

# Optimal Control *Via* Self-Generated Stochasticity

Michail Zak\*

*Jet Propulsion Laboratory California Institute of Technology, Pasadena, CA 91109, USA*

**Abstract:** Stochastic approach to maximization of a functional constrained by governing equation of a controlled system is introduced and discussed. The idea of the proposed algorithm is the following: represent the functional to be maximized as a limit of a probability density governed by the appropriately selected Liouville equation. Then the corresponding ODE become stochastic, and that sample of the solution which has the largest value will have the highest probability to appear in ODE simulation. Application to optimal control is discussed. Two limitations of optimal control theory - local maxima and possible instability of the optimal solutions - are removed. Special attention is paid to robot motion planning.

## 1. INTRODUCTION

Optimal control theory is a mathematical method for deriving control policies. It deals with the problem of finding a control law for a given system such that a certain optimality criterion is achieved. The optimal control can be derived using Pontryagin's maximum principle (a necessary condition), or by solving the Hamilton-Jacobi-Bellman equation (a sufficient condition). From the viewpoint of analytical mechanics, the first approach is based upon canonical Hamilton's ODE, while the second approach can be linked to Hamilton-Jacobi PDE. Both approaches impose some restrictions upon the functional to be optimized, and one of them is its differentiability with respect to all the state variables. In case of a singular control when the maximum principle is indecisive, the existence of the second derivative is required. Notwithstanding undisputable progress in this area, there are still many limitations of modern approach, and one of them is the problem of local maxima. Mathematical roots of local maxima are the same as those for a much simpler problem of finding global maximum of a multi-dimensional function, but even this problem, in general, is not yet solved. The second unsolved problem is stability: even if the optimal trajectory is found, there is no guarantee that it is stable. This problem has not been solved even for the simpler case when the optimal trajectory is found from the classical Lagrange-Euler method. Strictly speaking, the criteria of stability (if they are available) should be included into a list of unilateral constraints imposed upon the functional to be minimized, and that makes the problem even harder.

This paper introduces a fundamentally new approach to optimal control based upon two new ideas.

The first idea is the following: represent the functional to be maximized as a limit of a probability density governed by the appropriately selected Liouville equation. Then the corresponding ODE become stochastic, and that sample of the solution which has the largest value will have the highest probability to appear in ODE simulation. This idea has been introduced and discussed in [1] being applied to finding the

global maximum of a multi-dimensional function subject to equality and inequality constraints. The main advantages of the stochastic approach are that it is not sensitive to local maxima, the function to be maximized must be only integrable, but not necessarily differentiable, and global equality and inequality constraints do not cause any significant obstacles.

The second idea is to remove possible instability of the optimal solution by equipping the control system with a self-stabilizing device.

Thus, the main challenge of this paper is to remove the problem of local maxima, and remove the problem of instability of optimal solution.

## Dynamical Model for Simulations

We will start this section with a brief review of models introduced and discussed in [1]. For mathematical clarity, we will consider a one-dimensional motion of a unit mass under action of a force  $f$  depending upon the velocity  $x$  and time  $t$

$$\dot{x} = f(x, t) \quad (1)$$

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

$$\rho_0 = \rho_0(X), \quad \text{where } \rho \geq 0, \quad \text{and} \quad \int_{-\infty}^{\infty} \rho dX = 1 \quad (2)$$

while  $\rho$  is a single-valued function, then the evolution of this density is expressed by the corresponding Liouville equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial X}(\rho f) = 0 \quad (3)$$

The solution of this equation subject to initial conditions and normalization constraints (2) determines probability density as a function of  $X$  and  $t$

$$\rho = \rho(X, t) \quad (4)$$

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

\*Address correspondence to this author at the Jet Propulsion Laboratory California Institute of Technology, Pasadena, CA 91109, USA; E-mail: Michail.Zak@jpl.nasa.gov

$$f(x,t) = \varphi[\rho(x,t)] \tag{5}$$

and analyze the motion after substituting the force (5) into Eq. (1)

$$\dot{x} = \varphi[\rho(x,t)], \tag{6}$$

**Remark**

Theory of stochastic differential equations makes distinction between the random variable  $x(t)$  and its values  $X$  in probability space.

This is a fundamental step in our approach. Although 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, as shown in [1], such a coupling leads to non-Newtonian properties of the underlying model. Indeed, substituting the force  $f$  from Eq. (5) into Eq. (3), one arrives at the *nonlinear* and, in general, *non-reversible* equation for evolution of the probability density

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial X} \{ \rho \varphi[\rho(X,t)] \} = 0 \tag{7}$$

Now we will demonstrate the destabilizing effect of the feedback (6). For that purpose, it should be noted that the derivative  $\partial \rho / \partial x$  must change its sign, at least once, within the interval  $-\infty < x < \infty$ , in order to satisfy the normalization constraint (2). But since

$$Sign \frac{\partial \dot{x}}{\partial x} = Sign \frac{d\varphi}{d\rho} Sign \frac{\partial \rho}{\partial x} \tag{8}$$

there will be regions of  $x$  where the motion is unstable, and this instability generates randomness with the probability distribution guided by the Liouville equation (7).

Let us consider Eqs. (1) and (7) defining  $f$  as the following function of probability density

$$f = \frac{1}{\rho(x,t)} \int_{-\infty}^x [\rho(\xi,t) - \rho^*(\xi)] d\xi \tag{9}$$

With the feedback (9), Eqs. (1) and (7) take the form, respectively

$$\dot{x} = \frac{1}{\rho(x,t)} \int_{-\infty}^x [\rho(\xi,t) - \rho^*(\xi)] d\xi \tag{10}$$

$$\frac{\partial \rho}{\partial t} + \rho(t) - \rho^* = 0 \tag{11}$$

The last equation has the analytical solution

$$\rho = (\rho_0 - \rho^*)e^{-t} + \rho^* \tag{12}$$

Subject to the initial condition

$$\rho(t = 0) = \rho_0 \tag{13}$$

this solution converges to a prescribed, or target, stationary distribution  $\rho^*(x)$ . Obviously the normalization condition for  $\rho$  is satisfied if it is satisfied for  $\rho_0$  and  $\rho^*$ .

Substituting the solution (12) in to Eq. (10), one arrives at the ODE that simulates the stochastic process with the probability distribution (12)

$$\dot{x} = \frac{e^{-t}}{[\rho_0(x) - \rho^*(x)]e^{-t} + \rho^*(x)} \int_{-\infty}^x [\rho_0(\xi) - \rho^*(\xi)] d\xi \tag{14}$$

As notices above, the randomness of the solution to Eq. (14) is caused by instability that is controlled by the corresponding Liouville equation. It should be emphasized that in order to run the stochastic process started with the initial distribution  $\rho_0$  and approaching a stationary process with the distribution  $\rho^*$ , one should substitute into Eq. (14) *analytical expressions* for these functions.

It is reasonable to assume that the solution (12) starts with sharp initial condition

$$\rho_0(X) = \delta(X) \tag{15}$$

As a result of that assumption, all the randomness is supposed to be generated *only* by the controlled instability of Eq. (14). Substitution of Eq. (15) into Eq. (14) leads to two different domains of  $x$ :  $x \neq 0$  and  $x = 0$ . The solution for the first domain is

$$\int_{-\infty}^x \rho^*(\xi) d\xi = \frac{C_1}{e^{-t} - 1}, \quad x \neq 0 \tag{16}$$

Indeed,

$$\dot{x} = \frac{e^{-t}}{\rho^*(x)(e^{-t} - 1)} \int_{-\infty}^x \rho^*(\xi) d\xi = \frac{e^{-t}}{\rho^*(x)(e^{-t} - 1)} \int_{-\infty}^x \rho^*(\xi) d\xi$$

whence

$$\frac{\rho^*(x)}{\int_{-\infty}^x \rho^*(\xi) d\xi} dx = \frac{e^{-t}}{e^{-t} - 1} dt$$

Therefore,

$$\ln \int_{-\infty}^x \rho^*(\xi) d\xi = \ln \frac{C}{e^{-t} - 1} \quad \text{and that leads to Eq. (16).}$$

The solution for the second domain is

$$x = 0 \tag{17}$$

Eq. (17) represents a singular solution, while Eq. (16) is a regular solution that includes arbitrary constant  $C$ . The regular solutions is unstable at  $t=0, |x| \rightarrow 0$  where the Lipschitz condition is violated

$$\frac{d\dot{x}}{dx} \rightarrow \infty \quad \text{at} \quad t \rightarrow 0, |x| \rightarrow 0 \quad (18)$$

and therefore, an initial error always grows generating randomness.

