JOURNAL OF MATHEMATICAL PHYSICS **46**, 043508 (2005)

Constructing and exploring wells of energy landscapes

Jean-Pierre Aubin[a)]
Réseau de Recherche Viabilité, Jeux, Contrôle, 14 rue Domat, F-75005 Paris

Annick Lesne[b)]
Laboratoire de Physique Théorique des Liquides, Case 121, Université Pierre et Marie Curie, 4 place Jussieu, F-75252 Paris

(Received 18 May 2004; accepted 10 January 2005; published online 24 March 2005)

Landscape paradigm is ubiquitous in physics and other natural sciences, but it has to be supplemented with both quantitative and qualitatively meaningful tools for analyzing the topography of a given landscape. We here consider dynamic explorations of the relief and introduce as basic topographic features "wells of duration T and altitude y." We determine an intrinsic exploration mechanism governing the evolutions from an initial state in the well up to its rim in a prescribed time, whose finite-difference approximations on finite grids yield a constructive algorithm for determining the wells. Our main results are thus (i) a quantitative characterization of landscape topography rooted in a dynamic exploration of the landscape, (ii) an alternative to stochastic gradient dynamics for performing such an exploration, (iii) a constructive access to the wells, and (iv) the determination of some bare dynamic features inherent to the landscape. The mathematical tools used here are not familiar in physics: They come from set-valued analysis (differential calculus of set-valued maps and differential inclusions) and viability theory (capture basins of targets under evolutionary systems) that have been developed during the last two decades; we therefore propose a minimal Appendix exposing them at the end of this paper to bridge the possible gap. © *2005 American Institute of Physics.*
[DOI: 10.1063/1.1874332]

I. INTRODUCTION

A. The landscape paradigm in natural sciences

The general notion of landscape is encountered in many different domains, for instance in physics, neural networks (Hopfield nets[22]) and learning processes, molecular biology,[13,19] ecology and evolutionary biology,[23] or optimization problems, to cite but a few. From the mathematical viewpoint, a *landscape* is simply a function $V: X \mapsto \mathbb{R} \cup \{+\infty\}$ (more precisely, an extended[30] function since it might take infinite values $+\infty$) associating a real value $V(x)$ to each state $x \in X$ of the system. From a physical viewpoint, the status and definition of V strongly depend on the scale at which the system is described, reflecting in the choice of the space of states X.

Let us give some examples to sustain our exposition. In statistical physics and molecular biology, $V(x)$ can be the *energy landscape* if x is the (high-dimensional) microscopic configuration of the considered system: atomic coordinates in a glass,[15] spin orientations in a spin glass,[17] tridimensional conformation of the hundred or more amino acids forming a protein,[18] spatial positions of bead centers in a granular medium.[16] It can also be a (mesoscopic) *free energy landscape* if x is the value of a (low-dimensional) order parameter describing the global state of the system: spatially average density, overall magnetization, conformational parameter(s) for a macromolecule (as, for instance, its radius of gyration). At a still more macroscopic level, x can be

a)Electronic mail: J.P.Aubin@wanadoo.fr
b)Electronic mail: lesne@lptl.jussieu.fr

a reaction coordinate measuring the progress along a path representing some transformation of the system and inscribed on an *effective energy landscape*. In quite different contexts, cost functions encountered in optimization problems are close analogs to energy landscapes,[2] whereas fitness landscapes encountered in ecology and evolutionary biology can be cast in the frame of free or effective energy landscapes, up to a sign change (namely, by considering the opposite of the fitness). (See Ref. 28 for an introductory review.)

Energy or free energy landscapes are currently exploited in stochastic gradient methods accounting for the interplay between thermal motion and interaction forces (effective forces in the case of a free energy landscape) deriving from the potential $x \mapsto V(x)$. In complex systems (glasses, spin glasses, proteins, for instance) the landscape V typically presents a large number of local optima around which the solution of a stochastic gradient method is trapped and travels a long time before going away and visiting other local minima. This dynamical behavior has been advocated by Giorgio Parisi to encapsulate a meaning of complexity and rugged landscapes are often seen as a mark of complex systems. (See, for instance, Ref. 25.)

Although the landscape $V(x)$ is thus endowed with different status and interpretations in varying contexts, understanding and controlling the system behavior requires in any case a quantitative knowledge of the landscape topography. It is thus of the utmost importance to design efficient tools allowing a dynamical analysis of local minima of such a function $x \mapsto V(x)$. We emphasize that it is not just an academic issue since actual energy or free energy landscapes of real systems are available through either of the following:

(i) a *theoretical access* from first principles (e.g., molecular interactions, spin–spin interactions) and/or modeling hypotheses, allowing us to write an explicit formula for $V(x)$;
(ii) an *experimental access*, for instance, for proteins (indirect kinetic or spectral measurements) (Ref. 19);
(iii) a *numerical access*, either through molecular dynamics at an atomic scale, yielding the energy landscape, either through Monte Carlo sampling of the configuration space according to the Boltzmann distribution, yielding free energy landscapes for the relevant order parameter(s) of the system.[20]

B. Dynamical analysis of a landscape topography

Here we propose a theoretical and algorithmic analysis allowing us to determine quantitatively the landscape relief of a function V, e.g., location of wells, location and heights of the barriers associated with a given dynamics for exploring the landscape.[31] It gives access to a hierarchical picture of the landscape and allows us to determine the nesting of wells and barriers at different scales.

Given a dynamic exploration mechanism (such as a stochastic gradient dynamics), we define the "wells of velocity λ, duration T, and altitude y" as the sets of initial states $x \in X$ "below the level y," i.e., $V(x) \leqslant y$, from which at least one evolution governed by the exploration mechanism, and of velocity bounded by λ, reaches the rim y of the well at the prescribed time T. When the well is not empty, we then evidence intrinsic dynamics governing the evolutions from an initial state of the well up to its rim y at prescribed time T. This intrinsic exploration mechanism is characterized from *the time derivative of the well, regarded as a set-valued map associated with the prescribed duration T and the altitude y the elements of the well*. Both the wells and their intrinsic exploration mechanism can be approximated by finite-difference approximations on finite grids, *which allows us to implement a constructive algorithm*.

In this study we offer an alternative to stochastic gradient-type exploration mechanisms. In quite a similar way of thought, second-order exploration mechanisms of the graph of an energy landscape function has been proposed in Ref. 1. Here we suggest starting the landscape exploration with a universal mechanism, independent of the energy function, allowing us to look at any possible velocity with prescribed norm λ and retaining its intrinsic exploration dynamics as a good candidate for a dynamical system exploring the given energy landscape. The stochastic gradient

method is thus replaced by a differential inclusion involving the time derivative of the well, but allowing in the same spirit the system state to escape the trap of local minima, while being quantitatively influenced by their depth.

The mathematical tools we use are quite novel in physics: They come from set-valued analysis (differential calculus of set-valued maps and differential inclusions) and viability theory (capture basins of targets under evolutionary systems) that have been developed during the last two decades.

The resulting quantitative topographic description by wells rooted in a constructive dynamic exploration of the landscape and the associated determination of the statistical properties of its relief can then be exploited for the following:

(1) *performing a quantitative characterization of the landscape*, for comparison or classification purposes. It allows us to investigate bifurcations, more currently called phase transitions in many-particle systems (Ref. 24);
(2) *providing a quantitative access to the landscape hierarchical structure* and allowing us to estimate its ruggedness, which yields a tentative measure of the system complexity;
(3) *defining macrostates and macroscopic variables* to be used in coarse-grained descriptions of the system. The relevance of such an approach is to provide an intrinsic determination of macrostates, founded upon the identification of macroscopic features with slow modes and slowly evolving properties (Ref. 21).

