2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Optimal control subsumes harmonic control

Amine Boumaza
LORIA Campus Scientifique B.P. 239
54506 Vandœuvre-lès-Nancy, France
Email: amine.boumaza@loria.fr

Bruno Scherrer
LORIA Campus Scientifique B.P. 239
54506 Vandœuvre-lès-Nancy, France
Email: bruno.scherrer@loria.fr

Abstract— **We consider trajectory planning within the frameworks of optimal control and harmonic control. We present a formal evidence, in the continuous domain and in a standard discretization, that harmonic control is the limit case of some optimal control problem in which we make the noise level tend to infinity. In other words we show that optimal control subsumes harmonic control. We discuss properties of both paradigms and present simulations that illustrate this relationship.**

I. Introduction

The use of harmonic functions (solutions to the Laplace equation) as potential fields for trajectory planning was proposed by Connolly et al. [1] and Akishita et al. [2] in 1990. In such an approach, obstacles in the configuration space correspond to maxima of the potential, while goals correspond to minima; control algorithms then reduce to *locally* descend the potential until they reach the *global* minimum. The harmonic approach for control (we will now on write "harmonic control") has had some impact on the robotics community [3]–[10]. However, the use of potential functions derived from partial differential equations (PDE), does not date from these seminal works on harmonic functions. Optimal control and reinforcement learning [11] are based on the computation of a potential: the value function. In a continuous domain, the theory of dynamic programming, which was pioneered in the 1950s by Bellman [12], implies that the value function satisfies a PDE, the Hamilton-Jacobi-Bellman equation. Similarly to the harmonic case, the *globally* optimal control is *locally* derived from the value function.

The motivation of this paper is to emphasize a strong relationship between harmonic and optimal control. We argue that the former is a special case of the latter: we formally show, in the continuous domain as well as in a standard finite difference discretization counterpart, that harmonic control is the limit case of trajectory planning with optimal control in a space with isotropic noise[1], when the noise level tends to infinity. Our aim is not to propose some new control algorithms but to *explain* that many practitioners of harmonic control do optimal control without necessarily knowing it.

To our knowledge, the articles that discuss both harmonic and optimal control do not highlight such a specific-to-general relationship. Some articles consider both approaches as complementary alternatives [13]–[15], whereas in others the connection between harmonic control and optimal control is close but not stated by the authors. Connolly et al. [16] incorporate non-holonomic constraints in harmonic control by using Neumann conditions that constrain the system to satisfy its real degrees of freedom. Though this is what is done in optimal control through the state dynamics' function (see section II), no link with optimal control is made explicit. Masoud et al. [3] suggest a way to enforce directional constraints in some parts of the state space by introducing an anisotropic harmonic potential. This is similar to having a constant drift (in the dynamics function) and a non trivial diffusion matrix in optimal control.

The work presented in this article closely relates to [17] where Connolly provides an analysis of harmonic functions in terms of collision probabilities. He shows that, on a grid, the discrete harmonic function at one point is related to the probability of colliding with an obstacle given that the process follows a random walk from that point. He explains that a reinforcement learning algorithm, `TD(0)`, can calculate such a harmonic function (in this case, the harmonic function is indeed equal to the value function of the random uniform policy). Though close, the link with optimal control is not explicitly stated: 1) This analysis does not clarify to which extent "descending the gradient of the value function of the random uniform policy" makes sense. 2) The relaxation technique for computing the harmonic function is presented as a new rapid technique for computing the value function even though it is equivalent to the Value Iteration algorithm. 3) Last but not least, only a discrete version of the problem is studied and the continuous setting, from which the harmonic function comes, is not addressed. The work we present in this paper addresses these issues: we show that a harmonic controller is the optimal controller of some navigation problem on a (finite difference) discretized version, as well as on the (original) continuous domain.

The organization of the paper is as follows. Section II gives a brief introduction to the optimal control framework. Section III shows how to perform trajectory planning with isotropic noise within such a framework. Section IV describes our main result: control with harmonic function corresponds to the case where, in this planning approach, we make the noise level tend to infinity. Finally, section V provides a discussion that clarifies some properties that are shared by the specific (harmonic) case and the general (optimal control) case and others that are not (interesting particular properties of the

[1]Isotropic noise means that the noise is the same in every direction.

specific case over the general case).

