SAPIENZA
UNIVERSITÀ DI ROMA

Sapienza University of Rome

Dipartimento di Matematica Guido Castelnuovo
PhD in Mathematics

Thesis For The Degree Of Doctor Of Philosophy

Dynamics and Learning in Cortical Neuronal Networks

Thesis Advisor
Elena Agliari

Correlatore
Maurizio Mattia

Candidate
Luca Falorsi
1620638

Academic Year MMXXV-MMXXVI (XXXVIII cycle)

♡

Abstract

The brain is a large many-body system whose function emerges from collective neuronal activity, yet current theoretical approaches often split between bottom-up statistical mechanics (biophysically grounded but analytically hard at the behavioral scale) and top-down machine learning (algorithmically powerful but weakly constrained by biological implementation). This thesis develops mathematical tools that operate between these perspectives, advancing the theory of stochastic neural populations while proposing dynamical mechanisms for rapid, biologically plausible adaptation.

First, we establish a rigorous spectral theory for the Fokker–Planck operators arising in population density approaches to integrate-and-fire neurons with fire-and-reset boundary conditions. Using the framework of boundary eigenvalue operator functions, we provide a functional-analytic foundation for spectral decompositions. A central result is the resolution of the completeness problem: by proving Birkhoff regularity we establish completeness of the biorthogonal system of root functions. This rigorous setting also enables a systematic characterization of exceptional points of the spectrum as defective eigenvalues and motivates a smooth regularization of mode projections near spectral singularities.

Second, we develop a theory of non-stationary inter-spike-interval (ISI) statistics beyond renewal and quasi-renewal approximations. We formulate the dynamics on an augmented state space (v, τ) (membrane potential and time since last spike), deriving a two-dimensional Fokker–Planck equation that unifies voltage-based and age-structured descriptions. From this framework we obtain exact ISI statistics far from stationarity, derive a tractable hierarchy of moment equations and develop an analitical linear response theory for the inter-spike intervals.

Finally, we shift to a computational perspective and study how neuronal circuits may adapt rapidly without synaptic weight updates. Building on reservoir computing and latent-variable readout parametrizations, we propose gain-modulated recurrent architectures that implement gradient-descent-like learning dynamics in real time through neural activity modulation. This provides a concrete biophysical hypothesis for in-context adaptation in cortical circuits and a bridge between modern machine-learning principles and biological constraints.

Keywords: neural populations; population density approach; Fokker–Planck PDE; fire–reset boundaries; Integrate and Fire models; refractory period; spectral decomposition; non-self-adjoint operators; Birkhoff regularity; completeness (root functions); exceptional points; defective eigenvalues; semigroup methods; dissipativity; sectorial operators; non-stationary Inter-spike Intterval; linear response; transfer functions; coupled populations; reservoir computing; recurrent networks; latent readouts; gain modulation; in-context learning.

Contents

Chapter 1

Introduction

> Quando ho pianto per un'ora perché mi avevano detto di sognare, e non sapevo da dove cominciare
>
> — Tutte le cose inutili, *Bombe Carta*

The human brain is a large, many-body complex system composed of approximately 10^{10} interacting cells of different types and functions, each integrating information received from tens of thousands of connections [23, 98]. Over the past century, research in biology has accumulated a large body of evidence on how this system is anatomically structured across multiple scales: from single neuron structure to detailed neural circuitry to macroscopic organization at the level of cortical and subcortical areas[119, 204]. In recent decades, experimental neuroscience has developed increasingly complex recording techniques, ranging from single-neuron calcium imaging and multi-electrode array recordings to electroencephalography (EEG) and whole-brain functional magnetic resonance imaging (fMRI). Yet despite this vast amount of data, our perspective remains constrained. We are still, in some sense looking to the ssytem through a keyhole. In fact, while we can observe local spiking activity or global blood-oxygen-level-dependent signals, understanding how the brain operates as a unified system remains fundamentally challenging. Given the complexity of this problem, the effort is inherently multidisciplinary, with each discipline approaching it through its own distinct framework and methods.

Current theoretical approaches often fracture along the levels of analysis proposed by David Marr [121]. On one hand, *bottom-up approaches* rooted in statistical mechanics prioritize physical realism, deriving macroscopic laws from the coarse-graining of microscopic models [30, 126]. While rigorous, these methods face significant analytical hurdles when scaling to complex behaviors. On the other hand, *top-down approaches* driven by Machine Learning (ML), specifically Artificial Neural Networks (ANNs), have proven exceptionally successful at emulating brain function at the algorithmic level [17, 75, 118, 211]. However, these models often operate as "black boxes," lacking a transparent link to the biological substrate.

This thesis operates in the space between these disciplines, developing mathematical instruments that refine our understanding on both sides of the divide. We advance the spectral theory of neural populations to better describe its dynamics, and we ground machine learning paradigms in biological plausibility to better understand algorithmic adaptation.

Spectral Theory for Population Density Approach

Rooted in the statistical mechanics of non-equilibrium systems and stochastic processes, the Population Density Approach (PDA) has the primary goal of linking the microscopic dynamics of single neurons to macroscopic observables, such as the firing rate of neural populations. This approach, reviewed in Chapter 2, is motivated by the fundamental finding that complex cognitive functions do not reside in the trajectory of any single cell, but emerge from the collective activity of neural ensembles. Pioneering experimental studies [69, 70, 208] demonstrated that variables such as movement direction and spatial location are encoded in the joint activity of large populations: information that is unrecoverable from single-neuron recordings alone.

Here, the fundamental object is the probability density function over the state space of neuronal variables (primarily the membrane potential). In the thermodynamic limit, the evolution of the network state is described by a Fokker-Planck partial differential equation governing the neuron density over time.

A defining feature of this system is the fire-and-reset mechanism: neurons reach a threshold and are instantaneously reset. This imposes nonstandard boundary conditions that break the detailed balance of the system. Unlike the self-adjoint operators found in quantum mechanics or equilibrium thermodynamics, the operator governing neural populations describes a system driven far from equilibrium. Furthermore, the equation becomes inherently nonlinear due to the self-consistent coupling of the population firing rate with the diffusion moments.

In the literature, the analytical characterization of population dynamics has mostly focused on steady states [145] and perturbative expansions via linear response theory [30]. To overcome the limitations of local analyses, spectral decomposition methods were introduced [101, 126], projecting the population density onto a basis formed by the eigenfunctions of the Fokker-Planck operator. While effective as a dimensionality reduction tool, the spectral properties of the operator were never fully characterized in a rigorous functional analytic setting. Most notably, a proof of the completeness of the biorthonormal system of eigenfunctions is still missing. The problem cannot be reduced to the study of a self-adjoint operator, and the spectral theorem cannot be applied due to the boundary conditions. Ultimately, this mathematical gap is a consequence of the broken detailed balance in the physical system.

In this thesis, we address these foundational gaps by developing a rigorous spectral theory for the population density equation. The contributions are structured as follows:

Rigorous Functional Analytic Framework and Refractory Period. In Chapter 4, we ground the spectral theory in a rigorous functional analytic setting using boundary eigenvalue operator functions [131]. Furthermore, we show how to extend this framework to incorporate the presence of an absolute refractory period, a feature often overlooked in the literature analyzing neural population dynamics via Fokker-Planck PDEs. Despite being frequently regarded as negligible, the few studies that incorporate it [68] observe that the absolute refractory period represents an important source of nonlinearity in current-to-rate transfer functions, which in turn gives a fundamental contribution to the presence of stable equilibrium points in the dynamics.

Characterization of Exceptional Points. Our rigorous treatment allows us to study and properly characterize the presence of exceptional points in the spectrum of the operator, where two branches of real eigenvalues meet, giving rise to complex conjugate eigenvalues. These points were often overlooked in previous work, yet they are a source of numerical instability in projection

methods. We characterize them as defective eigenvalues, requiring the inclusion of associated root functions for the basis to be complete.

Completeness, Regularity, and Smooth Projections. In the next chapter 5, we formally prove the completeness of the biorthogonal system of root functions by establishing Birkhoff regularity. This allows us to lay the spectral decomposition on solid mathematical grounds. We also prove that the operator is analytic, laying out the foundations for a rigorous study of the existence and regularity of solutions through semigroup theory. Finally, we show how to remove the singularity from eigenmode projections in time-dependent coupled dynamics. We derive an alternative basis, based on symmetric functions of eigenvalue clusters, that is smooth with respect to the singularities introduced by the exceptional points in the spectrum.

Non-Stationary Temporal Statistics

Beyond the mean firing rate, neural information is encoded in the precise timing of spikes. The distribution of Inter-Spike Intervals (ISI) serves as the fundamental metric for this temporal variability, traditionally analyzed via renewal theory as a First Passage Time problem [192]. While successful for constant inputs, this classical framework collapses in realistic biological regimes where populations are driven by rapidly varying stimuli or intrinsic network oscillations. In such non-stationary conditions, the standard "quasi-renewal" approximations, such as the Spike Response Model [72–74], which assume that instantaneous spiking statistics match those of a steady state, fail to capture the complex, history-dependent interplay between membrane potential dynamics and spike timing. This creates a disconnect in current theoretical models: voltage-based approaches (like the standard Fokker-Planck equation) track membrane potentials but lose explicit ISI information, while age-based models track spike timing but rely on invalid stationary approximations for voltage dynamics.

To bridge this gap, in Chapter 3 we introduce a novel framework that rigorously unifies voltage and timing dynamics. We formulate the population evolution on an augmented state space, deriving a two-dimensional Fokker-Planck equation for the joint density of membrane potential v and time since the last spike τ. This formalism allows us to correctly compute exact ISI statistics even far from the stationary state. From this general framework, we derive a hierarchy of moment equations to quantify spike timing variability during transients, and we develop a linear response theory that extends standard frequency analysis to higher-order temporal statistics. Finally, we numerically validate the proposed framework against simulations of oscillating excitatory-inhibitory networks, demonstrating that our approach accurately predicts the ISI distributions when the system is far from the stationary state.

Learning Families of Dynamical Systems

In the second part of the thesis, we shift to a macroscopic and computational perspective, modeling cortical areas as interacting Recurrent Neural Networks (RNNs) of neural populations. Here, we leverage *Reservoir Computing* (RC), a machine learning paradigm rooted in complex systems and dynamical systems theory, widely used to model nonlinear dynamical systems. A random RNN serves as a computational substrate, extracting spatio-temporal features of the trajectories in a rich state space, which are then used to predict the future evolution of the dynamics [92, 95, 114]. While

RC represents a state-of-the-art, efficient system for data-driven modeling of complex and chaotic dynamical systems, it typically lacks flexibility, often requiring additional supervised contextual signals to represent multiple distinct dynamics within a single network.
In Chapter 6, we propose an augmented architecture that integrates RC with *Latent Variable Models*. In this framework, the network does not learn a single input-output mapping but rather a continuous *family* of dynamical systems. The "readout" weights are constructed dynamically as linear combinations of prototypes, where the mixing coefficients are inferred latent variables. This allows the network to learn a low-dimensional static representation of the problem geometry, effectively encoding distinct dynamical behaviors within a single fixed substrate. This approach is inspired by the "conceptors" framework [94] [93] and connects it to manifold learning, demonstrating how a single neural circuit can reuse its dynamics to represent a continuum of tasks.

Biological Plausibility of In-Context Learning

Finally, in Chapter 7, we bridge the gap between abstract machine learning principles and biological implementation. Inspired by the computational mechanisms developed in our Latent Variable Models, we translate the abstract concept of dynamic readout adaptation into a biologically plausible framework for in-context learning. Modern "Transformer" architectures in deep learning have demonstrated the remarkable ability to perform *in-context learning*, adapting to new tasks during inference without updating synaptic weights. Recent theoretical work suggests that these models implicitly implement gradient descent steps within their forward pass [200]. However, how such meta-learning capabilities could be implemented in biological neural networks remains an open question.
Addressing this, we propose that biological networks can achieve a similar form of "in-context" adaptation through *gain modulation*. Just as our latent variable model modifies the system's effective output by manipulating latent states, we show that modulating the gain of neuronal transfer functions, a mechanism consistent with dendritic integration in pyramidal neurons [105], allows a fixed recurrent network to implement gradient-based learning dynamics in real time. This provides a concrete biophysical hypothesis for how cortical circuits can flexibly traverse a manifold of cognitive tasks without the need for immediate synaptic rewiring.

Chapter 2

Population Density Dynamics

Fuori in realtà non c'era cambiamento,
è il morbo stagionato che mi sottrae alle strade:
dentro di me è cresciuto e mi ha corrotto gli occhi
e tutti gli altri sensi: e il mondo arriva
come una citazione.
Tutto è accaduto ormai, ma io dov'ero?
Quando è avvenuta la grande distrazione?
Dove si è slegato il filo, dove si è aperto
il crepaccio, qual è il lago
che ha perso le sue acque
e mutando il paesaggio
mi scombina la strada?

— Patrizia Cavalli, *Poesie*

2.1 From single neuron dynamics to the Fokker-Planck equation

2.1.1 Single neuron dynamics

The biophysical dynamics of a single neuron are classically described by the conductance-based Hodgkin-Huxley formalism ([88]). While this nonlinear system provides an accurate description of ionic channel kinetics, its complexity poses significant challenges for the analytical treatment of large-scale networks. To study the collective properties of interacting neural populations, it is therefore necessary to adopt a reduced model that retains the essential features of excitability: the subthreshold integration of synaptic inputs and the threshold-dependent generation of action potentials. To this end, we utilize the framework of **Integrate-and-Fire (IF) neurons** ([192]), which has also been shown to quantitatively reproduce both spiking activity and subthreshold membrane dynamics observed in biological neurons [13, 156, 190].
In this framework, the state of the neuron is characterized by the membrane depolarization $V(t)$ at the soma. The subthreshold evolution is governed by the following stochastic differential equation

(SDE):

$$\tau_m dV(t) = b(V(t))dt + dI(t), \quad V(t) \in (\alpha, \theta). \tag{2.1}$$

Here, τ_m is the membrane time constant, and θ represents the firing threshold. When the depolarization reaches θ, a spike is emitted, and $V(t)$ is instantaneously reset to a reset potential $H \in (\alpha, \theta)$. This reset is followed by an absolute refractory period τ_0, during which the dynamics are frozen. We explicitly introduce a lower bound for the membrane potential, $V(t) \geq \alpha$. Physically, this barrier models the inhibitory reversal potential, preventing the potential from hyperpolarizing indefinitely [68, 162]. Mathematically, we impose a reflecting boundary condition at $V = \alpha$. This ensures the state space is bounded, a property that facilitates the spectral analysis of the associated operators in later chapters.