Outline of the paper. In Sec. II, we shall define wells, introduce some mathematical features of their relief, and reformulate their characterization in terms of the "capture basin of a target," a key concept of viability theory that finds here an unexpected, yet natural, application. In Sec. III, we present the algorithm allowing us to construct explicitly these wells and the intrinsic exploration mechanism on which it is based. In Sec. IV, we introduce the notion of complete wells, matching more closely with physical landscape features. After a conclusive summary in Sec. V, the essential notions of viability theory needed for this paper are presented in an Appendix. ■

II. WELLS OF AN ENERGY LANDSCAPE

A. An efficient alternative to stochastic exploration

In order to provide both a quantitatively meaningful and quantitative topographic analysis of a landscape V on a space X, we introduce "wells of duration T and depth y." Given a dynamical system, allowing upward steps of velocity bounded by a parameter λ, these wells are the sets $\mathbf{P}_V(\lambda;t,y)$ of initial states "below the level y" [i.e., of states $x \in X$ such that $V(x) \leqslant y$] from which at least one[32] evolution reaches the upper level y (what we call the *rim of the well*) at time T. In other words, given some tolerance λ allowing upwards exploration, and some level y, the wells and their depth might be dynamically (the experimentally meaningful and operational way) determined according to the trapping time T.

For exploratory purposes, here we implement an alternative to stochastic gradient dynamics and replace stochastic differential equations encountered in physics by differential inclusions of the form[33]

$$\forall t \geqslant 0, \quad x'(t) \in F(\lambda;x(t)),$$

where $x \rightsquigarrow F(\lambda;x(t))$ is some set-valued function on X [i.e., $F(\lambda;x)$ is a subset of X] parametrized by a parameter $\lambda \in \mathbb{R}$. Compared to a differential equation, the solution of a differential inclusion is less constrained since the full specification of the derivative $x'(t)$ at each time t is replaced by a constraint on the region $F(\lambda;x(t))$, where it has to lie. Such a tolerance is highly valuable and quite realistic in the modeling of an actual system, since the experimentally available knowledge about its dynamics generally provides only bounds (or more generally viability constraints) on the kinetic rates, rather than explicit pointwise expressions of these rates as a function of the system state. These bounds might nevertheless vary with the system state $x(t)$, hence defining a specific

set $F(\lambda;x(t))$at each time t. For instance, in the case when the function V is differentiable, a close analog to stochastic gradient dynamics is provided by

$$x'(t) \in -\nabla V(x(t)) + \lambda B,$$

where B denotes the unit ball of the finite-dimensional vector space X. Indeed, the gradient dynamics $x'(t)=-\nabla V(x(t))$ governs evolutions decreasing along the function V, but stopping at the first encountered local minimum. To overcome this stalling situation, a natural idea is to perturb the gradient equation either by a stochastic noise, as currently implemented in simulated annealing methods, or, as we suggest here, by a "*tychastic*" one. Indeed, differential inclusion $x'(t) \in -\nabla V(x(t))+\lambda B$ is the "tychastic version" of the stochastic differential equation $dx=-\nabla V(x(t))dt+\lambda\, dW(t)$ (see Ref. 8 for the links between stochastic and tychastic viability).

However, we have to overcome the fact that the function V is usually not differentiable, if obtained through experimental measures or simulations and no longer analytically defined. Hence the concept of gradient disappears (when the observable or simulated configuration space is discrete), or has to be approximated by gradients of functions interpolating in one way or another the experimental data. Any method allowing to bypass these obstacles and to deal with graphs of such functions may be worthy of being investigated.

Another suggestion is to leave open the choice of the directions of exploration by looking for any way to climb the landscape V to reach a given level y at a given time T. For that purpose, we can choose $F(\lambda;x) := \lambda B$, stating that any velocity of norm λ is *a priori* an eligible candidate to apply for such a mission. We shall provide below the way of further selecting the most efficient (subset of) velocity(ies), i.e., achieving the most thoroughly and the most efficiently from a numerical viewpoint the quantitative exploration of the landscape relief. The same type of strategy has been used in previous works for constructing an algorithm that is also of relevance for landscapes. This so-called Montagnes Russes Algorithm converges to global minima of an extended function jumping over local minima, which amounts to using the gradient algorithm to the smallest of the exponential Lyapunov functions above the energy function for the differential inclusion $x'(t) \in \lambda B$. (See Refs. 10 and 11.) But whereas this algorithm was devoted to the search of global minima, we are here looking for exploratory tools providing a complete hierarchical picture of the landscape.

B. Definition and characterization of wells

From now on, we assume that the set-valued map $x \rightsquigarrow F(\lambda;x)$ governing the exploration dynamics is given. We denote by $y \in \mathbb{R}$ the altitude of the well we wish to study. $y=0$ is set through the (arbitrary) choice of a base level (or, if known and finite, by the lower bound on V). Usually, the relevant altitudes are the values of the *local maxima* or *saddle points* of the function V. We shall associate with it the concept of well $\mathbf{P}(\lambda;T,y)$ of duration T and altitude y defined as follows.

Definition II.1: Consider an extended function $V: X \mapsto \mathbb{R} \cup \{+\infty\}$ and a differential inclusion

$$\forall\, t \geqslant 0, \quad x'(t) \in F(\lambda;x(t)).$$

Denote by

$$\mathbf{S}(V,y) := \{x \in X \text{ such that } V(x) \leqslant y\} \text{ and } \mathbf{S}_0(V,y) := \{x \in X \text{ such that } V(x) = y\}$$

the level sets of the function V and by $\mathcal{S}_\lambda(x)$ the set of solutions to the above differential inclusion starting at x. The well $\mathbf{P}_V(\lambda;T,y) \subset \mathbf{S}(V;y)$ of duration T and altitude y of the function V is defined by the set of initial states $x \in \mathbf{S}(V;y)$ such that there exists at least one solution $x_\lambda(\cdot) \in \mathcal{S}_\lambda(x)$ such that

$$\text{(i) } V(x_\lambda(T)) = y,$$

$$\text{(ii)}\ \ \forall\, t \in [0,T], \quad V(x_\lambda(t)) \leqslant y.$$

We observe that $\mathbf{P}_V(\lambda;0,y)=\mathbf{S}_0(V,y):=\{x\in X \text{ such that } V(x)=y\}$. In other words, the well $\mathbf{P}_V(\lambda;T,y)$ is the set of initial conditions x in the well from which there exists at least one evolution $x_\lambda(\cdot)$ staying below the level y during a duration T and reaching the level y at exactly time T. This does not exclude the fact that for some earlier time $t^*\leqslant T$ (or some later time $t^*\geqslant T$), the evolution reaches the level y. This just means that x belongs to the intersection $\mathbf{P}_V(\lambda;T,y)\cap\mathbf{P}_V(\lambda;t^*,y)$ of wells of several durations. This point can be made more explicit: considering the initial state x of the system as a variable and the time to reach the level y as the result, we can define the *reaching function* $(\lambda,x,y)\mapsto\xi(\lambda,x,y)$ by

$$\xi(\lambda,x,y) := \inf_{x\in\mathbf{P}_V(\lambda;T,y)} T,$$

providing the first instant when one evolution starting from x reaches the level y.

We can also regard the same object by introducing the set-valued map $(\lambda;T,x)\rightsquigarrow\mathbf{P}_V^{-1}(\lambda;T,x)$ associated with the parameter λ, the duration T, and the initial state x the altitude y of the well, the rim of which can be reached at time T by at least one evolution governed by differential inclusion $x'(t)\in F(\lambda;x(t))$.

Turning back to the initial definition, the *maximal depth* $\delta_V(\lambda;T,y)$ of the well $\mathbf{P}_V(\lambda;T,y)$ is defined by

$$\delta_V(\lambda;T,y) := \sup_{x\in\mathbf{P}_V(\lambda;T,y)} (y - V(x)).$$

The knowledge of the wells provides some physical characteristics of the landscape V, thus bridging the above mathematical definitions with a more traditional description of landscapes. We observe, for instance, that $\xi(\lambda,x,y)$ is the *escape time* for the given dynamics, also called the *first passage time*, from above a barrier of top y when the velocity is bounded by λ. Its inverse $[\xi(\lambda,x,y)]^{-1}$ has the meaning of a kinetic constant.

