

Planning of operations and propulsion maneuvers in satellites based on digital twins and multidisciplinary optimization

Gianni Pecora

Supervisor: Daniel González-Arribas (UC3M)

Advisor: Hodei Urrutxua (URJC) • Advisor: Manuel Sanjurjo-Rivo (UC3M)

PhD Aerospace Engineering UC3M
Doctoral Meetings 2026

The Guidance Gap in Electric Propulsion

7,000+

operational EP satellites

$I_{sp} \geq 1,000 \text{ s}$

Hall-effect thrusters

~6 months

GTO → GEO all-electric (Eutelsat 172B)

THE PROBLEM

- ▶ Onboard guidance must be closed-loop, lightweight, and certifiable.
- ▶ Adaptive weight scheduling improves Q-law *but breaks* its Lyapunov certificate.
- ▶ No existing framework couples adaptive guidance with realistic power dynamics.

Goal: Establish asymptotic Lyapunov stability for state-dependent Q-law scheduling with integrated power/thruster coupling.

Industrial PhD with IENAI SPACE – all development integrated into the i360™ mission-analysis suite.

State of the Art & Open Gaps

Q-law Petropoulos 2004	NSGA-II tuning Varga 2016	RL-scheduled Q-law Holt et al. 2021
APPROACH Lyapunov feedback	APPROACH Offline multi-obj. optimization	APPROACH PPO actor schedules $W(x)$
STRENGTH Closed-form	STRENGTH Multi-objective, eclipse-aware	STRENGTH State-dependent, adaptive
LIMITATION Static weights	LIMITATION Mission-specific, non-adaptive	LIMITATION No asymptotic stability guarantee
GAP 1 No formal asymptotic stability for state-dependent Q-law weight scheduling		RQ1 Provable Lyapunov stability with RL-scheduled weights
GAP 2 Adaptive guidance decoupled from power dynamics (eclipse, SOC, variable I_{sp})		RQ2 Native coupling with realistic power and thruster physics
		RQ3 Benchmark validation without performance loss

Q-law & the stability problem

Proximity quotient

$$Q(\mathbf{x}, \mathbf{W}) = W_a q_a + W_e q_e + W_i q_i$$

$$q_j = \frac{(x_j - x_j^*)^2}{\dot{x}_{j,xx}^2}$$

Each penalty q_j normalized by max achievable element rate

Optimal thrust

$$\hat{\mathbf{u}}^* = -\frac{\mathbf{c}(\mathbf{x})}{\|\mathbf{c}(\mathbf{x})\|} \quad \mathbf{c} = \mathbf{G}(\mathbf{x})^\top \nabla_{\mathbf{x}}(Q)$$

Closed-loop steering law – analytical

Descent (fixed $\mathbf{W}$)

$$\dot{Q}_{min} = -a_{th} \|\mathbf{c}\| \leq 0$$

Asymptotic convergence guaranteed with constant weights

⚠ When $\mathbf{W} = \mathbf{W}(t)$

$$\dot{Q} = \sum_j \frac{\partial Q}{\partial x_j} \dot{x}_j + \sum_j \frac{\partial Q}{\partial W_j} \dot{W}_j$$

impulsive at each decision epoch

Lyapunov Stability

- Q positive definite
- $\dot{Q} \leq 0$ along trajectories
- $\mathbf{x}(t) \rightarrow \mathbf{x}^*$

⚠ $\mathbf{W} = \mathbf{W}(t) \Rightarrow \dot{Q} \leq 0$ not guaranteed

— Q no longer valid Lyapunov function

Separate the certificate: freeze V, free the control weights

- Freeze a state-only Lyapunov candidate $\dot{V} = a_{th} c_V^T \hat{u}$
- Let RL steer with its own weights $\hat{u}^* = -c_Q / \|c_Q\|$

Construction:

$$V(x) = Q(x, W^0) = W_a^0 q_a + W_e^0 q_e + W_i^0 q_i$$

W^0 = frozen reference weights (anchor)

