

# Review of Research Methods for Hypersonic Vehicle Reentry Trajectory Planning

Qingzhuang Ma, Yongqi Yin, Xueqian Shi

Rocket Force University of Engineering, Xi'an, 710025, China

**Abstract:** Hypersonic vehicle has many advantages, such as wide range of maneuver, strong penetration ability, high strike accuracy, and so on. Reentry trajectory planning is one of the key technologies to support hypersonic vehicle systems. It is necessary to plan feasible or optimal trajectory under the process constraints such as heat flux, dynamic pressure, overload, and terminal constraints such as altitude and velocity. At present, traditional methods are difficult to meet the task requirements of trajectory planning and online trajectory generation under complex conditions with multiple constraints. As an artificial intelligence method, reinforcement learning has strong robustness and the characteristics of "offline training and online deployment", which can make up for the shortcomings of traditional methods and show great potential in trajectory planning. This paper introduces the current research status of traditional trajectory planning methods and reinforcement learning methods, and proposes that the reentry trajectory planning methods will be intelligent in the future.

**Keywords:** Hypersonic flight vehicle; Reentry trajectory planning; Reinforcement Learning; Imitation learning.

## 1. Introduction

Hypersonic vehicle has many advantages, such as wide maneuvering range, strong penetration ability, high strike accuracy, etc. At present, it has become a key research object in the aerospace field. Reentry trajectory planning is one of the key technologies to support hypersonic vehicle systems. It is necessary to plan feasible or optimal trajectory under the process constraints such as heat flux, dynamic pressure, overload, and terminal constraints such as altitude and velocity. At present, traditional methods are difficult to meet the task requirements of trajectory planning and online trajectory generation under complex conditions with multiple constraints. As an artificial intelligence method, reinforcement learning has strong robustness and the characteristics of "offline training and online deployment", which can make up for the shortcomings of traditional methods and show great potential in trajectory planning. This paper introduces the current application and research of trajectory planning, and discusses and puts forward the development direction of future trajectory planning methods towards intelligence.

## 2. Current research status of hypersonic vehicle reentry trajectory planning method

The reentry phase of hypersonic vehicle is the general term of reentry initial phase and reentry glide phase. During reentry flight, there are many challenges, such as: the control system is challenged by the uncertainty of environmental model and the dramatic change of aerodynamic parameters, the thermal protection system is challenged by a large amount of aerodynamic heat, the huge aerodynamic force generated by high-speed flight, and the structural strength of aircraft is challenged by overload [2]. According to the different principles of reentry trajectory planning methods, it can be divided into reentry corridor-based trajectory planning methods and optimization/intelligent algorithm-based trajectory planning methods.

### 2.1. Section Headings

It can make reentry trajectory planning efficiently by using reentry corridor, and can also meet the real-time requirements of online trajectory generation. The online trajectory planning algorithm can make full use of the orbit and attitude information before departure, so compared with the nominal trajectory planned offline, it can significantly reduce the initial deviation of the entry point. The online generation of hypersonic vehicle reentry trajectory is difficult, mainly due to the following two reasons:

- (1) The control variables in the glide phase are hidden in the nonlinear time-varying differential motion equations, and it is difficult to obtain high-precision three degree of freedom trajectory through analytical method.
- (2) The hypersonic glide process is affected by many uncertain factors, such as large changes in altitude and Mach number, drastic changes in aerodynamic parameters and so on, which easily leads to large errors in online trajectory generation.

The representative methods of online trajectory planning using reentry corridors are Quasi Equilibrium Gliding Condition (QEGC) and Evolved Acceleration Guidance Logic for Entry (EAGLE). Z Shen [8] and others proposed the QEGC method, which combines the process constraints such as heat flow density, dynamic pressure, overload and quasi equilibrium glide conditions into the constraints on the pitch angle in the height velocity profile. Combined with the designed piecewise linear pitch angle profile, it effectively reduces the amount of calculation and realizes the online generation of trajectory. Mease [9] et al. divided the drag acceleration velocity profile planning into two subproblems: trajectory length planning and trajectory curvature planning, and used the alternating iteration of longitudinal motion and lateral motion to plan the required three-dimensional trajectory. Leavitt [10] et al. proposed an EAGLE method based on Mease to plan the three-dimensional trajectory with large lateral maneuver. Y Zhang [11] proposed a trajectory planning method based on 3D drag acceleration profile, and deeply discussed the trajectory planning and guidance method

of standard profile focusing on large-scale lateral maneuver missions.

## 2.2. Trajectory planning method based on optimization/intelligent algorithm

Trajectory planning often needs to make the aircraft meet process constraints and terminal constraints to plan a trajectory with the best index, that is, trajectory optimization problem. The most commonly used method for trajectory optimization is numerical method. According to different solving principles, it can be divided into indirect method and direct method.

