Optimality in Robot Motion: Optimal Versus Optimized Motion
Jean-Paul Laumond, Nicolas Mansard, Jean-Bernard Lasserre

To cite this version:
Jean-Paul Laumond, Nicolas Mansard, Jean-Bernard Lasserre. Optimality in Robot Motion: Optimal Versus Optimized Motion. Communications of the ACM, ACM, 2014, 57 (9), pp.82 - 89. 10.1145/2629535. hal-01376749

HAL Id: hal-01376749
https://hal.archives-ouvertes.fr/hal-01376749
Submitted on 7 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Optimality in Robot Motion: Optimal versus Optimized Motion

Jean-Paul Laumond  Nicolas Mansard  Jean-Bernard Lasserre

Abstract

The paper emphasizes the distinction between an optimal robot motion and a robot motion resulting from the application of optimization techniques. Most of the time, optimal motions do not exist and when they exist they are difficult to compute. A clear distinction should be made between optimal motion planning and motion optimization.

1 Introduction

The first book dedicated to “Robot Motions” was published in 1982 with the subtitle “Planning and Control” [5]. The distinction between motion planning and motion control has mainly historical roots. Sometimes motion planning refers to geometric path planning, sometimes it refers to open loop control; sometimes motion control refers to open loop control, sometimes it refers to close loop control and stabilization; sometimes planning is considered as an off-line process whereas control is real-time. From a historical perspective, robot motion planning arose from the ambition to provide robots with motion autonomy: the domain was born in the computer science and artificial intelligence communities [22]. Motion planning is about deciding on the existence of a motion to reach a given goal and computing one if this one exists. Robot motion control arose from manufacturing and the control of manipulators [30] with rapid effective applications in automotive industry. Motion control aims at transforming a task defined in the robot workspace into a set of control functions defined in the robot motor space: a typical instance of the problem is to find a way for the end-effector of a welding robot to follow a predefined welding line.

What kind of optimality is about in robot motion? Many facets of the question are treated independently in different communities ranging from control and computer science, to numerical analysis and differential geometry, with a large and diverse corpus of methods including e.g. the maximum principle, the applications of Hamilton-Jacobi-Bellman equation, quadratic programming, neural networks, simulated annealing, genetic algorithms, or Bayesian inference. The ultimate goal of these methods is to compute a so-called optimal solution whatever the problem is. The objective of the paper is not to overview this entire corpus that follows its own routes independently from robotics, but rather to emphasize the distinction between “optimal motion” and “optimized motion”. Most of the time, robot algorithms aiming at computing an optimal motion provide in fact an optimized motion which is not optimal at all, but is the output of a given optimization method. Computing an optimal motion is mostly a challenging issue as it can be illustrated by more than twenty years of research on wheeled mobile robots (Section 8).

Note that the notion of optimality in robot motion as it is addressed in this paper is far from covering all the dimensions of robot motion [7]. It does account neither for low level dynamical control, nor for sensory-motor control, nor for high level cognitive approaches to motion generation (e.g., as developed in the context of robot soccer or in task planning).

2 What is Optimal in Robot Motion Planning and Control?

Motion planning explores the computational foundations of robot motion, by facing the question of the existence of admissible motions for robots moving in an environment populated with obstacles: how to transform the continuous problem into a combinatorial one?

This research topic [22, 26] evolved in three main stages. In the early 80’s, Lozano Perez first transformed the problem of moving bodies in the physical space into a problem of moving a point in some so-called configuration space [28]. In doing so, he initiated a well-defined mathematical problem: planning a robot motion is equivalent to searching for connected components in a configuration space. Schwartz and Sharir then show that the problem is decidable as soon as we can prove that the connected components are semi-algebraic sets [35]. Even if a series of papers from computational geometry explored various instances of the problem, the general “piano mover” problem remains intractable [14]. Finally by relaxing the completeness exigence for the benefit of probabilistic completeness, Barraquand and Latombe introduced in the early 90’s
Motion planning solves a point-to-point problem in the configuration space. Whereas the problem is a difficult computational challenge that is well understood, optimal motion planning is a much more difficult challenge. In addition to finding a solution to the planning problem (i.e. a path that accounts for collision-avoidance and kinematic constraints if any), optimal motion planning refers to finding a solution that optimizes some criterion. These can be the length, the time or the energy (which are equivalent criteria under some assumption), or more sophisticated ones, as the number of maneuvers to park a car.