II. Optimal control: a brief introduction

We begin by a brief introduction to optimal control. One considers a system defined at time t by its state $x(t) \in \bar{\Omega}$ (the state space) where $\bar{\Omega} \subset \mathbb{R}^n$ is the closure of an open set Ω and $\partial\Omega$ is its boundary ($\bar{\Omega} = \Omega \cup \partial\Omega$). This system is controlled by $u(t) \in U$ where U is a compact set (the control space). The dynamics of such a system is governed by a stochastic differential equation (SDE):

$$dx = f(x(t), u(t))\, dt + \sigma(x(t), u(t))\, dw, \qquad (1)$$

where w corresponds to a m-dimensional Wiener process or Brownian motion and σ is a $n \times m$ diffusion matrix. We consider the case of infinite time horizon. For any initial state x_0, any control law $u(.)$ and trajectory $x(.)$, we note τ the exit time of $x(.)$ from Ω, with the convention that $\tau = \infty$ when the trajectory stays infinitely within Ω. We define the discounted cost functional J as the expected cost over all possible trajectories:

$$J(x, u(\cdot)) = E\left[\int_0^\tau \gamma^t c(x(t), u(t))\, dt + \gamma^\tau C(x(\tau))\right], \quad (2)$$

where c is the instantaneous cost in state $x \in \Omega$, $C(x)$ is the terminal cost at $x \in \partial\Omega$ and $\gamma \in (0,1)$ is a discount factor which guarantees that J remains bounded. We define the optimal value function J^*, function of the initial state x, as the minimum of the cost functional J for all control laws $u(.)$:

$$J^*(x) = \min\nolimits_{u(.)} J(x, u(\cdot)).$$

We let $a = \sigma \cdot \sigma^T$ and note ∇J^* the gradient of J^*. Under reasonable assumptions[2], it can be proved [18] that J^* is $\mathcal{C}^2$ and satisfies the following PDE known as the Hamilton-Jacobi-Bellman (HJB) equation:

$$J^*(x)\ln(\gamma) + \min_{u \in U}\left\{ c(x,u) + \nabla J^*(x) \cdot f(x,u) + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial^2 J^*}{\partial x_i \partial x_j}(x) \right\} = 0, \quad (3)$$

for $x \in \Omega$, with boundary conditions $\forall x \in \partial\Omega$, $J(x) = C(x)$. The optimal value function is particularly interesting since it enables to compute a deterministic optimal controller. For every state x the optimal control $u^*(x)$ is the argument u for which the min in eq. 3 is attained, which is also known as the greedy controller with respect to J^*:

$$u^*(x) = \arg\min_{u \in U}\left\{ c(x,u) + \nabla J^*(x) \cdot f(x,u) + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial^2 J^*}{\partial x_i \partial x_j}(x) \right\}. \quad (4)$$

[2] Sufficient conditions are: 1) the matrix $a = \sigma \cdot \sigma^T$ satisfies a "uniform parabolicity assumption": there exists a $c > 0$ such that $\forall x \in \bar{\Omega}$, $\forall u \in U$ and $\forall y \in \mathbb{R}^n$, $\sum_{i=1}^n \sum_{j=1}^n a_{ij}(x,u) y_i y_j \geq c\|y\|^2$; 2) f, σ and the boundary $\partial\Omega$ are of class $\mathcal{C}^2$, c and C are Lipschitzian.

III. Applying optimal control to trajectory planning

In order to make the connection with harmonic control, we now focus on a simple instance of optimal control. Consider that the state space $\bar{\Omega}$ is an environment in which an agent navigates. For simplicity of exposition, we consider the 2-dimensional case $\bar{\Omega} \subset \mathbb{R}^2$; note however that our result generalizes directly to any dimension n. The boundary of the environment is decomposed into 2 sets $\partial\Omega = \mathcal{O} \cup \mathcal{G}$: $\mathcal{O}$ is the set of obstacles and $\mathcal{G}$ the set of goals. At each time $t > 0$, the agent is characterized by its coordinates $x(t) = (x_1(t), x_2(t))$ and its dynamics is related to a command direction $\theta(t)$ according to the following SDE:

$$dx = \overrightarrow{u}(\theta(t))\, dt + \sigma\, dw.$$

where $\overrightarrow{u}(\theta)$ is a unit vector in the direction $\theta \in [0, 2\pi]$, and σ is a positive constant. In other words, we consider an optimal control problem with a unit speed control ($u = \theta$ and $f(x,u) = \overrightarrow{u}(\theta)$) and a constant isotropic noise ($\sigma \in \mathbb{R}$). In this case it amounts to replace the matrix a by a real number σ^2. Furthermore, we consider that there are no intermediate costs ($\forall x \in \Omega, c(x) = 0$), only terminal costs C on the boundary: $C(x) = 1$ on obstacle boundaries ($\forall x \in \mathcal{O}$) and $C(x) = -1$ on goal boundaries ($\forall x \in \mathcal{G}$). Note ΔJ^* the Laplacian of J^*. In this specific case, the HJB equation (eq. 3) reduces to:

$$J^*(x)\,\ln(\gamma) + \min_{\theta \in [0,2\pi]}\{\nabla J^*(x) \cdot \overrightarrow{u}(\theta)\} + \frac{\sigma^2}{2}\,\Delta J^*(x) = 0 \quad (5)$$

