial \rho}{\partial t} + \nabla \bullet \{\rho \mathbf{F}[\rho(\mathbf{V})]\} = 0 \qquad (5)$$

and therefore, the system (2),(3) departs from Newtonian dynamics. However although it has the same topology as quantum mechanics (since now the equation of motion is coupled with the equation of continuity of probability density as it does in the Madelung version of the Schrödinger equation), it does not belong to it either. Indeed Eq. (2) is more general than the Hamilton-Jacoby equation: it is not necessarily conservative, and **F** is not necessarily the quantum potential although further we will impose some restriction upon it that links **F** to the concept of information. The relation of the system (2), (3) to Newtonian and quantum physics is illustrated in Fig.1.

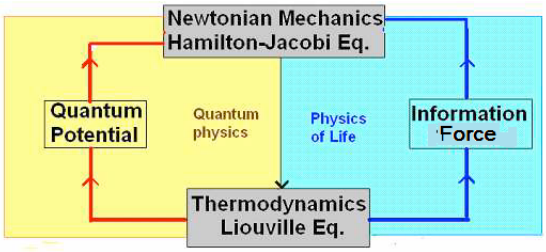


Figure 1. Classic Physics, Quantum Physics and Physics of Life

Remark. Here and below we make distinction between the random *variable v(t)* and its *values V* in probability space.

2. Selected self-supervised dynamical system.

In this section we will concentrate on a special type of the self-supervised system Eqs. (2),(3).

We will start with derivation of an auxiliary result that illuminates departure from Newtonian dynamics. For mathematical clarity, we will consider here a one-dimensional motion of a unit mass under action of a force *f* depending upon the *velocity v* and time *t* and present it in a dimensionless form

$$\dot{v} = f(v,t) \qquad (6)$$

referring all the variables to their representative values $v_0, t_0, etc.$

If initial conditions are not deterministic, and their probability density is given in the form

$$\rho_0 = \rho_0(V), \quad where\, \rho \geq 0, \; and \quad \int_{-\infty}^{\infty} \rho\, dV = 1 \qquad (7)$$

while ρ is a *single- valued* function, then the evolution of this density is expressed by the corresponding Liouville equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial V}(\rho f) = 0 \qquad (8)$$

The solution of this equation subject to initial conditions and normalization constraints (7) determines probability density as a function of *V* and *t*:

$$\rho = \rho(V,t) \qquad (9)$$

In order to deal with the constraint (7) let us integrate Eq. (8) over the whole space assuming that $\rho \to 0$ at $|V| \to \infty$ and $|f| < \infty$. Then

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} \rho\, dV = 0, \qquad \int_{-\infty}^{\infty} \rho\, dV = const, \qquad (10)$$

Hence, the constraint (7) is satisfied for $t > 0$ if it is satisfied for $t = 0$.

Let us now specify the force f as a feedback from the Liouville equation

$$f(v,t) = \phi[\rho(v,t)] \qquad (11)$$

and analyze the motion after substituting the force (11) into Eq.(6)

$$\dot{v} = \phi[\rho(v,t)], \qquad (12)$$

Although the theory of ODE does not impose any restrictions upon the force as a function of space coordinates, the Newtonian physics does: equations of motion are never coupled with the corresponding Liouville equation. Moreover, it can be shown that such a coupling leads to non-Newtonian properties of the underlying model. Indeed, substituting the force f from Eq. (11) into Eq. (8), one arrives at the *nonlinear* equation of evolution of the probability density

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial V}\{\rho\phi[\rho(V,t)]\} = 0 \qquad (13)$$

Let us now demonstrate the destabilizing effect of the feedback (11). For that purpose, it should be noticed that the derivative $\partial\rho/\partial v$ must change its sign at least once, within the interval $-\infty < v < \infty$, in order to satisfy the normalization constraint (7).

But since

$$Sign\frac{\partial \dot{v}}{\partial v} = Sign\frac{d\phi}{d\rho}\, Sign\frac{\partial \rho}{\partial v} \qquad (14)$$

there will be regions of v where the motion is unstable, and this instability generates randomness with the probability distribution guided by the Liouville equation (13). It should be noticed that the condition (14) may lead to exponential or polynomial growth of v (in the last case the motion is called neutrally stable, however, as will be shown below, it causes the emergence of randomness as well if prior to the polynomial growth, the Lipchitz condition is violated).