Let us analyze the behavior of the solution (16) in more details. As follows from this solution, all the particular solutions intersect at the same point  $x=0$  at  $t=0$ , and that leads to non-uniqueness of the solution due to violation of the Lipschitz condition. Therefore, the same initial condition  $x=0$  at  $t=0$  yields infinite number of different solutions forming a family (16); each solution of this family appears with a certain probability guided by the corresponding Liouville equation (11). For instance, in cases plotted in Fig. (1a) and Fig. (1b), the “winner” solution is, respectively,

$$x_1 = \varepsilon \rightarrow 0, \quad \rho(x_1) = \rho_{\max}, \quad \text{and} \quad x = x_2, \quad \rho(x_2) = \sup\{\rho\} \quad (19)$$

since it passes through the maximum of the probability density (11). However, with lower probabilities, other solutions of the family (13) 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 (11). It should be emphasized that in the ideal case, when no noise is present, the choice of displaying a certain solution is made by the system only once, at  $t=0$ , i.e. when it departs from a deterministic to a random state; since then, it stays with this solution as long as the Liouville feedback is present. However, strictly speaking, an actual realization of the trajectory may be affected by a non-

Lipschitz-originated instability at  $t=0$ ; as a result, small initial errors may grow exponentially, and the motion will be randomly deviated from the theoretical trajectory in such a way that a moving particle visits all the possible trajectories with the probability prescribed by the Liouville equation, (Fig. (1d)).

The approach is generalized to n-dimensional case simply by replacing  $x$  with a vector  $x = x_1, x_2, \dots, x_n$  since Eq. (11) does not include space derivatives.

**Remark**

The function  $f$  has been chosen in a special form (9) to remove the space derivative from the Liouville equation. The same procedure can be repeated for each variable, so Eq. (11) will formally remain the same with the only difference that for a multi-dimensional case  $\rho$  will stand for the joint probability density.

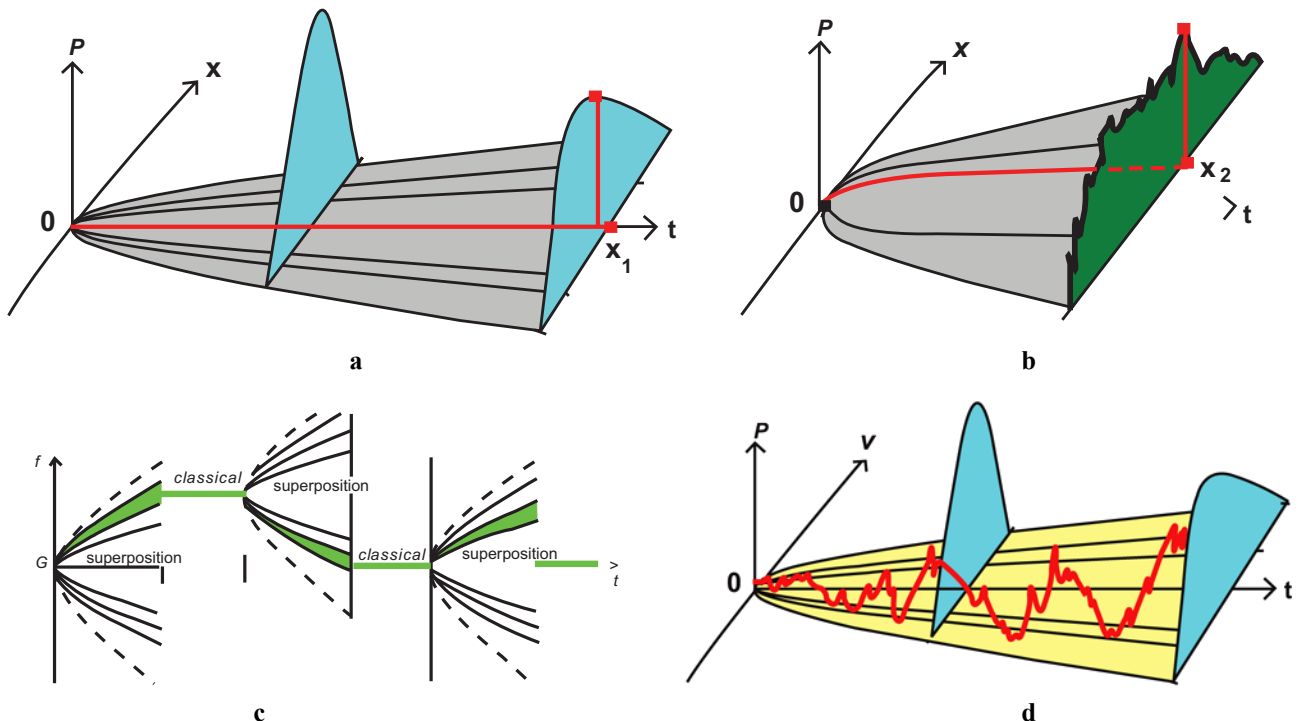
The emergence of randomness due to violation of Lipschitz condition described above can be identified as self-generated stochasticity.

**Examples**

Let us start with the following normal distribution

$$\rho^*(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (20)$$

Substituting the expression (20) and (15) into Eq. (16) at  $X=x$ , one obtains



**Fig. (1).** a. Stochastic process and probability density. b. Global maximum. c. Switching between superposition and classical states in ideal (no-noises) case. d. Random trajectory destabilized by noise.

$$x = \operatorname{erf}^{-1}\left(\frac{C_1}{e^{-t} - 1}\right), \quad x \neq 0 \tag{21}$$

Turning again to the solution (16), let us choose the target density  $\rho^*$  as the Student's distribution, or so called power law distribution

$$\rho^*(X) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{X^2}{\nu}\right)^{-(\nu+1)/2} \tag{22}$$

Substituting the expression (22) into Eq. (16) at  $X=x$ , and  $\nu=1$ , one obtains

$$x = \tan\left(\frac{C}{e^{-t} - 1}\right) \quad \text{for } x \neq 0 \tag{23}$$

Let us now choose the final density as a uniform distribution

$$\rho^*(X) = \begin{cases} 1/(b-a) & \text{if } -a \leq X \leq b, \\ 0 & \text{otherwise} \end{cases}, a > 0 \tag{24}$$

Then

$$\int_{-\infty}^x \rho^*(\xi)d\xi = \begin{cases} (x-a)/(b-a) & \text{if } a \leq x < b \\ 1 & \text{if } x \geq b \end{cases} \tag{25}$$

Substituting Eq. (25) into Eq. (16) at  $X=x > 0$  yields

$$x = a + \frac{C}{e^{-t} - 1}, 0 > C > a - b, a \leq x \leq b, \quad t \geq \ln \frac{b-a}{C+b-a} \tag{26}$$

Qualitative behavior of the solutions' family is shown in Fig. (2a, b).

**Finding Global Maximum**

One of the oldest (and still unsolved) problems in optimization theory is to find a global maximum of a multi-dimensional function. Almost all the optimization problems, one way or another, can be reduced to this particular one. However, even under severe restrictions imposed upon the function to be maximized (such as existence of the second derivatives), the classical methods cannot overcome the

blem of local maxima. The situation becomes even worse if the function is not differentiable since then the concept of gradient ascend cannot be applied.