for $x \in \Omega$ with boundary conditions $\forall x \in \partial\Omega, J^*(x) = C(x)$. Under this model, it is easy to see that the optimal control $\theta^*(x)$ in state x is colinear with $\frac{-\nabla J^*(x)}{\|\nabla J^*(x)\|}$: indeed, in eq. 5, the only term involved in the min is $\nabla J^*(x) \cdot \overrightarrow{u}(\theta)$, which in turn is minimal when θ is in the opposite direction of the gradient $\nabla J^*(x)$. Furthermore, for every $\gamma \in (0,1)$, the value function of the controller $\theta(\cdot)$ (eq. 2) reduces to:

$$\begin{aligned} J(x, \theta(\cdot)) = {} & E[\gamma^\tau | x(\tau) \in \mathcal{O}]\, Pr[x(\tau) \in \mathcal{O}] \\ & - E[\gamma^\tau | x(\tau) \in \mathcal{G}]\, Pr[x(\tau) \in \mathcal{G}]. \end{aligned} \qquad (6)$$

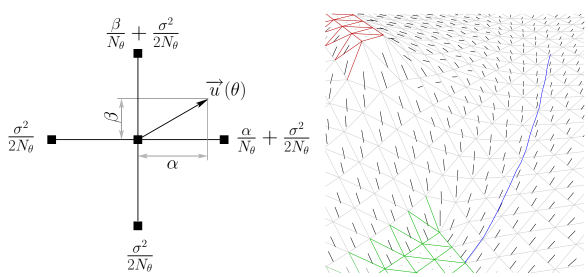
Note that the integral term in eq. 2 vanishes because there are no instantaneous costs $c(\cdot)$. Also recall that the optimal controller minimizes this value function. Let us now interpret what this means. In order to minimize $J(x, \theta(\cdot))$, the optimal controller tries both to minimize $E[\gamma^\tau | x(\tau) \in \mathcal{O}]\, Pr[x(\tau) \in \mathcal{O}]$ and to maximize $E[\gamma^\tau | x(\tau) \in \mathcal{G}]\, Pr[x(\tau) \in \mathcal{G}]$. Since $\tau \mapsto \gamma^\tau$ is a decreasing function of τ, this intuitively means that the optimal controller both tries to maximize the time of hitting an obstacle and to minimize the time of reaching the goal. In the simple deterministic case[3] when $\sigma = 0$, and for any value of $\gamma \in (0,1)$, the optimal controller which consists in following the gradient of the optimal value function, is the one which minimizes the path length (and the time) to reach the goal. In practice, one cannot compute exact analytical solutions to eq. 5 and a usual approach, which we follow here, consists in

[3] In this case, the analysis of the HJB is a bit more complex, involving viscosity solutions (see [19]) which we do not address in details here.

building a finite difference scheme. This is what we describe now. To do so, we follow the lines of [20] and build a discrete-time discrete-space controlled Markov chain. Given a resolution $\delta > 0$, we build a grid Σ^δ and its border $\partial\Sigma^\delta$ on the domain of the problem. For simplicity we assume that any point of the border of the discretized domain belongs to the border of the initial domain: $\partial\Sigma^\delta \subset \partial\Omega$. Given a grid resolution δ, the function J is approximated by a function J^δ defined on $\Sigma^\delta \cup \partial\Sigma^\delta$. We define $\cos^+$, $\cos^-$, $\sin^+$ and $\sin^-$ as the positive and negative parts of cos and sin: $\cos^\pm(\theta) = \max(\pm\cos\theta, 0)$ and $\sin^\pm(\theta) = \max(\pm\sin\theta, 0)$. Notice that $\cos\theta = (\cos^+\theta - \cos^-\theta)$ and $|\cos\theta| = (\cos^+\theta + \cos^-\theta)$. Furthermore, we define the following transitions probability from grid point (x, y) to its four neighbors when going in direction of θ:

$$\begin{cases} p_\theta\left[(x,y),(x\pm\delta,y)\right] = \frac{1}{N_\theta}\left[\frac{\sigma^2}{2} + \delta\cos^\pm\theta\right] \\ p_\theta\left[(x,y),(x,y\pm\delta)\right] = \frac{1}{N_\theta}\left[\frac{\sigma^2}{2} + \delta\sin^\pm\theta\right] \end{cases} \quad (7)$$

where $N_\theta = \delta(\cos^+\theta + \cos^-\theta + \sin^+\theta + \sin^-\theta) + 4\sigma^2/2 = \delta(|\cos\theta| + |\sin\theta|) + 2\sigma^2$ can be regarded as a normalizing factor that ensures that the the transition probabilities $p_\theta((x,y),(x',y'))$ sum to 1. These transition probabilities on the discretized grid have a natural geometric interpretation: all of them have the same noise component ($\frac{\sigma^2}{2N_\theta}$), and two of them (the ones that are in the control direction θ) have non-zero weights (see figure 1(a)). We finally write $\tau(\theta) = \frac{\delta^2}{N_\theta}$,