Denoting $\Omega_V(\lambda;T,y)$ the number of the connected components of well $\mathbf{P}_V(\lambda;T,y)$, its logarithm is the *configurational entropy*. (See Refs. 29 and 16 for its meaning and use in physics, respectively, for glasses and granular media):

$$\sigma_V(\lambda;T,y) := \log(\Omega_V(\lambda;T,y)).$$

In summary, what we are basically looking for is the subset of (x,y,λ,T) such that either $x\in\mathbf{P}_y(\lambda;T,y)$ or $T>\xi(\lambda;x,y)$ or $y\in\mathbf{P}_V^{-1}(\lambda;T,x)$. As detailed in the next section, we shall give a mathematical characterization of this set as a "capture basin of a target under an auxiliary system," allowing us to implement a constructive algorithm. We choose here the representation of this set through the above concept of well $x\in\mathbf{P}_V(\lambda;T,y)$.

C. Viability characterization of wells

The next step of our investigation is to translate the above topographically meaningful features in terms of capture basins for which many properties have been established and constructive algorithms are available (See the Appendix and for further details, Refs. 3–6.)

Proposition II.2: Consider an extended function $V:X\mapsto\mathbb{R}\cup\{+\infty\}$ and a differential inclusion

$$\forall t\geqslant 0, \quad x'(t)\in F(\lambda;x(t)).$$

We associate with it the auxiliary system of differential inclusions

$$\text{(i)}\ x'(t)\in F(\lambda(t);x(t)),$$

$$\text{(ii)}\ y'(t)=0,$$

$$\text{(iii)}\ \lambda'(t) = 0,$$

$$\text{(iv)}\ \tau'(t) = -1 \tag{1}$$

the constrained set $\mathcal{K}$ and the target $\mathcal{C}$ defined by

$$\mathcal{K} := \mathcal{E}_p(V) \times \mathbb{R}_+ \times \mathbb{R}_+ \text{ and } \mathcal{C} := \text{Graph}(V) \times \mathbb{R}_+ \times \{0\},$$

where Graph(V) and $\mathcal{E}p(V) \subset X \times \mathbb{R} \cup \{+\infty\}$ are, respectively, the graph and epigraph of V (see the Appendix for a precise definition). Then

$$\mathbf{P}_V(\lambda; T, y) = \{x \in X \text{ such that } (x, y, \lambda, T) \in \text{Capt}_{(1)}(\mathcal{K}, \mathcal{C})\},$$

where $\text{Capt}_{(1)}(\mathcal{K}, \mathcal{C})$ is the capture basin of the target $\mathcal{C}$ under evolutionary system (1) and under the constraint of remaining in $\mathcal{K}$ (see Definition A.2 below).

Proof: Indeed, to say that $(x, y, \lambda, T) \in \text{Capt}_{(1)}(\mathcal{K}, \mathcal{C})$ amounts to saying that there exist one evolution $x_\lambda(\cdot) \in \mathcal{S}_\lambda(x)$ and a time $t^\star \geqslant 0$ such that the associated auxiliary evolution,

$$t \to (x(t), y(t); \lambda(t); \tau(t)) = (x(t), y, \lambda, T - t),$$

starting from (x, y, λ, T) at $t=0$, reaches the target $\mathcal{C}$ at time $t^\star$ while staying meanwhile in $\mathcal{K}$:

$$\text{(i)}\ (x(t^\star), y, \lambda, T - t^\star) \in \mathcal{C},$$

$$\text{(ii)}\ \forall\, t \in [0, t^\star], \quad (x(t), \lambda, y, T - t) \in \mathcal{K}.$$

The first condition is equivalent to both equations $t^\star = T$ and $V(x(T)) = y$. The second equation means that for every $t \in [0, T]$, $V(x(t)) \leqslant y$. These are the very properties stating that x belongs to the well $\mathbf{P}_V(\lambda; T, y)$, or, equivalently, that $\xi(\lambda; x, y) \leqslant T$. ■

Therefore, the graph of the set-valued map $(\lambda, T, y) \rightsquigarrow \mathbf{P}_V(\lambda; T, y)$ inherits the properties of capture basins. For instance, it can be shown (using Theorem A.6 given in the Appendix) that the well satisfies a kind of dynamical programming principle that can be stated in the following way:

Proposition II.3: The set-valued map $\mathbf{P}_V$ is the unique set-valued map $(\lambda, T, y) \rightsquigarrow \mathbf{P}(\lambda; T, y)$ satisfying the initial condition

$$\mathbf{P}(\lambda; 0, y) := \mathbf{S}_0(V, y) := \{x \in X \text{ such that } V(x) = y\}$$

the constraints

$$\mathbf{P}(\lambda; T, y) \subset \mathbf{S}(V; y)$$

and the "tracking property:" for any $x \in \mathbf{P}(\lambda; T, y)$, any evolution $x_\lambda(\cdot) \in \mathcal{S}_\lambda(x)$ starting from x at time 0 climbing the well until it reaches the rim at time T satisfies

$$\text{(i)}\ \forall\, t \in [0, T], \quad x(t) \in \mathbf{P}(\lambda; T - t, y),$$

$$\text{(ii)}\ \forall\, s \geqslant T \text{ such that } \forall\, t \in [T, s], \quad V(x(t)) \leqslant y, \quad \text{then } x(t) \in \mathbf{P}(\lambda; t - T, y).$$

D. Time derivative of the well as an *a posteriori* exploratory dynamical system

Since we have related the well of a landscape function to capture basins, the basic viability theorems provide tangential characterization of the wells, allowing us to find the underlying dynamical system governing the evolutions of differential inclusion climbing the wells up to their rims. This can be done to the price of using differential calculus of set-valued maps (invented in

the beginning of the 1980s for this purpose): Knowing the "derivatives" with respect to time of the set-valued map $t \rightsquigarrow \mathbf{P}_V(\lambda;t,y)$ (see Definition A.10 for a rigorous definition), we obtain an intrinsic exploration mechanism of the well.

Proposition II.4: For any $x \in \mathbf{P}_V(\lambda;T,y)$, those evolutions $x_\lambda(\cdot) \in \mathcal{S}_\lambda(x)$ starting at x and climbing the well $\mathbf{P}_V(\lambda;T,y)$ in the sense that $V(x_\lambda(t)) \leqslant y$ for any $t \in [0,T]$ and $V(x_\lambda(T)) = y$ are governed by differential inclusion

$$x'(t) \in -\frac{\partial \mathbf{P}_V(\lambda;T-t,y)}{\partial t} \cap F(\lambda;x(t)).$$

In particular, taking for initial exploration mechanism the set-valued map $F(\lambda;x) := \lambda B$ independent of the energy function V instead of exploration mechanisms $F(\lambda;x) := -\nabla V(x(t)) + \lambda B$ already dependent of V, we obtain a more intrinsic exploration mechanism.

Theorem A.12 stated in the Appendix gives a technically precise meaning to this symbolic statement. In other words, the underlying dynamical system governing the evolutions climbing the wells up to their rims is the set of velocities $v \in F(\lambda;x)$ pointing to the time derivative of the well in order to climb it from $-T$ to 0 in order to reach the rim of the well at altitude y. The associated mathematical problem to comfort this intuitive result starts with the definition of the time derivative and the proof of this result is based on results of viability theory. Let us just mention the following informal version of Theorem A.12 stated in the Appendix:

Proposition II.5: The set-valued map $\mathbf{P}_V$ is the unique "Frankowska solution to the partial differential inclusion"

$$\forall t > 0, x \in \mathbf{P}(\lambda;T,y), \quad 0 \in \frac{\partial \mathbf{P}(\lambda;T,y)}{\partial t} + F(\lambda;x)$$

satisfying the initial condition

$$\mathbf{P}(\lambda;0,y) := \mathbf{S}_0(V,y) := \{x \in X \text{ such that } V(x) = y\}$$

and the constraints

$$\mathbf{P}(\lambda;T,y) \subset \mathbf{S}(V;y).$$

We propose now to check the same statement in the discrete case, which allows us to define an algorithm providing the wells under discrete dynamics and the exploratory mechanisms.