The idea of the proposed algorithm is very simple: based upon QIM model (10), (11), introduce a *positive* function  $\psi(x_1, x_2, \dots, x_n), |x_i| < \infty$  to be maximized as the probability density  $\rho^*(x_1, x_2, \dots, x_n)$  to which the solution of Eq. (10) 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 (11), as an analog devise

$$\dot{x}_i = \frac{e^{-t}}{[\rho_0(\{x\}) - \rho^*(\{x\})]e^{-t} + \rho^*(\{x\})} \int_{-\infty}^{x_i} [\rho_0(\{\xi\}) - \rho^*(\{\xi\})]d\xi_i, \quad i = 1, 2, \dots, n \tag{27}$$

2. Normalize the function to be maximized

$$\bar{\psi}(\{x\}) = \frac{\psi(\{x\})}{\int_{-\infty}^{\infty} \psi(\{x\})d\{x\}} \tag{28}$$

3. Using Eq. (12), evaluate time  $\tau$  of approaching the stationary process to accuracy  $\varepsilon$

$$\tau \approx \ln \frac{1 - \bar{\psi}}{\varepsilon \bar{\psi}} \tag{29}$$

4. Substitute  $\bar{\psi}$  instead of  $\rho^*$  into Eqs. (27) and run the system during the time interval  $\tau$ .
5. The solution will "collapse" into one of possible solutions with the probability  $\bar{\psi}$ . Observing (measuring) the corresponding values of  $\{x^*\}$ , 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

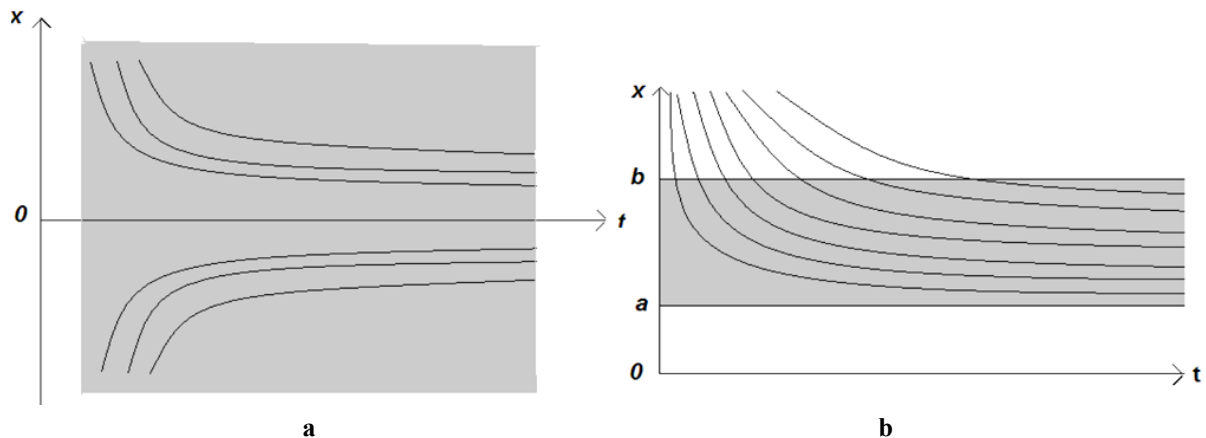


Fig. (2). a. Simulation of power law distribution. b. Simulation of uniform distribution.

the optimal solution to become a winner. Indeed, the probability of success  $\rho_s$  and failure  $\rho_f$  after the first trial are, respectively

$$\rho_s = \bar{\psi}_1, \quad \rho_f = 1 - \bar{\psi}_1 \quad (30)$$

Then the probability of success after M trials is

$$\rho_{sM} = 1 - (1 - \bar{\psi})^M \rightarrow 1 \quad \text{at} \quad M \rightarrow \infty \quad (31)$$

Therefore, after *polynomial* number of trials, one arrived at the solution to the problem (unless the function  $\psi$  is flat).

The main advantage of the proposed methodology is in a weak restriction imposed upon the *space structure* of the function  $\bar{\psi}(\{x\})$ : it should be only integrable since there is no space derivatives included in the model (27). This means that  $\bar{\psi}(\{x\})$  is not necessarily to be differentiable. For instance, it can be represented by a Weierstrass-like function  $f(x) = \sum_0^\infty a^n \cos(b^n \pi x)$ , where  $0 < a < 1$ ,  $b$  is a positive

odd integer, and  $ab > 1 + 1.5\pi$ , (Fig. (1b)). In a particular case when  $\bar{\psi}(\{x\})$  is twice-differentiable, the algorithm is insensitive to local maxima because it is driven *not by gradients*, but by the *values* of this function.

The restriction requiring  $\psi(\{x\})$  being positive can be removed by adding to this function a large positive number: it will not change the optimal value  $\{x^*\}$ ; the only side effect of that may be in longer sequence of Bernoulli trials to obtain the same accuracy since the augmented function will become relatively flatter, and therefore, the global maximum will be less dominating over the rest values of the function to be maximized.

The most significant restriction imposed upon  $\psi(\{x\})$  is the requirement of its normalization. As follows from Eq. (28), the normalization includes computing n-dimensional definite integral, and that procedure is exponentially complex being executed with deterministic algorithms. Indeed, in this case, the number of computations is given by the following formula

$$L \approx \varepsilon^{-n} \quad (32)$$

where  $\varepsilon$  is the assigned accuracy, and  $n$  is the dimensionality of the problem.

However, the Monte-Carlo methods (for which computing definite integrals was the main solid success) provide independence of the number of computations upon  $n$

$$L \approx \frac{1}{\varepsilon^2} \quad (33)$$

It should be noticed that the Monte-Carlo-based computing of definite integrals requires generation of uniformly distributed random numbers, and for that purpose, *the same proposed device* (27) can be exploited. Indeed, prior to executing the step 2, one should run the device (27) in the mode described by Eqs. (24) and (27) for producing uniformly distributed random numbers. In order to generate these numbers in one measurement procedure, one has to apply switches to initial state and back.

### Finding Constrained Maximum

In this section we will discuss several types of constraints imposed upon the function to be maximized that will include local constraints, global constraints, and integer constraints. Local constraints can be associated with a function defined for a piece-wise continuous arguments

$$\psi = \psi(x_1, \dots, x_n) \text{ at } x_i \neq \alpha_{ij}, \quad \beta_{ij} < \alpha_{ij} < \gamma_{ij}, \quad (34)$$

Imposing upon the function (37) the condition

$$G_i(x_1, \dots, x_m) \leq 0, \quad (35)$$

one arrives at a global constraint, (Fig. (3a)). Both of these constraints do not cause any significant obstacles to applying the dynamical system Eq. (27) for simulation the constraint maximum. The only difference with the unconstrained case is in the integration procedure that should exclude the areas where the function  $\psi$  is not defined (see the blue and the red areas in Fig. 3a).

The integer constraint can be introduced using a delta-function representation of an integer

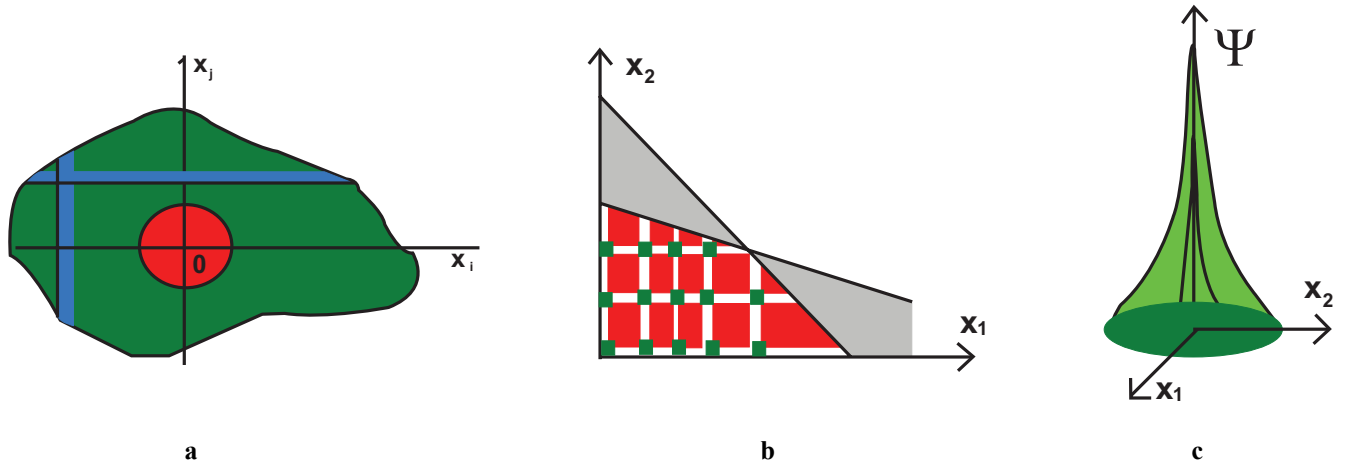


Fig. (3). a. Local and global constraints. b. Integer constraint. c. Continuous representation of an integer.

$$\psi = \psi \left[ \sum_{m \in x} \delta(x_1 - n_j, \dots, x_m - m_j) \right]$$

In addition to that, this function can be subjected to the constraints (34) and (35).