(a) (b)

Fig. 1. (a) Transition probabilities $p_\theta[(x,y),(x',y')]$ in the discrete model: neighbors in direction θ have a component that is proportional to the coordinates of the unit vector $\vec{u}(\theta)$. (b) A discrete representation of J^δ, illustrating the orientation of the gradients, and an (approximate) optimal trajectory computed by following the gradient.

which is a quantity that can be interpreted as the time needed to go from one grid point to another when following direction θ. Then, an approximation of J is obtained by computing the unknown J^δ in the following system:

$$J^\delta(x,y) = \gamma^{\tau(\theta^{J^\delta}_{x,y})} \sum_{(x',y')\in\Sigma^\delta} p_{\theta^{J^\delta}_{x,y}}\left[(x,y),(x',y')\right] J^\delta(x',y'), \quad (8)$$

for $(x, y) \in \Sigma^\delta$ and $J^\delta = C$ on $\partial\Sigma^\delta$, and where $\theta^{J^\delta}_{x,y}$ is the angle that corresponds to the steepest slope direction at (x, y) when considering a piecewise linear interpolation of J^δ (see figure 1(b)). The notation $\theta^{J^\delta}_{x,y}$ may seem a bit heavy, but it is important to remember that the optimal angle depends on the coordinate (x, y) and the value function J^δ. Introducing the following operator B^δ on Σ^δ:

$$B^\delta[W](x,y) = \gamma^{\tau(\theta^W_{x,y})} \sum_{(x',y')\in\Sigma^\delta} p_{\theta^W_{x,y}}\left[(x,y),(x',y')\right] W(x',y')$$

for $(x, y) \in \Sigma^\delta$ and $B^\delta[W] = C$ on $\partial\Sigma^\delta$, eq. 8 becomes $J^\delta = B^\delta[J^\delta]$. Since $\gamma < 1$ and $\tau(\cdot) > \frac{\delta}{\delta+2\sigma^2} > 0$, the operator B^δ satisfies a contraction property (with contraction factor at least $\gamma^{\frac{\delta}{\delta+2\sigma^2}}$). Therefore J^δ is unique, and can be computed by *relaxation*, that is as the limit of the sequence $J^\delta_{n+1} \leftarrow B^\delta[J^\delta_n]$ when n tends to infinity. Once J^δ is computed, the (approximate) optimal control is the direction $\theta^{J^\delta}_{x,y}$ that descends the gradient of (the linear interpolation of) J^δ. It can be proved [20] that such a finite difference scheme is convergent: J^δ uniformly tends to J^* and the approximate optimal controller ∇J^δ tends to the optimal controller ∇J^* when the discretization resolution δ tends to 0.

IV. Harmonic control is a limit case of optimal control

Now that we have introduced an instance of optimal control problem for trajectory planning, we turn back to the main claim of this study: we demonstrate that when the noise level σ tends to infinity, the optimal control of our simple model tends to harmonic control.

As we are interested in the dependence on σ, we denote J_σ the value function J^* of eq. 5. In order to study the behavior of J_σ when $\sigma \to \infty$, we regard it as a function of the Sobolev space $H^1(\Omega)$ with its norm $\|f\|_{H^1(\Omega)} = (\int_\Omega f^2 + \int_\Omega(\nabla f)^2)^{1/2}$ [21]. When $\sigma \to \infty$, we show that J_σ tends to a harmonic function, that is a function Φ that satisfies Laplace equation $\Delta\Phi = 0$.

Theorem 1: Let Φ be the unique function satisfying $\Delta\Phi = 0$ on Ω and $\Phi = C$ on $\partial\Omega$. Then the family of solutions $(J_\sigma)_{\sigma>0}$ of equation 5 satisfies

$$\lim_{\sigma\to\infty} \|J_\sigma - \Phi\|_{H^1(\Omega)} =$$

$$\lim_{\sigma\to\infty} \left[\int_\Omega (J_\sigma - \Phi)^2\, dx + \int_\Omega (\nabla J_\sigma - \nabla\Phi)^2\, dx\right]^{1/2} = 0$$

Notice that this convergence is strong in the sense that 1) J_σ tends to Φ *and* that 2) its gradient ∇J_σ tends to $\nabla\Phi$.