III. THE SAINT-PIERRE CAPTURE BASIN ALGORITHM

The Saint-Pierre Capture Basin Algorithm provides both the set-valued map $\mathbf{P}_V$ and for any $x \in \mathbf{P}_V(\lambda;T,y)$, the evolutions climbing the well up to its rim under a given duration.

Let us consider any discrete time approximation $\Phi(\lambda;x)$ of $F(\lambda;x)$ governing the evolution of sequences $\vec{x} \in \vec{\mathcal{S}}_\lambda(x)$, governed by

$$x_{n+1} \in \Phi(\lambda;x_n).$$

[For instance, $\Phi(x): x + hF_h(\lambda;x)$, where h is a time step and F_h is an approximation of F in the sense that the graph of F_h converges to the graph of F in the Painlevé–Kuratowski sense]. The discrete version of a well defined by Definition II.1 for continuous time systems becomes as follows:

Definition III.1: Consider an extended function $V: X \mapsto \mathbb{R} \cup \{+\infty\}$ and a set-valued map $(\lambda,x) \rightsquigarrow \Phi(\lambda;x)$. The discrete time well $\vec{\mathbf{P}}_V(\lambda;N,y) \subset \mathbf{S}(V;y)$ of duration T and depth y of the function V is the subset of initial states $x \in \mathbf{S}(V;y)$ such that there exists one sequence $\vec{x} \in \mathcal{S}_\lambda(x)$ such that

$$\text{(i) } V(x_N) = y,$$

$$\text{(ii)}\ \ \forall\, n \in \{0,N\}, \quad V(x_n) \leqslant y.$$

In the discrete time, we obtain the intrinsic exploration mechanism under mere inspection:

Proposition III.2: Knowing the well $\vec{\mathbf{P}}_V$, the discrete dynamical system

$$x_{n+1} \in \Phi(\lambda;x_n) \cap \vec{\mathbf{P}}_V(\lambda;N-n,y)$$

governs the evolutions starting from $x \in \vec{\mathbf{P}}_V(\lambda;N,y)$ and arriving at step N at some $x_N \in \vec{\mathbf{P}}_V(\lambda;0,y)=\mathbf{S}_0(V,y)$ of the rim of the well $\vec{\mathbf{P}}_V(\lambda;N,y)$.

In the discrete case, the discrete well is obtained by the Capture Basin Algorithm:

Proposition III.3: The Saint-Pierre Capture Basin Algorithm yields the discrete well as the intersection of the following subsets defined recursively by

$$\text{(i)}\ \vec{\mathbf{P}}_V(\lambda;0,y) = \mathbf{S}_0(V,y),$$

$$\text{(ii)}\ \ \forall\, N \geqslant 0, \quad \vec{\mathbf{P}}_V(\lambda;N+1,y) = \Phi(\lambda;\cdot)^{-1}(\vec{\mathbf{P}}_v(\lambda;N,y)) \cap \mathbf{S}(V,y).$$

When $\Phi(\lambda;x) := x+\lambda B$, this algorithm can be written

$$\text{(i)}\ \vec{\mathbf{P}}_V(\lambda;0,y) = \mathbf{S}_0(V,y),$$

$$\text{(ii)}\ \ \forall\, N \geqslant 0, \quad \vec{\mathbf{P}}_V(\lambda;N+1,y) = (\vec{\mathbf{P}}_v(\lambda;N,y) + \lambda B) \cap \mathbf{S}(V,y).$$

Proof: Indeed, we introduce the auxiliary system Ψ by

$$\Psi(x,y,\lambda,\tau) := \Phi(\lambda,x) \times \{y\} \times \{\lambda\} \times \{\tau - 1\},$$

governing the evolution of the sequence:

$$\begin{aligned} &\text{(i)}\ x_{n+1} \in \Phi(\lambda;x_n),\\ &\text{(ii)}\ y_{n+1} = y_n,\\ &\text{(iii)}\ \lambda_{n+1} = \lambda_n,\\ &\text{(iv)}\ \tau_{n+1} = \tau_n - 1, \end{aligned} \tag{2}$$

and the constrained set $\mathcal{K}$ and the target $\mathcal{C}$ defined by

$$\mathcal{K} := \mathcal{E}p(V) \times \mathbb{R}_+ \times \mathbb{R}_+ \quad \text{and} \quad \mathcal{C} := \text{Graph}(V) \times \mathbb{R}_+ \times \{0\}.$$

Then one can prove as in the continuous time case that

$$\vec{\mathbf{P}}_V(\lambda;N,y) = \{x \in X, \quad \text{such that} \quad (x,y,\lambda,N) \in \text{Capt}_{(2)}(\mathcal{K},\mathcal{C})\},$$

where the subscript (2) in $\text{Capt}_{(2)}(\mathcal{K},\mathcal{C})$ refers to the discrete evolutionary system (2).

The capture basin algorithm defines recursively a sequence of subsets $\mathcal{C}_n$ starting at $\mathcal{C}_0$ by

$$\mathcal{C}_{n+1} := \mathcal{K} \cap (\mathcal{C}_n \cup \Psi^{-1}(\mathcal{C}_n)),$$

which converges to the capture basin $\text{Capt}_{(2)}(\mathcal{K},\mathcal{C})$. ■

One can prove that whenever the discrete map $x \rightsquigarrow \Phi_h(x) := x + hF_h(\lambda;x)$ is a time discretization of the differential inclusion $x'(t) \in F(\lambda;x(t))$, the graph of the discrete well converges to the graph of the well in the Kuratowski–Painlevé sense (see Refs. 26 and 27 and see Ref. 14, among other references).

IV. COMPLETE WELLS

The concept of well we proposed in Definition II.1 is not restrictive enough to match its physical counterpart, in the sense that it does not require that all evolutions starting from a point of a well $\mathbf{P}_V(\lambda;T,y)$ to remain below the rim of the well before time T while one of them at least reaches its rim at time T.

Definition IV.1: Consider an extended function $V:X\mapsto\mathbb{R}\cup\{+\infty\}$ and a differential inclusion

$$\forall t \geqslant 0, \quad x'(t) \in F(\lambda;x(t)).$$

The complete well $\mathbf{W}_V(\lambda;T,y)\subset\mathbf{P}(V;y)$ of duration T and depth y of the function V is defined by the set of initial states $x\in\mathbf{S}(V;y)$ such that

(i) all solutions $x_\lambda(\cdot)\in\mathcal{S}_\lambda(x)$ satisfy

$$\forall t \in [0,T], \quad V(x_\lambda(t)) \leqslant y$$

(ii) at least one solution $x_\lambda(\cdot)\in\mathcal{S}_\lambda(x)$ satisfies

$$V(x_\lambda(T)) = y.$$

The complete wells can be characterized in terms of absorption and capture basins. (See, for instance, Refs. 3–6.)

Proposition IV.2: Consider an extended function $V:X\mapsto\mathbb{R}\cup\{+\infty\}$ and a differential inclusion

$$\forall t \geqslant 0, \quad x'(t) \in F(\lambda;x(t)).$$

We associate with it the auxiliary system of differential inclusions (1). The constrained set $\mathcal{K}$ and the targets $\mathcal{C}$ and $\mathcal{D}$ are defined by

$$\mathcal{K} := \mathcal{E}_p(V)\times\mathbb{R}_+\times\mathbb{R}_+ \text{ and } \mathcal{C} := \mathrm{Graph}(V)\times\mathbb{R}_+\times\{0\}$$

and

$$\mathcal{D} := \mathcal{E}_p(V)\times\mathbb{R}_+\times\{0\}.$$

Then

$$\mathbf{W}_V(\lambda;T,y) = \{x\in X \text{ such that } (x,y,\lambda,T)\in \mathrm{Capt}_{(1)}(\mathcal{K},\mathcal{C})\cap \mathrm{Abs}_{(1)}(\mathcal{K},\mathcal{D})\}.$$

Proof: Indeed, to say that $(x,y,\lambda,T)\in\mathrm{Capt}_{(1)}(\mathcal{K},\mathcal{C})\cap\mathrm{Abs}_{(1)}(\mathcal{K},\mathcal{D})$ amounts to saying the following:

(1) $(x,y,\lambda,T)\in\mathrm{Capt}_{(1)}(\mathcal{K},\mathcal{C})$, and thus, as we have seen, that $x\in\mathbf{P}_V(\lambda;T,y)$.