Stability condition:

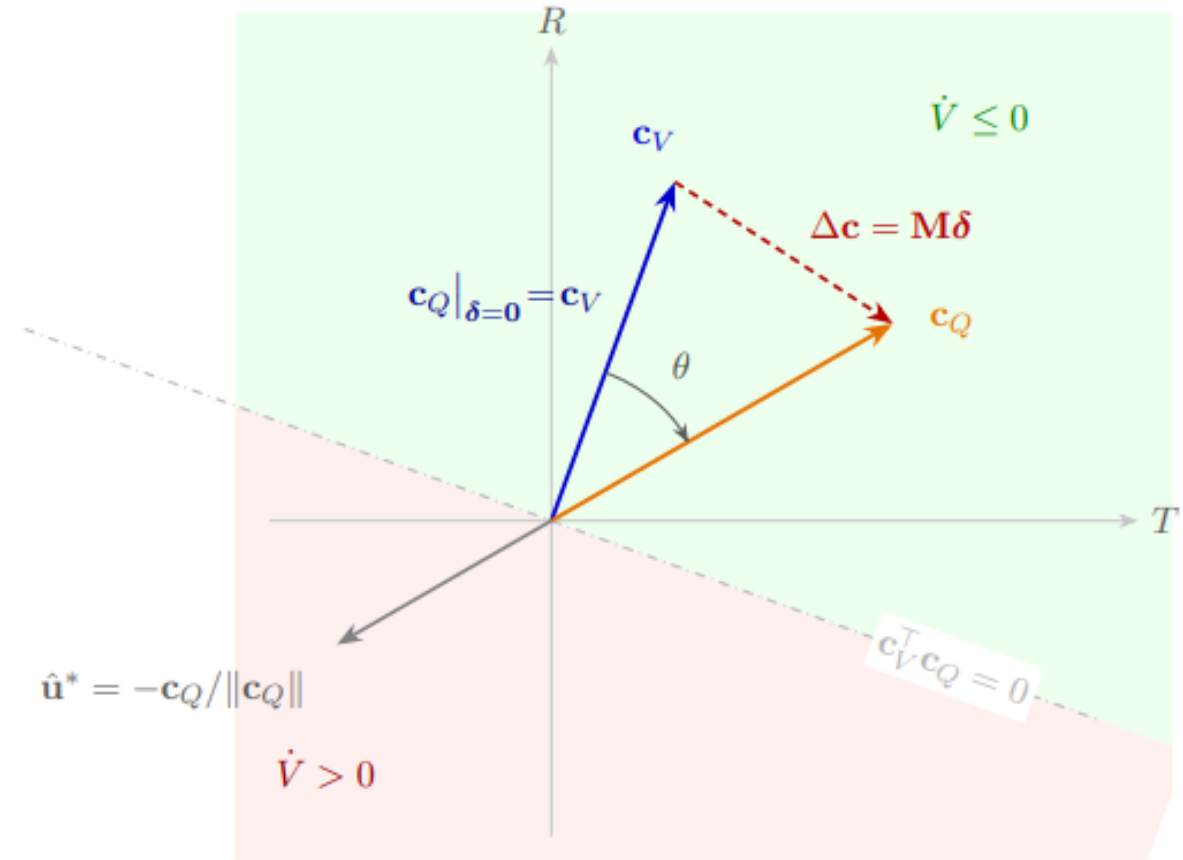
$$\dot{V}|_{\hat{u}^*} = -a_{th} \frac{c_V^T c_Q}{\|c_Q\|} \Rightarrow \dot{V} \leq 0 \Rightarrow c_V^T c_Q \geq 0$$

Perturbation decomposition ($\delta = W_k - W_0$)

$$c_Q - c_V = M(x)\delta$$

($M \in \mathbb{R}^{3 \times 3}$ perturbation matrix)

Key insight: Weight deviation δ rotates c_Q away from c_V via M .
Stability holds in green cone.
Global safe set W^* always contains anchor W^0 .