### Indirect method

The indirect method uses the classical variational method or Pontryagin minimum principle to transform the optimal control problem into a Hamilton Boundary Value Problem (HBVP), and uses the corresponding numerical method to solve it [12]. As early as 1968, Lewallen [13] used the indirect method to realize the iterative solution of the trajectory optimization problem. The advantage of the indirect method is that it has high accuracy, and the solution meets the necessary conditions of first-order optimality. However, it is difficult to derive the optimal solution, estimate the initial value of the conjugate variable, and require a lot of integral operations. The algorithm has poor real-time performance and reliability, and is rarely used in engineering practice.

### Direct method

The direct method usually first discretizes and parameterizes the optimal control problem, converts it into a nonlinear programming problem, and solves it with nonlinear programming algorithms such as Sequential Quadratic Programming (SQP) [14-15] or heuristic algorithms [16]. The direct method has the advantages of large convergence radius, less strict requirements on initial values, and good robustness, and is widely used in aircraft trajectory optimization. However, its solution does not necessarily satisfy the necessary condition of first-order optimality, and the effect of solving discontinuous problems is poor. The parameterization process of optimal control problem is to transform the optimal control problem described by differential equations into a static parameter optimization problem that can be solved by nonlinear programming method. Representative methods include direct shooting method, collocation method, pseudo spectral method, etc.

The direct shooting method only discretizes the control variables, and the state variables are obtained by numerical integration in a single interval from the motion equation. Hull [17] and Bruschi [18] gave the solution scheme of direct shooting method earlier and transformed it into a nonlinear programming problem. Fisch [19] used this method to study the trajectory optimization problem with path point constraints. The collocation method is a method to discretize control variables and state variables at the same time. A series of algebraic equations are used to replace the differential equations in the optimal control problem. In 1987, Hargraves [20] and others put forward the collocation method of using cubic Hermite interpolation approximation for state variables, and obtained the application of the famous trajectory optimization software OTIS, which solved the optimization problems such as the minimum climb of supersonic interceptors and the ascent trajectory of advanced boosters. Pseudo spectral method is a direct method that discretizes control variables and state variables simultaneously. Due to the characteristics of few parameters, high precision, fast

convergence and insensitivity to initial values, it has developed rapidly and been widely used in recent years. Elnagar [21] was first used to solve optimal control problems. The development of optimal control software DIDO [22-23] and GPOPS [24-25] also accelerated the application of pseudo spectral method.

After the optimal control problem is transformed into the corresponding parameter optimization problem through the discrete method, the nonlinear programming method can be used to solve it. In unconstrained optimization algorithms, classical algorithms such as conjugate gradient method, simplex method and BFGS have appeared. For constrained optimization problems, penalty function method and SQP method are used. The SQP method is currently the most effective method for solving constrained optimization problems, and has been made into SNOPT [26] and NPSOL [27] software packages.

Gradient based algorithms such as nonlinear programming methods are easy to fall into local optima and difficult to obtain global optima. Heuristic intelligent optimization algorithms such as particle swarm optimization algorithm and genetic algorithm [28] are insensitive to initial values and have strong robustness. They can effectively solve local optimization problems in the solution process, but the calculation speed cannot meet the requirements.

## 3. Track planning method based on reinforcement learning

Reinforcement learning is a machine learning method that can acquire data independently and interactively with the environment to train the neural network. Compared with the deep learning algorithm that requires a large number of sample data, it has obvious advantages in ease of use. By modeling the problem, the Markov decision process is constructed, and combined with reasonable reward function design, the intelligent decision model can be obtained by using reinforcement learning algorithm training. The deep reinforcement learning, which combines the deep learning and reinforcement learning, can realize the end-to-end learning method from the original input to the direct control output, and provides a new solution for the perceptual decision-making task of complex systems. In the aspect of guidance, Daseon Hong [42] proposed a reinforcement learning intelligent terminal guidance law based on DDPG algorithm, and verified that its robustness is better than proportional guidance law in the first-order delay system, and it can replace the existing missile guidance law. J Wang [43] proposed an intelligent cooperative interception guidance law based on Q-learning algorithm, which is better than the traditional proportional guidance law, for the problem of multi missile coordination. Kirk Hovell [44] used deep reinforcement learning to generate guidance instructions, and used traditional controllers to track, which realized the migration and deployment of spacecraft docking strategies trained in the simulation environment to the actual environment, and narrowed the gap between simulation and reality. In terms of path planning, R Zhi [45] proposed a three-dimensional A\* improved algorithm based on reinforcement learning to solve the problem of high requirements for real-time algorithm and optimal results in aircraft online path planning, which effectively improved the time performance of the algorithm while ensuring optimal path results. Y Lv [46] proposed a path planning method for hypersonic vehicles

based on Q-Learning algorithm, which verified the feasibility of Q-Learning algorithm for hypersonic vehicles and took less time than traditional algorithms.