(2) $(x,y,\lambda,T)\in\mathrm{Abs}_{(1)}(\mathcal{K},\mathcal{D})$ means that for all evolutions $x_\lambda(\cdot)\in\mathcal{S}_\lambda(x)$, there exists a time $t^\star\geqslant 0$ such that the associated auxiliary evolutions,

$$t\rightarrow(x(t),y(t);\lambda(t);\pi(t)) = (x(t),y,\lambda,T-t),$$

starting from (x,y,λ,T) at $t=0$ reaches the target $\mathcal{D}$ at time $t^\star$ while staying, meanwhile, in $\mathcal{K}$:

$$\text{(i) } (x(t^\star),y,\lambda,T-t^\star)\in\mathcal{D},$$

$$\text{(ii) } \forall\, t\in[0,t^\star], \quad (x(t),\lambda,y,T-t)\in\mathcal{K}.$$

The first condition is equivalent to both equation $t^\star=T$ and inequality $V(x(T))\leqslant y$. The second equation means that for every $t\in[0,T]$, $V(x(t))\leqslant y$.

These are the two properties stating that x belongs to the well $\mathbf{W}_V(\lambda;T,y)$. ■

V. CONCLUSIONS

Our objective in this investigation was to build exploration dynamics of a landscape V associating with a bound λ on the velocities of the exploration mechanism, a duration T, and an altitude y:

(1) The set $\mathbf{P}_V(\lambda;T,y)$ of initial states x below altitude y from which starts at least one evolution climbing the landscape in order to reach the altitude y at exactly the prescribed time T; altitude y might be either a reference level, thus providing access to the depth of the well, or chosen among the values of local maxima of the landscape function, thus providing access to the height of the barriers separating the well from the other ones.

(2) An underlying dynamical system governing the evolutions climbing the wells up to their rims the velocities of the exploration mechanism are consistently chosen among

$$x'(t) \in -\frac{\partial \mathbf{P}_V(\lambda;T-t,y)}{\partial t}.$$

Hence, the exploration mechanism is no longer an external stochastic modification of the gradient equation, but an intrinsic set-valued method involving the time derivative of the well.

This dynamic description of landscape topography has then been reformulated in the framework of viability theory, which provides a constructive algorithm to characterize quantitatively the landscape, built as an intrinsic exploration mechanism of energy landscapes; this mechanism could be either a perturbed gradient method or a universal mechanism independent of the energy function. The more refined notion of complete well, introduced in Sec. IV, allows us to bridge still more our mathematical definitions and exploration with the current landscape paradigm. As discussed in the Introduction (Sec. I B), our results can then be exploited for taxinomic purposes, to investigate phase transitions, to quantify the landscape hierarchical structure. It also proposes an alternative to standard stochastic gradient methods, namely differential inclusions, in modeling dynamics associated with an experimentally determined landscape.

ACKNOWLEDGMENTS

Jean-Pierre Aubin acknowledges the financial support provided through the European Community's Human Potential Programme under Contract No. HPRNCT-2002-00281 (Evolution Equations for Deterministic and Stochastic Systems).

APPENDIX: ELEMENTS OF VIABILITY THEORY

Let X be a finite-dimensional vector space. A *set-valued map* $F:X \rightsquigarrow X$ associates to any $x \in X$ a subset $F(x) \subset X$. The set-valued map F generates the evolutionary system $\mathcal{S}_F:X \rightsquigarrow \mathcal{C}(0,\infty;X)$ associating with any initial state $x_0 \in X$ the set $\mathcal{S}_F(x_0) \subset \mathcal{C}(0,\infty;X)$ of solutions to *differential inclusion* $x'(t) \in F(x(t))$ starting at x_0. We denote by

$$\mathrm{Graph}(F) := \{(x,y) \in X \times Y | y \in F(x)\} \subset X \times Y,$$

the *graph* of a set-valued map $F:X \rightsquigarrow Y$ and $\mathrm{Dom}(F) := \{x \in X | F(x) \neq \emptyset\}$ its *domain*.

We shall say that a subset $K \subset X$ *is locally viable under* F (or under $\mathcal{S}_F$) if from every $x \in K$ starts *at least one* solution $x(\cdot)$ to the differential inclusion $x' \in F(x)$ *viable in* K on the nonempty interval $[0,T_x[$ in the sense

$$\forall\, t \in [0,T_x[, \quad x(t) \in K,$$

and that K is *viable* if we can take $T_x = +\infty$ for any $x \in K$. Most of the results of viability theory are true whenever we assume that the dynamics are Marchaud:

Definition A.1: We shall say that the set-valued map $F:X \rightsquigarrow Y$ is a Marchaud map if

(i) the graph of F is closed in $X \times Y$,
(ii) the values of $F(x)$ of F are convex subsets of Y,
(iii) the growth of F is linear: $\exists c > 0 \mid \forall x \in X$, $\|F(x)\| := \sup_{v \in F(x)} \|v\| \leqslant c(\|x\| + 1)$.

We shall say that F is λ-Lipschitz if (set-valued extension of the standard Lipschitz property)

$$\forall x, x' \in X, \quad F(x) \subset F(x') + \lambda \|x - x'\| B,$$

where B is the unit ball in Y.

We shall also need some other prerequisites from the Viability Theory: among which capture and absorption basins.

Definition A.2: Let $C \subset K \subset X$ be two subsets, C being regarded as a target, K as a constrained set. The subset $\mathrm{Capt}_F(K,C)$ of initial states $x_0 \in K$ such that C is reached in finite time, without leaving K, by at least one solution $x(\cdot) \in \mathcal{S}_F(x_0)$ starting at x_0 is called the viable-capture basin of C in K (the solution might eventually leave K, but only after having reached C). The subset $\mathrm{Abs}_F(K,C)$ of initial states $x_0 \in K$ such that ***all evolutions*** $x(\cdot) \in \mathcal{S}_F(x_0)$ starting at x_0 are viable in K until they reach C in finite time is called the absorption basin of K with target C.

Obviously $\mathrm{Abs}_F(K,C) \subset \mathrm{Capt}_F(K,C)$. We recall the following result (Ref. 7) of bilateral fixed point property:

Theorem A.3: The viable-capture basin $\mathrm{Capt}(K,C)$ of a target C viability being with respect to the constrained set K) is the unique subset D satisfying $C \subset D \subset K$ and

$$D = \mathrm{Capt}_F(K,D) = \mathrm{Capt}_F(D,C)$$

and the absorption basin of K with target C is the unique subset A satisfying $C \subset A \subset K$ and

$$A = \mathrm{Abs}_F(A,C) = \mathrm{Abs}_F(K,A).$$

We also recall backward invariance:

Definition A.4: The subset K is locally backward invariant under F if for every $t_0 \in]0, +\infty[$, $x \in K$, for all solutions $x(\cdot)$ to the differential inclusion $x' \in F(x)$ arriving at x at time t_0, there exists a time $s \in [0, t_0[$ (depending on the solution) such that $x(\cdot)$ is viable in K on the interval $[s, t_0]$. The subset K is backward invariant under F if we can take $s = 0$ for all solutions.