In such a context many issues are concerned with optimization:

- For a given system, what are the motions optimizing some criteria? Do such motions exist? The existence of optimal motion may depend either on the presence of obstacles or on the criterion to be optimized (Section 3).

- When optimal motions exist, are they computable? If so, how complex is their computation? How to relax exactness constraints to compute approximate solutions? Section 4 addresses the combinatorial structure of the configuration space induced by the presence of obstacles and by the metric to be optimized. Time criterion is considered in Section 5, while Section 6 overviews practical approaches to optimize time along a predefined path.

- Apart from finding a feasible solution to a given problem, motion planning also wants to optimize this solution once it has been found. The question is particularly critical for the motions provided by probabilistic algorithms that introduce random detours. The challenge here is to optimize no more in the configuration space of the system, but in the motion space (Section 7).

In the rest of the paper, optimal motion planning is understood with the underlying hypothesis that the entire robot environment is known and the optimization criterion is given: the quest is to find a global optimum without considering any practical issue such as model uncertainties or local sensory feedback.

However, most of the time, robots do not have access to a global knowledge of their environment. So the problem of optimal motion planning becomes a local one. At a given time, we can only use the partial information available. This means that the optimal motions deal with solutions of local optimization problems. They may be computed on line giving rise to real-time implementations based on sensory feedback information, as we will discuss it in the companion paper [23].

3 Optimal Motion Existence

Before trying to compute an optimal motion, the first question to ask is about its existence. To give some intuition about the importance of this issue, consider a mobile robot moving among obstacles. For some right security reason, the robot cannot touch the obstacles. In mathematical language, the robot has to move in an open domain of the configuration space. Yet, an optimal motion to go from one place to another one located behind some obstacle will necessarily touch the obstacle. So this optimal motion is not a valid one. It appears as an ideal motion that cannot be reached. The best we can do is to get a collision-free motion whose length approaches the length of this ideal shortest (but non-admissible) motion. In other words, there is no optimal solution to the corresponding motion planning problem. The question here is of topological nature: combinatorial data structures (e.g. visibility graphs) may allow to compute solutions that are optimal in the closure of the free space, and that are not solutions at all in the open free space.

Even without obstacle, the existence of an optimal motion is far from being guaranteed. In deterministic continuous-time optimal control problems we usually search for a time-dependent control function that optimizes some integral functional over some time interval.

Addressing the issue of existence requires to resort to geometric control theory [18]: For instance, Fillipov’s theorem proves the existence of minimum-time trajectories1, whereas Prontryagin Maximum Principle (PMP) or Boltyanskii’s conditions give respectively necessary and sufficient conditions for a trajectory to be optimal. However it is usually difficult to extract useful information from these tools. If PMP may help to characterize optimal trajectories locally, it generally fails to give their global structure. Section 8 shows how subtle the question may be in various instances of wheeled mobile robots.

The class of optimal control problems for which the existence of an optimal solution is guaranteed, is limited. The minimum time problems for controllable linear systems with bounded controls belong to this class: optimal solutions exist and optimal controls are of bang-bang type. However the so-called Fuller problem may arise: it makes the optimal solution not practical at all as it is of bang-bang type with infinitely many switches. Other examples include the famous linear-quadratic-Gaussian problem (the cost is quadratic and the dynamics is linear in both control and state variables), and systems with a bounded input and with a dynamics that is affine in the control variables. In the former a closed loop optimal solution can be computed by solving algebraic Riccati equations, whereas in the latter the existence of an optimal control trajectory is guaranteed under some appropriate assumptions.

In more general cases, we can only hope to approx-
imitate as closely as desired the optimal value via a sequence of control trajectories. There is indeed no optimal solution in the too restricted space of considered control functions. This has already been realized in the nineteen-sixties. The limit of such a sequence can be given a precise meaning as soon as we enlarge the space of functions under consideration. For instance, in the class of problems in which the control is affine and the integral functional is the $L_1$-norm, the optimal control is a finite series of impulses and not a function of time (see e.g. [29]). In some problems such as the control of satellites, such a solution makes sense as it can approximately be implemented by gas jets. However, in general, it cannot be implemented because of the physical limitations of the actuators.