Analytical bound + pointwise enforcement

Theorem: Effectivity–Alignment Bound

Let $C_Q = \max'_v \|c_{Q(v')}\|$ and $\sigma^* = \sup_{v \in [0, 2\pi]} \|M(v)\|_2$

Then $c_V^T c_Q \geq 0$ is guaranteed wherever:

$$\eta(v) \geq \frac{2\sigma^* \|\delta\|}{C_Q}$$

Corollary (operational):

$$\|\delta_k\| \leq \delta^* = \frac{\eta_{\text{thresh}} C_Q}{(2\sigma^*)}$$

Key insight: η_{thresh} gates *both* propellant efficiency and Lyapunov stability — zero extra runtime cost.

Runtime Enforcement Pipeline

Worst-case alignment sweep (each epoch t_k):

- v : samples over $[0, 2\pi]$
- State: $3 \times 3 \times 3$ grid over reachable drift

$$\alpha_k^{wc} = \min \left(\frac{(c_V^T c_Q)}{\|c_V\| \|c_Q\|} \right)$$

Bisection projection (if $\alpha_k^{wc} < 0$):

- Convex half-space $\rightarrow$ unique crossing

$$W_k^{\text{safe}} = (1 - \lambda)W_k + \lambda W_0$$

✓ Performance Metrics

0 Lyapunov violations across all campaigns & environments
 ~2 ms median · ~13 ms worst-case · overhead <1% of propagation

PPO Scheduling with Stability-in-the-Loop

$\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, P, R, \gamma \rangle$ state · action · transition · reward · discount

State

- Orbital elements: **a** (SMA), **e** (eccentricity), **i** (inclination)
- **Time-optimal**: state only | **Mass-optimal**: + elapsed t, propellant m_p

Action

$u \in [-1, 1]^3 \rightarrow (W_e, W_i, \eta_{thresh})$

Reward

Key: reward uses frozen V_0 , not current $Q(x, W(t)) \rightarrow$ prevents Q-reshaping

Algorithm

PPO — on-policy actor-critic

 **Stability-in-the-Loop** — every W_k projected onto $W(x)$ *inside the integrator* $\rightarrow$ agent never trains on unstable trajectories

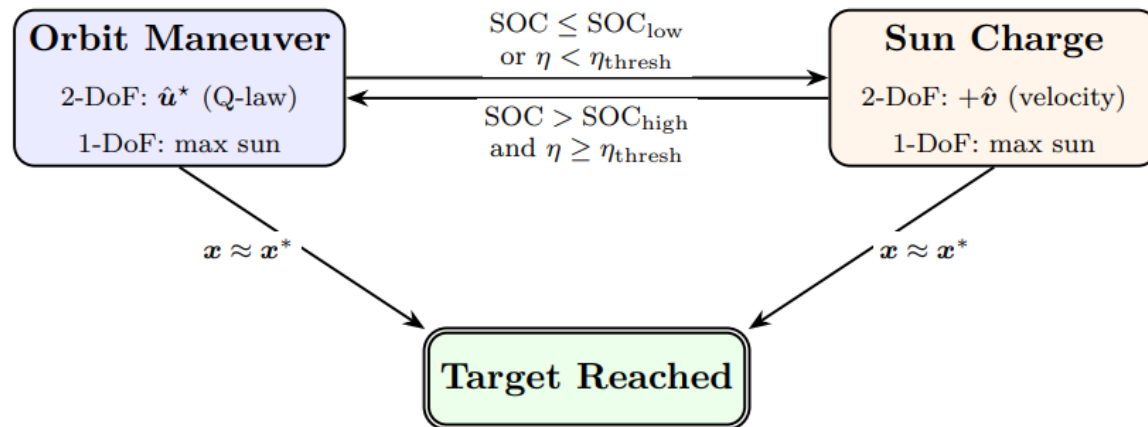
Power dynamics & throttling

BATTERY STATE-OF-CHARGE DYNAMICS

$$\dot{E} = P_{SA}(r, r_{sun}) - P_{cons}(\tau)$$

$$SOC(t) = E(t)/E_{max} \quad \tau = [0, 1]$$

- P_{SA} : solar array output (cylindrical eclipse model)
- P_{cons} : total consumption (PPU + thruster + housekeeping)
- $\tau \in [0, 1]$: throttle command