It is straightforward to check that backward evolutions $\theta \to z(\theta) = x(t_0 - \theta)$ are solutions of the differential inclusion $z'(\theta) \in -F(z(\theta))$ with initial condition $z(0) = x(t_0)$; we call them *backward solutions* (starting from $x(t_0)$ at time $\theta = 0$). It is noted that the (local) backward invariance of K is stronger than (local) viability of K under this backward evolution, since *all* solutions starting in a backward invariant subset K remain in K for a finite time (depending on each considered solution in case of the local version of the property), whereas the (local) viability of K only requires that for each point $x \in K$, at least *one* solution is (locally) viable in K.

We also introduce a weaker notion: A subset $D \subset K$ is *locally backward invariant relatively* to K if all backward solutions starting from D and viable in K (i.e. remaining in K for a finite time) are actually viable in D (i.e., remain in D for a finite time).

Definition A.5: A subset $R \subset X$ is a repeller under F if all solutions starting from R leave R in finite time.

Hence, R is not viable, but this does not exclude local viability. It is, moreover, obvious that any subset of a repeller is itself a repeller.

We can derive the following characterization of capture basin (see Ref. 4):

Theorem A.6: Let us assume that F is Marchaud and that the subsets $C \subset K$ and K are closed. If $K \backslash C$ is a repeller (this is for instance the case when K itself is a repeller), then the viable-capture basin $\mathrm{Capt}_F(K,C)$ of the target C under F is the unique closed subset D satisfying $C \subset D \subset K$ and[34]

(i) $D \backslash C$ is locally viable under F,
(ii) D is locally backward invariant relatively to K.

Definition A.7: The contingent cone $T_L(x)$ to $L \subset X$ at $x \in L$ is the set (obviously a closed cone)

of directions $v \in X$ such that there exist sequences $h_n > 0$ converging to 0 and v_n converging to v satisfying $x + h_n v_n \in L$ for every n (see, for instance, Ref. 9).

For instance, if L is a differentiable manifold in X, $T_L(x)$ coincides with the tangent space to L at point x. If the interior of L is nonempty, then $T_L(x) = X$ for any $x \in \mathrm{Int}(L)$.

We introduce the following Frankowska property that we need for deriving the system of Hamilton–Jacobi–Bellman equations of which the well is a solution:

Definition A.8: Let us consider a set-valued map $F: X \rightsquigarrow X$ and two subsets $C \subset K$ and K. We shall say that a subset D between C and K (i.e., $C \subset D \subset K$) satisfies the Frankowska property with respect to F if

$$\text{(i)} \quad \forall\, x \in D \setminus C, \quad F(x) \cap T_D(x) \neq \emptyset,$$

$$\text{(ii)} \quad \forall\, x \in D \text{ such that (Ref. 35) } -F(x) \cap T_K(x) \neq \emptyset \text{ then } -F(x) \subset T_D(x). \tag{A1}$$

When K is assumed further to be locally backward invariant (then $-F(x) \subset T_K(x)$ for any $x \in K$) the above conditions (A1) boil down to

$$\text{(i)} \quad \forall\, x \in D \setminus C, \quad F(x) \cap T_D(x) \neq \emptyset,$$

$$\text{(ii)} \quad \forall\, x \in D, \quad -F(x) \subset T_D(x) \tag{A2}$$

[The minus sign in front of F arises when considering backward evolution, governed by the differential inclusion, $z'(\theta) \in -F(z(\theta))$.]

Theorem A.6 and the Viability[36] and Invariance Theorems imply

Theorem A.9: Let us assume that F is Marchaud, that K and $C \subset K$ are closed subsets, and that $K \setminus C$ is a repeller. Then the capture basin $\mathrm{Capt}_F(K, C)$ is

(1) the largest closed subset D satisfying $C \subset D \subset K$ and

$$\forall x \in D \setminus C, \quad F(x) \cap T_D(x) \neq \emptyset. \tag{A3}$$

Furthermore, the evolutions $x(\cdot) \in \mathcal{S}_F(x)$ viable in K until they reach C are governed by the differential inclusion

$$x'(t) \in F(x(t)) \cap T_D(x(t)).$$

(It roughly means that these trajectories point into D or are taugent to D at any point where they reach the boundary of D, thus ensuring their viability until they reach C.)

(2) if F is Lipschitz, the unique closed subset D satisfying the Frankowska property (A1).

The absorption basin $\mathrm{Abs}_F(K, C)$ is the largest closed subset D satisfying $C \subset D \subset K$ and

$$\forall x \in D \setminus C, \quad F(x) \subset T_D(x). \tag{A4}$$

We shall apply Theorem A.9 to the case when subsets $K := \mathrm{Graph}(F)$ and $C := \mathrm{Graph}(H)$ are graphs of set-valued maps from X to X and when we decide to regard D as the graph of a set-valued map $G: \mathbb{R} \times X \rightsquigarrow Y$. We then interpret the contingent cone to the graph as the graph of the contingent derivative. We obtain set-valued solutions to systems of Hamilton–Jacobi inclusions that this unknown function G should satisfy in order that its graph yields the desired capture basin. We refer to Refs. 5 and 6, Ref. 12, and their references for more details on this topic. Here, we recall the definition of contingent derivative of a set-valued map and translate Theorem A.9 in the framework of wells.

Definition A.10: Let us consider a set-valued map $G: \mathbb{R} \times X \rightsquigarrow Y$. The graph of the contingent derivative $DG(t, x, y)$ (a set-valued map defined from $\mathbf{R} \times X$ to Y) at a point $(t, x, y) \in \mathrm{Graph}(G)$ is equal to the contingent cone to the graph of G at (t, x, y):

$$T_{\mathrm{Graph}(G)}(t, x, y) = \mathrm{Graph}(DG(t, x, y)).$$

Consequently, to say that $w \in Y$ belongs to the contingent derivative $DG(t,x,y)(\pm 1,v)$ of G at (t,x,y) in the direction $(\pm 1,v) \in \mathbf{R} \times X$ means that

$$\liminf_{h \to 0+, v' \to v} d\left(w, \frac{G(t \pm h, x + hv') - y}{h}\right) = 0,$$

where d is any distance in Y. Since the contingent cone is a closed subset, the graph of a contingent derivative is always closed and positively homogeneous (this is what remains of the required linearity of the derivative in classical analysis, but, fortunately, we can survive pretty well without linearity).

When $g: \mathbf{R} \times X \mapsto Y$ is single-valued, we set $Dg(t,x) := Dg(t,x,g(t,x))$. We see at once that $Dg(t,x)(\pm 1,v) = \pm \partial g(t,x)/\partial t + \partial g(t,x)/\partial x \cdot v$ whenever g is differentiable at (t,x). The above definition (A.10) generalizes to set-valued maps a property obviously valid for differentiable maps, hence provides a consistent extension of the differentiation to set-valued maps, coinciding with the plain notion for smooth single-valued maps. Moreover, it is to note that when g is Lipschitz on a neighborhood of (t,x) and when the dimension of X is finite, the domain of $Dg(t,x)$ is not empty. Furthermore, the Rademacher Theorem stating that a locally Lipschitz single-valued map is almost everywhere differentiable implies that $x \rightsquigarrow Dg(t,x)$ is almost everywhere single valued. However, in this case, equality $Dg(t,x)(-1,-v) = -Dg(t,x)(1,v)$ is not true in general. We refer to Ref. 9 for more details.

Remark: This is how Fermat defined in 1637 the derivative of a function as the slope of the tangent to its graph. Leibniz and Newton provided the characterization in terms of limits of difference quotients. Here, too, *the graph of the contingent derivative* $DG(t,x,y)$ *is the upper Painlevé–Kuratowski limit of the graphs of difference quotients* $\nabla_h G(t,x,y)$ of G at $(t,x,y) \in \mathrm{Graph}(G)$, defined by

$$(\lambda, v) \mapsto \nabla_h G(t,x,y)(\lambda, v) := \frac{G(t + \lambda h, x + hv) - y}{h}.$$

Indeed, we observe that

$$\mathrm{Graph}(\nabla_h G(t,x,y)) = \frac{\mathrm{Graph}(G) - (t,x,y)}{h} (\subset \mathbf{R} \times X \times Y),$$

so that the contingent cone to the graph of G, being the upper limit of the graphs of the difference quotients, is equal by definition to the graph of the upper graphical limit of the difference quotients.