3. Emergene of self-generated stochasticity.

In order to illustrate mathematical aspects of the concepts of Liouville feedback in systems under consideration as well as associated with it instability and randomness, let us take the feedback (11) in the
following form

$$f = \frac{\xi}{\rho(v,t)} \int_{-\infty}^{v} [\rho(\eta,t) - \rho^*(\eta)] d\eta \qquad (26)$$

Here $\rho^*(v)$ is a preset probability density satisfying the constraints (13), and ξ is a positive constant with dimensionality [1/sec]. As follows from Eq. (26), f has dimensionality of a force per unit mass that depends upon the probability density ρ, and therefore, it can be associated with the concept of information, so we will call it the *information force*. In this context, the coefficient ξ can be associated with the Planck constant that relates Newtonian and *information* forces. But since we are planning to deal with systems that belong to the macro-world, ξ must be of order of a viscose friction coefficient.

With the feedback (26), Eqs. (7) and (8) take the form, respectively

$$\dot{v} = \frac{\xi}{\rho(v,t)} \int_{-\infty}^{v} [\rho(\eta,t) - \rho^*(\eta)] d\eta \qquad (27)$$

$$\frac{\partial \rho}{\partial t} + \xi[\rho(t) - \rho^*] = 0 \qquad (28)$$

The last equation has the analytical solution

$$\rho = [(\rho_0 - \rho^*)e^{-\xi t} + \rho^*] \qquad (29)$$

Subject to the initial condition

$$\rho(t=0) = \rho_0 \qquad (30)$$

that satisfies the constraint (13).

This solution converges to a preset stationary distribution $\rho^*(V)$. Obviously the normalization condition for ρ is satisfied if it is satisfied for ρ_0 and $\rho*$. Indeed,

$$\int_{-\infty}^{\infty} \rho dV = [\int (\rho_0 - \rho^*) dV] e^{-\xi t} + \int_{-\infty}^{\infty} \rho^* dV = 1 \qquad (31)$$

Rewriting Eq. (29) in the form

$$\rho = \rho_0 e^{-\xi t} + \rho^*(1 - e^{-\xi t}) \qquad (32)$$

one observes that $\rho \geq 0$ at all $t \geq 0$ and $-\infty < V < \infty$.

As follows from Eq. (29), the solution of Eq. (28) has an attractor that is represented by the preset probability density $\rho^*(V)$. Substituting the solution (29) into Eq. (27), one arrives at the ODE that simulates the stochastic process with the probability distribution (29)

$$\dot{v} = \frac{\zeta e^{-\xi t}}{[\rho_0(v) - \rho^*(v)]e^{-\xi t} + \rho^*(v)} \int_{-\infty}^{v} [\rho_0(\eta) - \rho^*(\eta)]\ d\eta \tag{33}$$

It is reasonable to assume that the solution (4) starts with sharp initial condition

$$\rho_0(V) = \delta(V) \tag{34}$$

As a result of that assumption, all the randomness is supposed to be generated *only* by the controlled instability of Eq. (33). Substitution of Eq. (34) into Eq. (33) leads to two different domains of *v*: $v \neq 0$ and *v*=0 *where* the solution has two different forms, respectively

$$\int_{-\infty}^{v} \rho^*(\xi) d\xi = \left(\frac{C}{e^{-\xi t} - 1}\right)^{1/\xi}, \quad v \neq 0 \tag{35}$$

$$v \equiv 0 \tag{36}$$

Indeed, $\dot{v} = \dfrac{\zeta e^{-\xi t}}{\rho^*(v)(e^{-\xi t} - 1)} \int_{-\infty}^{v} \rho^*(\eta)] d\eta$

whence $\dfrac{\rho^*(v)}{\int_{-\infty}^{v} \rho^*(\eta)\, d\eta} dv = \dfrac{\zeta e^{-\xi t}}{e^{-\xi t} - 1} dt$. Therefore, $\ln \int_{-\infty}^{v} \rho^*(\eta) d\eta = \ln\left(\dfrac{C}{e^{-\xi t} - 1}\right)^{1/\xi}$

and that leads to Eq. (35) that presents an implicit expression for *v* as a function of time since ρ^* is the known function. Eq. (36) represents a singular solution, while Eq. (35) is a regular solution that includes arbitrary constant C. The regular solutions is discontinuous:

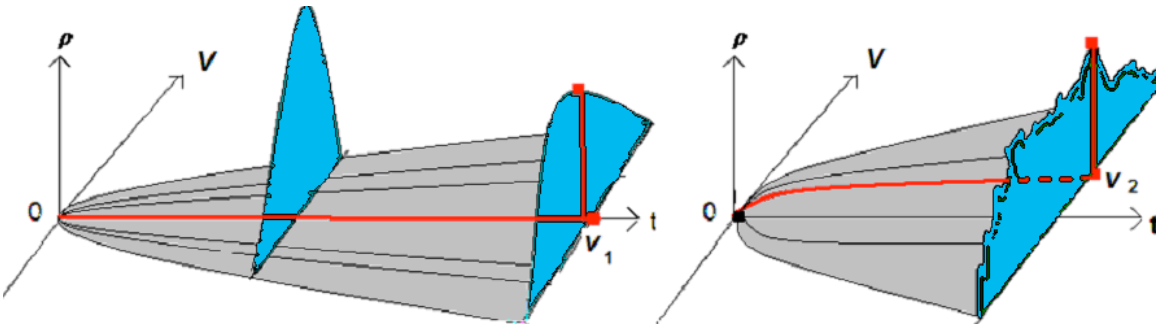
$$v \to \infty \quad at \quad t \to 0, \quad v = 0 \quad at \quad t = 0 \tag{37}$$

the Lipschitz condition is violated

$$\left|\frac{\partial \dot{v}}{\partial v}\right| \to \infty \qquad at \qquad t \to 0\ , \quad |v| \to 0 \tag{38}$$

and therefore, the uniqueness of the solution is lost thereby generating *randomness*.

As follows from Eq. (35), all the particular solutions for different values of *C* possess the same property (37), and that leads to non-uniqueness of the solution due to violation of the Lipchitz condition. Therefore, the same initial condition at $t \to 0$ yields infinite number of different solutions forming a family (35); each solution of this family appears with a certain probability guided by the corresponding Liouville equation (28). For instance, in cases plotted in Fig.2, a) and Fig.2, b), the "winner" solution is, respectively,



a) **b)**

Figure 2. Stochastic processes and their attractors.

$$v_1 = \varepsilon \to 0, \quad \rho(v_1) = \rho_{\max}, \; and \quad v = v_2, \qquad \rho(v_2) = \sup\{\rho\}$$

since it passes through the maximum of the probability density . However, with lower probabilities, other solutions of the same family can appear as well. Obviously, this is a non-classical effect. Qualitatively, this property is similar to those of quantum mechanics: the system keeps all the solutions simultaneously and displays each of them "by a chance", while that chance is controlled by the evolution of probability density (29).

Let us emphasize the connections between solutions of Eqs. (27) and (28): the solution of Eq. (27) is an one-parametrical family of trajectories (35), and each trajectory occurs with the probability described by the solution (32) of Eq. (28).). It should be recalled that the choice of displaying a certain solution is made only once, at *t=0*, i.e. when it departs from the deterministic to a random state; since than, it stays with this solution as long as the Liouville feedback is present.

Example1. Let us start with the following normal distribution

$$\rho^*(V) = \frac{1}{\sqrt{2\pi}} e^{-\frac{V^2}{2}} \tag{39}$$

Substituting the expression (39) and (34) into Eq. (35) at *V=v*, and $\xi = 1$ one obtains

$$v = erf^{-1}\left(\frac{C_1}{e^{-t} - 1}\right), \qquad v \neq 0 \tag{40}$$

Example 2. Let us choose the target density ρ^* as the Student's distribution, or so-called power law distribution

$$\rho^*(V) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{V^2}{\nu}\right)^{-(\nu+1)/2} \tag{41}$$

Substituting the expression (41) and (34) into Eq. (35) at *V=v*, ν=1, and $\xi = 1$ one obtains

$$v = \cot\left(\frac{C}{e^{-t} - 1}\right) \; for \qquad v \neq 0 \tag{42}$$

The 3D plot of the solutions of Eqs.(40) and (42), are presented in Figures 3a, and 3b, respectively.