The term $b(V)$ represents the deterministic drift of the membrane potential. We will always assume $b \in \mathrm{C}^\infty[\alpha, \theta]$, throughout the thesis we will use the two main working examples.

1. **Leaky Integrate-and-Fire (LIF):** The drift includes a linear leak term driving the potential toward a resting state, $b(v) = -v$.
2. **Bounded Perfect Integrate-and-Fire (BPIF)** The neuron acts as a perfect integrator, i.e., b(v)=0. More generally, b(v) is a negative constant; we set it to 0 without loss of generality since it can be absorbed into the drift term. This model was introduced for $\alpha = H$ (as a VLSI neuron) in [68], and later generalized to the case $\alpha < H$ by [160] under the name CLIFF.

The term $dI(t)$ in (2.1) represents the stochastic forcing from afferent synaptic currents.[1] If the neuron receives inputs from C presynaptic sources, the current is modeled as a superposition of Dirac delta distributions:

$$dI(t) = \tau_m \sum_{k=1}^{C} J_k \sum_i \delta(t - t_i^k - \delta_k)dt, \tag{2.2}$$

where t_i^k denotes the i-th spike time of the k-th presynaptic neuron, J_k is the synaptic efficacy (dimensionless voltage jump), and δ_k is the transmission delay.

2.1.2 Diffusion approximation and Kramers-Moyal expansion

When the number of presynaptic contacts is large and synaptic efficacies are small relative to the distance to threshold, we approximate the afferent current with a continuous diffusion process. This approximation allows for the application of stochastic calculus and has been a cornerstone of theoretical neuroscience since the pioneering works of Capocelli and Ricciardi [35].

Consider a neuron receiving inputs from $N \to \infty$ presynaptic contacts. We assume the presynaptic neurons belong to distinct populations indexed by $\sigma \in \Sigma$ (e.g., excitatory $\sigma = E$ and inhibitory $\sigma = I$). The neuron receives $K_\sigma^{(N)}$ connections from population σ, such that $\sum_\sigma K_\sigma^{(N)} = N$. The presynaptic spike trains are modeled as independent renewal processes $X_\sigma^{(j)}$ (where $j = 1, \ldots, K_\sigma^{(N)}$).

[1] We assume the membrane resistance $R = 1$ is absorbed into the synaptic weights.

The total input current is given by:

$$dI^{(N)}(t) = \tau_m \sum_{\sigma\in\Sigma} \sum_{j=1}^{K_\sigma^{(N)}} J_{\sigma,j}^{(N)} dX_\sigma^{(j)}(t - \delta_{\sigma,j}). \tag{2.3}$$

Here, $J_{\sigma,j}^{(N)}$ and $\delta_{\sigma,j}$ represent the synaptic efficacy and delay of the j-th connection from population σ. We assume these are i.i.d. random variables drawn from distributions with finite moments. If we further model the spike trains $X_\sigma^{(j)}$ as non-homogeneous Poisson processes with rates $\nu_\sigma(t)$, the evolution of the probability density $p_N(I,t)$ of the current $I^{(N)}$ is governed by a master equation:

$$\frac{1}{\tau_m}\frac{\partial p_N(I,t)}{\partial t} = \sum_{\sigma\in\Sigma} \sum_{j=1}^{K_\sigma^{(N)}} \nu_\sigma(t - \delta_{\sigma,j}) \left[p_N(I - J_{\sigma,j}^{(N)}, t) - p_N(I,t)\right]. \tag{2.4}$$

Using the **Kramers-Moyal expansion**, we can rewrite Equation (2.4) as an infinite-order partial differential equation:

$$\frac{1}{\tau_m}\frac{\partial p_N(I,t)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial I^n} \left(M_n^{(N)}(t) p_N(I,t)\right), \tag{2.5}$$

where the n-th infinitesimal moment is defined as:

$$M_n^{(N)}(t) = \sum_{\sigma\in\Sigma} K_\sigma^{(N)} \langle (J_\sigma^{(N)})^n \rangle \langle \nu_\sigma(t - \delta_\sigma) \rangle. \tag{2.6}$$

Convergence conditions and the balanced state

To obtain a valid diffusion limit as $N \to \infty$, one must ensure the process converges to a continuous path diffusion. As rigorously established by Ricciardi and colleagues [35, 161, 162], convergence requires that the first two infinitesimal moments tend to finite limits while all higher-order moments vanish. Specifically, we require:

$$\lim_{N\to\infty} \tau_m M_1^{(N)}(t) = \mu(t), \tag{2.7}$$

$$\lim_{N\to\infty} \frac{\tau_m}{2} M_2^{(N)}(t) = D(t), \tag{2.8}$$

$$\lim_{N\to\infty} M_4^{(N)}(t) = 0. \tag{2.9}$$

In the context of cortical networks, these conditions are satisfied in the so-called **balanced state** [30, 195]. In this regime, synaptic weights are assumed to scale as $J \sim 1/\sqrt{N}$ while the number of connections N becomes large. Without inhibition, such scaling would lead to a divergent mean input current of order $\mathcal{O}(\sqrt{N})$. The balanced state hypothesis posits that recurrent inhibitory feedback dynamically adjusts to cancel the excitation at leading order, leaving a residual mean input $\mu(t)$ of order $\mathcal{O}(1)$. Crucially, because variances sum (for independent inputs), the fluctuations remain significant ($\mathcal{O}(1)$) even as $N \to \infty$. This results in a *fluctuation-driven regime* where the mean potential remains subthreshold, and firing is driven by random fluctuations in the input current. This mechanism explains the irregular, Poisson-like firing patterns observed in cortical experiments, which cannot be accounted for by standard mean-driven integrators.

However, rigorously proving the diffusion limit for fully recurrent, balanced networks remains a significant open challenge in mathematical physics. In such networks, the assumption of independent Poisson inputs breaks down due to the emergence of correlations and non-trivial temporal structure in the spike trains [157, 167]. While recent progress has been made using mean-field methods for Hawkes processes [41, 54], a complete derivation of the diffusion limit from microscopic dynamics in the balanced regime is not the focus of this thesis.
Instead, we adopt the diffusion approximation as our starting ansatz. We assume that the resulting membrane potential dynamics are effectively described by the SDE:

$$\tau_m dV(t) = [b(V(t)) + \mu(t)]\, dt + \sqrt{2\tau_m D(t)} dW(t), \quad V(t) \in (\alpha, \theta), \tag{2.10}$$

where the drift $\mu(t)$ and diffusion $D(t)$ coefficients are given by:

$$\mu(t) := \tau_m \sum_{\sigma \in \Sigma} K_\sigma \langle J_\sigma \rangle (\nu_\sigma * p_{\delta_\sigma})(t), \tag{2.11}$$

$$D(t) := \frac{\tau_m}{2} \sum_{\sigma \in \Sigma} K_\sigma \langle J_\sigma^2 \rangle (\nu_\sigma * p_{\delta_\sigma})(t). \tag{2.12}$$

Here, we define $D(t)$ as τ_m times half the second moment ($\tau_m M_2/2$), such that it appears in the Fokker-Planck equation without the prefactor 1/2. With p_{δ_σ} we denote the distribution of delays.

Time normalization

The characteristic timescale of the subthreshold dynamics is determined by the membrane time constant τ_m. To simplify the mathematical formulation and notation in the subsequent chapters, it is convenient to introduce a dimensionless time variable $t' = t/\tau_m$.
Applying this transformation to the stochastic differential equation derived above, we obtain the normalized SDE:

$$dV(t') = \left[b(V) + \mu(t')\right] dt' + \sqrt{2D(t')} dW(t'). \tag{2.13}$$

In the remainder of this thesis, we will drop the prime notation and assume, without loss of generality, that time is measured in units of the membrane time constant (effectively setting $\tau_m = 1$). Accordingly, all temporal parameters, including synaptic delays δ and the refractory period τ_0, are expressed in these normalized units.

2.2 The Fokker-Planck equation

Consider an ensemble of $N \to \infty$ indistinguishable neurons evolving according to the normalized SDE derived in Section 2.1.2. The state of the system is fully characterized by the probability density function $p_t(v)$, which describes the distribution of membrane potentials at time t.
Let $\boldsymbol{\gamma}(t) := \{\mu(t), D(t)\}$ represent the infinitesimal moments of the afferent current, and let $b : [\alpha, \theta] \to \mathbb{R}$ be a smooth function describing the membrane leakage. The collective behaviour of the

network is described by the **Fokker-Planck (FP) equation**:

$$\partial_t p_t(v) = \mathcal{L}_{\boldsymbol{\gamma}(t)} p_t(v) = -\partial_v \mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](v), \quad (t, v) \in [0, \infty) \times (\alpha, \theta), \tag{2.14}$$

which governs the evolution of the probability density $p_t(v)$. Here, $\mathcal{S}_{\boldsymbol{\gamma}}[p](v)$ denotes the probability current (flux):

$$\mathcal{S}_{\boldsymbol{\gamma}}[p](v) := [b(v) + \mu]\, p(v) - D\partial_v p(v). \tag{2.15}$$

The value of the flux at the threshold potential gives the fraction of neurons emitting an action potential per unit time, i.e., the **firing rate** of the population:

$$\nu(t) := \mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](\theta). \tag{2.16}$$

2.2.1 Boundary conditions

To obtain a well-posed problem, we must impose boundary conditions at the limits of the domain $[\alpha, \theta]$. Following Feller's classification of boundaries for one-dimensional diffusion processes [60, 117]:

1. **Absorbing boundary at θ:** The firing threshold θ acts as a regular boundary that is accessible from the interior. Physically, trajectories reaching θ are instantaneously removed (fired) from the domain. To model this perfect absorption, we impose the condition that the probability density vanishes at the boundary:

$$p_t(\theta) = 0. \tag{2.17}$$

2. **Reflecting boundary at α:** The lower bound α (representing the inhibitory reversal potential) acts as a reflecting barrier. Physically, the driving force vanishes or reverses below α, preventing trajectories from crossing it. This corresponds to a zero-flux condition:

$$\mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](\alpha) = 0. \tag{2.18}$$

3. **Reinjection and reset dynamics**: The firing mechanism involves the re-entry of absorbed trajectories at the reset potential $H \in (\alpha, \theta)$ after a refractory period τ_0.
   We model the reset as a discontinuity in the probability flux at $v = H$, requiring that the incoming flux at the reset potential equals the outgoing flux at the threshold (delayed by τ_0):

$$\mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](H^+) - \mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](H^-) = \nu(t - \tau_0). \tag{2.19}$$

   This is supplemented by the continuity of the probability density itself:

$$p_t(H^+) = p_t(H^-). \tag{2.20}$$

Alternatively, we can separate the continuous evolution from the singular reinjection source. Let $\mathcal{L}^A_{\boldsymbol{\gamma}(t)}$ denote the Fokker-Planck operator associated with the absorbing boundary conditions (2.17),

(2.18) *without* the internal reinjection flux. The full dynamics can then be written as:

$$\partial_t p_t(v) = \mathcal{L}^A_{\boldsymbol{\gamma}(t)} p_t(v) + \nu(t-\tau_0)\delta(v-H). \tag{2.21}$$

Here, the Dirac delta $\delta(v-H)$ explicitly represents the source of probability mass.
The mathematical well-posedness of this formulation has been recently established. In [111], Liu et al. provided a rigorous derivation of the Fokker-Planck equation for the Integrate-and-Fire model starting from the microscopic stochastic dynamics. They employed a novel **iteration perspective**. By decomposing the probability density based on the number of firing events (renewal cycles), they proved that the density of the diffusion-jump process converges to the unique classical solution of the Fokker-Planck equation with the reinjection boundary condition.

Extended mean-field coupling

In a fully recurrent network, the input statistics $\boldsymbol{\gamma}(t) = \{\mu(t), D(t)\}$ are not extrinsic parameters but depend on the collective activity of the network itself. As discussed in Section 2.1.2, a rigorous derivation of the diffusion limit for such coupled systems remains an open problem. Therefore, we proceed under an **extended mean-field approximation** [6]. We assume that the moments $\boldsymbol{\gamma}(t)$ are determined self-consistently by the past activity of the population, specifically the delayed firing rate $\nu(t-\delta)$. This closes the loop, coupling the Fokker-Planck equation to its own boundary flux and rendering the system **nonlinear** and non-local in time.
Despite the heuristic nature of this closure, this mean-field description has proven remarkably effective in capturing the qualitative and quantitative behavior of large neural networks. It successfully predicts rich dynamical phenomena observed in simulations and experiments, including transitions to multistability [27], the emergence of fast global oscillations [31], and finite-time blowup regimes [33].

2.2.2 Stationary solution and the gain function

We now investigate the stationary statistics of the population under constant input moments $\boldsymbol{\gamma} = \{\mu, D\}$. The stationary probability density, denoted by $\phi_0(v, \boldsymbol{\gamma})$, satisfies the time-independent Fokker-Planck equation subject to the boundary conditions defined in (2.17),(2.18),(2.19) and (2.20).
Theorem 2.2.1 (Stationary Distribution)**.** *Let $\boldsymbol{\gamma} = \{\mu, D\}$ with $D > 0$. The unique stationary probability density $\phi_0(v, \boldsymbol{\gamma})$ is given by:*

$$\phi_0(v, \boldsymbol{\gamma}) = \frac{\nu_0}{D}\, w_{\boldsymbol{\gamma}}(v) \int_{H\vee v}^{\theta} \frac{1}{w_{\boldsymbol{\gamma}}(u)}\, du, \tag{2.22}$$

where $H \vee v := \max(H, v)$ and the Wronskian $w_{\boldsymbol{\gamma}}(v)$ is defined as:

$$w_{\boldsymbol{\gamma}}(v) := \exp\left(\int_{\alpha}^{v} \frac{b(y)+\mu}{D}\, dy\right), \tag{2.23}$$

up to a normalization constant ν_0 that represents the stationary firing rate and is determined by the normalization condition.

Proof. We construct the general solution from two linearly independent fundamental solutions defined by their flux properties:

1. $y_1(v) = w_{\boldsymbol{\gamma}}(v)$, which carries zero flux ($\mathcal{S}_{\boldsymbol{\gamma}} y_1 = 0$).
2. $y_2(v) = \frac{1}{D} w_{\boldsymbol{\gamma}}(v) \int_v^\theta \frac{1}{w_{\boldsymbol{\gamma}}(u)}\, du$, which satisfies the absorbing boundary condition $y_2(\theta) = 0$ and carries unit flux ($\mathcal{S}_{\boldsymbol{\gamma}} y_2 = 1$).