The strong requirement of pointwise convergence of differential quotients involved in the usual derivatives can be weakened in (at least) two ways, each way sacrificing different groups of properties of these usual derivatives.

(i) *Distributional derivatives:* Fix the direction v and take the limit of the function $x \mapsto \nabla_h g(x)(v)$ in the weaker sense of distributions. The limit $D_v g$ may then be a distribution, and no longer a single-valued map. However, it coincides with the usual limit $(D_v g(x) = Dg(x) \cdot v)$ when g is Gâteaux differentiable. Moreover, one can define difference quotients of distributions, take their limit, and thus, differentiate distributions.
Distributions are no longer functions or maps defined on $\mathbf{R}^n$, so these distributional derivatives lose the pointwise character of functions and maps; on the other hand, this generalization retains the linearity of the operator $g \mapsto D_v g$, mandatory for using the theory of linear operator for solving partial differential equations.

(ii) *Graphical derivatives:* Fix the point x and take the limit of the function $v \mapsto \nabla_h g(x)(v)$ in the weaker sense of graphical convergence (the graph of the graphical limit being by definition the Painlevé–Kuratowski upper limit of the graphs). The limit $Dg(x)$ may then be a set-valued map, and no longer a single-valued map. However, it coincides with the usual limit when g is Gâteaux differentiable. Moreover, one can define difference quotients of

set-valued maps, take their graphical limit, and thus differentiate set-valued maps. These graphical derivatives keep the pointwise character of functions and maps, mandatory for implementing the Fermat Rule, proving inverse function theorems under constraints or using Lyapunov functions, for instance, but lose the linearity of the map $g \mapsto Dg(x)$.

In both cases, the approaches are similar: They use (different) convergences *weaker than the pointwise convergence* for increasing the possibility for the difference quotients to converge. But the price to pay is the loss of some properties by passing to these weaker limits (the pointwise character for distributional derivatives, the linearity of the differential operator for graphical derivatives). ■

Proposition II.2 related the graph of the well to the capture basin $P_v(\lambda;T,y) = \{x \in X \text{ such that } (x,y,\lambda,T) \in \text{Capt}_1(\mathcal{KC})\}$ under system (A1):

$$\begin{aligned} &\text{(i) } x'(t) \in F(\lambda(t);x(t)), \\ &\text{(ii) } y'(t) = 0, \\ &\text{(iii) } \lambda'(t) = 0, \\ &\text{(iv) } \tau'(t) = -1. \end{aligned} \tag{A1'}$$

At this point, we need to introduce the concepts of epigraph and epiderivative of extended numerical functions:

Definition A.11: Let $V: X \mapsto \mathbb{R} \cup \{+\infty\}$ be an extended function. Its epigraph $\mathcal{E}p(V)$ is the set of pairs $(x,y) \in X \times \mathbb{R}$ satisfying $V(x) \leqslant y$ [thus $\mathcal{E}p(V) \subset X \times \mathbf{R}$]. The contingent epiderivative $D_\uparrow V(x): X \mapsto \overline{\mathbb{R}}$ is defined through the relation

$$\mathcal{E}p(D_\uparrow V(x)) := T_{\mathcal{E}p(V)}(x, V(x)).$$

We can check that $D_\uparrow V(x)$ consistently coincide with the usual derivative $DV(x)$ when V is differentiable in x, and that for any $v \in X$,

$$D_\uparrow V(x)(v) = \liminf_{h \to 0+, v' \to v} \frac{V(x+hv') - V(x)}{h}$$

is a generalized limit of differential quotients.

We deduce from Proposition II.2 and Theorem A.9 the following characterization of the well as the unique solution to an initial-value problem of a partial differential inclusion satisfying viability constraints:

Theorem A.12: Assume that the set-valued map F is Marchaud and that the function V is continuous. Then the well $\mathbf{P}_V: \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \rightsquigarrow X$ is the largest set-valued map $\mathbf{P}: \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \rightsquigarrow X$ solution to the partial differential inclusion

$$\forall x \in \mathbf{P}(\lambda;T,y), \quad F(\lambda;x) \cap D\mathbf{P}(\lambda;T,y,x)(0,-1,0) \neq \emptyset$$

the initial condition

$$\mathbf{P}(\lambda;0,y) = \mathbf{S}_0(V,y)$$

and the viability constraint

$$\mathbf{P}(\lambda;T,y) \subset \mathbf{S}(V,y).$$

Furthermore, if F is Lipschitz, this solution is the unique solution satisfying

(i) $\forall x \in \mathbf{P}(\lambda;T,y)$, $F(\lambda;x) \cap D\mathbf{P}(\lambda;T,y,x)(0,-1,0) \neq \emptyset$,

(ii) $\forall x \in \mathbf{P}(\lambda;T,y)$ such that $\inf_{v \in F(\lambda;x)} D_{\uparrow}V(x)(-v) \leqslant 0$ then $-F(\lambda;x) \subset D\mathbf{P}(\lambda;T,y,x)(0,+1,0)$.

Proof: Theorem A.9 implies that the graph of the well $\mathbf{P}_V : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \rightsquigarrow X$, once transformed by the permutation $(\lambda,\tau,y,x) \to (x,y,\lambda,\tau)$ of the coordinates, is the largest subset $\mathcal{D}$ between $\mathcal{C}$ and $\mathcal{K}$ (i.e., $\mathcal{C} \subset \mathcal{D} \subset \mathcal{K} \subset X \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$), such that

$$\forall (x,y,\lambda,\tau), \quad (F(\lambda;x) \times \{0\} \times \{0\} \times \{-1\}) \cap T_{\mathcal{D}}(x,y,\lambda,\tau) \neq \emptyset.$$

This amounts to saying that the well $\mathbf{P}_V$ is the largest set-valued map $\mathbf{P}$ satisfying the initial condition $\mathbf{P}(\lambda;0,y) = \mathbf{S}_0(V,y)$, the constraint $\mathbf{P}(\lambda;T,y) \subset \mathbf{S}(V,y)$, and the contingent solution to the partial differential inclusion,

$$\forall x \in \mathbf{P}(\lambda;T,y), \quad F(\lambda;x) \cap D\mathbf{P}(\lambda;T,y,x)(0,-1,0) \neq \emptyset,$$

and that the evolutions $[t \mapsto (\lambda, T-t, y, x(t))]$ viable in the well until they reach its rim are governed by the differential inclusion

$$(0,-1,0,x'(t)) \in (\{0\} \times \{-1\} \times \{0\} \times F(\lambda;x(t))) \cap \mathrm{Graph}(D\mathbf{P}_V(\lambda;T-t,y,x(t))(0,-1,0)).$$

This can be written as

$$x'(t) \in F(\lambda;x(t)) \cap D\mathbf{P}_V(\lambda;T-t,y,x(t))(0,-1,0).$$

This is what we meant symbolically above as

$$x'(t) \in F(\lambda;x(t)) \cap -\frac{\partial \mathbf{P}_V(\lambda;T-t,y)}{\partial t}.$$

When F is Lipschitz [this is the case when $F(\lambda;x) := \lambda B$ (where B is the unit ball in X)], the graph of the well $\mathbf{P}_V$ (after permutation of the coordinates as above) is the unique subset $\mathcal{D}$, satisfying

$$\forall (x,y,\lambda,\tau), \quad (F(\lambda;x) \times \{0\} \times \{0\} \times \{-1\}) \cap T_{\mathcal{D}}(x,y,\lambda,\tau) \neq \emptyset,$$

and, whenever $(-F(\lambda;x) \times \{0\} \times \{0\} \times \{+1\}) \cap T_{\mathcal{K}}(x,y,\lambda,\tau) \neq \emptyset$, then

$$(-F(\lambda;x) \times \{0\} \times \{0\} \times \{+1\}) \subset T_{\mathcal{D}}(x,y,\lambda,\tau).$$