The Figs. (3a and 3b) illustrate the following example of global and integer constraints:

Find the maximum of the following function

$$\psi = 2x_1 + 3x_2 \tag{36}$$

subject to the global and integer constraints

$$x_1 + 3x_2 \leq 8.25, 2.5x_1 + x_2 \leq 8.75, x_1, x_2 > 0, \tag{37}$$

where  $x_1, x_2$  are integers.

However, in cases when the undefined areas cannot be expressed explicitly, a more general strategy is required. This strategy can be associated with Lagrange multipliers or penalty functions. In order to extend the stochastic approach discussed above, we will exploit the penalty function strategy.

## 2. MAXIMIZATION OF THE PENALIZED PERFORMANCE MEASURE

We will start with an equality constraint

$$f(x_1, \dots, x_n) = 0, \tag{38}$$

imposed upon the maximization of the function

$$\psi(x_1, \dots, x_m) \rightarrow Max, \tag{39}$$

As shown in [2], such a constrained maximization can be reduced to unconstrained maximization of the function (39) augmented by the weighted square of the constraint (38)

$$\Psi(x_1, \dots, x_n, \lambda) = \psi(x_1, \dots, x_m) - \lambda f^2(x_1, \dots, x_n) \rightarrow Max, \quad \lambda > 0 \tag{40}$$

The most difficult part of the problem is to assign an appropriate value of the weight  $\lambda$  that controls the contributions of the function (38) and the constraint (39) into the function (40). Obviously, in order to enforce the constraint (39) exactly, the weight must be infinitely large, i.e.

$$\lambda \rightarrow \infty \tag{41}$$

However such a solution would be highly unstable: a small deviation of  $f$  from zero completely suppresses contribution of the function  $\psi$  to be maximized. On the other hand, a small weight would cause an opposite effect: it will suppress the contribution of the constraint  $f$ . Therefore, there exists an optimal weight

$$0 < \lambda < \infty \tag{42}$$

that provides a reasonable, (but approximate) solution.

As suggested in [2], the optimal value of  $\lambda$  is found as a result of a search based upon successive approximations. Instead of that, in this paper we propose to extend the stochastic approach described above. For that purpose, we will consider  $\lambda$  as an additional variable to be optimized in the same way as the variables  $x_i$  are.

The same approach is extendable to the inequality constraint

$$G_{\min} \leq G(x_1, \dots, x_n) \leq G_{\max} \tag{43}$$

if one reduces it to an equality constraint via an additional 'slack' variable  $\xi > 0$  so that the constraint (43) becomes

$$(G_{\max} - G)(G - G_{\min}) - \xi^2 = 0 \tag{44}$$

Now we are ready to formulate the general result.

Consider an n-dimensional function

$$\psi = \psi(x_1, \dots, x_n) \rightarrow Max \tag{45}$$

to be maximized subject to  $m$  equalities

$$f_i(x_1, \dots, x_n) = 0, i = 1, 2, \dots, m \tag{46}$$

and  $k$  inequality

$$G^i_{\min} \leq G^i(x_1, \dots, x_n) \leq G^i_{\max}, \quad i = 1, 2, \dots, k \tag{47}$$

constraints.

The first step is reducing the constrained maximization to the unconstrained one by introduction of the following function of  $n+m+2k$  variables:

$$\Psi(x_1, \dots, x_n, \lambda_1, \dots, \lambda_{m+k}, \xi_1, \dots, \xi_k) = \tag{48}$$

$$= \psi(x_1, \dots, x_n) - \sum_{i=1}^m \lambda_i f_i^2(x_1, \dots, x_n) -$$

$$\sum_{i=m+1}^{m+k} \lambda_i [(G^i_{\max} - G^i)(G^i - G^i_{\min}) - \xi_i^2] \rightarrow Max.$$

where

$$0 < \xi_i < \frac{G^i_{\max} + G^i_{\min}}{4}, i = n + m + k, \dots, n + m + 2k, \tag{49}$$

Now the system (27) can be applied. In order to do that, the function (49) should be normalized

$$\bar{\psi}(\{y\}) = \frac{\psi(\{y\})}{\int_V \psi(\{y\}) d\{y\}} \tag{50}$$

and substituted into Eqs. (27)

$$\dot{x}_i = \frac{e^{-t}}{[\rho_0(\{x\}) - \bar{\Psi}(\{x\})]e^{-t} + \bar{\Psi}(\{x\})} \int_{x_i^0}^{x_i} [\rho_0(\{\xi\}) - \bar{\Psi}(\{\xi\})] d\xi_i, \tag{51}$$

$i = 1, 2, \dots, n + m + 2k$

Similar ODE should be written for the Lagrange multiplier by replacing  $x_i$  with  $\lambda_i$  or  $\xi_i$  in Eqs. (51).

## 3. APPLICATION TO OPTIMAL CONTROL

Optimal control theory is a mathematical method for deriving control policies. It deals with the problem of finding a control law for a given system such that a certain optimality criterion is achieved. The problem can be formulated as following.

Let governing equations of the object to be controlled are represented by a system of ODE

$$f_i(\dot{x}_1, \dots, \dot{x}_n, x_1, \dots, x_n, u_1, \dots, u_m) = 0, i = 1, 2, \dots, n. \tag{52}$$

subject to non-holonomic constraints

$$G^i_{\min} \leq G^i(\dot{x}_1, \dots, \dot{x}_n, x_1, \dots, x_n, u_1, \dots, u_m) \leq G^i_{\max}, \quad i = 1, 2, \dots, k \tag{53}$$

while the performance measure to be maximized is given by the functional

$$J = \int_{t_0}^{t_f} \psi(\dot{x}_1, \dots, \dot{x}_n, x_1, \dots, x_n, u_1, \dots, u_m) dt \rightarrow \text{Max} \tag{54}$$

Here  $u$  is the control vector, and  $[t_0, t_f]$  is the time interval where the measure  $J$  is defined.

The problem is to find the optimal trajectory

$$x_i = x_i(t), \quad i = 1, 2, \dots, n \tag{55}$$

and optimal control vector

$$u_i = u_i(t), \quad i = 1, 2, \dots, m \tag{56}$$

that deliver the global maximum to the performance measure ((54) while satisfying the equality (52) and inequality (53) constraints.

The idea of the proposed approach is based upon reducing the maximization of a functional to maximization of a function by approximating the sought solution (55) and (56) in the form of a weighted sum of basic functions. It is assumed that the system of basic functions is complete, and these functions as well as their linear combinations are “acceptable” for the functional. Introducing the sequence of such functions  $\varphi_1(t), \varphi_2(t), \dots, \varphi_s(t)$ , one can seek the solution (55) and (56) in the form

$$x_i = \sum_{j=1}^{j=s} a_{ij} \varphi_j, \quad i = 1, 2, \dots, n, \tag{57}$$

$$u_i = \sum_{j=1}^{j=s} b_{ij} \varphi_j, \quad i = 1, 2, \dots, m \tag{58}$$

where  $a_{ij}, b_{ij}$  are constant weight coefficients to be found, and  $s$  is the number of basic functions.

In particular, for homogenous boundary conditions  $x_i(t_0) = x_i(t_f) = 0, \quad u_i(t_0) = u_i(t_f) = 0$

one can choose  $\varphi_j = \sin \frac{j\pi t}{t_f}$  as an “acceptable” for the corresponding functional.

The method seeking the solution in the form of a sum of basic functions is known as the Ritz method, [3].