At present, most of the research on hypersonic vehicle trajectory planning using reinforcement learning methods focuses on the terminal guidance stage with short time series and few constraints. Sparse reward has no obvious effect on reinforcement learning algorithm, and the model is easy to converge. Due to the limitations of long flight time series, complex and severe constraints, and large amount of computation in the reentry phase, the application of reinforcement learning is mainly used to improve the traditional methods to reduce the complexity of the model and the difficulty in designing the reward function, and has also achieved significant results. J Zhu [47] proposed a multi constraint intelligent glide guidance strategy based on optimal guidance and reinforcement learning, which uses Q-Learning algorithm to intelligently adjust the maneuver amplitude in speed control without manually adjusting guidance parameters, thus improving the autonomy of guidance. T Wu [48] proposed a reentry guidance method based on depth reinforcement learning and altitude rate feedback, aiming at the periodic oscillation phenomenon of hypersonic aircraft with high lift drag ratio in reentry guidance, and established an altitude rate feedback mechanism based on DDPG algorithm, which can output pitch angle compensation command for predictive correction guidance according to the current state of the aircraft.

However, the above research is mainly based on traditional methods, and has not fully played the advantages of reinforcement learning intelligent decision-making. J Song [49] proposed a fault-tolerant integrated guidance and control design method for hypersonic vehicles based on Proximal Policy Optimization (PPO), aiming at the existence of various uncertain parameters and actuator failures in the control system, established an integrated guidance and control system for HSV, and used PPO algorithm to generate pitch angle commands according to the current state of the aircraft, The coupling relationship between guidance loop and control loop is fully utilized. J Hui [50] and others proposed an online generation method of "new quality" reentry flight corridor based on reinforcement learning, which avoids manually setting flight corridor parameters. The simulation results show that the new method meets the requirements of online guidance, and the flight path is different from the traditional way, which effectively enhances the penetration capability of the aircraft. At present, the reentry process trajectory planning method based on reinforcement learning shows great potential, but it still faces the problems of poor model convergence, difficult design of reward function, and long training time.

## 4. The future development direction of reentry trajectory planning methods

The development of traditional reentry trajectory planning methods is facing more and more obstacles, and intelligence is a better solution. In the future, better solutions can be found from the following three aspects.

### 4.1. Off-line reinforcement learning trajectory planning method

Due to the dramatic changes of atmospheric parameters during reentry flight, it is difficult to establish an effective

simulation environment, which also leads to the poor effect of reinforcement learning model in practical application. The "reality gap" can be effectively avoided by using offline reinforcement learning method to train the flight data collected from flight test on the basis of real data. So as to improve the model effect.

### 4.2. Imitation learning trajectory planning method

The reentry flight environment of the trajectory planning method based on imitation learning is harsh, the reward function of reinforcement learning is difficult to set, and the probability of the agent successfully reaching the destination is small, leading to low training efficiency. The trajectory planning method based on imitation learning avoids this problem and can effectively imitate the expert trajectory in the sample to make actions efficiently.

### 4.3. Reinforcement learning path planning method based on initialization of other learning methods

Offline reinforcement learning and imitation learning are limited by small sample data and low accuracy. Simple reinforcement learning is difficult to set reward function. The model trained by other learning methods as the initialization of reinforcement learning can effectively improve the probability of the agent reaching the end point, reduce the difficulty in designing the reward function, and thus improve the accuracy of the reinforcement learning model.

## 5. Conclusion

Weapon intelligence can effectively improve weapon performance. Reentry trajectory planning can use reinforcement learning as a new method to avoid the shortcomings of traditional methods, improve the unpredictability and robustness of trajectory, and improve weapon effectiveness.

## Acknowledgment

We thank the associate editor and the reviewers for their useful feedback that improved this paper.

## References

- [1] R He , "Study of all-course trajectory planning approach for hypersonic boost-glide vehicles," National University of Defense Technology, Changsha, Hunan, China.
- [2] Z Shen, P Lu, "On-board entry trajectory planning for sub-orbital flight," *Acta Astronautica*, 2005, 56(6): 573-591.
- [3] Mease K D, D Chen, Schonenberger H, "Reduced-order entry trajectory planning for acceleration guidance," *Journal of Guidance, Control, and Dynamics*, 2002,25(2):257-266.
- [4] Leavitt J, Saraf A, Chen D, "Performance of evolved acceleration guidance logic for entry," *Proceedings of AIAA Guidance, Navigation, and Control Conference and Exhibit*, AIAA-2002-4456,2002.
- [5] Y Zhang, "Research on entry trajectory generation for hypersonic glide vehicles based on three-dimensional profile," National University of Defense Technology, Changsha, Hunan, China.
- [6] C Xie, Y Zhang, M Gao, "Automatic control principle," Nanjing Southeast University Press,201812.268.