In the interval $[\alpha, H)$, the reflecting boundary implies the flux is zero, so the solution must be proportional to $y_1(v)$. In the interval $(H, \theta]$, the flux must equal the stationary firing rate ν_0, so the solution must be $\nu_0 y_2(v)$. Imposing continuity at the reset potential $v = H$ yields the unique stationary profile:

$$\phi_0(v) = \begin{cases} \nu_0 \frac{y_2(H)}{y_1(H)} y_1(v) & v \leq H \\ \nu_0 y_2(v) & v > H \end{cases} \tag{2.24}$$

Observing that for $v \leq H$, $y_2(v) = \frac{1}{D} w_{\boldsymbol{\gamma}}(v) \int_v^\theta w^{-1}$, we can unify the expression as:

$$\phi_0(v) = \nu_0 \frac{1}{D} w_{\boldsymbol{\gamma}}(v) \int_{H \vee v}^\theta \frac{1}{w_{\boldsymbol{\gamma}}(u)}\, du. \tag{2.25}$$

The firing rate ν_0 is finally determined by the normalization condition $\int \phi_0 + \nu_0 \tau_0 = 1$. □

The current-to-rate transfer function

The normalization condition implicitly defines the **current-to-rate transfer function** (or gain function) $\Phi(\mu, D)$. Integrating the density derived above:

$$\frac{1}{\nu_0} = \tau_0 + \frac{1}{D} \int_\alpha^\theta w_{\boldsymbol{\gamma}}(v) \left(\int_{H \vee v}^\theta \frac{1}{w_{\boldsymbol{\gamma}}(u)}\, du \right) dv. \tag{2.26}$$

Therefore, the firing rate is given by:

$$\nu_0 = \Phi(\mu, D) := \left(\tau_0 + \frac{1}{D} \int_\alpha^\theta w_{\boldsymbol{\gamma}}(v) \left(\int_{H \vee v}^\theta \frac{1}{w_{\boldsymbol{\gamma}}(u)}\, du \right) dv \right)^{-1}. \tag{2.27}$$

In the LIF case ($b(v) = -v$), this reduces to the Siegert formula [97]:

$$\Phi(\mu, D) = \left(\tau_0 + \sqrt{\pi} \int_{\frac{H-\mu}{\sqrt{2D}}}^{\frac{\theta-\mu}{\sqrt{2D}}} e^{x^2} (1 + \operatorname{erf}(x))\, dx \right)^{-1}. \tag{2.28}$$

This function provides the critical link between the microscopic fluctuations and the macroscopic observable. Physically, $\Phi(\mu, D)$ quantifies the excitability of the population: it determines the steady-state firing rate that emerges from a constant background drive μ and noise intensity D.
In the extended mean-field framework, the stationary activity of the coupled network is determined by the self-consistency condition[6, 9, 74]:

$$\nu_0 = \Phi\Big(\mu(\nu_0), D(\nu_0)\Big), \tag{2.29}$$

where the input moments $\mu(\nu_0)$ and $D(\nu_0)$ are linear functions of the rate ν_0, weighted by the

synaptic connectivity (see Eq. (2.11)). Finding the fixed points of this equation corresponds to identifying the stationary states of the network dynamics.

2.3 Spectral decomposition of the dynamics

In the literature, the analytical characterization of population dynamics has largely focused on steady states [145] or perturbative expansions via linear response theory [30]. While effective near equilibrium, these methods often fail to capture the full transient behavior of the network in the nonlinear regime. To overcome these limitations, a spectral decomposition method was introduced in [1, 100, 126]. This approach projects the population density onto a time-dependent basis formed by the eigenfunctions of the Fokker-Planck operator:

$$p_t(v) = \sum_{n=0}^{\infty} a_n(t)\phi_n(v, \boldsymbol{\gamma}(t)), \tag{2.30}$$

where the mode amplitudes are defined by the projection $a_n(t) := \langle \psi_n(\cdot\,, \boldsymbol{\gamma}(t)), p_t \rangle$ using the eigenfunctions $\{\psi_n\}$ of the adjoint operator.

This decomposition is particularly powerful because the spectrum typically displays a hierarchy of timescales, such that only a few slowest modes dominate the firing rate dynamics. This allows for a significant dimension reduction of the problem. However, despite its computational effectiveness and widespread use in theoretical neuroscience [198], the spectral properties of the Fokker-Planck operator $\mathcal{L}_{\boldsymbol{\gamma}}$ have never been fully characterized in a rigorous functional analytic setting. Standard treatments often rely on heuristic assumptions regarding the completeness of the basis and the discrete nature of the spectrum. Furthermore, a critical limitation of the standard theory is the assumption of a **vanishing refractory period** ($\tau_0 = 0$). The inclusion of a finite refractory period $\tau_0 > 0$ introduces delays that fundamentally alter the nature of the operator, complicating the spectral definition.

In this section, we present the classical derivation of the spectral basis for the "frozen" operator with fixed parameters $\boldsymbol{\gamma}$, maintaining the standard simplification $\tau_0 = 0$ to illustrate the method before addressing the rigorous challenges in later chapters.

2.3.1 Eigenvalues and Eigenfunctions

We now construct the spectral basis for the "frozen" operator $\mathcal{L}_{\boldsymbol{\gamma}}$ in the absence of a refractory period ($\tau_0 = 0$). The eigenvalue problem is defined by the ordinary differential equation:

$$\mathcal{L}_{\boldsymbol{\gamma}}\phi(v) = \lambda\phi(v), \quad v \in (\alpha, \theta). \tag{2.31}$$

Let $f_1(v, \lambda, \boldsymbol{\gamma})$ and $f_2(v, \lambda, \boldsymbol{\gamma})$ denote the two linearly independent fundamental solutions of Eq. (2.31). For convenience, we normalize these functions at the lower boundary α such that they form a canonical basis in phase space:

$$f_1(\alpha, \lambda, \boldsymbol{\gamma}) = 1, \quad \mathcal{S}_{\boldsymbol{\gamma}}[f_1](\alpha, \lambda, \boldsymbol{\gamma}) = 0, \tag{2.32}$$

$$f_2(\alpha, \lambda, \boldsymbol{\gamma}) = 0, \quad \mathcal{S}_{\boldsymbol{\gamma}}[f_2](\alpha, \lambda, \boldsymbol{\gamma}) = 1. \tag{2.33}$$

Here, $\mathcal{S}_{\boldsymbol{\gamma}}[f]$ denotes the probability flux associated with the function f. It is evident that f_1 corresponds to the solution satisfying the reflecting (zero-flux) boundary condition at α.
The general eigenfunction satisfying the boundary conditions can be constructed piecewise. In the sub-interval $(\alpha, H]$, the reflecting boundary condition $\mathcal{S}_{\boldsymbol{\gamma}}[\phi](\alpha) = 0$ implies that the solution must be proportional to f_1. In the sub-interval (H, θ), the solution is a general linear combination of the fundamental set. Thus, the eigenfunction takes the form:

$$\phi_\lambda(v) = \begin{cases} a_- \, f_1(v, \lambda, \boldsymbol{\gamma}) & v \in (\alpha, H] \\ a_+ \, f_1(v, \lambda, \boldsymbol{\gamma}) + b_+ \, f_2(v, \lambda, \boldsymbol{\gamma}) & v \in (H, \theta) \end{cases} \tag{2.34}$$

We find that non-trivial coefficients $(a_-, a_+, b_+) \neq \mathbf{0}$ exist satisfying the boundary conditions (flux continuity at H, absorption at θ) if and only if λ satisfies the transcendental characteristic equation:

$$\Delta_{\boldsymbol{\gamma}}(\lambda) := \frac{f_1(\theta, \lambda, \boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(\theta)} - \frac{f_1(H, \lambda, \boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(H)} = 0. \tag{2.35}$$

The adjoint problem and the characteristic equation

It is often more insightful to formulate this spectral problem using the **adjoint operator**:

$$\mathcal{L}^*_{\boldsymbol{\gamma}} f(v) = (b(v) + \mu)\partial_v f(v) + D\partial_v^2 f(v), \tag{2.36}$$

subject to the adjoint boundary conditions:

$$f(H) = f(\theta), \tag{2.37}$$

$$f'(\alpha) = 0. \tag{2.38}$$

Let $\psi_\lambda(v)$ denote the eigenfunctions satisfying $\mathcal{L}^*_{\boldsymbol{\gamma}} \psi_\lambda = \lambda \psi_\lambda$. We can construct these using the fundamental solutions of the forward problem, scaled by the Wronskian $w_{\boldsymbol{\gamma}}(v)$ to simplify the adjoint flux terms:

$$\psi_1(v, \lambda) = \frac{f_1(v, \lambda)}{w_{\boldsymbol{\gamma}}(v)}, \quad \psi_2(v, \lambda) = \frac{f_2(v, \lambda)}{w_{\boldsymbol{\gamma}}(v)}. \tag{2.39}$$

The condition $f'(\alpha) = 0$ selects ψ_1 as the relevant solution. The non-local boundary condition $f(H) = f(\theta)$ then leads directly to the characteristic equation:

$$\Delta(\lambda) := \psi_1(\theta, \lambda) - \psi_1(H, \lambda) = 0, \tag{2.40}$$

which is equivalent to Eq. (2.35). The roots of this equation $\{\lambda_n\}_{n=0}^{\infty}$ constitute the discrete spectrum of the operator. For each eigenvalue λ_n, the corresponding (unnormalized) adjoint eigenfunction is simply $\tilde{\psi}_n(v) = \psi_1(v, \lambda_n)$.

Example: The LIF spectrum

In the case of the Leaky Integrate-and-Fire (LIF) neuron, the fundamental solutions f_1 and f_2 can be expressed analytically in terms of Kummer's confluent hypergeometric functions $M(a, b, z)$. In

the limit $\alpha \to -\infty$, the solutions commonly used in the literature [1, 126] are:

$$f_2(x,\lambda) \propto M\left(\frac{1-\lambda}{2}, \frac{1}{2}, -x^2\right), \tag{2.41}$$

$$f_1(x,\lambda) \propto \frac{\Gamma(\frac{\lambda}{2})}{\Gamma(\frac{\lambda+1}{2})} M\left(\frac{1-\lambda}{2}, \frac{1}{2}, -x^2\right) + 2xM\left(1-\frac{\lambda}{2}, \frac{3}{2}, -x^2\right), \tag{2.42}$$

where $x = (v-\mu)/\sqrt{2D}$ is the dimensionless potential. The characteristic equation then becomes a condition on the values of these special functions at the reset and threshold potentials.

Biorthogonality and Normalization

Since the operator $\mathcal{L}_{\boldsymbol{\gamma}}$ is non-self-adjoint, the eigenfunctions $\{\phi_n\}$ are not mutually orthogonal. Instead, they form a **biorthogonal system** with the adjoint eigenfunctions $\{\psi_n\}$:

$$\langle \psi_m, \phi_n \rangle = \int_\alpha^\theta \psi_m(v)\phi_n(v)\, dv = \delta_{mn}. \tag{2.43}$$

We normalize the adjoint eigenfunctions such that $\psi_n(v) = Z_n^{-1}\tilde{\psi}_n(v)$. The normalization factor Z_n is the leading coefficient of the characteristic function at the eigenvalue (equivalently, the reciprocal of the residue of the resolvent), consistently with the general treatment of Chapter 4, and can be computed analytically as [198]:

$$Z_n = \lim_{s\to\lambda_n} \frac{\Delta(s)}{s-\lambda_n} = \Delta'(\lambda_n). \tag{2.44}$$

We further normalize the forward eigenfunctions ϕ_n to have unit flux at the threshold:

$$\mathcal{S}_{\boldsymbol{\gamma}}[\phi_n](\theta) = 1, \quad \forall n \geq 1. \tag{2.45}$$

Structure of the spectrum

To understand the nature of the population dynamics, it is instructive to analyze how the spectrum $\{\lambda_n\}$ evolves as a function of the input parameters μ and D (or $\sigma = \sqrt{2D}$). We numerically solve [11] the characteristic equation $\Delta(\lambda) = 0$ for the LIF neuron (with $\tau_0 = 0$) and observe the following general features (Figures 2.1 and 2.2).

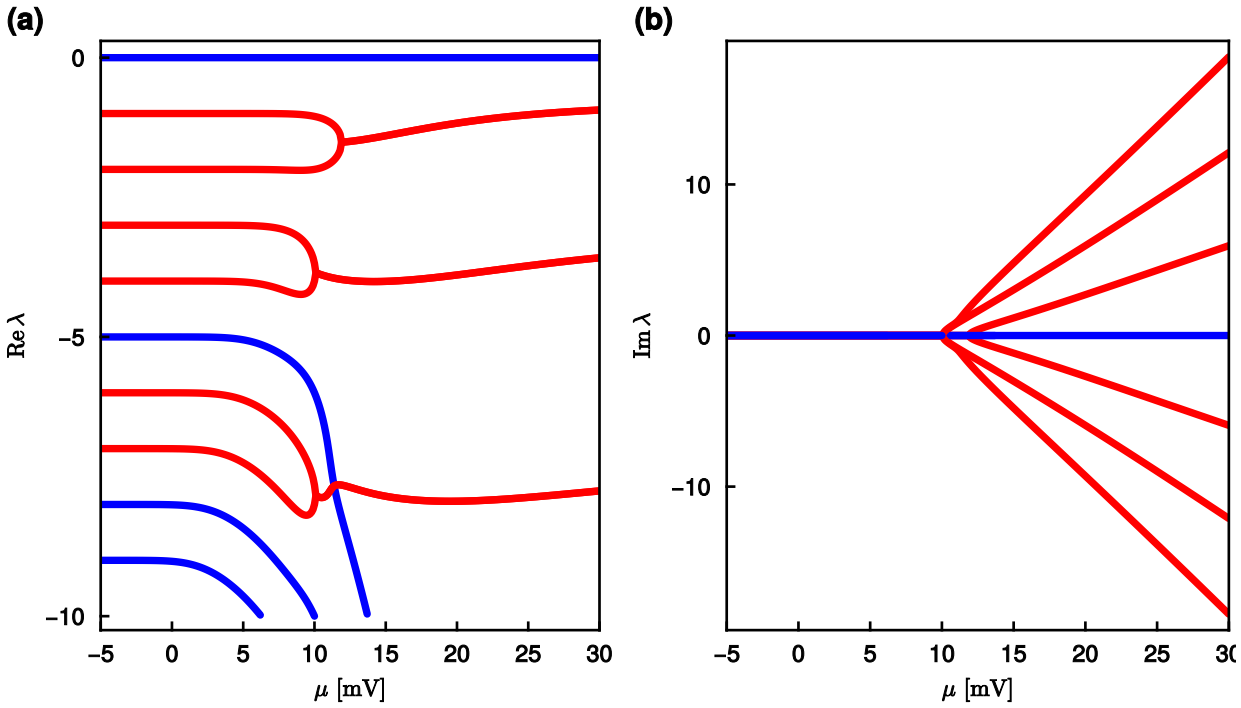

Figure 2.1: Low variance regime. (a) Real part and (b) imaginary part of the eigenvalues λ_n as a function of the mean input μ; red branches are complex (mixed), blue branches purely real. Parameters used: $D = 8\,\mathrm{mV}^2$, $H = 0$ mV, $\theta = 20$ mV, $\tau_0 = 0$, $\alpha = -5\theta = -100$ mV, $\tau_m = 1$.