Approximating the time derivatives as

$$\dot{x}_i(t_q) = \frac{x_i(t_{q+1}) - x_i(t_q)}{\Delta t_q},$$

and

$$\dot{u}_i(t_q) = \frac{u_i(t_{q+1}) - u_i(t_q)}{\Delta t_q}, \tag{59}$$

substituting the representations (57)-(59) into Eqs. (52), (53), and (54), and then integrating these equation with respect to time over the period  $[t_0, t_f]$  one arrives at the problem of maximization of the function

$$\psi(\{a\}, \{b\}) \rightarrow \text{Max}, \tag{60}$$

subject to equality

$$f_i(\{a\}, \{b\}) = 0, \quad i = 1, 2, \dots, m \tag{61}$$

and inequality

$$G^i_{\min} \leq G^i(\{a\}, \{b\}) \leq G^i_{\max}, \quad i = 1, 2, \dots, k \tag{62}$$

constraints.

Therefore, the problem is reduced to maximization of the unconstrained function similar to those described by Eq. (48)

$$J_s = \bar{\psi}(\{a\}, \{b\}, \lambda_1, \dots, \lambda_{m+k}, \xi_1, \dots, \xi_k) = \psi(\{a\}, \{b\}) - \sum_{i=1}^m \lambda_i f_i^2 - \tag{63}$$

$$\sum_{i=m+1}^{m+k} \lambda_i [(G^i_{\max} - G^i)(G^i - G^i_{\min}) - \xi_i^2] \rightarrow \text{Max}.$$

It is shown in [3] that if the functions (52), (53) as well as the functional (54) are continuous, and the system of the basic functions is complete, then

$$\lim J_s \rightarrow J \quad \text{at} \quad s \rightarrow \infty \tag{64}$$

However, in many practical cases it is sufficient to take  $s \leq 4$  to obtain an acceptable approximation.

#### 4. SELF-STABILIZATION DEVICE

##### Introductory Remarks

Optimal solution following from the Pontryagin's maximum principle, or from Hamilton-Jacobi-Bellman equation is not necessarily stable. There are many examples of instability of even Euler equation defining optimal solution. This limitation of optimal control theory is a consequence of more fundamental limitation of Newtonian mechanics: the Newton's laws do not discriminate between stable and unstable motions. (For instance, exact solution for the Navier-Stokes equation for a flow with a supercritical Reynolds number can never be observed because of its instability). That is why special mathematical methods for discrimination between stable and unstable motions are required. If the conditions of stability of ODE are expressed in the form of inequalities, they could be included into inequality constraints. However, such inequalities can be formulated in exceptionally few cases, namely: for conservative systems, and for linear ODE with constant coefficients of order below seven. But control systems are never conservative since the main idea of control is to manipulate with the input of external energy, and it is

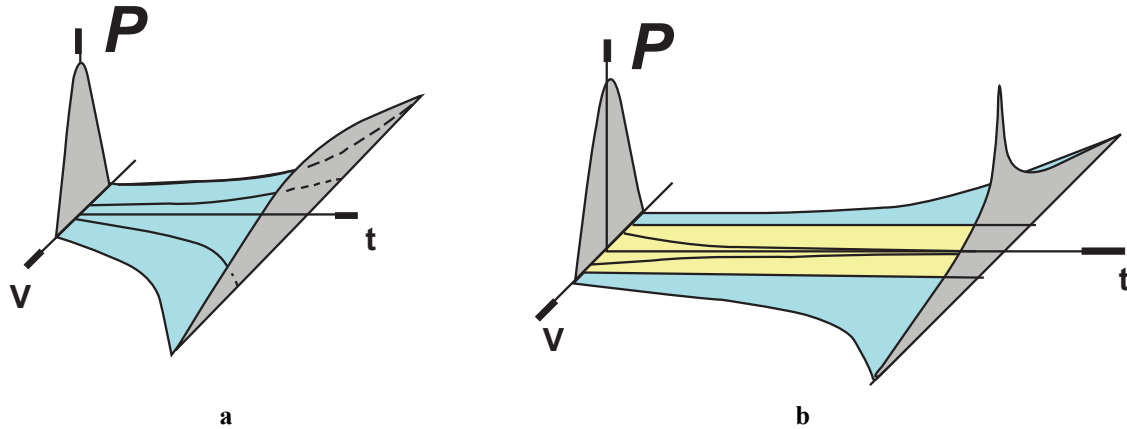


Fig. (4). Divergence of trajectories. a. Uncontrolled evolution. b. Controlled evolution.

very rear that they can be represented by ODE with constant coefficients. That is why the removal of the limitation of possible instability of optimal solution is still a challenge, and this challenge will be addressed below.

There are at least three fundamental types of instability in dynamics: the Lyapunov instability, the Hadamard (or blow up) instability, and non-Lipchitz instability. The difference between them is outlined in [4] and [5]. However, the basic invariant of these instabilities is divergence of neighboring trajectories corresponding to infinitely close initial conditions. The idea of our approach is to suppress this divergence without affecting the “target” trajectory that starts with the prescribed initial conditions.

In order to capture the divergence of trajectories, we will turn again to the Liouville equation and start with Eqs. (1), (2), and (3) following the ideas discussed in [6]. First we will specify a feedback that is different from those in Eq. (9)

$$f = -\sigma^2 \frac{\partial}{\partial v} \ln \rho \tag{65}$$

to obtain the following equation of motion

$$\dot{x} = -\sigma^2 \frac{\partial}{\partial x} \ln \rho, \tag{66}$$

This equation should be complemented by the corresponding Liouville equation (in this particular case, the Liouville equation takes the form of the Fokker-Planck equation)

$$\frac{\partial \rho}{\partial t} = \sigma^2 \frac{\partial^2 \rho}{\partial X^2} \tag{67}$$

The solution of Eq. (67) subject to the sharp initial condition is

$$\rho = \frac{1}{2\sigma\sqrt{\pi t}} \exp\left(-\frac{X^2}{4\sigma^2 t}\right) \tag{68}$$

Substituting this solution into Eq. (66) at  $X=x$  one arrives at the differential equation with respect to  $x(t)$

$$\dot{x} = \frac{x}{2t} \tag{69}$$

and therefore,

$$x = C\sqrt{t} \tag{70}$$

Here  $C$  is an arbitrary constant. Since  $x=0$  at  $t=0$  for any value of  $C$ , the solution (70) is consistent with the sharp initial condition for the solution (68) of the corresponding Liouville equation (67). The solution (70) describes the simplest irreversible motion: it is characterized by the “beginning of time” where all the trajectories intersect (that results from the violation of Lipschitz condition at  $t=0$ , Fig. 5), while the backward motion obtained by replacement of  $t$  with  $(-t)$  leads to imaginary values of velocities. One can notice that the probability density (68) possesses the same properties. For a fixed  $C$ , the solution (70) is *unstable* since

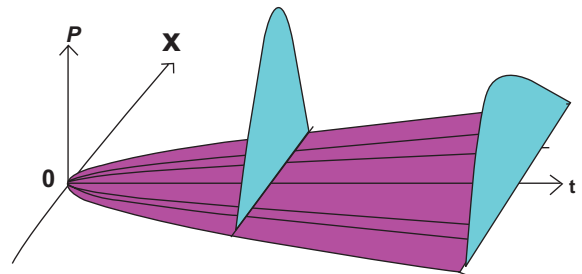


Fig. (5). Stochastic process and probability density.

$$\frac{d\dot{x}}{dx} = \frac{1}{2t} > 0 \tag{71}$$

and therefore, an initial error always grows generating *randomness*. Initially, at  $t=0$ , this growth is of infinite rate since the Lipschitz condition at this point is violated

$$\frac{d\dot{x}}{dx} \rightarrow \infty \quad \text{at} \quad t \rightarrow 0 \tag{72}$$

This is the same type instability that has been demonstrated in Eq. (14).