Changing the mathematical formulation of the problem (e.g., considering a larger space of control candidates) may allow the existence of an optimal solution. In the former case of satellite control, the initial formulation is coherent as an “ideal” impulse solution can be practically approximated by gas jets. However, in other cases the initial problem formulation may be incorrect as an ideal impulse solution is not implementable. Indeed, if we “feel” that a smooth optimal solution should exist in the initial function space considered and if in fact it does not exist, then either the dynamics and/or the constraints do not reflect appropriately the physical limitations of the system or the cost functional is not appropriate to guarantee the existence of an optimal solution in that function space. To the best of our knowledge, this issue is rarely discussed in textbooks or courses in optimal control.

4 Optimal Path Planning

Considering that a motion is a continuous function of time in the configuration (or working) space, the image of a motion is a path in that space. The “piano mover” problem refers to the path planning problem, i.e. the geometric instance of robot motion planning. The constraint of obstacle avoidance is taken into account (see Fig. 1). In that context, optimality deals with the length of the path without considering time and control. The issue is to find a shortest path between two points.

Depending on the metric that equips the configuration space, a shortest path may be unique (e.g. for the Euclidean metric) or not unique (e.g. for the Manhattan metric). All configuration space metrics are equivalent from a topological point of view (i.e. if there is a sequence of Euclidean path linking two points, then there is also a sequence of Manhattan paths linking these two points). However, different metrics induce different combinatorial properties in the configuration space. For instance, for a same obstacle arrangement, two points may be linked by a Manhattan collision-free path, while they cannot by a collision-free straight line segment: both points are mutually visible in a Manhattan metric, while they are not in the Euclidean one. So, according to a given metric, there may or may not exist a finite number of points that “watch” the entire space [25]. These combinatorial issues are particularly critical to devise sampling based motion planning algorithms.

Now, consider the usual case of a configuration space equipped with an Euclidean metric. Exploring visibility graph data structures easily solves the problem of finding a bound on the length of the shortest path among polygonal obstacles. This is nice, but this is no longer true if we consider three-dimensional spaces populated with polyhedral obstacles. Indeed finding the shortest path in that case becomes a NP-Hard problem [14]. So, in general, there is no hope to get an algorithm that computes an optimal path in presence of obstacles, even if the problem of computing an optimal path in the absence of obstacle is solved and even if we allow the piano-robot to touch the obstacles.

As a consequence of such poor results, optimal path planning is usually addressed by means of numerical techniques. Among the most popular ones are the discrete search algorithms operating on bitmap representations of work or configuration spaces [3]. The outputs we only obtain are approximately optimal paths, i.e. paths that are “not so far” from a hypothetical (or ideal) estimated optimal path. Another type of methods consists in modeling the obstacles by repulsive potential. In doing so, the goal is expressed by an attractive potential, and the system tends to reach it by following a gradient descent [21]. The solution is only locally optimal. Moreover, the method may get stuck in a local minimum without finding a solution, that a solution actually exists or not. So it is not complete. Some extensions may be considered. For instance exploring harmonic potential fields [8] or devising clever navigation functions [34] allow providing globally optimal solutions; unfortunately, these methods require an explicit representation of obstacles in the configuration space which is generally not an available information. At this stage, we can see how the presence of obstacles makes optimal path planning a difficult problem.
5 Optimal Motion Planning

In addition to obstacle avoidance, constraints on robot controls or robot dynamics add another level of difficulties. The goal here is to compute a minimal-time motion that goes from a starting state (configuration and velocity) to a target state while avoiding obstacles and respecting constraints on velocities and acceleration. This is the so-called kinodynamic motion planning problem [12]. The seminal algorithm is based on discretizations of both the state space and the workspace. It gave rise to many variants including nonuniform discretization, randomized techniques, and extensions of A* algorithms (see [26]). They are today the best algorithms to compute approximately optimal motions.

Less popular in the robot motion planning community are numerical approaches to optimal robot control [11]. Numerical methods to solve optimal control problems fall into three main classes. Dynamic programming implements the Bellman optimality principle saying that any sub-motion of an optimal motion is optimal. This leads to a partial differential equation (the so-called Hamilton-Jacobi-Bellman equation in continuous time) whose solutions may sometimes be computed numerically. However dynamic programming suffers from the well-known curse of dimensionality bottleneck. Direct methods constitute a second class. They discretize in time both control and state trajectories so that the initial optimal control problem becomes a standard static non-linear programming (optimization) problem of potentially large size, for which a large variety of methods can be applied. However, generally, local optimality is the best one can hope for. Moreover, potential chattering effects may appear hidden in the obtained optimal solution when there is no optimal solution in the initial function space. Finally, in the third category are indirect methods based on optimality conditions provided by the PMP and for which, ultimately, the resulting two-point boundary value problem to solve (e.g., by shooting techniques) may be extremely difficult. In addition, the presence of singular arcs requires specialized treatments. So direct methods are usually simpler than indirect ones even though the resulting problems to solve may have very large size. Indeed, their structural inherent sparsity can be taken into account efficiently.

