

IAF ASTRODYNAMICS SYMPOSIUM (C1)
Attitude Dynamics (1) (1)

Author: Mr. Ahmed Mahfouz
University of Luxembourg, Luxembourg , Ahmed.Mahfouz@uni.lu

Mr. Ayrat Valiullin
Russian Federation, imjobjob@gmail.com
Mr. Alexey Lukashevichus
Russian Federation, lukashevichus.aa@yandex.ru
Dr. Dmitry Pritykin
Russian Federation, dpritykin@rambler.ru

REINFORCEMENT LEARNING FOR ATTITUDE CONTROL OF A SPACECRAFT WITH
FLEXIBLE APPENDAGES

Abstract

This study explores the reinforcement learning (RL) approach to constructing attitude control strategies for a LEO satellite with flexible appendages. Attitude control system actuated by a set of three reaction wheels is considered. The satellite is assumed to move in a circular low Earth orbit under the action of gravity-gradient torque, random disturbance torque, and oscillations excited in flexible appendages. The control policy for rest-to-rest slew maneuvers is learned via the Proximal Policy Optimization (PPO) technique. The robustness of the obtained control policy is analyzed and compared to that of conventional controllers.

The first part of the study is focused on problem formulation in terms of Markov Decision Processes, analysis of different reward-shaping techniques, and finally training the RL-agent and comparing the obtained results with the state-of-the-art RL-controllers as well as with the performance of a commonly used quaternion feedback regulator (Lyapunov-based PD controller). We then proceed to consider the same spacecraft with flexible appendages added to its structure. Equations of excitable oscillations are appended to the system and coupling terms are added describing the interactions between the main rigid body and the flexible structures. The dynamics of the rigid spacecraft thus becomes coupled with that of its flexible appendages and the control strategy should change accordingly in order to prevent actions that entail excitation of oscillation modes. Again PPO is used to learn the control policy for rest-to-rest slew maneuvers in the extended system. All in all, the proposed reinforcement learning strategy is shown to converge to a policy that matches the performance of the quaternion feedback regulator for a rigid spacecraft. It is also shown that a policy can be trained to take into account the highly nonlinear dynamics caused by the presence of flexible elements that need to be brought to rest in the required attitude. We also discuss the advantages of the reinforcement learning approach such as robustness and ability of online learning pertaining to the systems that require a high level of autonomy.