: Fixed-threshold Variable-power
 Thrust during maneuver : $\tau = 1$ (full power) $\tau = f(SOC)$

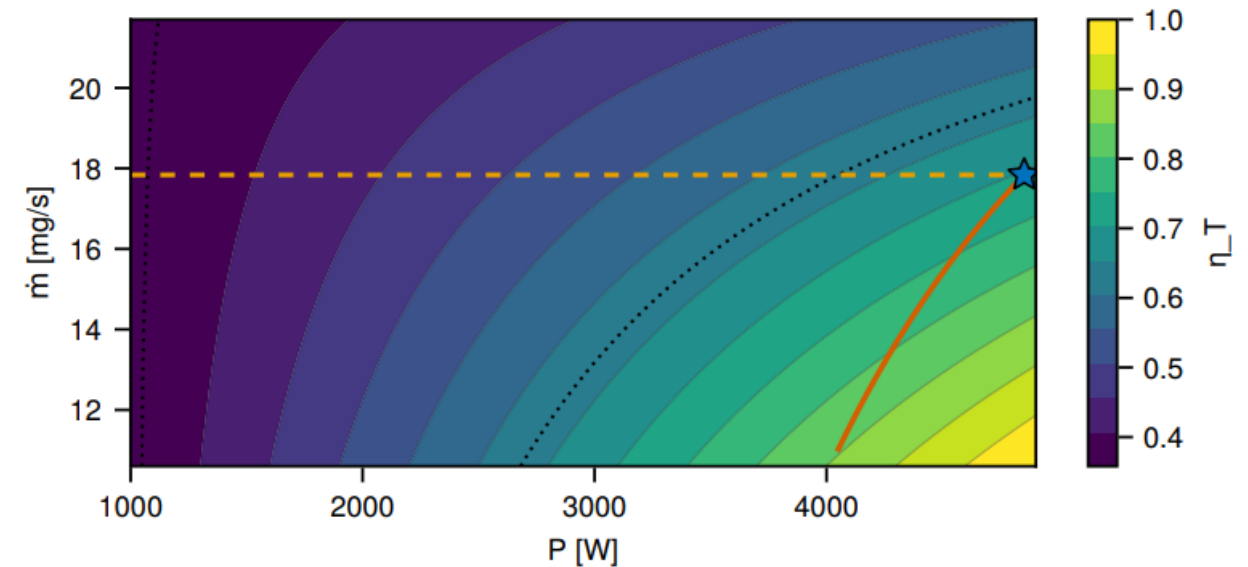
THROTTLING STRATEGIES

Strategy 1 – Fixed-Threshold

Strategy 2 – Variable-Power ★

$$\tau = \begin{cases} 1 & \text{(Maneuver)} \\ 0 & \text{(Charge)} \end{cases}$$

$$\tau = f(SOC)$$



--- $\dot{m} \equiv \dot{m}_0$ (time-optimal)

★ $(P_{max}, \dot{m}_0)$ design point

— $\eta_T \cdot P = K$ (mass-optimal)

..... Extrapolation bounds (η Table 1)

Data extrapolated from: D.L. Brown, C.W. Larson, B.E. Beal, A.D. Gallimore, *Methodology and historical perspective of a Hall thruster efficiency analysis*, Journal of Propulsion and Power, vol. 25, no. 6, pp. 1163–1177, DOI: 10.2514/1.38092, 2009.