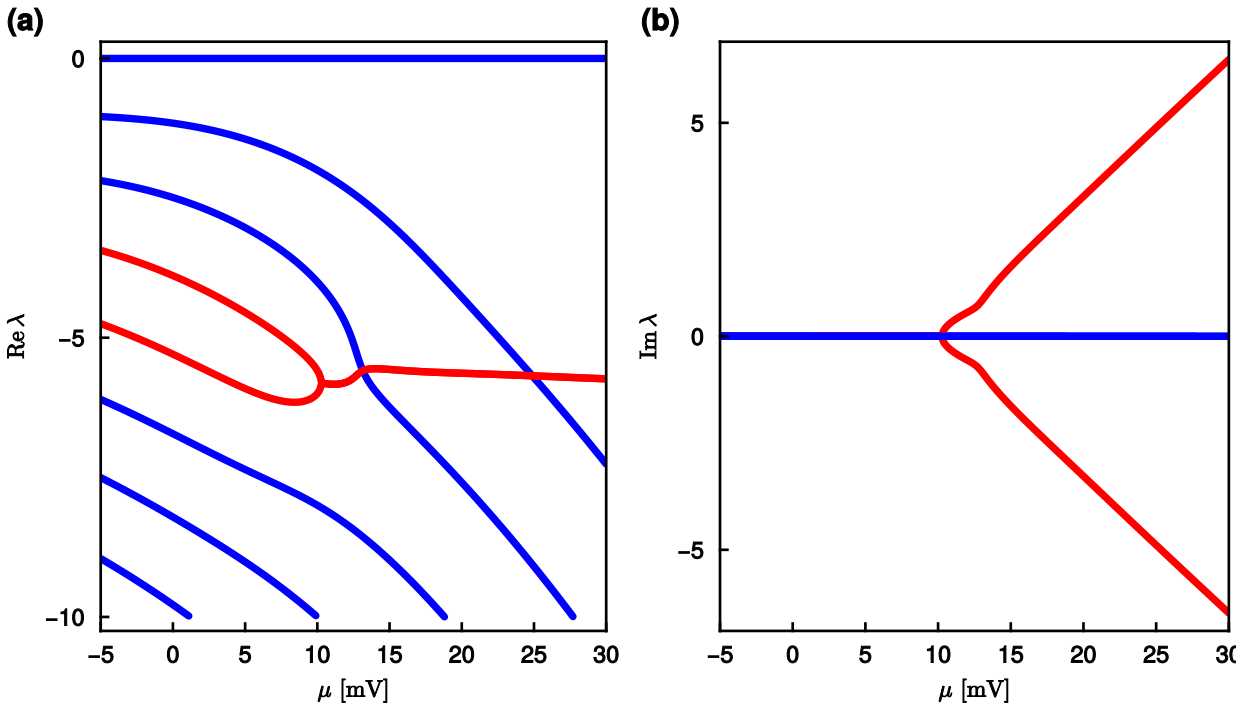

Figure 2.2: High variance regime. (a) Real part and (b) imaginary part of the eigenvalues λ_n as a function of the mean input μ; red branches are complex (mixed), blue branches purely real. Parameters used: $D = 60.5\,\mathrm{mV}^2$, $H = 0$ mV, $\theta = 20$ mV, $\tau_0 = 0$, $\alpha = -5\theta = -100$ mV, $\tau_m = 1$.

First, we observe that $\lambda_0 = 0$ is always an eigenvalue, corresponding to the unique stationary distribution ϕ_0. The remaining non-zero eigenvalues determine the transient relaxation and can be

categorized into two distinct groups:

- **Real eigenvalues:** These modes correspond to purely exponential decay. The coefficients $a_n(t)$ associated with real eigenvalues decay to zero monotonically with a characteristic time constant $|\lambda_n|^{-1}$. When real eigenvalues dominate the spectrum (i.e., have the largest real parts), the population firing rate converges to its steady-state value $\Phi(\mu, D)$ without oscillations.
- **Complex eigenvalues:** These appear in complex conjugate pairs with non-trivial imaginary parts. The associated mode coefficients decay while oscillating at a frequency proportional to $|\mathrm{Im}(\lambda_n)|$. Consequently, when complex eigenvalues dominate, the firing rate exhibits damped oscillations as it converges to the stationary state.

Varying the mean input μ while keeping the noise intensity fixed reveals that the eigenvalues organize into specific "branches" [11]:

- **Real branches (blue):** These consist of eigenvalues that remain real for all values of μ. Typically, these eigenvalues decrease for suprathreshold currents ($\mu > \theta$), tending to $-\infty$ as $\mu \to \infty$. The rate at which they retreat depends heavily on the noise level; for high noise, they remain closer to the origin for a wider range of μ.
- **Mixed branches (red):** These are formed by pairs of eigenvalues that are real at low μ but merge and bifurcate into a complex conjugate pair as μ increases. The merger occurs at a singular point (typically between the threshold and reset potentials); numerically, for $\tau_0 = 0$, the merger occurs for $\mu > (H + \theta)/2$ [127]. Beyond this point, as μ increases further, the real part of the pair tends to stabilize while the imaginary part grows linearly, leading to higher-frequency oscillations.

The dynamical regime of the network is determined by which of these branches is "dominant" (i.e., closest to the imaginary axis). In the **low noise regime** (e.g., small D), the mixed branches are dominant even for subthreshold currents ($\mu \approx (\theta + H)/2$). This implies that the population will exhibit oscillatory relaxation even when the mean drive is weak. In contrast, in the **high noise regime** (e.g., large D), the real branches dominate for low μ. The system behaves like an overdamped oscillator. Oscillations only emerge at suprathreshold inputs ($\mu > \theta$) when the real branches have decreased sufficiently, allowing the complex branches to emerge as the leading modes. As one of the fundamental contributions of this thesis, we will later show how this structure, i.e. the organization of the spectrum into real and mixed branches, emerges rigorously from the asymptotic analysis of the boundary eigenvalue problem developed in Chapter 5.

2.3.2 Dynamics in the spectral basis

We now leverage the spectral properties of the Fokker-Planck operator to decompose the full time-dependent dynamics of the population. Assuming the completeness of the biorthogonal basis formed by the eigenfunctions $\{\phi_n\}_{n=0}^{\infty}$ and their adjoints $\{\psi_n\}_{n=0}^{\infty}$, we can expand the probability density $p_t(v)$ as:

$$p_t(v) = \sum_{n=0}^{\infty} a_n(t)\phi_n(v, \boldsymbol{\gamma}(t)). \tag{2.46}$$

It is crucial to note that in the coupled regime, the basis functions themselves evolve in time. The input moments $\boldsymbol{\gamma}(t) = \{\mu(t), D(t)\}$ are determined by the network firing rate $\nu(t)$ (via the mean-field closure), rendering the eigenfunctions $\psi_n(v, \boldsymbol{\gamma}(t))$ and eigenvalues $\lambda_n(\boldsymbol{\gamma}(t))$ time-dependent.
The dynamics of the system are fully captured by the evolution of the projection coefficients $a_n(t)$. Using the biorthogonality relation $\langle \psi_m, \phi_n \rangle = \delta_{mn}$, we define the amplitudes as:

$$a_n(t) = \langle \psi_n(\cdot\ , \boldsymbol{\gamma}(t)), p_t \rangle. \tag{2.47}$$

The zeroth mode corresponds to the stationary distribution. Since the total probability mass is conserved ($\int p_t(v)dv = 1$) and the stationary mode is the only one with non-zero mass ($\int \phi_0 = 1, \int \phi_{n\geq 1} = 0$), the first coefficient is constant:

$$a_0(t) = \langle \psi_0, p_t \rangle = \langle 1, p_t \rangle = 1. \tag{2.48}$$

For the excited modes ($n \geq 1$), we differentiate the projection with respect to time:

$$\dot{a}_n(t) = \langle \psi_n, \partial_t p_t \rangle + \langle \partial_t \psi_n, p_t \rangle \tag{2.49}$$

$$= \langle \psi_n, \mathcal{L}_{\boldsymbol{\gamma}(t)} p_t \rangle + \sum_{m=0}^{\infty} a_m(t) \langle \partial_t \psi_n, \phi_m \rangle. \tag{2.50}$$

Using the adjoint property $\langle \psi_n, \mathcal{L}_{\boldsymbol{\gamma}} p \rangle = \langle \mathcal{L}^*_{\boldsymbol{\gamma}} \psi_n, p \rangle = \lambda_n \langle \psi_n, p \rangle$, the first term simplifies to $\lambda_n(t) a_n(t)$. The second term captures the non-adiabatic transitions induced by the time-varying basis. Formally expanding the time derivative via the chain rule $\partial_t = \dot{\boldsymbol{\gamma}} \cdot \nabla_{\boldsymbol{\gamma}}$, we obtain:

$$\dot{a}_n(t) = \lambda_n(t) a_n(t) + \sum_{m=0}^{\infty} a_m(t) \left[\dot{\mu}(t) \langle \partial_\mu \psi_n, \phi_m \rangle + \dot{D}(t) \langle \partial_D \psi_n, \phi_m \rangle \right]. \tag{2.51}$$

The instantaneous firing rate of the population is given by the flux at the threshold. Using the expansion of the density and the normalization condition for the flux ($\mathcal{S}[\phi_0] = \nu_0$, $\mathcal{S}[\phi_{n\geq 1}] = 1$), we have:

$$\nu(t) = \mathcal{S}_{\boldsymbol{\gamma}}[p_t](\theta) = \sum_{n=0}^{\infty} a_n(t) \mathcal{S}_{\boldsymbol{\gamma}}[\phi_n](\theta) = \Phi(\boldsymbol{\gamma}(t)) + \sum_{n=1}^{\infty} a_n(t). \tag{2.52}$$

Matrix formulation and Coupling Coefficients

It is convenient to rewrite the dynamical system in matrix notation [126]. Let $\mathbf{a}(t) = (a_1(t), a_2(t), \dots)^\top$ be the vector of excited mode amplitudes. The system becomes:

$$\begin{cases} \dot{\mathbf{a}}(t) = \boldsymbol{\Lambda}(t)\mathbf{a}(t) + \dot{\boldsymbol{\gamma}} \cdot \mathbf{C}^{\boldsymbol{\gamma}} \mathbf{a}(t) + \dot{\boldsymbol{\gamma}} \mathbf{c}^{\boldsymbol{\gamma}} \\ \nu(t) = \Phi(\boldsymbol{\gamma}(t)) + \mathbf{f} \cdot \mathbf{a}(t) \end{cases} \tag{2.53}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \dots)$ is the diagonal matrix of eigenvalues, and $\mathbf{f} = (1, 1, \dots)^\top$ sums the excited modes. The interaction between modes is mediated by the **coupling coefficients**, $\mathbf{C}^{\boldsymbol{\gamma}} =$

$\left(\mathbf{C}^{\mu}, \mathbf{C}^{D}\right)$ defined as:

$$(\mathbf{C}^{\chi})_{nm} := \langle \partial_{\chi} \psi_n, \phi_m \rangle, \quad n, m \geq 1, \ \chi \in \{\mu, D\} \tag{2.54}$$

$$(\mathbf{c}^{\chi})_{n} := \langle \partial_{\chi} \psi_n, \phi_0 \rangle, \quad n \geq 1, \ \chi \in \{\mu, D\} \tag{2.55}$$

For notational convenience, in this thesis, we will use the notation:

$$\dot{\boldsymbol{\gamma}} \cdot \mathbf{C}^{\boldsymbol{\gamma}} = \dot{\mu}\mathbf{C}^{\mu} + \dot{D}\mathbf{C}^{D} \tag{2.56}$$

The coupling coefficients quantify how changes in the input parameters μ and D transfer energy between different spectral modes. Specifically, the vectors $\mathbf{c}^{\mu}$ and $\mathbf{c}^{D}$ represent the direct excitation of transient modes from the stationary distribution due to parameter shifts, while the matrices $\mathbf{C}^{\mu}$ and $\mathbf{C}^{D}$ describe the mixing among transient modes.
In the extended mean-field approximation, the parameter dynamics $\dot{\boldsymbol{\gamma}}$ are themselves driven by the firing rate history (e.g., $\dot{\mu} \propto \dot{\nu}$ in the simplest case or via convolution with delay kernels). This closes the system, yielding a self-consistent set of nonlinear ODEs for the spectral amplitudes.

Open mathematical problems

While the spectral approach described above provides a compelling physical intuition and a powerful computational scheme, its rigorous mathematical foundation in this context remains incomplete. A major theoretical gap is the lack of a Spectral Theorem for the non-self-adjoint Fokker-Planck operator $\mathcal{L}_{\boldsymbol{\gamma}}$. The completeness of the eigenbasis $\{\phi_n\}$ is generally assumed in the physics literature but has not been rigorously proven. Furthermore, the standard analysis relies on the assumption that eigenvalues are simple and discrete, yet the precise structure of the spectrum for the Fokker-Planck operator with reinjection boundary conditions requires careful analysis. Another significant limitation is the simplifying assumption of a vanishing refractory period ($\tau_0 = 0$); incorporating a finite $\tau_0 > 0$ introduces delays that fundamentally alter the nature of the operator, complicating the spectral definition. Finally, a systematic treatment of the appropriate function spaces and the convergence properties of the spectral expansion is largely missing.
In the chapters that follow, we will tackle these problems directly. We will provide a rigorous characterization of the spectrum of the Fokker-Planck operator with reinjection, extend the theory to handle finite refractory periods, and establish the convergence properties of the spectral expansion, thereby placing the spectral decomposition of population dynamics on solid mathematical ground.

Chapter 3

Non-stationary Inter-Spike Interval Dynamics

> Come per fargli assaporare il gusto della terra
> La sorpresa della prima volta che senti dolore
> — Tutte le cose inutili, *Bombe Carta*

In this chapter, we study the joint dynamics of membrane potential and time since the last spike in a population of integrate-and-fire neurons. By extending the population density framework introduced in the previous chapter, we derive a two-dimensional Fokker–Planck equation that captures the evolution of the full neuronal state. This formalism, introduced in [58], allows us to characterize the time-dependent inter-spike interval (ISI) distribution even when the population is far from stationarity, such as under time-varying external input or during network oscillations. Furthermore, by performing a perturbative expansion around the stationary state, we derive an analytic expression for the linear response of the ISI distribution to weak input modulations.