We can prove a similar result for the standard finite difference approximations of J_σ and Φ. For a resolution $\delta > 0$ and the grid $\Sigma^\delta \cup \partial\Sigma^\delta$ introduced in section III, let Φ^δ be the standard 5-point finite difference approximation of Φ, that is the unique solution to:

$$\Phi^\delta(x,y) = \frac{1}{4}\left[\Phi^\delta(x+\delta,y) + \Phi^\delta(x-\delta,y) + \Phi^\delta(x,y+\delta) + \Phi^\delta(x,y-\delta)\right]$$

for all $(x, y) \in \Sigma^\delta$ and $\Phi^\delta = C$ on $\partial\Sigma^\delta$. When the noise level σ tends to infinity, the approximate value function J^δ_σ tends to the approximate harmonic function Φ^δ:

Theorem 2: Fix a discretization step $\delta > 0$. Let Φ^δ be the standard 5-point finite difference approximation of the harmonic function Φ. Then the family $(J^\delta_\sigma)_{\sigma>0}$ of discretizations of the value function defined in section III satisfies

$$\lim_{\sigma\to\infty} J^\delta_\sigma = \Phi^\delta \text{ uniformly on } \Sigma^\delta \cup \partial\Sigma^\delta.$$

The proofs of theorems 1 and 2 can be found in [22]. On one hand, the proof of theorem 2 is rather straightforward (notice that it is specific to the finite difference schemes used to discretize both problems). On the other hand, if the result of theorem 1 seems intuitive (in eq. 5, we expect that when the term $\frac{\sigma^2}{2}$ before ΔJ^* tends to ∞, the other terms get negligible), the proof we make relies on a few technical manipulations of eq. 5 and non-trivial properties of the Sobolev space $H^1(\Omega)$. In fact, our proof of theorem 1 in [22] is more general that what we really need here: we show that the value function J_σ and its gradient ∇J_σ tend to the harmonic function Φ and its gradient $\nabla\Phi$ for any bounded control function f and any bounded cost c (recall the general optimal control equations 1 and 3); in our simple instance model, we take $u = \theta$, $f(x, u) = \overrightarrow{u}(\theta)$ and $c = 0$ so that the optimal controller matches the opposite direction of the gradient ∇J_σ.

Overall, we can summarize *our* convergence results (when $\sigma \to \infty$) and the standard numerical schemes convergences (when $\delta \to 0$) in the following diagram:

$$\begin{array}{ccc} (J_\sigma, \nabla J_\sigma) & \stackrel{\sigma\to\infty}{\longrightarrow} & (\Phi, \nabla\Phi) \\ {\scriptstyle \delta\to 0} \uparrow & & \uparrow {\scriptstyle \delta\to 0} \\ (J_\sigma^\delta, \nabla J_\sigma^\delta) & \stackrel{\sigma\to\infty}{\longrightarrow} & (\Phi^\delta, \nabla\Phi^\delta) \end{array}$$

When the noise level σ tends to infinity, the value function becomes harmonic, and the corresponding optimal control therefore consists in descending the gradient of this harmonic function. This limit case of optimal control is strictly equivalent to what Connolly [1] and Akishita [2] proposed in 1990: path planning is done by gradient descent of the function Φ. Furthermore, our formal analysis gives some insight on the discrete analysis of harmonic control by Connolly [17] discussed in the introduction. If harmonic control happens to be equivalent to "descending the gradient of the value function of a random uniform policy", this is because the agent's actions become negligible when the noise level gets infinitely large: at the limit, its movements are mostly due to the isotropic noise. Also, as we said that theorem 1 is general (it is true for any control function f and any cost c), the connection we make has a corollary that may be of interest to harmonic control practitionners: in the limit case, and for general $f(x, u)$ and $c(x, u)$, the optimal control theory prescribes to take the controller given by equation 4, which reduces here to:

$$u^*(x) = \arg\min_u \{c(x, u) + \nabla\Phi(x).f(x, u)\},$$

(and this is only equal in general to the gradient direction when $u = \theta$, $f(x, u) = \overrightarrow{u}(\theta)$ and $c = 0$). Recall that $f(x, u)$ describes the real dynamics for action u in state x while $c(x, u)$ can be seen as a bias to prevent form taking action u in state x. Though such a general control law might generate some complications (the controller may in general get stuck[4]), it constitutes a natural candidate for incorporating kinematical and dynamical constraints such as non-holonomicity and/or a notion of cost within the harmonic control framework.

[4] A non-zero c can block the trajectory away from the goal (since in practice the gradient of harmonic functions can be very small). If $c = 0$, then a sufficient condition for always getting to a goal is that f allows to move in any direction (possibly with different speeds) in the state space.