Time-optimal GTO→GEO transfer

- This work: 137.34 d — analytical stability bound
- In-environment policies (RL-Ecl, RL-Pow, RL-Var)
- RL-Base zero-shot: ≤ 1 d penalty across Groups B–D
→ **Stability anchor gives cross-env. robustness**

Benchmark mission

GTO→GEO ($m_0 = 2000$ kg, $T = 350$ mN, $I_{sp} = 2000$ s)

#	Method	Ecl	Pow	Var	Time (d)	Fuel (kg)
<i>Group A — Tested on: baseline (no eclipse, no power)</i>						
1	RL-Base	-	-	-	137.34	211.75
16	RL-Base Q-Jac (rbf) [†]	-	-	-	137.45	211.93
14	GA-Base	-	-	-	137.87	212.03
17	RL-Base Q-Jac (tri) [†]	-	-	-	137.90	212.62
3	Time-opt classical	-	-	-	144.00	222.03
<i>Group B — Tested on: eclipse environment</i>						
18	RL-Ecl Q-Jac (rbf) [†]	Yes	-	-	142.04	213.30
5	RL-Ecl	Yes	-	-	142.09	213.50
19	RL-Base Q-Jac (rbf) [†]	Yes	-	-	142.40	213.62
2	RL-Base	Yes	-	-	142.48	213.67
22	GA-Ecl	Yes	-	-	142.61	213.86
21	RL-Base Q-Jac (tri) [†]	Yes	-	-	142.79	214.10
20	RL-Ecl Q-Jac (tri) [†]	Yes	-	-	143.27	214.54
4	Time-opt classical	Yes	-	-	151.20	224.32
<i>Group C — Tested on: power environment</i>						
12	RL-Var	-	Yes	-	143.13	213.38
8	RL-Pow	-	Yes	-	143.16	213.23
15	GA-Pow	-	Yes	-	143.82	213.08
6	RL-Base	-	Yes	-	144.19	212.73
7	Time-opt classical	-	Yes	-	156.83	222.00
<i>Group D — Tested on: power + variable thrust</i>						
13	RL-Var	-	Yes	Yes	142.31	215.62
9	RL-Pow	-	Yes	Yes	142.35	215.58
10	RL-Base	-	Yes	Yes	142.99	216.09
11	Time-opt classical	-	Yes	Yes	155.11	231.42

[†] Literature reference (Holt et al.); not simulated in this work.

Mass-optimal GTO→GEO transfer

- This work: 191.96 kg — analytical stability bound
- In-environment policies (RL-Ecl, RL-Pow, RL-Var)
- Zero-shot: adapts to mass-opt., +5–7 d over deadline
Stability works, deadline-pacing needs in-env. training

RL-Var (variable-power on iso- η_T P locus):
Saves 0.82–1.80 kg vs RL-Pow (fixed-threshold)