Considering first Eq. (70) at fixed  $C$  as a sample of the underlying stochastic process (68), and then varying  $C$ , one



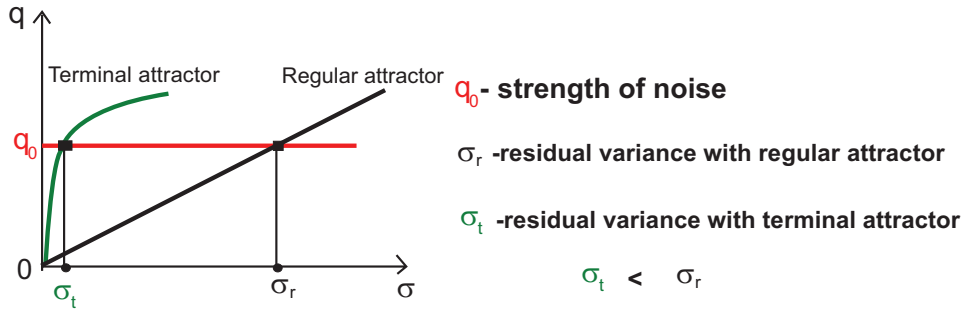


Fig. (6). Effect of terminal attractor.

arrives at the whole ensemble characterizing that process, (see Fig. 6). One can verify that, as follows from Eq. (68), (Risken., 1989) the expectation and the variance of this process are, respectively

$$MX = 0, \quad DX = 2\sigma^2 t, \quad (73)$$

The same results follow from the ensemble (70) at  $-\infty \leq C \leq \infty$ . Indeed, the first equality in (73) results from symmetry of the ensemble with respect to  $x=0$ ; the second one follows from the fact that

$$DX \propto x^2 \propto t \quad (74)$$

It is interesting to notice that the stochastic process (70) is an alternative to the following Langevin equation, (Risken., 1989)

$$\dot{x} = \Gamma(t), \quad M\Gamma = 0, \quad D\Gamma = \sigma \quad (75)$$

that corresponds to the *same* Fokker-Planck equation (67). Here  $\Gamma(t)$  is the Langevin (random) force with zero mean and constant variance  $\sigma$ .

**Negative Diffusion**

Let us consider the Langevin equation (75) and couple it with the corresponding Liouville equation in the same fashion as in Eq. (66)

$$\dot{x} = \Gamma(t) + \alpha \frac{\partial}{\partial x} \ln \rho, \quad \Gamma(t)\Gamma(t') = q\delta(t - t') \quad (76)$$

If one chooses  $\alpha = q^2$ , then the corresponding Liouville equation (that takes the form of the Fokker-Planck equation) will change from Eq. (67) to the following

$$\frac{\partial \rho}{\partial t} = q^2 \frac{\partial^2 \rho}{\partial X^2} - \frac{\partial}{\partial X} \left[ \rho \frac{q^2}{\rho} \frac{\partial \rho}{\partial X} \right] = 0, \quad \rho = \rho_0(X) = const. \quad (77)$$

Thus, the Liouville feedback stops the diffusion. However, the feedback force can be even more effective: it can reverse the diffusion process and push the probability density back to the sharp value in *finite* time. Indeed, suppose that in the Liouville feedback

$$\alpha = q^2 \exp \sqrt{D}., \quad \text{where } D(t) = \int_{-\infty}^{\infty} \rho X^2 dX. \quad (78)$$

Then the Fokker-Planck equation takes the form

$$\frac{\partial \rho}{\partial t} = [q^2(1 - \exp \sqrt{D})] \frac{\partial^2 \rho}{\partial X^2}. \quad (79)$$

Multiplying Eq. (79) by  $X^2$ , then integrating it with respect to  $X$  over the whole space, one arrives at ODE for the variance  $D(t)$

$$\begin{aligned} \dot{D} &= 2q^2(1 - \exp \sqrt{D}), \text{ i.e. } \dot{D} \leq 0 \\ \text{if } D &\geq 0 \end{aligned} \quad (80)$$

Thus, as a result of *negative* diffusion, the variance  $D$  monotonously vanishes regardless of the initial value  $D(0)$ . It is interesting to note that the time  $T$  of approaching  $D = 0$  is finite

$$T = \frac{1}{2q^2} \int_{D(0)}^0 \frac{dD}{1 - \exp \sqrt{D}} \leq \frac{1}{2q^2} \int_0^{\infty} \frac{dD}{\exp \sqrt{D} - 1} = \frac{\pi}{6q^2} \quad (81)$$

This terminal effect is due to violation of the Lipchitz condition at  $D = 0$

Let us turn to a linear version of Eq. (79)

$$\frac{\partial \rho}{\partial t} = -q^2 \frac{\partial^2 \rho}{\partial X^2}. \quad (82)$$

and discuss a negative diffusion in more details. As follows from the linear equivalent of Eq. (80)

$$\frac{d\dot{D}}{dD} = -2q^2, \text{ i.e. } D = D_0 - 2q^2 t < 0 \quad \text{at} \quad t > D_0 / 2q^2 \quad (83)$$

Thus, eventually the variance becomes negative, and that disqualifies Eq. (82) from being meaningful. It has been shown in [7] that the initial value problem for this equation is ill-posed: its solution is not differentiable at any point. (Such an ill-posedness expresses the Hadamard instability similar to those studied in [8]). Therefore, a *negative diffusion must be nonlinear* in order to protect the variance from becoming negative. One of possible realization of this condition is placing a terminal attractor [9], at  $D=0$ .

It should be emphasized that negative diffusion represents a major departure from Newtonian formalism.

**5. MODIFIED NEWTONIAN FORMALISM**

**General Model**

In this section, based upon the stabilizing effect of negative diffusion considered above, we will introduce a general approach to modeling postinstability dynamics. The modi-

fied Newtonian formalism is based upon coupling the classical governing equations with the corresponding Liouville equation and suppression exponential divergence of trajectories by the effect of negative diffusion introduced above. The idea of the proposed approach is in introduction such a Liouville feedback that as a fictitious force acts only upon the erratic component of a trajectory without affecting its “expected” value. For that purpose, introduce a system of  $n$  first order ODE with  $n$  unknowns

$$\dot{x}_i = f_i[\{x(t)\}, t], \quad \{x\} = x_1, \dots, x_n, \quad i = 1, 2, \dots, n. \quad (84)$$

subject to initial conditions

$$x_i(0) = x_i^0 \quad (85)$$

Due to finite precision, the values (85) are not known exactly, and we assume that the error possesses some joint distribution

$$Err(X_i^0) = \rho(X_1^0, \dots, X_n^0) = \rho_0. \quad (86)$$

It is reasonable to assume that the initial conditions (85) coincide with the initial expectations i.e. that

$\rho_0$  has a maximum at  $X_i^0 = x_i^0, i = 1, 2, \dots, n$ . This means that

$$\frac{\partial \rho_0}{\partial X_0} = 0, \quad \frac{\partial^2 \rho_0}{\partial X_i \partial X_j} < 0, \quad i = 1, 2, \dots, n. \quad (87)$$

This is true for any symmetric initial density (for instance, the normal distribution) when the expected values have the highest probability to occur. The Liouville equation describing the evolution of the joint density  $\rho$  is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho f) = 0, \quad f = f_1 \dots f_n, \quad (88)$$

$$f_i = f_i(\{X\}, t), \quad \rho = \rho(\{X\}, t).$$

Its formal solution

$$P = P_0 \exp\left(-\int_0^t \nabla \cdot f d\tau\right) \quad (89)$$

suggests that the flattening of the error distribution is caused by the divergence of the trajectories of the governing equations (84) from the target trajectory that starts with the prescribed initial conditions (85), (Fig. (4)).

Let us introduce the following Liouville feedback

$$F_i = \alpha_i \frac{\partial}{\partial x_i} \ln \rho, \quad \alpha_i > 0 \quad (90)$$

Then the system (84) is modified to the following one

$$\dot{x}_i = f_i + \alpha_i \frac{\partial}{\partial x_i} \ln \rho, \quad (91)$$

that should be complemented by the corresponding Liouville equation

$$\frac{\partial \rho}{\partial t} + \sum_i \alpha_i \frac{\partial}{\partial X_i} (\rho f_i + \frac{\partial \rho}{\partial X_i}) = 0 \quad (92)$$

Now we will summarize mathematical aspects of the system (91), (92).