From neurophysiology to mathematical models

Neurons communicate through stereotyped, spike-like electrical impulses generated when their membrane potentials exceed a threshold. The intervals between consecutive spikes, known as inter-spike intervals (ISIs), are highly irregular in the cortex, even under identical stimulus or environmental conditions [177, 184]. Such variability gives rise to rich ISI statistics and non-Poisson distributions [15, 116, 189], suggesting that information transmission by cortical neurons may be fundamentally constrained by noise [178]. Nevertheless, stimulus-dependent modulations of ISI statistics can be faithfully represented at the population level, enabling temporally precise encoding of information [14, 178]. From this perspective, developing a theoretical framework that explains how ISI statistics unfold in neuronal populations is crucial for understanding the principles of neural information processing and encoding in the brain.

Variability in neuronal spike patterns arises from the incessant fluctuations of input currents driven by both continuous changes in the sensed environment and intrinsic brain activity, even under constant external stimulation [10, 57]. These fluctuations represent an unavoidable source of noise that can be further amplified by neurons acting as excitable systems [56], where stochastic resonance

may shape and enrich the ISI distribution [112]. During sleep or general anesthesia, when the unconscious brain enters an isolated and synchronized state, slow waves generate a quasi-periodic alternation between high-firing and silent phases [170], leading to distinctive ISI distributions [202] that differ from those observed during wakefulness across cortical layers and regions [49, 180]. The diversity of ISI statistics has been theoretically linked to the presence of metastable dynamics in neuronal systems, both during spontaneous activity [110, 206] and during the performance of specific cognitive tasks [47, 214]. Therefore, a comprehensive theory of ISI dynamics must account for neuronal populations operating far from equilibrium.

The study of ISI distributions in stationary neuronal populations has long been a cornerstone of stochastic analysis in neuroscience [71, 162, 193]. At the single-neuron level, ISIs naturally arise within the framework of the *first-passage time* (FPT) problem, which characterizes the distribution of times required for the membrane potential $V(t)$ to reach threshold. Under the diffusion approximation, the subthreshold dynamics of an integrate-and-fire (IF) neuron model is governed by a stochastic differential equation, and the corresponding membrane potential density $p^{\mathrm{A}}(v,t)$ satisfies the Fokker-Planck (FP) equation. In the classical formulation, the ISI distribution is obtained by solving the FP equation with an absorbing barrier, where the ISI density corresponds to the flux of realizations crossing the threshold. For stationary currents, exact ISI probability densities exist only for a few simplified models [71, 163, 193]. However, Siegert's recursion formula [182, 193] has enabled the derivation of the first two moments of the ISI for a broader class of IF neurons [30, 68, 109, 162], even in the presence of multiplicative noise, as in conductance-based models [138, 171]. These results have been partially recovered through the spectral decomposition of the Fokker-Planck operator [154, 163, 198]. Collectively, they indicate a degree of universality in ISI distributions across IF models [148, 197].

ISI and population dynamics

Although this framework is well established for stationary inputs, its extension to *non-stationary* or *self-consistent* regimes introduces substantial difficulties. As discussed in the previous chapter, in mean-field descriptions of spiking neuron networks [6], the mean and variance of synaptic currents depend on the population firing rate, which is intimately tied to ISI statistics, rendering the evolution operator both time-dependent and nonlinear [31, 126]. Self-consistency also necessitates modeling the reset of the membrane potential after spike emission. This is achieved in the *population density approach*[31, 99, 101, 126, 141], detailed in Chapter 2 by reinjecting the outgoing flux at the reset potential. While this approach preserves individual realizations and maintains a normalized probability density, the continuous coupling of subthreshold voltage dynamics to spike-time history transcends the renewal approximation, consequently losing the explicit link to single-neuron ISI statistics.

Alternative approaches linking population dynamics to single-neuron ISIs include the *spike response model* (SRM) formalism [72] and the *refractory density method* (RDM) [44]. The SRM formulation computes population activity by filtering past firing rates with a state-dependent ISI distribution. The RDM, in contrast, explicitly evolves the population density structured by the age τ (time since last spike), employing a hazard function derived from the stationary ISI distribution. To address non-stationarity, both frameworks break the strict renewal hypothesis by modulating these renewal quantities based on the instantaneous input statistics. However, this relies on a *quasi-renewal*

approximation: it assumes that at any instant, the spiking statistics match those of a stationary state driven by the current inputs. Far from steady state, this approximation fails because the true hazard rate and ISI statistics depend nontrivially on the full voltage dynamics, which are shaped by the history of time-varying, self-consistent inputs. Consequently, the actual non-stationary quantities differ from the stationary renewal ones, and, as we will show, it is impossible to fully decouple the age-dependent dynamics from the underlying voltage evolution.
To address all these issues, here we formulate the neuronal dynamics in an extended state space that jointly tracks the membrane potential V and the time since last spike τ. This construction leads naturally to a two-dimensional FP equation for the joint density $q(v, \tau, t)$, enabling the analysis of ISI statistics in non-stationary, self-consistent regimes beyond the reach of classical renewal formulations and their approximated extensions.

Chapter outline

The chapter is organized as follows. We first review the classical theory for stationary regimes, discussing both the forward-in-time evolution of the probability density and the backward-in-time approach that leads to the Siegert recursive relations. We then introduce the core contribution of this work: the population density equation in the extended state space, clarifying its relationship with age-structured dynamics. From this general framework, we derive a hierarchy of equations for the moments of the ISI distribution, which constitutes a primary result of this study. Subsequently, we apply this formalism to obtain analytical results regarding the relaxation to the stationary state and the linear response to weak input modulations via a first-order perturbation theory. We conclude by validating our theoretical predictions against numerical simulations of spiking neural networks, specifically exploring ISI distributions in excitatory and inhibitory populations exhibiting limit cycle oscillations.

3.1 ISI from the evolution problem without reinjection

We consider a single neuron with membrane voltage $v \in (\alpha, \theta)$, where α is the reversal potential and θ is the spiking threshold. For a given initial voltage $y \in (\alpha, \theta)$, we study the probability density evolution under the Fokker-Planck equation, without the reinjection, modelling its dynamics until spike generation.
The evolution of the probability density $p^{\mathrm{A}}(v, t)$ is governed by the absorbing Fokker-Planck equation:

$$\partial_t p^{\mathrm{A}}(v, t) = \mathcal{L}^A_{\boldsymbol{\gamma}} p^{\mathrm{A}}(v, t), \quad v \in (\alpha, \theta), \quad t > 0, \tag{3.1}$$

with initial condition $p^{\mathrm{A}}(v, 0) = \delta(v - y)$ for $y \in (\alpha, \theta)$. The superscript "A" denotes that we are only considering absorbing (at θ) and reflecting (at α) boundary conditions at the interval endpoints. The probability density of the inter-spike interval (ISI), i.e., the time at which the neuron first reaches threshold starting from initial state y, is given by:

$$\rho(t|y) = \mathcal{S}_{\boldsymbol{\gamma}(t)}[p^A(\cdot, t)](\theta) = -D(t)\partial_v p^{\mathrm{A}}(v, t)\Big|_{v=\theta^-}, \quad t > 0. \tag{3.2}$$

We formally define $\mathcal{L}_{\boldsymbol{\gamma}}^A$ as the unbounded operator on the Hilbert space $\mathrm{L}^2(\alpha,\theta)$:

$$\mathcal{L}_{\boldsymbol{\gamma}}^A : \mathbb{D}_{\boldsymbol{\gamma}}^A \subset \mathrm{L}^2(\alpha,\theta) \to \mathrm{L}^2(\alpha,\theta), \quad \mathcal{L}_{\boldsymbol{\gamma}}^A p = \mathcal{L}_{\boldsymbol{\gamma}} p \quad \text{for } p \in \mathbb{D}_{\boldsymbol{\gamma}}^A, \tag{3.3}$$

with domain:

$$\mathbb{D}_{\boldsymbol{\gamma}}^A = \left\{ p \in \mathrm{W}^{2,2}(\alpha,\theta) : p(\theta) = 0, \quad \mathcal{S}_{\boldsymbol{\gamma}}[p](\alpha) = 0 \right\}, \tag{3.4}$$

where $\mathrm{W}^{2,2}(\alpha,\theta)$ denotes the Sobolev space of functions with weak derivatives up to order 2 in $\mathrm{L}^2(\alpha,\theta)$. The boundary condition at θ is Dirichlet (absorbing), while at α the condition enforces zero total probability flux, cf. Eq. (2.18).

Theorem 3.1.1. *The operator $\mathcal{L}_{\boldsymbol{\gamma}}^A$ defined on $\mathbb{D}_{\boldsymbol{\gamma}}^A$ is self-adjoint in the weighted space $\mathrm{L}^2_{w^{-1}}(\alpha,\theta)$, with weight $w^{-1} > 0$ given by the* reciprocal *of the Wronskian $w = w_{\boldsymbol{\gamma}}$ of Eq. (2.23).*
For all $p \in \mathbb{D}_{\boldsymbol{\gamma}}^A$:

$$\langle p, \mathcal{L}_{\boldsymbol{\gamma}}^A p \rangle_{w^{-1}} \leq 0. \tag{3.5}$$

Consequently, all eigenvalues of $\mathcal{L}_{\boldsymbol{\gamma}}^A$ satisfy $\mu_n < 0$ for all $n \in \mathbb{N}$.

Proof. We verify symmetry: let $p \in \mathbb{D}_{\boldsymbol{\gamma}}^A$, $q \in \mathrm{W}^{2,2}(\alpha,\theta)$. Using $\mathcal{L}_{\boldsymbol{\gamma}}^A p = -\partial_v \mathcal{S}_{\boldsymbol{\gamma}}[p]$ and integration by parts:

$$\langle \mathcal{L}_{\boldsymbol{\gamma}}^A p, q \rangle_{w^{-1}} = -\left[\mathcal{S}_{\boldsymbol{\gamma}}[p] \overline{q}\, w^{-1} \right]_\alpha^\theta - \frac{1}{D} \int_\alpha^\theta \mathcal{S}_{\boldsymbol{\gamma}}[p] \overline{\mathcal{S}_{\boldsymbol{\gamma}}[q]}\, w^{-1}\, dv. \tag{3.6}$$

The integral term is manifestly symmetric in p and q. The surface term vanishes iff $q \in \mathbb{D}_{\boldsymbol{\gamma}}^A$. Thus, the operator is self-adjoint. The inequality (3.5) follows directly from (3.6) with $p = q \in \mathbb{D}_{\boldsymbol{\gamma}}^A$.
Since $\mathcal{L}_{\boldsymbol{\gamma}}^A$ is self-adjoint, its eigenvalues are real, and (3.5) applied to an eigenfunction p with $\mathcal{L}_{\boldsymbol{\gamma}}^A p = \mu p$ gives $\mu \|p\|^2_{w^{-1}} \leq 0$, hence $\mu \leq 0$. The inequality is in fact strict. Suppose $\mu = 0$, i.e. $\mathcal{L}_{\boldsymbol{\gamma}}^A p = 0$ for some $p \in \mathbb{D}_{\boldsymbol{\gamma}}^A$. Setting $p = q$ in (3.6), the surface term vanishes because $\mathcal{S}_{\boldsymbol{\gamma}}[p](\alpha) = 0$ and $p(\theta) = 0$, leaving

$$\frac{1}{D} \int_\alpha^\theta \left| \mathcal{S}_{\boldsymbol{\gamma}}[p](v) \right|^2 w^{-1}(v)\, dv = 0\,. \tag{3.7}$$

As $w^{-1} > 0$ on (α,θ), this forces $\mathcal{S}_{\boldsymbol{\gamma}}[p] \equiv 0$. The zero-flux solutions of $\mathcal{S}_{\boldsymbol{\gamma}}[p] = 0$ are exactly the multiples of $w_{\boldsymbol{\gamma}}$, so $p = c\, w_{\boldsymbol{\gamma}}$ for some $c \in \mathbb{C}$; the absorbing condition $p(\theta) = 0$ together with $w_{\boldsymbol{\gamma}}(\theta) > 0$ then gives $c = 0$, i.e. $p \equiv 0$. Hence 0 is not an eigenvalue and $\mu_n < 0$ for all $n \in \mathbb{N}$. □

By Corollary 4.7 of [82], the operator $\mathcal{L}_{\boldsymbol{\gamma}}^A$ is sectorial of angle $\pi/2$. Here we adopt the convention of Engel and Nagel for sectorial operators [55]. Since b is continuous on the bounded interval $[\alpha,\theta]$, the weight w^{-1} is bounded above and below by strictly positive constants, so the norms of $\mathrm{L}^2_{w^{-1}}(\alpha,\theta)$ and $\mathrm{L}^2(\alpha,\theta)$ are equivalent; consequently the operator is also sectorial on the original unweighted $\mathrm{L}^2(\alpha,\theta)$ space. Consequently, $\mathcal{L}_{\boldsymbol{\gamma}}^A$ generates an analytic semigroup.
In the following, we suppress the subscript $\boldsymbol{\gamma}$ to lighten the notation, and we denote stationary quantities with the subscript 0. For stationary inputs $(\mu(t), D(t)) = (\mu_0, D_0)$, this problem admits

a closed-form representation in Laplace space. Taking the Laplace transform yields:

$$(s - \mathcal{L}_0^A)\tilde{p}^{\mathrm{A}}(v, s) = \delta_y(v), \tag{3.8}$$

$$\Rightarrow \tilde{p}^{\mathrm{A}}(v, s) = G^A(v, y; s), \tag{3.9}$$

where the solution is expressed in terms of the Green's function $G^{\mathrm{A}}(v, y; s)$ for the absorbing boundary-value problem. The Laplace transform of the stationary ISI density ρ_0 reads:

$$\tilde{\rho}_0(s|y) = -D_0\, \partial_v G^{\mathrm{A}}(v, y; s)\big|_{v=\theta} = \frac{\psi_1(y; s)}{\psi_1(\theta; s)}, \tag{3.10}$$

where $\psi_1(y; s)$ is the solution for the homogeneous problem $\mathcal{L}_0^+\psi_1(\cdot, s) = s\psi_1(\cdot, s)$ with the boundary conditions $\psi_1(\alpha; s) = 1$ and $\partial_y\psi_1(y; s)\big|_{y=\alpha} = 0$. All stationary ISI moments then follow as:

$$m_n^0(y) = (-1)^n\, \partial_s^n \tilde{\rho}_0(s|y)\big|_{s=0}. \tag{3.11}$$

3.1.1 Details on the Green function

The Green function $G^{\mathrm{A}}(x, y; s)$ for the Fokker-Planck operator $\mathcal{L}_0^A$ with absorbing boundary conditions can be expressed as a function of the solutions of the homogeneous equations:

$$\mathcal{L}_0 f_i(\cdot; s) = s f_i(\cdot; s)$$

Defined as in Equation (2.32),
then the Green's function has the form:

$$G^{\mathrm{A}}(x, y; s) = \begin{cases} f_1(x; s) \cdot a_-(y, s) + f_2(x; s) \cdot b_-(y, s) & x < y, \\ \\ f_1(x; s) \cdot a_+(y, s) + f_2(x; s) \cdot b_+(y, s) & x > y. \end{cases} \tag{3.12}$$

The coefficients $a_\pm(y, s), b_\pm(y, s) \in \mathbb{C}$, can be found by imposing the conditions:

$$\mathcal{S}\big[G^{\mathrm{A}}(\cdot, y; s)\big](\alpha) = 0, \quad G^{\mathrm{A}}(\theta, y; s) = 0, \quad (s - \mathcal{L}_0)G^{\mathrm{A}}(\cdot, y; s) = \delta_y. \tag{3.13}$$

Solving the associated linear system, we find:

$$G^{\mathrm{A}}(x, y; s) = \frac{1}{f_1(\theta; s)w(y)} \begin{cases} f_1(x; s)u_+(y; s) & x < y, \\ \\ f_1(y; s)u_+(x; s) & x > y. \end{cases}$$

where $u_+(x; s) = f_2(x; s)f_1(\theta; s) - f_1(x; s)f_2(\theta; s)$, such that $u_+(\theta; s) = 0$ and $\mathcal{S}[u_+](\theta) = w(\theta)$. Thus:

$$\mathcal{S}[G^A(\cdot, y; s)](\theta) = \frac{f_1(y)w(\theta)}{f_1(\theta)w(y)} = \frac{\psi_1(y; s)}{\psi_1(\theta; s)} \tag{3.14}$$

proving (3.10).