Benchmark mission

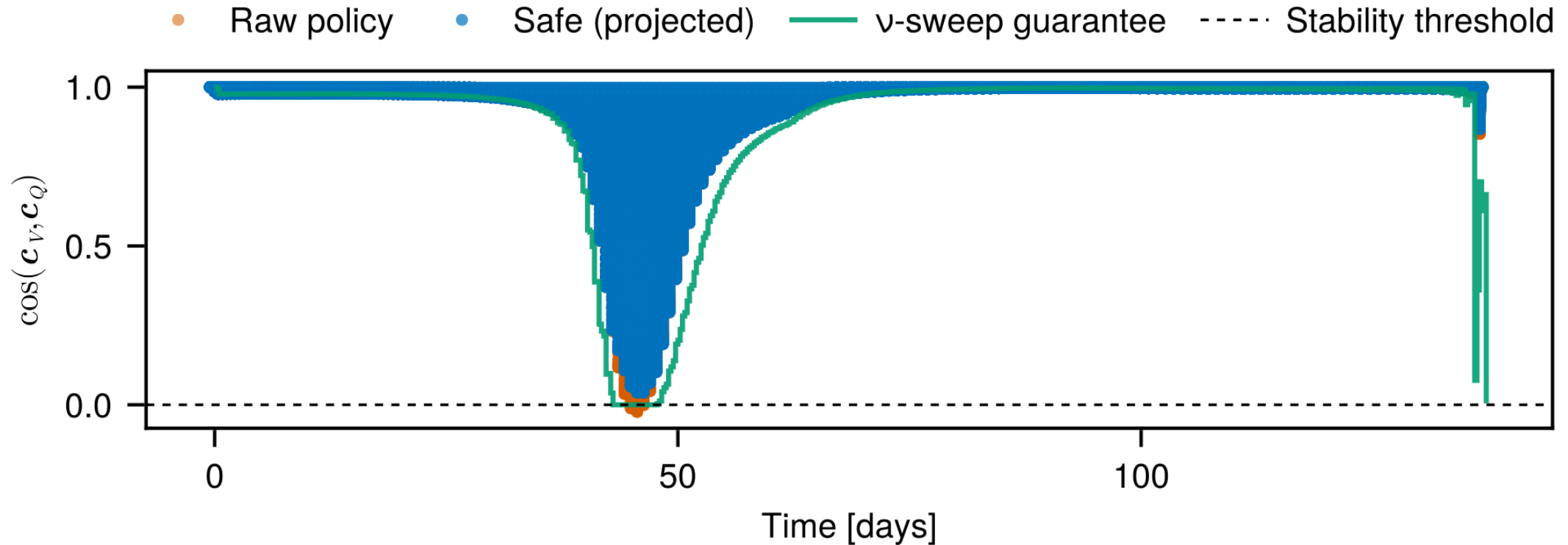
GTO→GEO ($m_0 = 2000$ kg, $T = 350$ mN, $I_{sp} = 2000$ s)

#	Method	Ecl	Pow	Var	Time (d)	Fuel (kg)
<i>Group A — Tested on: baseline (no eclipse, no power)</i>						
6	RL-Base Q-Jac (tri) [†]	-	-	-	149.76	191.65
5	RL-Base Q-Jac (rbf) [†]	-	-	-	149.73	191.82
1	RL-Base	-	-	-	150.09	191.96
2	GA-Base	-	-	-	150.12	194.70
<i>Group B — Tested on: eclipse environment</i>						
5	RL-Base [‡]	Yes	-	-	154.99	195.40
2	RL-Ecl	Yes	-	-	150.16	201.32
11	GA-Ecl	Yes	-	-	148.51	205.60
<i>Group C — Tested on: power environment</i>						
6	RL-Base [‡]	-	Yes	-	157.24	195.06
9	RL-Var	-	Yes	-	149.77	203.04
3	RL-Pow	-	Yes	-	150.01	203.86
4	GA-Pow	-	Yes	-	148.75	206.43
<i>Group D — Tested on: power + variable thrust</i>						
8	RL-Base [‡]	-	Yes	Yes	156.53	194.06
0	RL-Var	-	Yes	Yes	150.05	201.20
7	RL-Pow	-	Yes	Yes	149.67	203.00

[†] Literature reference (Holt et al.); not simulated in this work.

[‡] Cross-environment deployment: policy trained on a different environment and tested here without retraining. Transfer time exceeds $T_{bud} = 150$ d, so propellant figures are not equal-horizon comparable with in-environment entries.