At this stage, we can conclude that exact solutions for optimal motion planning remain today out of reach. Only numerical approximate solutions are conceivable.

6 Optimal Motion Planning along a Path

A pragmatic way to by-pass (not overcome) the intrinsic complexity of the above kinodynamic and numerical approaches is to introduce a decoupled approach that solves the problem in two stages: first, an (optimal) path planning generates a collision-free-path; then a time-optimal trajectory along the path is computed while taking into account robot dynamics and control constraints. The resulting trajectory is of course not time-optimal in a global sense; it is just the best trajectory for the predefined path. From a computational point of view, the problem is much simpler than the original global one because the search space (named phase plane) is reduced to two dimensions: the curvilinear abscissa along the path and its time-derivative. Many methods have been developed since the introduction of dynamic programming approaches by Shin and McKay [36] in configuration space and simultaneously by Bobrow et al [4] in the Cartesian space. Many variants have been considered including the improvement by Pfeiffer and Johann [31] that combine forward and backward integrations, until the recent work by Verscheure et al [39] who transform the problem into a convex optimization one.

7 Optimization in motion space

Extensions of path tracking methods may be considered as soon as we allow the deformations of the supporting paths. In this section we assume that some motion planner provides a first path (or trajectory). Depending on the motion planner, the path may be far from being optimal. For instance, probabilistic motion planners introduce many useless detours. This is the price to pay for their effectiveness. So, the initial path has to be reshaped, i.e., optimized with respect to certain criteria. Geometric paths require to be shortened according to a given metric. The simplest technique consists in picking pairs of points on the path and linking them by a shortest path: if the shortest path is collision-free, it replaces the corresponding portion of the initial path. Doing so iteratively, the path becomes shorter and shorter. The iterative process stops as soon as it does not improve the quality of the path significantly. The technique gives good results in practice.

Beside this simple technique, several variational methods operating in the trajectory space have been introduced. Among the very first ones, Barraquand and Ferbach [2] propose to replace a constrained problem by a convergent series of less constrained subproblems increasingly penalizing motions that do not satisfy the constraints. Each sub-problem is then solved using a standard motion planner. This principle has been successfully extended recently to humanoid robot motion planning [9].

Another method introduced by Quinlan and Khatib consists in modeling the motion as a mass-spring system [32]. The motion then appears as an elastic band that is reshaped according to the application of an energy function optimizer. The method applies for nonholonomic systems as soon as the nonholonomic metric is known [16] as well as for real-time obstacle avoidance in dynamic environments [6].

Recently, successful improvements have been introduced by following the same basic principle of optimiz-
\[
\begin{bmatrix}
\dot{x} \\
\dot{y} \\
\dot{\theta}
\end{bmatrix} = 
\begin{bmatrix}
\cos \theta & \sin \theta & 0 \\
\sin \theta & -\cos \theta & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
v \\
0 \\
\omega
\end{bmatrix}
\]

Fig. 2: A car (logo of the European Project Esprit 3 PROMotion in the 90’s) together with the unicycle model equations.

8 What we know and what we do not know about optimal motion for wheeled mobile robots

Mobile robots constitute a unique class of systems for which the question of optimal motion is best understood. Since the seminal work by Dubins in the nineteen-fifties [13], optimal motion planning and control for mobile robots has attracted a lot of interest. We briefly review how some challenging optimal control problems have been solved and which problems still remain open.

Let us consider four control models of mobile robots based on the model of a car (Fig. 2). Two of them are simplified models of a car: the so-called Dubins (Fig. 3) and Reeds-Shepp (Fig. 4) cars respectively. Dubins car moves only forward. Reed-Shepp car can moves forward and backward. Both of them have a constant velocity of unitary absolute value. Such models account for a lower bound on the turning radius, i.e. the typical constraint of a car. Such a constraint does not exist for a two-wheel differentially driven mobile robot. This robot may turn on the spot while a car can’t. Let us consider two simple controls schemes of a two-driving wheel mobile robot:\footnote{The distance between the wheels is supposed to be 2.}: in the first one (Hilare-1), the controls are the linear velocities of the wheels; in the second one (Hilare-2), the controls are the accelerations (i.e. the second system is a dynamic extension of the first).