Green's function for the driftless BPIF

We consider the driftless Bounded Perfect Integrate-and-Fire (BPIF) model on the interval $(\alpha, \theta) = (0, \pi)$ with Neumann boundary conditions at α and Dirichlet at θ. The Green's function for the resolvent equation with $\lambda = -z^2$ is given by:

$$G^A(x, y; -z^2) = \frac{1}{z\cos(z\pi)} \begin{cases} \cos(zx)\sin(z(\pi - y)) & x < y, \\ \cos(zy)\sin(z(\pi - x)) & y < x. \end{cases}$$

To show the operator is sectorial, we derive a uniform bound for λ in a sector avoiding the negative real axis. Let $\epsilon \in (0, \pi)$ and define

$$\Sigma_{\pi-\epsilon} := \{\lambda \in \mathbb{C} \setminus \{0\} : \ |\arg\lambda| < \pi - \epsilon\}.$$

Assume $\lambda \in \Sigma_{\pi-\epsilon}$ and write $\lambda = -z^2$ with a fixed branch of the square root (e.g. $\operatorname{Re} z \geq 0$). Then $\arg(-\lambda) = \arg\lambda + \pi \pmod{2\pi}$ stays at distance ϵ from π and hence there exists $c_\epsilon > 0$ such that

$$|\operatorname{Im}(z)| \geq c_\epsilon |z|, \qquad \text{and} \qquad |z| = |\lambda|^{1/2}.$$

We use the standard asymptotic estimates for large complex argument w (uniform in closed subsectors): $|\sin(w)|, |\cos(w)| \sim e^{|\operatorname{Im} w|}$, which yields (uniformly for $\lambda \in \Sigma_{\pi-\epsilon}$) a constant $C_\epsilon > 0$ such that for $x < y$ gives

$$|G^A(x, y; -z^2)| \leq \frac{C_\epsilon}{|z|} e^{|\operatorname{Im}(z)|x} e^{|\operatorname{Im}(z)|(\pi - y)} e^{-\pi|\operatorname{Im}(z)|} = \frac{C_\epsilon}{|z|} e^{-|\operatorname{Im}(z)|(y-x)} = \frac{C_\epsilon}{|z|} e^{-|\operatorname{Im}(z)||x-y|}.$$

The same estimate holds for $y < x$, hence

$$|G^A(x, y; -z^2)| \leq \frac{C_\epsilon}{|z|} e^{-|\operatorname{Im}(z)||x-y|}, \qquad x, y \in (0, \pi), \ \lambda = -z^2 \in \Sigma_{\pi-\epsilon}. \tag{3.15}$$

We now pass from the pointwise kernel bound to an L^2-operator norm bound for the resolvent $(\lambda - \mathcal{L}^A)^{-1}$, viewed as an integral operator on $\mathrm{L}^2(0, \pi)$:

$$\big((\lambda - \mathcal{L}^A)^{-1} f\big)(x) = \int_0^\pi G^A(x, y; \lambda)\, f(y)\, dy.$$

From (3.15),

$$\sup_{x\in(0,\pi)} \int_0^\pi |G^A(x, y; -z^2)|\, dy \leq \frac{C_\epsilon}{|z|} \sup_{x\in(0,\pi)} \int_0^\pi e^{-|\operatorname{Im}(z)||x-y|}\, dy \leq \frac{C_\epsilon}{|z|} \cdot \frac{2}{|\operatorname{Im}(z)|},$$

and similarly

$$\sup_{y\in(0,\pi)} \int_0^\pi |G^A(x, y; -z^2)|\, dx \leq \frac{C_\epsilon}{|z|} \cdot \frac{2}{|\operatorname{Im}(z)|}.$$

By Schur's test, it follows that $(\lambda - \mathcal{L}^A)^{-1}$ is bounded on $\mathrm{L}^2(0, \pi)$ and

$$\|(\lambda - \mathcal{L}^A)^{-1}\|_{\mathrm{L}^2} \leq \frac{2C_\epsilon}{|z|\,|\operatorname{Im}(z)|}.$$

Using $|\mathrm{Im}(z)| \geq c_\epsilon |z|$ and $|z|^2 = |\lambda|$, we conclude that there exists $M_\epsilon > 0$ such that

$$\|(\lambda - \mathcal{L}^A)^{-1}\| \leq \frac{M_\epsilon}{|\lambda|}, \qquad \lambda \in \Sigma_{\pi-\epsilon}, \tag{3.16}$$

proving sectoriality of angle $\pi/2$.

3.2 ISI from the backward Kolmogorov equation

The inter-spike interval corresponds to the random time required for the process, starting at y, to exit the domain (α, θ) through the threshold θ. We formally define the first hitting time random variable $T(y)$ as:

$$T(y) = \inf\{t > 0 : V(t) = \theta \mid V(0) = y\}. \tag{3.17}$$

Let $\rho_0(t|y)$ denote the probability density function of $T(y)$. The survival probability function, $I(t|y) = \mathbb{P}(T(y) > t)$, represents the probability that the neuron has not spiked up to time t. This can be expressed as the probability mass remaining within the domain:

$$I(t|y) = \int_t^\infty \rho_0(\tau|y)\, d\tau = \int_\alpha^\theta p^{\mathrm{A}}(v, t|y)\, dv, \tag{3.18}$$

where $p^{\mathrm{A}}(v, t|y)$ is the transition density evolving under the absorbing boundary conditions. Since the transition density satisfies the backward Kolmogorov equation with respect to the initial coordinate y, the survival probability inherits this dynamics:

$$\partial_t I(t|y) = \mathcal{L}_0^* I(t|y), \quad y \in (\alpha, \theta), \quad t > 0. \tag{3.19}$$

By definition, the ISI density is the negative time derivative of the survival probability, $\rho_0(t|y) = -\partial_t I(t|y)$. Differentiating the evolution equation for $I(t|y)$ with respect to time, and assuming the interchangeability of the differential operator $\mathcal{L}_0^*$ and the time derivative, we find that the ISI density itself satisfies the backward Kolmogorov equation:

$$\partial_t \rho_0(t|y) = \mathcal{L}_0^* \rho_0(t|y). \tag{3.20}$$

This partial differential equation provides the foundation for deriving the moments of the distribution without explicit knowledge of the time-dependent solution. Defining the n-th moment as $m_n^0(y) = \int_0^\infty t^n \rho_0(t|y)\, dt$, we multiply the backward equation by t^n and integrate over $\mathbb{R}^+$. Integration by parts with respect to t yields the recursive Siegert hierarchy [182]:

$$-\mathcal{L}_0^* m_n^0(y) = n\, m_{n-1}^0(y), \quad n \geq 1, \tag{3.21}$$

subject to the boundary conditions $m_n^0(\theta) = 0$ and $\partial_y m_n^0(\alpha) = 0$, with the convention $m_0^0(y) \equiv 1$.

The system (3.21) can be solved iteratively. For $n = 1$, this recovers the mean first passage time. For arbitrary n, the solution admits an integral representation utilizing the Wronskian $w = w_{\boldsymbol{\gamma}}$ of

the homogeneous Fokker–Planck problem, defined in Eq. (2.23) [182]:

$$m_n^0(y) = \frac{n}{D_0} \int_y^\theta \frac{1}{w(z)} \left(\int_\alpha^z w(v) m_{n-1}^0(v)\, dv \right) dz. \tag{3.22}$$

3.3 ISI distribution in time-varying neural populations

We now turn to the central contribution of this work: the characterization of the ISI distribution in a neural population subject to time-varying diffusive dynamics.

Analyzing the ISI distribution in non-stationary regimes requires tracking the history of the process, which is explicitly lost in the standard population density approach. The classical Fokker-Planck equation for $p(v,t)$ describes the evolution of the membrane potential but aggregates all trajectories regardless of their previous spike times. The reinjection flux at $v = H$ mixes neurons that have just spiked with the existing population, erasing the temporal correlations required to compute the interval statistics. To resolve the ISI distribution, we must disentangle these trajectories by explicitly tracking the time elapsed since the last reset.

We formulate the neuronal dynamics in an extended state space (V, τ), where τ represents the time since the last spike. The coupled stochastic dynamics is given by:

$$\begin{cases} dV(t) = b(V(t))\, dt + \mu(t)\, dt + \sqrt{2D(t)}\, dW(t) \\ d\tau(t) = dt \end{cases}, \tag{3.23}$$

subject to the reset conditions: if $V(t^-) = \theta$, then $V(t) = H$ and $\tau(t) = 0$. The domain is bounded below by a reflecting barrier at $v = \alpha$.

This construction leads to a continuity equation in the extended phase space. The joint density $q(v, \tau, t)$ evolves according to the two-dimensional Fokker-Planck equation:

$$\begin{aligned} \partial_t q(v,\tau,t) =& \mathcal{L}^A_{\boldsymbol{\gamma}(t)}\, q(v,\tau,t) - \partial_\tau q(v,\tau,t) \\ &+ \delta_H(v)\, \delta_0(\tau)\, \nu(t), \end{aligned} \tag{3.24}$$

where $\mathcal{L}^A_{\boldsymbol{\gamma}(t)}$ acts solely on the voltage variable v. The system satisfies the absorbing boundary condition $q(\theta, \tau, t) = 0$ for all $t > 0$ and $\tau \geq 0$, and the reflecting condition at $v = \alpha$. The source term $\delta_H(v)\delta_0(\tau)\nu(t)$ enforces the immediate reinjection of firing neurons at the reset potential H with age zero.

We define the partial flux through the threshold for neurons with age τ as[1]:

$$\nu_q(\tau, t) := -D(t)\, \partial_v q(v,\tau,t)\big|_{v=\theta}. \tag{3.25}$$

The total population firing rate is recovered by integrating the partial flux over all possible ages:

$$\nu(t) = \int_0^\infty \nu_q(\tau, t)\, d\tau. \tag{3.26}$$

The instantaneous ISI distribution at time t is then defined as the normalized flux of neurons exiting

[1] In general, given a function $f(v, \cdot)$ vanishing at the boundary $v = \theta$, we define $\nu_f(\cdot) := -D(t)\partial_v f(v,\cdot)|_{v=\theta}$ as the partial flux at the boundary.

the system at age τ:

$$\rho(\tau, t) = \frac{\nu_q(\tau, t)}{\nu(t)}. \tag{3.27}$$

Remark. In this section, we adopt a formal approach. As this extended formulation represents a novel modelling tool, our primary objective is to investigate the structure of Eq. (3.24) and demonstrate its utility in computing ISI distributions far from stationarity, rather than establishing the functional analytic properties of the solution.

3.3.1 Relationship with age-structured population dynamics

Our approach clarifies the relationship between the population density method and age-structured population dynamics, providing a systematic framework for assessing the validity of the quasi-renewal approximation used in these models.

We can establish a direct connection between the introduced formalism and the Spike Response Model (SRM) [72]. To this end, we decompose the full density $q(v, \tau, t)$ by normalizing the reinjection flux at the reset conditions $(v, \tau) = (H, 0)$:

$$q(v, \tau, t) = g(v, \tau, t)\nu(t - \tau) + \ell(v, \tau, t). \tag{3.28}$$

Here, the function $\ell(v, \tau, t)$ accounts for neurons that have not fired during the observation interval $[0, t]$, with support restricted to $\tau \geq t$. Conversely, $g(v, \tau, t)$ describes the conditional probability density of a neuron that last fired at time $t - \tau > 0$ and is thus defined only for $\tau < t$. The initial conditions are given by $g(v, \tau, 0) = 0$ and $\ell(v, \tau, 0) = q(v, \tau, 0)$.

Substituting the ansatz $q(v, \tau, t) = g(v, \tau, t)\nu(t-\tau)$ (valid for $\tau < t$) into the master equation (3.24) and utilizing the identity $\partial_t \nu(t - \tau) = -\partial_\tau \nu(t - \tau)$, we find that the conditional kernel g satisfies the partial differential equation:

$$\begin{aligned} \partial_t g(v, \tau, t) =& \mathcal{L}^A_{\boldsymbol{\gamma}(t)}\, g(v, \tau, t) - \partial_\tau g(v, \tau, t) \\ &+ \delta_H(v)\delta_0(\tau) \quad \forall \tau < t. \end{aligned} \tag{3.29}$$

Crucially, g depends on the past firing history only through the modulation of the synaptic current moments $\mu(t)$ and $D(t)$. This allows us to express the instantaneous population firing rate $\nu(t)$ as an integral equation:

$$\nu(t) = \int_0^t \nu_g(\tau, t)\nu(t - \tau)\, d\tau + \rho_\ell(t), \tag{3.30}$$

where $\nu_g(\tau, t)$ is the flux generated by the kernel g. The term $\rho_\ell(t) = \int_0^\infty \nu_\ell(\tau, t)d\tau$ represents the contribution from the initial population (neurons firing for the first time since $t = 0$) and progressively vanishes as t increases. In the standard SRM framework, the kernel $\nu_g(\cdot, t)$ is typically approximated by a stationary ISI density. In contrast, our formulation provides an exact expression, albeit at the computational cost of solving the two-dimensional evolution of g.