- [7] Hull, David G, "Conversion of Optimal Control Problems into Parameter Optimization Problems," *Journal of guidance, control, and dynamics*, 1997, 20(1): 57-60.
- [8] Fahroo F, Ross I M, "Direct Trajectory Optimization by a Chebyshev Pseudospectral Method," *American Control Conference* 2000.
- [9] Vinh N X, Lu P, "Chebyshev minimax problems for skip trajectories," *Journal of the astronautical sciences*, 1988.
- [10] Barron R L, Chick C M, "Improved Indirect Method for Air-Vehicle Trajectory Optimization," *Journal of Guidance Control & Dynamics*, 2006, 29(3): 643-652.
- [11] Hull, David G, "Conversion of Optimal Control Problems into Parameter Optimization Problems," *Journal of guidance, control, and dynamics*, 1997, 20(1): 57-60.
- [12] Betts J T, Frank P D, "A sparse nonlinear optimization algorithm," *Journal of Optimization Theory & Applications*, 1994, 82(3): 519-541.
- [13] Fisch F, Sewerin F, Holzapfel F, "Approach Trajectory Optimization including a Tunnel Track Constraint," *Aiaa Atmospheric Flight Mechanics Conference*, 2011.
- [14] Hargraves C R, Paris S W, "Direct trajectory optimization using nonlinear programming and collocation," *Journal of guidance, control, and dynamics*, 1986, 10(4): 338-342.
- [15] Elnagar G, Kazemi M A, Razzaghi M, "The pseudospectral Legendre method for discretizing optimal control problems," *IEEE Transactions on Automatic Control*, 2002, 40(10): 1793-1796.
- [16] Fahroo F, Ross I M, "Direct Trajectory Optimization by a Chebyshev Pseudospectral Method," *American Control Conference*, 2000.
- [17] Fahroo F, Doman D, "A Direct Method for Approach and Landing Trajectory Reshaping with Failure Effect Estimation," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2004.
- [18] Corrigendum, "GPOPS, a MATLAB software for solving multiple phase optimal control problems using the gauss pseudospectral method," *Acm Transactions on Mathematical Software*, 2011, 38(2): 1-2.
- [19] Benson D A, Huntington G T, Thorvaldsen T P, "Direct Trajectory Optimization and Costate Estimation via an Orthogonal Collocation Method," *Journal of Guidance Control & Dynamics*, 2006, 29(6): 1435-1439.
- [20] Gill P E, Murray W, Saunders M A, "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM review*, 2005, 47(1): 99-131.
- [21] Gill P E, Murray W, Saunders M A, "NPSOL (Version 4.0): A Fortran Package for Nonlinear Programming," *User's Guide*, 1986.
- [22] G Tang, Y Luo, E Yong, "Theory, method and application of spacecraft trajectory optimization", *Science Press*, 2012.
- [23] Hong D, Kim M, Park S, "Study on Reinforcement Learning-Based Missile Guidance Law," *Applied Sciences*, 2020, 10(18): 6567.
- [24] J Wang, R Su, L Liu, "Cooperative Interception Guiding Law Based on Reinforcement Learning of Q-learning," *Navigation Positioning & Timing*, 2022, 9(05):84-90.
- [25] Hovell K, Ulrich S. Deep reinforcement learning for spacecraft proximity operations guidance[J]. *Journal of spacecraft and rockets*, 2021, 58(2): 254-264.
- [26] Z Ren, D Zhang, S Tang, "Improved three-dimensional A\* algorithm of real-time path planning based on reinforcement learning", *Systems Engineering and Electronics*, 1-11.
- [27] Y Lv, "Research on Q-Learning Path Planning Method for Hypersonic Vehicles," *University of Electronic Science and Technology of China, Chengdu, Sichuan, China*.
- [28] J Zhu, C Zhao, X Li, "Multi constraint optimal intelligent gliding guidance via reinforcement learning", *Journal of National University of Defense Technology*, 2022,44(04):116-124.
- [29] T Wu, H Wang, Y Liu, "A Reentry Guidance Algorithm Based on Deep Reinforcement Learning and Altitude Rate Feedback", *Unmanned Systems Technology*, 2022,5(04):1-13.
- [30] J Song, Y Luo, M Zhao, "Fault-Tolerant Integrated Guidance and Control Design for Hypersonic Vehicle Based on PPO," *Mathematics*, 2022, 10(18): 3401.
- [31] J Hui, R Wang, Q Yu, "Research of generating new quality flight corridor for reentry aircraft based on reinforcement learning," *Acta Aeronautica et Astronautica Sinica*, 2022,43(09):623-635.