8.1 Time-optimal trajectories

Car-like robots of Figures 3 and 4 represent two examples of non linear systems for which we know exactly the structure of the optimal trajectories. Note that in both examples the norm of the linear velocity is assumed to be constant. In those cases, time-optimal trajectories are supported by the corresponding shortest paths. Dubins solved the problem for the car moving only forward [13]. More than thirty years after, Reeds and Shepp [33] solved the problem for the car moving both forwards and backwards. The problem has been completely revisited with the modern perspective of geometric techniques in optimal control theory [38, 37]: the application of PMP shows that optimal trajectories
are made of arcs of a circle of minimum turning radius (bang-bang controls) and of straight line segments (singular trajectories). The complete structure is then derived from geometric arguments that characterize the switching functions. Dubins and Reeds-Shepp cars are among the few examples of non linear systems for which optimal control is fully understood. The same methodology applies for velocity-based controlled differential drive vehicles (Hilare-1 in Fig. 6). In that case, optimal trajectories are bang-bang, i.e. made of pure rotations and straight line segments. The switching functions are also fully characterized [1]. This is not the case for the dynamic extension of the system, i.e. for acceleration-based controlled differential drive vehicle (Hilare-2 in Fig. 7). Only partial results are known: optimal trajectories are bang-bang (i.e. no singular trajectory appears) and are made of arcs of clothoid and involutes of a circle [15]. However, the switching functions remain unknown. The synthesis of optimal control for Hilare-2 system is still an open problem.

While the existence of optimal trajectories is proven for the four systems above, a last result is worth mentioning. If we consider the Reeds-Shepp car optimal control problem in presence of an obstacle, even if we allow the car touching the obstacles, it has been proven that a shortest path may not exist [10].

8.2 Motion planning

The results above are very useful for motion planning in the presence of obstacles. In Figures 3 and 4 we display the reachable domain for both Dubins and Reeds-Shepp cars. While the reachable set of Reeds-Shepp car is a neighborhood of the origin, it is not the case for Dubins car. Stated differently, Reeds-Shepp car is small-time controllable, while Dubins car is only controllable. The consequence in terms of motion planning is important. In the case of Reeds and Shepp car, any collision-free – not necessarily feasible – path can be approximated by a sequence of collision-free feasible paths. Optimal paths allow to build the approximation, giving rise to an efficient motion planning algorithm [24]. Not only such algorithm does not apply for Dubins car, but also we even do not know today whether the motion planning problem for Dubins car is decidable or not.
In [24], we prove that the number of maneuvers to park a car varies as the inverse of the square of the clearance. This result is a direct consequence of the shape of the reachable sets. So, the combinatorial complexity of (nonholonomic) motion planning problems is strongly related to optimal control and the shape of the reachable sets in the underlying (sub-Riemannian) geometry [17].

9 Conclusion

When optimal solutions cannot be obtained for theoretical reasons (e.g., non existence) or for practical ones (e.g., untractability), we have seen how the problem can be reformulated either by considering a discrete representation of space and/or time, or by slightly changing the optimization criterion, or by resorting to numerical optimization algorithms. In all these cases, the resulting solutions are only approximated solutions of the original problem.

In conclusion to this rapid overview, it appears that the existence of optimal robot motions is rarely guaranteed. When it is, finding a solution has never been proven to be a decidable problem as the motion planning problem is. So, “optimal motion” is most of the time an expression that should be understood as “optimized motion”, i.e. the output of an optimization numerical algorithm. However, motion optimization techniques follow progress in numerical optimization with effective practical results on real robotic platforms, if not with new theoretical results.

The distinction between optimal and optimized motions as it is addressed in this paper is far from covering all facets of optimality in robot motion. In the companion paper [23] we consider the issue of motion optimality as an action selection principle and we discuss its links with machine learning and recent approaches to inverse optimal control.

Acknowledgments

The paper benefits from comments by Quang Cuong Pham, from a careful reading by Joel Chavas, and above all, from the quality of the reviews. The work has been partly supported by ERC Grant 340050 Actanthrope, by a grant of the Gaspar Monge Program for Optimization and Operations Research of the Fédération Mathématique Jacques Hadamard (FMJH) and by the grant ANR 13-CORD-002-01 Entracte.

References