Thanks to the definition of the contingent epiderivative and the fact that $\mathcal{K} := \mathcal{E}p(V) \times \mathbb{R}_+ \times \mathbb{R}_+$, we infer that

$$(-v,0,0,+1) \in T_{\mathcal{D}}(x,y,\lambda,\tau),$$

if and only if $D_{\uparrow}V(x)(-v) \leqslant 0$. This concludes the proof. ■

[1] Attouch, H., Goudou, X., and Redont, P., "The heavy ball with friction method, I. The continuous dynamical system," Commun. Contemp. Math. **2**, 1–34 (2000).
[2] Attouch, H., and Soubeyran, A., "A landscape approach of satisficing by incremental exploration–exploitation," unpublished.
[3] Aubin, J.-P., *Viability Theory* (Birkhäuser, Boston, 1991).
[4] Aubin, J.-P., "Viability kernels and capture basins of sets under differential inclusions," SIAM J. Control Optim. **40**, 853–881 (2002).
[5] Aubin, J.-P., "Boundary-value problems for systems of Hamilton–Jacobi–Bellman Inclusions with constraints," SIAM J. Control Optim. **41**, 425–456 (2002).
[6] Aubin, J.-P., Bayen, A., Bonneuil, N., and Saint-Pierre, P., *Viability, Control and Games, Regulation of Complex Evolutionary Systems Under Uncertainty and Viability Constraints* (Springer-Verlag, Berlin, in press).
[7] Aubin, J.-P., and Catte, F., "Fixed-Point and Algebraic Properties of Viability Kernels and Capture Basins of Sets," Set-Valued Anal. **10**, 379–416 (2001).
[8] Aubin, J.-P., and Doss, H., "Characterization of Stochastic Viability of any Nonsmooth Set Involving its Generalized Contingent Curvature," Stoch. Anal. Appl. **25**, 951–981 (2003).
[9] Aubin, J.-P., and Frankowska, H., *Set-Valued Analysis* (Birkhäuser, Boston, 1990).

[10] Aubin, J.-P., and Najman, L., "L'algorithme des montagnes russes pour l'optimisation globale," C. R. Acad. Sci., Ser. I: Math. **319**, 631–636 (1994).
[11] Aubin, J.-P., and Najman, L., "The Russian Mountain Algorithm for global optimization," Math. Methods of Operations Research (Special issue on 'Set-valued optimization'), 1998, Vol. 48, pp. 153–168.
[12] Aubin, J.-P., and Frankowska, H., "Hyperbolic systems of partial differential inclusions," Annali Scuola Normale di Pisa, 18, 541–562 (1992).
[13] Becker, O. M., and Karplus, M., "The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics," J. Chem. Phys. **106**, 1495–1517 (1997).
[14] Cardaliaguet, P., Quincampoix, M., and Saint-Pierre, P., "Set-valued numerical methods for optimal control and differential games," in *Stochastic and Differential Games. Theory and Numerical Methods*, Annals of the International Society of Dynamical Games (Birkhäuser, Base 1, 1999), pp. 177–247.
[15] De Benedetti, P. G., and Stillinger, F. H., "Supercooled liquids and the glass transition," Nature (London) **410**, 259–267 (2001).
[16] Edwards, S. F., "The role of entropy in the specification of a powder." in *Granular Matter: An Interdisciplinary Approach*, edited by A. Mehta (Springer-Verlag, New York, 1994).
[17] Fischer, H. K., and Hertz, J. A., *Spins Glasses* (Cambridge University Press, Cambridge, 1991).
[18] Frauenfelder, H., "Proteins: paradigms of complexity," Proc. Natl. Acad. Sci. U.S.A. **99**, 2479–2480 (2002).
[19] Frauenfelder, H., Macmahon, B. H., Austin, R. H., Chu, K., and Groves, J. T., "The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin," Proc. Natl. Acad. Sci. U.S.A. **98**, 2370–2374 (2001).
[20] Frenkel, D., and Smit, B., *Understanding Molecular Simulation* (Academic, New York, 2001).
[21] Gaveau, B., Lesne, A., and Schulman, L. S., "Spectral signatures of hierarchical relaxation," Phys. Lett. A **258**, 222–228 (1999).
[22] Hopfield, J. J., "Neural networks and physical systems with emergent collective computational abilities," Proc. Natl. Acad. Sci. U.S.A. **79**, 2554–2558 (1982).
[23] Kauffman, S. A., *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, Oxford, 1993).
[24] Laguës, M., and Lesne, A., *Invariances d'Échelle*(Belin, Paris 2003), Chap. 9, Sec. I.
[25] Mézard, M., Parisi, G., and Virasoro, M. A., *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[26] Saint-Pierre, P., "Approximation of the viability kernel," Appl. Math. Optim. **29**, 187–209 (1994).
[27] Saint-Pierre, P., *Viable Capture Basin for Studying Differential and Hybrid Games*, à paraître dans International Game Theory Review (World Scientific, Singapore, 2003).
[28] Sherrington, D., "Landscape paradigms in physics and biology: Introduction and overview," Physica D **107**, 117–121 (1997).
[29] Stillinger, F. H., . "Exponential multiplicity of inherent structures," Phys. Rev. E **59**, 48–51 (1999).
[30] We call V an "extended function" in the sense that it can take the value $+\infty$. This allows us to encapsulate state constraints $x \in K$ [where K is the subset of "viable" (acceptable, feasible) states] in the definition of the extended function, equal to $+\infty$ outside K. A well-known example in physics is hard-core interaction potential.
[31] Care has to be taken not to confuse exploration mechanism (valid whatever the interpretation of V is) and actual evolution (whose modeling is indissociable from the status and definition of V). We here mainly consider exploration dynamics. Only at the end of the paper, it is suggested that the intrinsic exploration mechanism determined in Sec. III could be a relevant alternative to stochastic gradient dynamics for modeling actual dynamic features of the system.
[32] or, if needed, all evolutions.
[33] At odds with stochastic models currently used in physics, differential inclusions provide another way to translate uncertainty in mathematical terms. A full corpus of results have been steadily accumulated since their introduction in the early 1930s by Marchaud and Zaremba, and next, by the Polish and Russian schools, around Ważewski and Filippov, who laid the foundations of the mathematical theory of differential inclusions after the 1950s. The confrontation of evolutions governed by differential inclusion and viability constraints for reaching targets in finite time began in the early 1980s to provide regulation maps and feedbacks. Their use for translating mathematically uncertainty "against nature" as an alternative to stochastic differential equations began at the same period, but is taking a new start when it was realized recently that they yield much more general results than the stochastic paradigm: Differential inclusions allow us to represent uncertainty by state-dependent maps. As a name needed to refer to such a viewpoint and to underline its difference with stochastic differential equations, we called these *tychastic systems* when the velocities depend upon both the state and a parameter that is called "tyche" [meaning "chance" in classical Greek, from the Goddess Tyche of (good and bad) fortune], which play the role of perturbations or disturbances on which actors and decision makers have no control. Tyches could in fact have been more plainly called "random variables" if this vocabulary were not already confiscated by probabilists. This is why we borrow the term of *tychastic evolution* to the American philosopher Charles Peirce who introduced it in a paper published in 1893 *under the title "Evolutionary love:" "Three modes of evolution have thus been brought before us: evolution by fortuitous variation, evolution by mechanical necessity, and evolution by creative love. We may term them tychastic evolution, or tychasm, anancastic evolution, or anancasm, and agapastic evolution, or agapasm.*
[34] The subset $K \backslash C$ denotes the intersection of K and the complement of C, i.e., is the set of elements of K that do not belong to C.
[35] This is always satisfied when $x \in \mathrm{Int}(K)$.
[36] See for instance Theorems 3.2.4, 3.3.2 and 3.5.2 of Ref. 3.