Firstly, the force  $F_i$  makes the Liouville equation nonlinear, while ODE becomes dependent upon PDE. Secondly, this force introduces to PDE a negative diffusion that changes the type of the PDE from the hyperbolic to the parabolic one. At the same time, the behavior of the solution to Eq. (92) is fundamentally different from its Fokker-Planck analog.

Thirdly, as follows from Eq. (90), the force  $F_i$  does not affect the motion along that trajectory  $x_i = x_i^*$  which has the maximum probability of occurrence since

$$\frac{\partial \rho}{\partial x_i} (x_i = x_i^*) = 0 \quad (93)$$

and that property makes this force fictitious.

Before formulating the proposed model in the final form, we will consider a trivial, but instructive example.

**Example**

Let us consider an unstable linear ODE

$$\dot{x} = \varepsilon x, \quad \varepsilon \ll 1 \quad (94)$$

In this particular case, the expected trajectory is known in advance:

$$\bar{x} = 0 \quad (95)$$

However, any small error in initial conditions leads to a different trajectory that diverge exponentially

$$x = x_0 \varepsilon \exp t \quad (96)$$

Similar result follows from the corresponding Liouville equation:

$$\frac{\partial \rho}{\partial t} = -\varepsilon \frac{\partial}{\partial x} (\rho X) \quad (97)$$

$$X = X_0 \varepsilon \exp t \quad (98)$$

Let us introduce now the fictitious force as

$$F = \sqrt{D} \frac{\partial}{\partial x} \ln \rho, \quad (99)$$

where  $D$  is the variance

$$D(t) = \int_{-\infty}^{\infty} X^2 \rho(X, t) dX \quad (100)$$

and obtain the following modified version of Eq. (94)

$$\dot{x} = \varepsilon x + \sqrt{D} \frac{\partial}{\partial x} \ln \rho \quad (101)$$

Due to this Liouville feedback, Eq. (97) is modified to the following Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -\varepsilon \frac{\partial}{\partial X} (\rho X) - \sqrt{D} \frac{\partial^2 \rho}{\partial X^2} \quad (102)$$

Multiplying Eq. (102) by  $V$ , then using partial integration, one obtains for expectations the same Eq. (95) and its solution Eq. (98).

Similarly one obtains for variances

$$\dot{D} = -2\sqrt{D} - \varepsilon D \approx -2\sqrt{D} \quad (103)$$

For the initial condition

$$D = D_0 \text{ at } t = 0 \quad (104)$$

the solution to Eq. (103) is

$$D = (\sqrt{D_0} - t)^2 \quad \text{for } t < \sqrt{D_0}, \quad (105)$$

$$\text{and } D = 0 \text{ for } t \geq \sqrt{D_0}$$

It is easily verifiable that the Lipschitz condition is violated since

$$\frac{\partial \dot{D}}{\partial D} = -\frac{1}{\sqrt{D}} \rightarrow \infty \quad \text{at } D \rightarrow 0 \quad (106)$$

As will be shown later, this property of the solution is of critical importance for multi-dimensional case.

Now the solution to the nonlinear version of the Fokker-Planck Eq. (102) can be approximated by the first term in the Gram-Charlier series represented by the normal distribution with the variance  $D$ . For the case close to a sharp initial value at  $X=0$

$$\rho = \frac{1}{D\sqrt{2\pi}} \exp\left(-\frac{X^2}{2D}\right), \quad D > 0 \quad (107)$$

Substituting Eq. (107) (with reference to the solution (105)) into Eq. (101) one obtains

$$\dot{x} = x - (\sqrt{D_0} - t) \frac{x}{D^2} = x \left[ 1 - \frac{1}{\sqrt{D_0} - t} \right] \quad (108)$$

whence for  $x = x_0$  at  $t=0$  the solution is

$$x = \frac{x_0}{D_0} e^{t(\sqrt{D_0} - t)^2} \quad 0 \leq t \leq D_0, \quad x = 0 \quad t > D_0. \quad (109)$$

For sufficiently small variance of initial error distribution  $D_0 \ll 1$ , an exponential growth of initial error  $x_0$  is totally eliminated after  $t > \sqrt{D_0}$ , Fig. (7).

It should be noticed that a finite time of approaching equilibrium are special properties of the terminal attractors discussed in [4]. One has to recall again that although the example we just considered is trivial, the stabilization mechanism performed by the negative-diffusion-based Liouville-feedback forces is the same. It is also important to learn from this example that the true expected solution is given by Eq. (109) rather than by Eq. (98) despite the fact that Eq. (98) directly follows from the Liouville equation (97). Indeed, the solution (98) is identical to the original solution (96), and any initial error will grow exponentially. This means that both of these solutions are unstable in the class of differentiable functions. But the same physical phenomenon

described by Eq. (109) is stable in the enlarged class of functions that includes stochastic components [6].

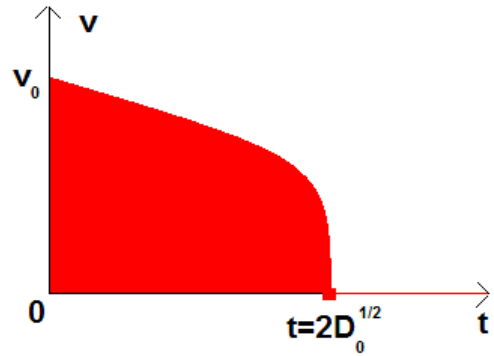


Fig. (7). Suppression of instability.

### Final form of Modified Newtonian Formalism

Based upon the example considered above, we can now specify the coefficients  $\alpha_i$  in Eqs. (91), and (92)

$$\dot{x}_i = f_i + \sqrt{D_{ii}} \frac{\partial}{\partial v_i} \ln \rho, \quad (110)$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial X_i} (\rho f_i + \sum_i \sqrt{D_{ii}} \frac{\partial \rho}{\partial X_i}) = 0 \quad (111)$$

where  $D_{ii}$  are principal variances

$$D_{ii} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (X_i - \bar{X}_i)^2 \rho(dX)^n \quad (112)$$

In order to verify the stabilizing effect of negative diffusion for n-dimensional case, let us linearize Eqs. (110) with respect to the initial state  $x_i = 0$ . Then the linearized versions of Eqs. (110) and (111) will be, respectively

$$\dot{x}_i = a_{ij} x_j + \sqrt{D_{ii}} \frac{\partial}{\partial x_i} \ln \rho, \quad a_{ij} = \left( \frac{\partial f_i}{\partial x_j} \right)_{x_j=0} \quad (113)$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial X_i} (\rho a_{ij} X_j + \sum_i \sqrt{D_{ii}} \frac{\partial \rho}{\partial X_i}) = 0 \quad (114)$$

An n-dimensional analog of Eq. (103) can be obtained by multiplying Eq. (114) by  $V_i$  and then using partial integration

$$\dot{D}_{ij} = -a_{ii} D_{ij} - a_{jj} D_{ij} - \sqrt{D_{ij}} \quad (115)$$

Let us first analyze the effect of terminal attractor and, turning to Eq. (115), start with the matrix  $[\partial \dot{D}_{ij} / \partial D_{lk}]$ . Its diagonal elements become infinitely negative when the variances vanish

$$\frac{\partial \dot{D}_{ij}}{\partial D_{ij}} = \left( -2a_{ij} - \frac{1}{2\sqrt{D_{ij}}} \right) \rightarrow -\infty \quad \text{at } D_{ii} \rightarrow 0 \quad (116)$$

while the off-diagonal elements are bounded. Therefore, due to the effect of terminal attractor (116), the system Eqs. (115) has infinitely negative characteristic roots, i.e. it is infinitely stable with respect to small errors regardless of the parameters  $a_{ij}$  of the original dynamical system (113). In addition to that, the terminal attractor (as well as any attractor) guarantees “impenetrability” of the state  $D_{ii} = 0$ , i.e. if the principle variances initially were non-negative, they will never become negative, and that prevent ill-posedness of the problem for the PDE (111).

Thus, all the properties of the modified model discovered in one-dimensional case are preserved in the n-dimensional case, namely: a simultaneous solution of the coupled ODE-PDE system (110) and (111) describes a stable “expected” motion regardless of the original instability.