We now present some simulations in order to illustrate what happens when the noise level tends to infinity. We have

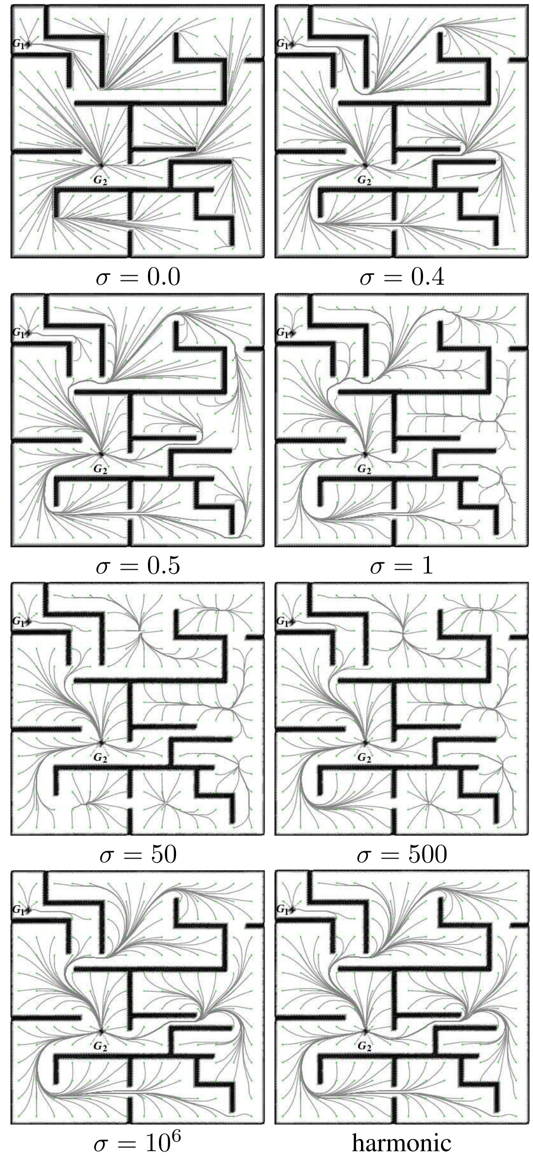


Fig. 2. Evolution of trajectories with respect to the noise level σ. Each figure illustrates different trajectories leading to goals G_1 and G_2, and starting at equally distributed grid points in the environment. All trajectories where computed using optimal control (with J_σ^δ) except the last one which is derived from harmonic control (with Φ^δ).

considered a maze environment containing two goals (G_1 and G_2 in figure 2). Dark lines represent obstacles, white represents free space and gray lines represent trajectories computed along the value function's gradient. Goal G_1 at the top left of the environment is accessible through a narrow corridor, and goal G_2 is located in a fairly uncluttered area. Figure 2 illustrates, for five increasing values of the noise level σ, the trajectories computed starting from equally spread initial positions.

In the case where the dynamics is deterministic ($\sigma =$

0), the optimal trajectories are the shortest; most of the trajectories starting in the upper part of the environment go toward goal G_1 whereas all others go toward goal G_2. When we increase the noise level ($\sigma = 0.4$), the trajectories stray away from the walls and start to become smoother. With little more noise ($\sigma = 0.5$), the corridor leading to goal G_1 becomes hazardous and most of the trajectories go toward goal G_2. At $\sigma = 1$, trajectories are smoother, and we begin to see some of them halting before reaching a goal (trajectories starting on the right side): paths through narrow corridors becoming more dangerous, the controller maintains the agent in safer areas away from obstacles. This behavior intensifies at $\sigma = 50$. A new phase begins at $\sigma = 500$: the noise being elevated, strategies consisting in staying in safe areas are not efficient anymore (the noise inevitably leads towards an obstacle). Then, as a kamikaze that would know he is going to die anyway, it becomes more interesting to start moving again towards a goal. This behavior is more noticeable when we go on increasing the noise level. Eventually observe that the last simulation ($\sigma = 10^6$) experimentally confirms our theoretical result: the trajectories are close to the trajectories computed using a harmonic control.

V. Discussion

We believe that harmonic control has had a significant impact in the literature because as a paradigm, it has several nice properties: *completeness*, *incrementality*, *flexibility*, and *parallel implementation*; we discuss them in this section. Morevover, as we showed that harmonic control is a special case of optimal control, it is interesting to check which of these properties are also shared by the general framework of optimal control. Conversely, we will highlight specific advantages of harmonic control over optimal control.

A fundamental property of harmonic functions is the absence of local optima. Whatever the starting position, planned trajectories always lead to a goal: such trajectories as said to be complete. This property makes harmonic function particularly interesting compared to other local potential field navigation [23] which might exhibit local optima. In general, the value function of an optimal control problem can have local optima. In the model we described in this paper, the underlying control law is complete in two cases: in the deterministic case ($\sigma = 0$) and, obviously, in the limit (harmonic) case ($\sigma = \infty$). In an environment with a large amount of noise, the control law may not attain a goal if this proves riskier (in the sense of the criterion defined in equation 6) than moving towards the goal.

Computation of harmonic functions using relaxation methods is performed iteratively, which allows incremental updates of the environment. Newly detected obstacles may be integrated in the model as new boundary conditions during computation. This property permits to use them in dynamic environments or in environments explored on-line, as for instance in [4]–[6], [8]. Value functions, which can also be computed using relaxation, also permit incremental updating of the environment's model: new obstacles and goals may be integrated during computation. More generally, updating incomplete or dynamic models of the environment within the framework of optimal control is known as "*indirect* reinforcement learning". Moreover, research on "*direct* reinforcement learning" shows that incremental planning can be done even without maintaining an explicit model of the environment (see [11] for a general introduction, and [24] for a study of the continuous time case).