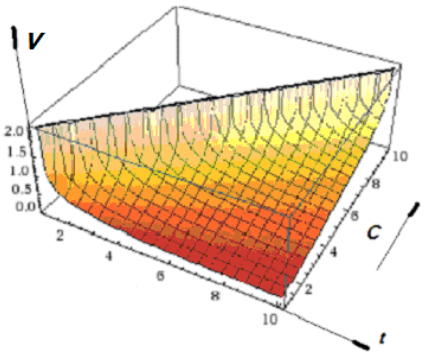


Figure 3a. Dynamics driving random events to normal distribution.

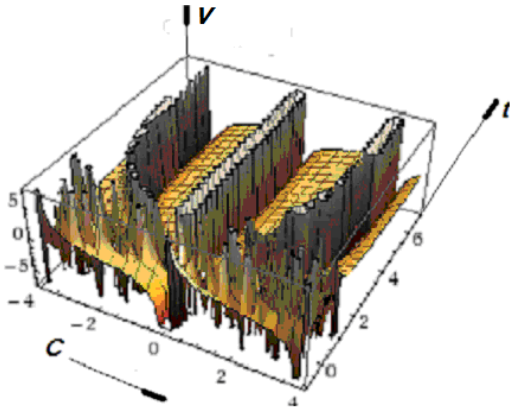


Figure 3b. Dynamics driving random events to power law.

The approach is generalized to *n*-dimensional case simply by replacing *v* with a vector $v = v_1, v_2, ...v_n$ since Eq. (28) does not include space derivatives

$$\dot{v}_i = \frac{\xi}{n\rho(\{v\},t)} \int_{-\infty}^{v_i} [\rho(\{\eta\},t) - \rho^*(\{\eta\})] d\eta_i \qquad (43)$$

$$\frac{\partial \rho(\{V\},t)}{\partial t} + \xi\rho(\{V\}t) - \rho^*(\{V\}) = 0 \qquad \textbf{(44)}$$

The idea of the proposed algorithm in more details is the following: introduce a *positive* function $\psi(v_1, v_2, ...v_n)$, $|v_i| < \infty$ to be maximized as the probability density $\rho^*(v_1, v_2, ...v_n)$ to which the solution of Eq. (44) is attracted.

Then the larger value of this function will have the higher probability to appear. The following steps are needed to implement this algorithm:

1. Build and implement the *n*-dimensional version of the model Eqs. (43), and (44), as an ***analog*** devise

$$\dot{v}_i = \frac{e^{-t}}{n\{[\rho_0(v) - \rho^*(v)]e^{-t} + \rho^*(v)\}} \int_{-\infty}^{v_i} [\rho_0(\xi) - \rho^*(\xi)]\, d\xi \quad , \qquad i = 1,2,..n. \quad (45)$$

2. Normalize the function to be maximized

$$\overline{\psi}(\{v\}) = \frac{\psi(\{v\})}{\int_{-\infty}^{\infty} \psi(\{v\}) d\{v\}} \qquad (46)$$

3. Using Eq. (32), evaluate time τ of approaching the stationary process to accuracy ε

$$\tau \approx \ln \frac{1-\bar{\psi}}{\varepsilon \bar{\psi}} \qquad (47)$$

4. Substitute $\bar{\psi}$ instead of ρ^* into Eqs. (45) and run the system during the time interval τ.

5. The solution will "collapse" into one of possible solutions with the probability $\bar{\psi}$. Observing (measuring) the corresponding values of $\{v^*\}$, find the first approximation to the optimal solution.

6. Switching the device to the initial state and then starting again, arrive at the next approximations.

7. The sequence of the approximations represents Bernoulli trials that exponentially improve the chances of the optimal solution to become a winner. Indeed, the probability of success ρ_s and failure ρ_f after the first trial is, respectively

$$\rho_s = \bar{\psi}_1, \qquad \rho_f = 1-\bar{\psi}_1 \qquad (48)$$

Then the probability of success after M trials is

$$\rho_{sM} = 1-(1-\bar{\psi})^M \to 1 \qquad at \qquad M \to \infty \qquad (49)$$

Therefore, after *polynomial* number of trials, one arrived at the solution to the problem. As follows from the above, the execution of the algorithm ***does not*** depend upon the dimensionality of the problem.

However despite several computational advantages of this algorithm over existing algorithms, the basic problem in question is the implementability of ***analog simulations*** using Newtonian/quantum resources. Indeed, the model described by Eq. (45) does not belong to physical space, as we know it: it belongs to the expanded quantum space, (see Fig. 1). This means that, in principle, the pure analog simulation of this algorithm is impossible unless some digital device is included, (see [1]). More fundamental mathematical approach that briefly described in Section 2 was performed in [8].