### 6. GENERAL FORMULATION OF OPTIMAL CONTROL

Now we are ready to combine all the previous results and formulate the general strategy of optimal control.

First we have to replace the governing equations of the object to be controlled (see Eqs. (52)) by a system of ODE written in the form (110)

$$f_i = \dot{x}_i + \sqrt{D_{ii}} \frac{\partial}{\partial v_i} \ln \rho, \tag{117}$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial X_i} (\rho f_i + \sum_i \sqrt{D_{ii}} \frac{\partial \rho}{\partial X_i}) = 0 \tag{118}$$

subject to non-holonomic constraints

$$G^i_{\min} \leq G^i(\dot{x}_1, \dots, \dot{x}_n, x_1, \dots, x_n, u_1, \dots, u_m) \leq G^i_{\max}, \tag{119}$$

$i = 1, 2, \dots, k$

while the performance measure to be maximized is given by the functional

$$J = \psi(\dot{x}_1, \dots, \dot{x}_n, x_1, \dots, x_n, u_1, \dots, u_m) dt \rightarrow Max \tag{120}$$

First the problem is to be reduced to maximization of the unconstrained function similar to those described by Eq. (63)

$$J_s = \bar{\psi} - \sum_{i=1}^m \lambda_i f_i^2 - \sum_{i=m+1}^{m+k} \lambda_i [(G^i_{\max} - G^i)(G^i - G^i_{\min}) - \xi_i^2] - \sum_{j=1}^N \lambda_j Q_j^2(c_j) \rightarrow Max. \tag{121}$$

while the system of  $N$  algebraic equations

$$Q_j = 0, j = 1, 2, \dots, N \tag{122}$$

represents space-time-discretized version of the PDE (118) with parameters  $c_j$  that plays the same role as parameters  $a_i$  in Eqs. (57).

The final form of the dynamical system simulating the components of the optimal trajectory in a compressed form are written below

$$\dot{y}_i = \frac{e^{-t}}{[\rho_0\{y\} - J_s\{y\}]e^{-t} + J_s\{y\}} \int_{y_i}^{y_{ii}} (\rho_0\{y\} - J_s\{y\}) d\xi_i, \tag{123}$$

Here

$$\{y\} = \{a, b, c, \lambda, \xi\} \tag{124}$$

is a symbolic variable that consists of all the state variable of the system.

It is expected that

$$\lim_{s \rightarrow \infty} J_s \rightarrow J \quad at \quad s \rightarrow \infty \tag{125}$$

As soon as the optimal values of  $y_i$  are determined, one can find the optimal trajectory, and optimal control forces using Eqs. (57) and (58).

### 7. APPLICATION TO ROBOT MOTION PLANNING

The optimal control *via* self-generated stochasticity proposed above can be directly applied to robot motion planning for no obstacle case; the problem of finding collision-free trajectories among obstacles requires some modification of the proposed approach that is out of scope of this paper.

The problem is to find a motion  $(x(t), u(t)), t \in [0, t_f]$  satisfying the equations of motion

$$u = M(q)\ddot{q} + C(q, \dot{q})\dot{q} + g(q) \tag{126}$$

such that  $x(0) = x_{start}, x(t_f) = x_{goal}$ .

Here  $C(q, \dot{q})$  is a vector of velocity product terms with the  $n_Q \times n_Q$  matrix ( $C(q, \dot{q})$  is linear in  $\dot{q}$ ),  $g(q)$  is a vector of gravitational forces,  $M(q)$  is an  $n_Q \times n_Q$  symmetric, positive definite inertia matrix,  $q = [q_1, \dots, q_{n_Q}]$  is a vector of generalized coordinates representing the configuration of the system on the  $n_Q$  configuration space, and  $u = [u_1, \dots, u_{n_Q}]^T$  is the vector of generalized forces acting on the generalized coordinates.

Typical cost function for robots is represented by a functional expressing the global minimum of the control input energy

$$J = \frac{1}{2} \int_0^{t_f} u^T u dt \tag{127}$$

Both Eqs. (126) and (127) are particular cases of Eqs. (117) and (120), respectively, and therefore, the proposed methodology is applicable to general case of robot no-obstacle trajectory planning.

It should be noticed that a particular case of underactuated systems (when the number of actuators is fewer than the number of degrees of freedom) leads to a portion of equations in (126) that do not include actuators. Such equations should be considered as constraints being included in Eq. (121).

## 8. DISCUSSION

### Computational Challenge

We will start our discussion with the computational challenge that has not been addressed in the proposed approach: how to *simulate* ODE with failed Lipchitz conditions. Indeed, although we obtained *closed form* solutions for that case (see Eqs. (16) and (70)), direct simulations for the ODE (14) and (69), respectively, is a fundamental challenge. Even the cases when Lipchitz conditions are only close to failure, i.e. when the ratio  $\dot{x}/dx$  is large, (but *bounded*), a special computational strategy (known as a stiff integration) is required. The main idea of this strategy is to suppress fast-growing errors. In case of failure of Lipchitz conditions, the ratio  $\dot{x}/dx$  becomes *unbounded*, the errors start growing in a super-exponential rate, and the stiff-integration-strategies do not work. However, in the non-Lipchitz case, the whole idea of suppression of errors is wrong. Indeed, as illustrated in Fig. (1a), the failure of Lipchitz condition at the point  $x=0$  leads to loss of uniqueness of the solution, while the multiple solution is represented by a one-parametric family of trajectories. Therefore, errors can be exploited as a vehicle for the transition from the deterministic state to a stochastic state: being amplified by the non-Lipchitz instability, they trigger this transition. It should be recalled that each trajectory appears with the probability prescribed by the corresponding Liouville equation (see Eqs. (12), and (68), respectively) since Eqs. (14) and (69) have been derived from the ODEs that are coupled with their own Liouville equations. However a detailed solution to this computational problem is beyond the scope of this paper.

### Implementation of Self-Stabilizing Device

Another problem that has not been addressed in details in the paper is implementation of self-stabilizing device. Analytical expression for the stabilizing force is presented by the last term in Eqs. (117) and (118). The components of this force are defined by the parameters  $\{c\}$  that appear as a result of space-time discretization of the Liouville equation (118) being a subject of optimization (see Eqs. (123) and (124)). Probably VLSI is the best way to implement self-stabilizing force as an additional component of the control system.

## 9. CONCLUSION

Thus, a general approach to optimal control has been introduced. The approach targets two major obstacles of

existing theory: local maxima and instability of optimal trajectory. The first idea is the following: represent the functional to be maximized as a limit of a probability density governed by the appropriately selected Liouville equation. Then the corresponding ODE become stochastic, and that sample of the solution which has the largest value will have the highest probability to appear in ODE simulation. This idea has been introduced and discussed in [1] being applied to finding the global maximum of a multi-dimensional function subject to equality and inequality constraints. The main advantages of the stochastic approach are that it is not sensitive to local maxima, the function to be maximized must be only integrable, but not necessarily differentiable, and global equality and inequality constraints do not cause any significant obstacles. The second idea is to remove possible instability of the optimal solution by equipping the control system with a self-stabilizing device.

## ACKNOWLEDGEMENTS

Copyright 2009 California Institute of Technology. Government sponsorship acknowledged.

The research described in this paper was performed at Jet Propulsion Laboratory California Institute of Technology under contract with National Aeronautics and Space Administration.

## REFERENCES

- [1] Zak M. Quantum-inspired maximizer. JMP 2008; 49: 042702.
- [2] Pierre DA. Optimization theory with applications. N.Y. Dover Publications 1986.
- [3] Ritz W. Ueber eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. J Reine Angew Math (1908 or 1909); 135: 1.
- [4] Zak M. Terminal model of Newtonian dynamics. Int J Theor Phys 1992; 32: 159-90.
- [5] Zak M. Postinstability models in dynamics. Int J Theor Phys 1994; 33(11): 2215-2280.
- [6] Zak M. Turbulence revisited. Chaos Solitons Fractals 2009; 41: 1136-49.
- [7] Zak M. From reversible thermodynamics to life. Chaos Solitons, Fractals 2005; 26: 1019-33.
- [8] Zak M. Expectation-based intelligent control. Chaos Solitons Fractals 2006; 28: 616-626.