Trajectories generated by harmonic control can be quite flexible. Using Dirichlet boundary conditions, one can generate safe trajectories that have the tendency to stray away from obstacles. This is due to the fact that the potential flow is orthogonal to the boundaries. Using Neumann conditions, the potential flow is tangential to the boundary which permits to have riskier trajectories. It is also possible to combine both boundary conditions [13], [25] which allows to have different intermediate behaviors. In optimal control, trajectories may also be refined in numerous ways. Contrary to the harmonic case, we can express precisely, via the cost functions c and C, the relative severity of hitting certain obstacles compared to others, the relative importance of attaining certain goals compared to others and the relative importance of attaining goals rather than hitting obstacles. In the same way, the laws of dynamics f, and the noise parameter σ, are parameters that may further explicitly influence the nature of the produced trajectories.

Another interesting property of harmonic functions that has been advocated is their inherent distribution of computation. The relaxation methods used to solve Laplace's equation are naturally distributable; computing grid values rely only on the local information at neighboring cells. Low level hardware implementation were proposed by Trassenko and Blake [26] implemented as resistive grid arrays. Furthermore parallelization makes this model a potential candidate to explain the computations of control undergone in the brain. For instance, Connolly and Burns [27] argue that the basal ganglia could compute harmonic function for motor control. The computation of the value function is also a natural candidate for parallelization. Iterative methods such as Gauss-Seidel or Jacobi, which solve the discrete version of the problem, may be implemented completely asynchronously using massively parallel architectures [28]. Though it is not clear whether these algorithms can be implemented in *very low level* architectures such as resistive grids, it is possible to implement them on parallel processors with communication delays and no synchronization [29]. Finally, as it is the case for harmonic functions, such parallel implementations motivated biological analogies with what may happen within the brain [30].

As we have just seen, most of the interesting properties of harmonic control are shared with optimal control, except the possibility of *very low level* implementation such as resistive grids. There is, however, another important property which is characteristic of harmonic control: Laplace equation (which characterizes the harmonic function) is a linear PDE for which one can derive analytical fundamental solutions *in any dimension* whereas this is generally not the case for the Hamilton-Jacobi-Bellman PDE. This does not make a real difference when one uses a relaxation technique for iteratively computing the potential. However, the linearity and the existence of such fundamental solution was exploited by Viéville et al. [31] to tackle the curse of dimensionality: the harmonic function is decomposed in a sum of fundamental (possibly high-dimensional) harmonic functions. With such an approach, the authors can compute trajectories for a 10-degrees of freedom virtual robotic arm in the presence of obstacles. Extending this idea to optimal control in general does not seem obvious and constitutes future research.

VI. Conclusion

This paper presented a link between harmonic control and optimal control: we argued that optimal control subsumes harmonic control. The results presented show formally that for a well chosen optimal control problem, the value function becomes harmonic when the noise level tends to infinity and that in both cases, control consists in following the gradient of these functions. The convergence is showed in the continuous domain and in a standard finite difference scheme. We believe that such an analysis provides more insight on harmonic control. We discussed different properties of both theories, and showed that many interesting properties of harmonic control are shared by optimal control. However, a specific power of harmonic controllers lies in the ability to implement them on very low level hardware as resistive arrays, and this does not seem to be the case for the Hamilton-Jacobi-Bellman equation. It would be interesting to further investigate the possibilities for such hardware implementations of optimal control. Furthermore, harmonic control relies on the computation of a simple linear PDE, which has some analytical fundamental solutions; this has recently been exploited to tackle the curse of dimensionality [31]. The transposition of this work to the optimal control framework is currently under investigation.

References

[1] C. I. Connolly, J. B. Burns, and R. Weiss, "Path planning using laplace's equation," in *Proceedings of the 1990 IEEE International Conference on Robotics and Automation*. IEEE, May 1990, pp. 2102–2106.

[2] S. Akishita, S. Kawamura, and K. Hayashi, "Laplace potential for moving obstacle avoidance and approach of a mobile robot," in *Japan-USA Symposium on Flexible Automation, A Pacific Rim Conference*, 1990, pp. 139–142.

[3] A. M. S.A. Masoud, "Motion planning in the presence of directional and obstacle avoidance constraints using nonlinear anisotropic, harmonic potential fields: A physical metaphor," *IEEE Transactions on Systems, Man, & Cybernetics, Part A: systems and humans*, vol. 32, no. 6, pp. 705–723, November 2002.

[4] J. Zelek, "Complete real-time path planning during sensor-based discovery," in *IEEE/RSJ International Conference on Intelligent Robots and systems*, 1998.