Marginalizing the population density equation (3.24) over the membrane potential yields the Re-

fractory Density Method (RDM) formulation [44, 73, 175]:

$$\partial_t r(\tau,t) = -\partial_\tau r(\tau,t) - \nu_q(\tau,t) + \nu(t)\delta_0(\tau), \tag{3.31}$$

where $r(\tau,t) := \int_\alpha^\theta q(v,\tau,t)\,dv$ is the refractory density. We can rewrite this conservation law in terms of a dynamic hazard rate:

$$\partial_t r(\tau,t) = -\partial_\tau r(\tau,t) - h(\tau,t)r(\tau,t) + \nu(t)\delta_0(\tau), \quad \forall\tau < t, \tag{3.32}$$

$$h(\tau,t) := \frac{\nu_g(\tau,t)}{I(\tau,t)}. \tag{3.33}$$

Here, $h(\tau,t)$ represents the instantaneous firing probability of neurons with age τ. The restriction $\tau < t$ is what makes the substitution $\nu_q = h\,r$ exact: on that domain the initial population ℓ does not contribute, so $\nu_q = \nu_g\,\nu(t-\tau)$ and $r = I\,\nu(t-\tau)$, whence $\nu_q/r = \nu_g/I = h$. For $\tau \geq t$ the neurons that have not yet fired since $t = 0$ carry a distinct, initial-condition-dependent hazard; their contribution vanishes as t increases. The quantity $I(\tau,t) = \int_\alpha^\theta g(v,\tau,t)\,dv$ defines the conditional survival function, governed by:

$$\partial_t I(\tau,t) = -\partial_\tau I(\tau,t) - \nu_g(\tau,t) + \delta_0(\tau), \quad \forall\tau < t. \tag{3.34}$$

Unlike the RDM, which closes the age-structured conservation law by approximating $h(\tau,t)$ via a quasi-stationary hazard approximation [42, 43], our approach retains the full hazard dynamics.

Hazard closures in refractory-density models. In refractory-density formulations, the main modeling step is therefore the specification of an effective hazard (or equivalently, an effective threshold flux) as a low-dimensional function of the conditional mean state, rather than computing it from a full (v,τ)-resolved evolution. In the conductance-based refractory density approach of Chizhov and Graham, the hazard function is obtained by solving an auxiliary first-passage problem for subthreshold voltage fluctuations under a Gaussian approximation [42, 43].
For white noise, h is constructed from two analytically tractable limits (slowly varying and rapidly varying mean membrane potential) and interpolated between them; for colored noise, an analogous construction is performed using a quasi-stationary solution of the corresponding Kolmogorov–Fokker–Planck equation. In both settings the resulting hazard is expressed as a low-dimensional function of the mean membrane potential and its speed, with noise statistics entering through a small set of effective parameters.
A complementary perspective on non-stationary interval statistics is provided by the mapping approach of Schwalger [174], who derives an approximate *escape-noise* representation for a leaky integrate-and-fire neuron driven by *colored* input noise. There, the first-passage problem is reformulated as a level-crossing problem with a moving boundary, yielding a local-in-time hazard rate (link function) that depends on the instantaneous distance to threshold and its speed (and, in a second-order approximation, an additional low-dimensional memory variable accounting for correlations between level crossings).This construction is specifically designed to close renewal/refractory-density descriptions efficiently and to predict non-stationary ISI densities and population activity without solving a full time-dependent first-passage problem [174].
Our formulation addresses the same conceptual bottleneck: namely, that the classical one-dimensional

population density $p(v,t)$ discards the time-since-last-spike information required to resolve interval statistics, but does so at the level of the diffusive population equation. By lifting the dynamics to (v,τ) we obtain an age-resolved flux $\nu_q(\tau,t)$ and hence a time-dependent ISI density $\rho(\tau,t) = \nu_q(\tau,t)/\nu(t)$ directly, with the past drive entering only through the (possibly) time-dependent moments $\mu(t)$ and *diffusion intensity* $D(t)$. In particular, the explicit dependence on $D(t)$ allows one to treat genuinely non-stationary (heteroscedastic) diffusions, where transient changes in noise amplitude can modulate the interval statistics even at fixed mean drive, an effect that, in hazard-based reductions, typically requires an explicit re-derivation of the effective link function under time-dependent noise statistics.

A further advantage of the present setting is that it yields a non-stationary hazard $h(\tau,t)$ as an *output* (via the age-resolved threshold flux and the corresponding survival probability), which can be used as a principled reference to assess low-dimensional hazard closures: under identical inputs $\mu(t)$ and $D(t)$ one may directly compare $h(\tau,t)$ to approximate local-in-time link functions proposed in escape-noise mappings, such as those obtained by level-crossing theory in [174].

It is also useful to view (3.24) as a *generative* description rather than a proposal for direct numerical simulation: a full two-dimensional PDE is indeed expensive (and often impractical) to solve at the resolution required for network-scale applications. However, the extended formulation provides a systematic starting point for controlled reductions. In the next sections, we show how it yields closed evolution equations for ISI *moments* and how it enables analytical first-order perturbation theory.

3.3.2 Moment equations

The population density equation (3.24) can be simplified by taking the Laplace transform with respect to the age variable τ. The time evolution of the Laplace transform $\tilde{q}(v,s,t) = \int_{0^-}^{\infty} e^{-s\tau} q(v,\tau,t)\, d\tau$ for $s \in \mathbb{C}$ is governed by:

$$\partial_t \tilde{q}(v,s,t) = \mathcal{L}^A_{\boldsymbol{\gamma}(t)}\, \tilde{q}(v,s,t) - s\tilde{q}(v,s,t) + \delta_H(v)\nu(t). \tag{3.35}$$

A key feature of Eq. (3.35) is the decoupling of frequencies: for each fixed Laplace variable s, the evolution of $\tilde{q}(v,s,t)$ is independent of other values $s' \neq s$, with the population firing rate $\nu(t)$ acting as the sole external drive.

By performing a Taylor expansion of Eq. (3.35) around $s = 0$, we derive a recursive hierarchy of one-dimensional partial differential equations for the unnormalized partial moments $T_n(v,t) := \int_{0^-}^{\infty} \tau^n q(v,\tau,t)\, d\tau$:

$$\partial_t T_0(v,t) = \mathcal{L}^A_{\boldsymbol{\gamma}(t)}\, T_0(v,t) + \delta_H(v)\nu(t), \tag{3.36}$$

$$\partial_t T_n(v,t) = \mathcal{L}^A_{\boldsymbol{\gamma}(t)}\, T_n(v,t) + nT_{n-1}(v,t), \quad \forall n \geq 1, \tag{3.37}$$

subject to the absorbing boundary condition at the threshold, $T_n(\theta,t) = 0$, and the reflecting boundary condition at the lower bound, $\partial_v T_n(\alpha,t) = 0$. Note that $T_0(v,t)$ is identical to the standard voltage probability density $p(v,t)$.

The flux of $T_n(\cdot,t)$ across the threshold yields the higher-order statistics of the firing intervals.

Specifically, the n-th moment of the ISI distribution for neurons spiking at time t is given by:

$$m_n(t) := \mathbb{E}_{\tau\sim\rho(\cdot,t)}[\tau^n] = \frac{-D(t)\partial_v T_n(v,t)\big|_{v=\theta}}{\nu(t)} = \frac{\nu_{T_n}(t)}{\nu(t)}. \tag{3.38}$$

This hierarchy constitutes one of the central results of this thesis: it reduces the computation of time-dependent ISI moments (even far from stationarity) to solving a sequence of coupled, one-dimensional PDEs, avoiding the computational complexity of the full age-structured population density equation.

3.4 Stationary state

We begin by analyzing the stationary solution $q^0(v,\tau)$ of the population density equation (3.24). In the stationary regime, the firing rate is constant and denoted by ν_0, and the density factorizes as:

$$q^0(v,\tau) = \nu_0\, g^0(v,\tau). \tag{3.39}$$

The function g^0 satisfies a time-homogeneous renewal equation in the variables (v,τ):

$$\partial_\tau g^0(v,\tau) = \mathcal{L}_0^A\, g^0(v,\tau) + \delta_0(\tau)\,\delta_H(v). \tag{3.40}$$

Taking the Laplace transform of g^0 with respect to τ yields:

$$(s - \mathcal{L}_0^A)\,\tilde{g}^0(v,s) = \delta_H(v), \tag{3.41}$$

$$\Rightarrow \quad \tilde{g}^0(v,s) = G^{\mathrm{A}}(v,H;s), \tag{3.42}$$

where G^{A} is the Green's function for the absorbing problem defined previously. Evaluating the flux of $\tilde{g}^0$ at the threshold provides the stationary ISI distribution:

$$\nu_{\tilde{g}^0}(s) = \tilde{\rho}_0(s). \tag{3.43}$$

Thus, as expected, the flux of g^0 through the threshold recovers the Laplace-transformed stationary ISI density.

Stationary moment equations

We now turn to the stationary hierarchy of ISI moments obtained from Eq. (3.36). Let $T_n^0(v)$ denote the n-th stationary marginal moment. These satisfy the recursive system:

$$T_0^0(v) = \phi_0(v), \tag{3.44}$$

$$-\mathcal{L}_0^A\, T_n^0(v) = n\, T_{n-1}^0(v), \tag{3.45}$$

where $\phi_0(v)$ is the stationary voltage distribution. Solving this hierarchy using the Green's function G^{A} yields the integral representation:

$$\begin{aligned} T_n^0(v) &= n \int_\alpha^\theta G^{\mathrm{A}}(v, y; 0)\, T_{n-1}^0(y)\, dy \\ &= \frac{n}{D_0}\, w(v) \int_\alpha^\theta T_{n-1}^0(y) \left[\int_{v \vee y}^\theta w^{-1}(z)\, dz \right] dy \\ &= \frac{n}{D_0}\, w(v) \int_v^\theta w^{-1}(y) \left[\int_\alpha^y T_{n-1}^0(z)\, dz \right] dy, \end{aligned} \tag{3.46}$$

where w denotes the Wronskian associated with the homogeneous operator $\mathcal{L}_0$. Applying the boundary flux operator $-D_0\, \partial_v|_{v=\theta}$ to Eq. (3.46) yields an explicit formula for the stationary ISI moments:

$$m_n^0 = \frac{\nu_{T_n^0}}{\nu_0} = \frac{n}{\nu_0} \int_\alpha^\theta T_{n-1}^0(y)\, dy. \tag{3.47}$$

This provides a compact integral characterization of all stationary ISI moments in terms of solutions of the recursive hierarchy for T_n^0.

3.4.1 Connection with the backward equation approach

Historically, the ISI distribution has been studied using a "forward-in-time" approach: for a neuron starting at a given state v, one computes the distribution of the first passage time to a threshold θ using the backward Kolmogorov operator $\mathcal{L}_0^+$. This leads to the Siegert relations for the moments $m_n^0(v)$ of the waiting time until the next spike.

In a non-stationary setting, this forward-looking approach requires knowledge of the future input current, making it difficult to extend to systems, such as coupled neural networks, where the input is determined self-consistently. In contrast, our framework adopts a "backward-in-time" perspective, describing the inter-spike interval distribution of the neurons currently firing based on their history.

We now show that for stationary currents, these two formal descriptions are equivalent. While $m_n^0(v)$ describes the n-th moment of the time *until* the next spike for a neuron at v, the function $T_n^0(v)$ is related to the n-th moment of the time *since* the last spike. Under stationarity, these quantities satisfy the following relationship:

$$\frac{1}{\nu_0} \int_\alpha^\theta T_{n-1}^0(v)\, dv = \frac{m_n^0}{n} = \frac{1}{\nu_0} \int_\alpha^\theta m_{n-1}^0(v) \phi_0(v)\, dv. \tag{3.48}$$

To see this, we start from (3.21) and integrate by parts:

$$\begin{aligned} n \int_\alpha^\theta m_{n-1}^0(v) \phi_0(v)\, dv &= -\int_\alpha^\theta \mathcal{L}_0^+ m_n^0(v) \phi_0(v)\, dv \\ &= \nu_0 m_n^0(H) = \nu_0 m_n^0 \\ &= n \int_\alpha^\theta T_{n-1}^0(v)\, dv, \end{aligned}$$

where the second equality uses $\mathcal{L}_0\phi_0 = -\nu_0\delta_H$ and the boundary conditions of m_n^0, and the last equality follows from Eq. (3.47). Similarly, we can establish the relationship between the densities:

$$\int_\alpha^\theta q^0(v,\tau)\,dv = r^0(\tau) = \nu_0 \mathbf{L}^{-1}\left\{\frac{1-\tilde{\rho}_0(s)}{s}\right\}(\tau)$$

where $\mathbf{L}^{-1}$ denotes the inverse Laplace transform and $r^0(\tau)$ is the stationary refractory density of Eq. (3.31); the right-hand side is ν_0 times the stationary survival function.