Lyapunov stability – empirical verification



Paper 2 – Toward a Real Digital Twin

Primary goal – 6-DoF digital twin

- Couple ADCS torques + thruster misalignment to stable Q-law
- Full spacecraft dynamics (attitude + orbit)

Other extensions under consideration

- Extend stability proof to J2, SRP & drag perturbations
- Transfer families: LEO→GEO, MEO injection

Current work

- Controllers: MRP-Feedback, PD/PID, EffectiveTorque
- Actuators: RW (imbalance), MTB (B-dot), Thrusters
- Multi-rate FSW

Validation (ongoing)

- Basilisk, 42: multiple scenarios

Achievements

CONFERENCES & TALKS

Conferences attended

- **MBSE 2024** — Bremen, May 2024
- **JuliaCon 2024** — July 2024
- **IAC 2024** — Milan, Oct 2024

IAC 2025 — Presented

- Sydney, 2025
- “Low-Thrust Trajectory Design with Power Constraints and Attitude Dynamics Integration”
- “Efficient Mission Analysis for Small Satellites: Maneuver Planning, Constellation Deployment, and Propulsion Optimization”

PUBLICATIONS

Paper 1 (To be submitted to Acta Astronautica)

- “Stability-Guaranteed Adaptive Q-Law Guidance with Integrated Power Management for Low-Thrust Orbit Transfers”

Paper 2 (in preparation)

- 6-DoF ADCS-coupled adaptive guidance with stability guarantees
- Extension to perturbations (J2, SRP, drag) + LEO

Thank you. Questions?

B1-Construction of c_V — From OE to RTN Thrust Direction

$$\mathbf{c}_V = \begin{pmatrix} \frac{\partial V}{\partial a} G_{a,R} + \frac{\partial V}{\partial e} G_{e,R} + \frac{\partial V}{\partial \omega} G_{\omega,R} \\ \frac{\partial V}{\partial a} G_{a,T} + \frac{\partial V}{\partial e} G_{e,T} + \frac{\partial V}{\partial \omega} G_{\omega,T} \\ \frac{\partial V}{\partial i} G_{i,N} + \frac{\partial V}{\partial \omega} G_{\omega,N} \end{pmatrix}.$$

$$\frac{\partial V}{\partial a} = F_{aa} + F_{ea} + F_{ia},$$

$$\frac{\partial V}{\partial e} = F_{ae} + F_{ee} + F_{ie},$$

$$\frac{\partial V}{\partial i} = F_{ii},$$

$$\frac{\partial V}{\partial \omega} = F_{i\omega},$$

$$F_{jk} = \delta q_j / \delta x_k$$

G(x): Gauss Variational Equations matrix

$$G_{a,R} = \frac{2a^2 e \sin \nu}{h}, \quad G_{a,T} = \frac{2a^2 (1 + e \cos \nu)}{h},$$

$$G_{a,N} = 0, \quad G_{e,R} = \frac{p \sin \nu}{h},$$

$$G_{e,T} = \frac{(p+r) \cos \nu + re}{h}, \quad G_{e,N} = 0,$$

$$G_{i,R} = 0, \quad G_{i,T} = 0,$$

$$G_{i,N} = \frac{r \cos(\omega + \nu)}{h},$$

$$G_{\omega,R} = \frac{-p \cos \nu}{he}, \quad G_{\omega,T} = \frac{(p+r) \sin \nu}{he},$$

$$G_{\omega,N} = \frac{-r \sin(\omega + \nu) \cos i}{h \sin i}.$$

B2-Perturbation Matrix $\mathbf{M}(\mathbf{x})$

$$\Delta \mathbf{c} = \mathbf{c}_Q - \mathbf{c}_V = \mathbf{M}(\mathbf{x}) \boldsymbol{\delta},$$

where $\mathbf{M}(\mathbf{x}) \in \mathbb{R}^{3 \times 3}$ is the perturbation matrix whose columns are

$$\mathbf{m}_1 = \begin{pmatrix} F_{aa}G_{a,R} + F_{ae}G_{e,R} \\ F_{aa}G_{a,T} + F_{ae}G_{e,T} \\ 0 \end{pmatrix},$$

$$\mathbf{m}_2 = \begin{pmatrix} F_{ea}G_{a,R} + F_{ee}G_{e,R} \\ F_{ea}G_{a,T} + F_{ee}G_{e,T} \\ 0 \end{pmatrix},$$

$$\mathbf{m}_3 = \begin{pmatrix} F_{ia}G_{a,R} + F_{ie}G_{e,R} + F_{i\omega}G_{\omega,R} \\ F_{ia}G_{a,T} + F_{ie}G_{e,T} + F_{i\omega}G_{\omega,T} \\ F_{ii}G_{i,N} + F_{i\omega}G_{\omega,N} \end{pmatrix}.$$