[5] D. Alvarez, J. Alvarez, and R. Gonzalez, "Online motion planning using laplace potential fields," in *Proceedings of the IEEE International Conference Robotics and Automation, 2003.*, vol. 3, 2003, pp. 3347–3352.

[6] H. Feder and J. Slotine, "Real-time path planning using harmonic potentials in dynamic environments," in *Proceedings of the IEEE International Conference Robotics and Automation, 1997*, vol. 1, 1997, pp. 874–881.

[7] M. Huber, W. MacDonald, and R. Grupen, "A control basis for multi-legged walking," in *Proceedings of the IEEE International Conference Robotics and Automation, 1996*, vol. 4, 1996, pp. 2988–2993.

[8] M. Kazemi, M. Mehrandezh, and K. Gupta, "An incremental harmonic function-based probabilistic roadmap approach to robot path planning," in *Proceedings of the IEEE International Conference Robotics and Automation, 2005*, 2005, pp. 2136–2141.

[9] J. Sweeney, H. Li, R. Grupen, and K. Ramamritham, "Scalability and schedulability in large, coordinated, distributed robot systems," in *Proceedings of the IEEE International Conference Robotics and Automation, 2003*, vol. 3, 2003, pp. 4074–4079.

[10] Y. Wang and G. Chirikjian, "A new potential field method for robot path planning," in *Proceedings of the IEEE International Conference Robotics and Automation, 2000*, vol. 2, 2000, pp. 977–982.

[11] R. S. Sutton and A. G. Barto, *Reinforcement Learning : An Introduction*. MIT Press, 1998.

[12] R. E. Bellman, *Dynamic Programming*. Princeton University Press, 1957.

[13] S. P. Singh, A. G. Barto, R. A. Grupen, and C. I. Connolly, "Robust reinforcement learning in motion planning," in *Advances in Neural Information Processing Systems Conference*, 1994, pp. 655–662.

[14] J. A. Coelho Jr., R. Sitaraman, and R. A. Grupen, "Parallel optimization of motion controllers via policy iteration," in *Advances in Neural Information Processing Systems*, vol. 8, 1996, pp. 996–1002. [Online]. Available: citeseer.ist.psu.edu/coelho96parallel.html

[15] T. Perkins and A. Barto, "Lyapunov-constrained action sets for reinforcement learning," in *Proc. 18th International Conf. on Machine Learning*. Morgan Kaufmann, San Francisco, CA, 2001, pp. 409–416. [Online]. Available: citeseer.ist.psu.edu/perkins01lyapunovconstrained.html

[16] C. I. Connolly and R. A. Grupen, "Nonholonomic path planning using harmonic functions," Computer Science Department, University of Massachusetts, Tech. Rep. 94-50, May 1994.

[17] C. Connolly, "Harmonic functions and collision probabilities," *International Journal of Robotics Research*, vol. 16, no. 4, pp. 497–507, Aug 1997.

[18] N. Krylov, *Controlled Diffusion Processes*. New York: Springer-Verlag, 1980.

[19] R. Munos, "A study of reinforcement learning in the continuous case by the means of viscosity solutions," *Machine Learning*, vol. 40, no. 3, pp. 265–299, 2000.

[20] H. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer Verlag, 1992.

[21] L. Evans, *Partial Differential Equations*. AMS, 1991.

[22] A. Boumaza and B. Scherrer, "Optimal control subsumes harmonic control, INRIA Technical report inria-00119243," 2006. [Online]. Available: http://hal.inria.fr/inria-00119243/en/

[23] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *International Journal of Robotic Research*, vol. 5, no. 1, pp. 90–98, 1986.

[24] R. Munos and P. Bourgine, "Reinforcement learning for continuous stochastic control problems," in *Advances in Neural Information Processing Systems Conference*, 1997.

[25] C. Connolly and R. Grupen, "On the applications of harmonic functions to robotics," *Journal of Robotic and Systems*, vol. 10, no. 7, pp. 931–946, October 1993.

[26] L. Trassenko and A. Blake, "Analogue computation of collision-free paths," in *International Conference on Robotics and Automation*. IEEE, April 1991, pp. 540–545.

[27] C. Connolly and J. Burns, "The basal ganglia and the planning of actions," COINS Department, University of Massachusetts, Tech. Rep. 92-18, Feb 1992.

[28] D. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice hall, 1989.

[29] D. Bertsekas, "Distributed asynchronous computation of fixed points." *Math. Programming*, vol. 27, pp. 107–120, 1983.

[30] B. Scherrer, "Asynchronous neurocomputing for optimal control and reinforcement learning with large state space," *Neurocomputing*, vol. 63, pp. 229–251, 2005.

[31] T. Viéville and C. Vadot, "An improved biologically plausible trajectory generator," INRIA, Tech. Rep. RR-4539, September 2002.