4. Solving noncomputable problems.

Let us now reformulate the problem in the way that would lead us to solution of noncomputable problems. For that purpose, let us introduce the following Weierstrass functions [5]

$$f_a(x) = \sum_{k=1}^{\infty} \frac{\sin(\pi k^a x)}{\pi k^a} \qquad (50)$$

(originally defined for $a=2$) that is continuous but differentiable only on a set of points of measure zero. The plots below (Wolfram MathWorld) show $f_a(x)$ for $a=2$ (red), 3 (green), and 4 (blue).

The global maximum of these functions cannot be found by gradient methods since these functions are non-differentiable; it cannot be found by combinatorial methods since these functions are continuous; if these functions are approximated by smooth polynomials, sharp picks could be missed. Meanwhile many financial market performance curves look similar to Weierstrass functions.

Now we will demonstrate that the proposed algorithm, which is based upon quantum-classical hybrid, can solve the problem. For that purpose we will follow all the steps of the algorithm applied to the function(50).

First we rewrite the function (50) in new notations and in the normalized form

$$\rho_a^*(v) = \frac{\sum_{k=1}^{\infty} \sin(\pi k^a v) / \pi k^a}{\int_{-\infty}^{\infty} \sum_{k=1}^{\infty} [\sin(\pi k^a v) / \pi k^a] dv} \tag{51}$$

and then substitute it into Eq. (33)

$$\dot{v} = \frac{\zeta e^{-\xi t}}{[\rho(v) - \rho_a^*(v)] e^{-\xi t} + \rho_a^*(v)} \int_{-\infty}^{v} [\rho_0(\eta) - \rho_a^*(\eta)] \, d\eta \tag{52}$$

while $\rho_0(V) = \delta(V)$ (53)

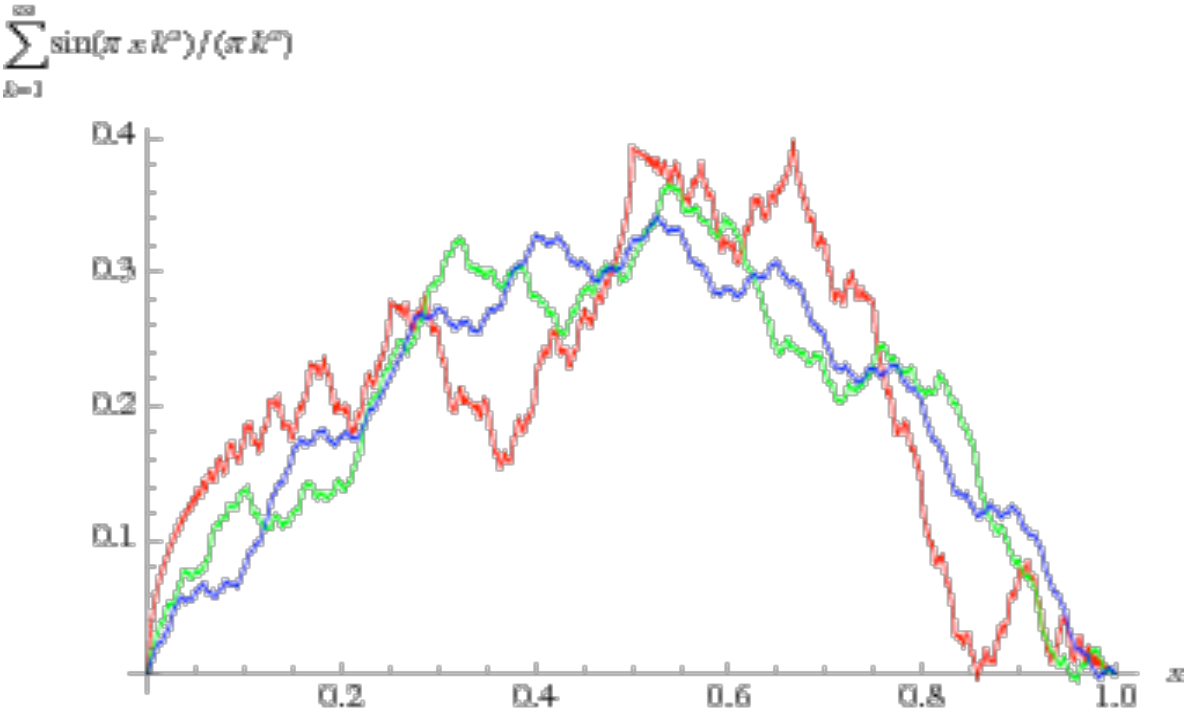


Figure 4. Weierstrass function.