B3-Effectivity-Alignment Bound

Thrust effectivity definition

$$\eta(\nu) = \frac{\|\mathbf{c}(\mathbf{x}(\nu))\|}{C_Q}, \quad C_Q \equiv \max_{\nu' \in [0, 2\pi]} \|\mathbf{c}(\mathbf{x}(\nu'))\|,$$

$$\text{thrust ON} \iff \eta(\nu) \geq \eta_{\text{thresh}}.$$

Theorem-Effectivity-Alignment Bound

$$\sigma^* \equiv \sup_{\nu \in [0, 2\pi]} \|\mathbf{M}(\nu)\|_2, \quad \eta_k(\nu) \geq \eta^*(\delta) \equiv \frac{2\sigma^* \|\delta\|}{C_Q}.$$

Proof

$$\mathbf{c}_V^T \mathbf{c}_Q \geq 0, \quad \mathbf{c}_Q = \mathbf{c}_V + \mathbf{M}\delta \implies \|\mathbf{c}_V\|^2 + \mathbf{c}_V^T \mathbf{M}\delta \geq 0$$

$$\text{Cauchy-Schwarz: } \|\mathbf{c}_V\| \geq \|\mathbf{M}\|_2 \cdot \|\delta\|$$

$$\text{Reverse } \Delta \text{ ineq.: } \|\mathbf{c}_V\| = \|\mathbf{c}_Q - \mathbf{M}\delta\| \geq \|\mathbf{c}_Q\| - \|\mathbf{M}\|_2 \|\delta\|$$

$$\text{Into } (\star): \|\mathbf{c}_Q\| \geq 2\|\mathbf{M}\|_2 \|\delta\| \implies \eta_k(\nu) \geq 2\sigma^* \|\delta\| / C_Q \equiv \eta^*(\delta)$$

PPO – Proximal Policy Optimization

Schulman et al., 2017 — OpenAI · The standard on-policy algorithm for continuous control

CORE IDEA

Policy gradient methods update the policy to increase the probability of **good actions** and decrease **bad actions**. But **how much** should we update? Too little → slow learning. Too much → catastrophic collapse.

PPO's solution: clipped surrogate

$$L^{CLIP}(\theta) = \mathbb{E}_t \left[\min(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t) \right]$$

r_t = how much the policy changed for this action

$\hat{A}_t$ = was this action better or worse than expected?

clip = don't let r_t go beyond $[1-\epsilon, 1+\epsilon]$ → bounded update, stable convergence

ARCHITECTURE & TRAINING

Actor-Critic

- Actor π_θ : state → action (what to do)
- Critic V_ψ : state → value (how good is this state?)
- Critic computes $\hat{A}_t$ (advantage) → tells Actor which actions were better than baseline

Training loop

- 1 – Collect: roll out π_{old} for N steps
- 2 – Evaluate: compute advantages $\hat{A}_t$ via Critic
- 3 – Optimize: update π_θ with clipped objective (K epochs)
- 4 – Sync: $\pi_{old} \leftarrow \pi_\theta$, repeat

B5-Throttling profiles – f(SOC)

Time-Optimal

$$f(\text{SOC}) = \sigma(\text{SOC}; C_1, C_2) = \frac{1}{1 + \exp[-C_1(\text{SOC} - C_2)]},$$

Mass-Optimal

$$f(\text{SOC}) = \tau_{\min} + (1 - \tau_{\min}) \text{SOC}^n, \quad n > 1,$$

$$\tau_{\min} = P_{\text{lo}}/P_{\text{max}}$$