It should be emphasized that the Weierstass functions are integrable , and that is why there is no problem with Eqs. (51) and (52).

As has been shown above, Eq. (52) has a one-parametric family of random solutions described by Eq.(35) and each trajectory occurs with the probability governed by the solution (29) of Eq. (28)

$$\rho(v) = \{[\delta(v) - \rho_a^*(v)] e^{-\xi t} + \rho_a^*(v)]\} \tag{54}$$

$$\frac{\partial \rho}{\partial t} + \xi[\rho(t) - \rho_a^*(v)] = 0 \tag{55}$$

As follows from Eqs.(54) and (55), the normalized Weierstrass function expressed by

Eq. (51) represents a static attractor in probability space: starting with the initial distribution in the form of delta-function, the probability distribution eventually approaches the Weierstrass function (51). But one should remember that both of

Eqs. (54) and (55) dwell in probability space and they are ***not*** a part of simulations: they only describe how Eq. (52) works: when Eq. (52) are run many times (in actual space), and statistics is collected, then Eq.(54) and its solution Eq. (55) demonstrate how this statistics changes in time. In addition to that, Eq. (55) guarantees uniqueness of the final probability distribution.

Following the steps of the proposed algorithm (see Eqs. (45) through (49)) one finds the optimal trajectory $v(\tau)$ that delivers global maximum to the target function Eq. (51) at

$t = \tau$, i.e.

$$\rho[v(\tau)] = \sup_v[\rho_a^*(v)] \qquad (56)$$

5. Conclusion.

The basic idea of this paper is to exploit a new kind of dynamical systems that would preserve superposition of random solutions, while allowing one to measure its state variables using classical methods. In other words, such a hybrid system would reinforce the advantages and minimize limitations of both quantum and classical aspects. The central point is the application of the self-supervised systems to finding global maximum of functions that are no-where differential, but everywhere continuous (such as Weierstrass functions). In the course of a brief discussion of the self-supervised systems, it was emphasized that the role of actuators there is played by the probability produced by the corresponding Liouville equation. Following the Madelung equation that belongs to this class, non-Newtonian properties such as randomness, entanglement, and probability interference typical for quantum systems have been described in [1]. It has been demonstrated there, that such systems exist in the ***mathematical*** world: they are presented by ODE coupled with their Liouville equation, but they belong neither to Newtonian nor to quantum physics. The idea of the proposed algorithm based upon self-supervised systems is the following: introduce a *positive* function $\psi(v_1, v_2, ...v_n), \ |v_i| < \infty$ to be maximized as the probability density $\rho^*(v_1, v_2, ...v_n)$ to which the solution of Eq. (44) is attracted in probability space. Then the larger value of this function will have the higher probability to appear. But since this probability is not necessarily 1, application of Bernoulli trail would lead to solution in polynomial number of steps. The advantages of this algorithm is the following: it does not depend upon the dimensionality of the problem; it does not require differentiability of the function to be maximized; it does not stop in local maxima. However, its main limitation is that this algorithm ***cannot be simulated*** neither by Newtonian, nor by quantum resources in the same way in which quantum algorithms cannot be simulated by Newtonian resources. However as shown in [1], the simulation of self-supervised systems as a quantum-classical hybrid can be arranged via quantum neural nets.

References

1. Zak,M., Non-Newtonian aspects of artificial intelligence. Foundations of Physics , Vol.46, May 2016
2. Zak, M., Entanglement in Livings, Journal of Quantum Information Science, 2012, 2, 66-77
3. Zak, M., Quantum-inspired maximizer, JOURNAL OF MATHEMATICAL PHYSICS **49,** 042702 ,2008
4. Zak, M., Randomness in physics revisited, Int. J. of quantum foundations, **2** (2016) 47 – 66
5. Zak,M.,Quantum-inspired maximizer, JOURNAL OF MATHEMATICAL PHYSICS **49,** 042702 ,2008