3.4.2 Relaxation to stationary state

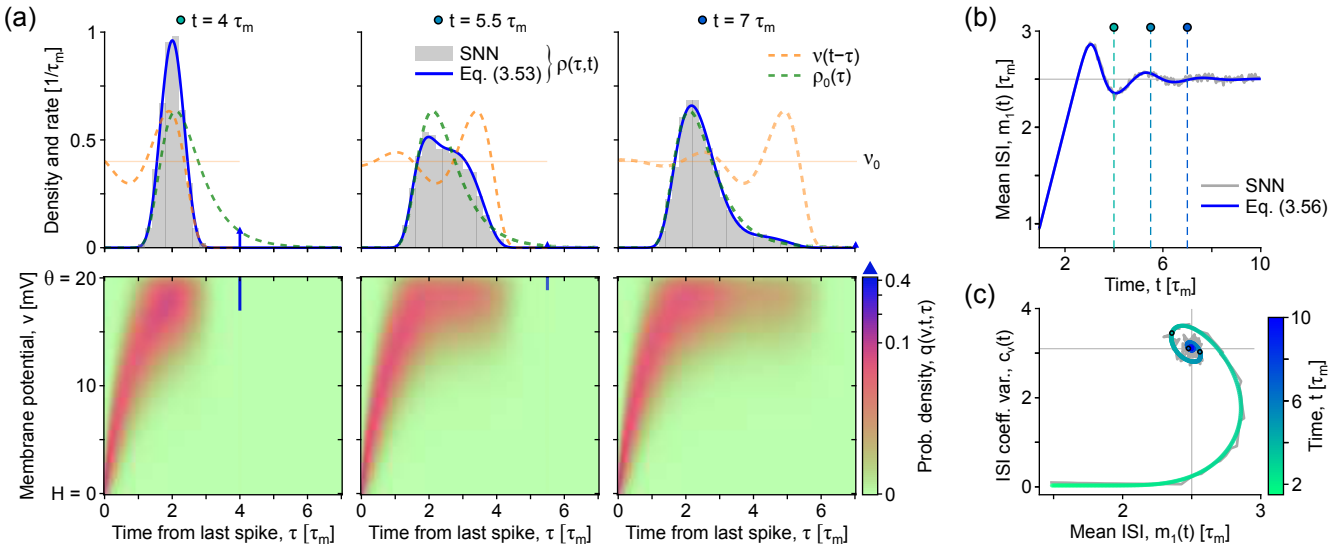

Figure 3.1: Relaxation to the stationary state in drift-dominated regime for a population of uncoupled leaky integrate-and-fire neurons (LIF, $b(v) = -v$): $\mu_0 = 21\,\mathrm{mV}/\tau_m$, $D_0 = 14.15\,\mathrm{mV}^2/\tau_m$, $H = 0\,\mathrm{mV}$, $\theta = 20\,\mathrm{mV}$. (a) Top: Time-dependent ISI distribution, comparing Eq. (3.54) with spiking neural network simulation (SNN $N = 10^4$ neurons). Bottom: Population density $q(v,\tau,t)$. (b) Temporal dynamics of mean ISI, comparison between theory and SNN simulation. (c) Temporal trajectory of mean ISI and coefficient of variation. Comparison between theory and SNN simulation.

We analytically solve the relaxation dynamics of an uncoupled neuronal population starting from an arbitrary initial state. For constant input moments (μ_0, D_0), the population dynamics in Eq. (3.24) can be solved explicitly. We first consider the elementary initial condition:

$$q(v,\tau,0) = \ell(v,\tau,0) = \delta_y(v)\,\delta_u(\tau). \tag{3.49}$$

Using the decomposition (3.28), the two flux contributions read:

$$\nu_\ell(\tau,t) = \delta_{u+t}(\tau)\,\rho_0(t|y), \tag{3.50}$$

$$\nu_g(\tau,t) = \rho_0(\tau), \qquad \tau < t, \tag{3.51}$$

from which the relaxation ISI distribution follows as:

$$\rho(\tau,t) = \frac{1}{\nu(t)} \begin{cases} \rho_0(\tau)\,\nu(t-\tau), & \tau < t, \\ \delta_{u+t}(\tau)\,\rho_0(t|y), & \tau \geq t. \end{cases} \tag{3.52}$$

For a generic initial condition $q(v,\tau,0) = q_*(v,\tau)$, linearity allows us to generalize this result:

$$\rho(\tau,t) = \frac{1}{\nu(t)} \begin{cases} \rho_0(\tau)\,\nu(t-\tau), & \tau < t, \\ \displaystyle\int_\alpha^\theta \rho_0(t|y)\,q_*(y,\tau-t)\,dy, & \tau \geq t. \end{cases} \tag{3.53}$$

We now focus on the specific case $(y,u) = (H,0)$, which corresponds to the relaxation of the population immediately after all neurons are synchronously reset to the potential $v = H$, for instance, following an externally triggered burst of activity. In this scenario, the time-dependent ISI distribution simplifies to:

$$\begin{aligned} \rho(\tau,t) &= \frac{\nu_q(\tau,t)}{\nu(t)} \\ &= \frac{1}{\nu(t)}\left(\rho_0(\tau)\nu(t-\tau) + \delta_t(\tau)\rho_0(\tau)\right), \end{aligned} \tag{3.54}$$

where $\rho_0(\tau) := \rho_0(\tau|H)$ is the stationary ISI distribution from renewal theory. Taking the double (unilateral) Laplace transform with respect to t (variable k) and τ (variable s); throughout, k denotes a unilateral and κ a bilateral transform variable in time, we obtain:

$$\nu_{\widehat{q}}(s,k) = \int_0^\infty \int_0^\infty e^{-s\tau - kt} \nu_q(\tau,t)\,d\tau\,dt = \frac{\tilde{\rho}_0(s+k)}{1-\tilde{\rho}_0(k)}. \tag{3.55}$$

As a function of k, this expression possesses poles at $\{k = \lambda_i\}$, with $\tilde{\rho}_0(\lambda_i) = 1$, and at $\{k = \mu_i - s\}$, where the μ_i are the poles of $\tilde{\rho}_0$, equivalently the roots of $\psi_1(\theta;\mu_i) = 0$. These correspond, respectively, to the eigenvalues of the Fokker–Planck operator under reinjection and under absorbing boundary conditions.

From Eq. (3.36), we also obtain a closed-form expression for the relaxation of the moments:

$$\nu_{\widehat{T_n}}(k) = (-1)^n \frac{\tilde{\rho}_0^{(n)}(k)}{1-\tilde{\rho}_0(k)}. \tag{3.56}$$

By evaluating the principal part of this expression at the poles and inverting the Laplace transform, we find a residue (spectral) expansion for the relaxation of the moment fluxes ν_{T_n}. The explicit expressions for the first two moments are:

$$m_1(t) = \frac{1}{\nu(t)} \sum_i \left(e^{\lambda_i t} - e^{\mu_i t}\right) \xrightarrow{t\to\infty} \frac{1}{\nu_0}, \tag{3.57}$$

$$\begin{aligned} m_2(t) = {} & \frac{2}{\nu(t)} \sum_i \left[\mathrm{Res}_{k=\lambda_i} \frac{\tilde{\rho}_0(k)}{(k-\lambda_i)^3} \mathrm{Res}_{k=\lambda_i} \frac{1}{1-\tilde{\rho}_0(k)} \right] e^{\lambda_i t} \\ & - \frac{2}{\nu(t)} \sum_i \left[t + \left(1 - \mathrm{Res}_{k=\mu_i} \frac{\tilde{\rho}_0(k)}{(k-\mu_i)}\right) \left(\mathrm{Res}_{k=\mu_i} \tilde{\rho}_0(k)\right)^{-1} \right] e^{\mu_i t}. \end{aligned}$$

3.5 Linear Response theory

Next, we analyse first-order perturbations of the stationary state in the frequency domain. We consider a small time-dependent modulation of the external current

$$(\mu(t), D(t)) = (\mu_0, D_0) + (\mu_1, D_1)\varepsilon(t).$$

In this section we derive the linear-response theory in the simplifying limit $\alpha \to -\infty$ and for a confining potential; in the next chapter (Section 4.7) we show how to treat the case of finite α in a fully rigorous way.
We can write the FP operator $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1\varepsilon(t)$ where $\mathcal{L}_1 = -\mu_1\partial_v + D_1\partial_v^2$. Whenever the perturbation is transferred onto a backward quantity by integration by parts, it acts through its formal adjoint $\mathcal{L}_1^* = \mu_1\partial_v + D_1\partial_v^2$. We start by considering the first order response of the function $g(v,\tau,t) = g^0(v,\tau) + g^1(v,\tau,t) + o(\varepsilon)$. From (3.29) we obtain:

$$\partial_t g^1(v,\tau,t) = \mathcal{L}_0^A g^1(v,\tau,t) - \partial_\tau g^1(v,\tau,t) + \varepsilon(t)\mathcal{L}_1 g^0(v,\tau),$$

Then, considering the Laplace transform in t and τ:

$$\left(s + \kappa - \mathcal{L}_0^A\right)\widehat{g}^1(v,s,\kappa) = \widehat{\varepsilon}(\kappa)\mathcal{L}_1\tilde{g}^0(v,s) \tag{3.58}$$

Which can be solved using the Green function G^A:

$$\begin{aligned}\mathrm{H}^{\mathrm{g}}(s,\kappa) =& \frac{\nu_{\widehat{g}^1}(s,\kappa)}{\widehat{\varepsilon}(\kappa)} = \int_\alpha^\theta \mathcal{L}_1^*\tilde{\rho}_0(s+\kappa|y)G^{\mathrm{A}}(y,H;s)dy = \\ =& \int_\alpha^\theta \frac{\mathcal{L}_1^*\psi_1(y;s+\kappa)}{\psi_1(\theta;s+\kappa)}G^{\mathrm{A}}(y,H;s)dy\end{aligned} \tag{3.59}$$

From this, we can get the response function for the flux ν_{q^1}:

$$\mathrm{H}^{\mathrm{q}}(s,\kappa) = \nu_0\mathrm{H}^{\mathrm{g}}(s,\kappa) + \tilde{\rho}_0(s+\kappa)\mathrm{H}^{\nu}(\kappa) \tag{3.60}$$

Where $\mathrm{H}^{\nu}(\kappa) = \nu^1(\kappa)/\varepsilon(\kappa)$ is the standard linear response function for the population firing rate. Evaluating Equation (3.60) at $s = 0$ we get an equivalent expression for $\mathrm{H}^{\nu}(\kappa)$:

$$\mathrm{H}^{\nu}(\kappa) = \int_\alpha^\theta \frac{\mathcal{L}_1^*\psi_1(y;\kappa)}{\psi_1(\theta;\kappa) - \psi_1(H;\kappa)}\phi_0(y)dy \tag{3.61}$$

Finally as $\nu_q = \nu_0\rho_0 + \nu_0\rho^1 + \nu^1\rho_0 + o(\varepsilon)$, the linear response of the ISI distribution is given by:

$$\mathrm{H}^{\rho}(s,\kappa) = \mathrm{H}^{\mathrm{g}}(s,\kappa) + \frac{1}{\nu_0}\mathrm{H}^{\nu}(\kappa)\left[\tilde{\rho}_0(s+\kappa) - \tilde{\rho}_0(s)\right]. \tag{3.62}$$

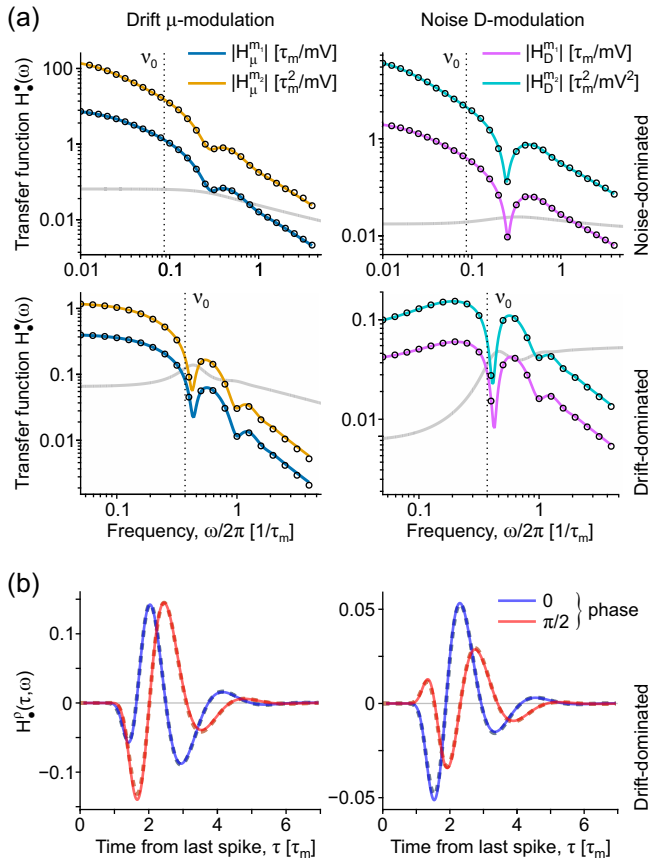

Figure 3.2: Linear response and transfer functions of first ISI moments. (a) Transfer function for the first two moments. Top: drift-dominated regime (same parameters as Fig. 3.1) and noise-dominated regime ($\mu_0 = 16.314\,\mathrm{mV}/\tau_m$, $D_0 = 14.25\,\mathrm{mV}^2/\tau_m$, $H = 10\,\mathrm{mV}$, $\theta = 20\,\mathrm{mV}$). Comparison between theory (solid lines) and moment hierarchy numerical integration (3.36). Solid grey line represents the firing rate transfer function H^ν. (b) Linear response theory for the ISI distribution in drift-dominated regime, $\omega/2\pi = 0.4/\tau_m$. Left: drift modulation, right: noise modulation. Comparison between theory and numerical integration of (3.24).

3.5.1 Linear response from the moment hierarchy

To derive the linear response of the ISI moments, we expand the moment hierarchy equations (3.36) to first order in the input modulation. Let

$$T_n(v,t) = T_n^0(v) + T_n^1(v,t) + o(\varepsilon)$$

denote the perturbed marginal moments, with $T_n^1 = O(\varepsilon)$ in the same convention used for g^1 above, and let $\widehat{T}_n^1(v,\kappa)$ be their bilateral Laplace transforms in time. The corresponding linear response of the ISI moments follows from the identity

$$\mathrm{H}^{m_n}(\kappa) = (-1)^n\, \partial_s^n \mathrm{H}^\rho(s,\kappa)\big|_{s=0}, \tag{3.63}$$

which links the moment response function to the ISI response function.
The linearized equations for the perturbations $\widehat{T}_n^1$ are obtained by applying the absorbing Green's function G^{A} of the stationary operator and collecting all first-order terms. For the zero-th element

of the hierarchy we obtain

$$\frac{\widehat{T}_0^1(v,\kappa)}{\widehat{\varepsilon}(\kappa)} = \int_\alpha^\theta G^{\mathrm{A}}(v,y;\kappa)\,\mathcal{L}_1\phi_0(y)\,dy + G^{\mathrm{A}}(v,H;\kappa)\,\mathrm{H}^\nu(\kappa), \tag{3.64}$$

For $n \geq 1$ the hierarchy propagates through the recursion

$$\frac{\widehat{T}_n^1(v,\kappa)}{\widehat{\varepsilon}(\kappa)} = \int_\alpha^\theta G^{\mathrm{A}}(v,y;\kappa)\,\mathcal{L}_1 T_n^0(y)\,dy + n\int_\alpha^\theta G^{\mathrm{A}}(v,y;\kappa)\,\frac{\widehat{T}_{n-1}^1(y,\kappa)}{\widehat{\varepsilon}(\kappa)}\,dy. \tag{3.65}$$

Substituting these expressions into the definition of the ISI moments yields the linear moment response function

$$\mathrm{H}^{m_n}(\kappa) = \frac{1}{\nu_0}\int_\alpha^\theta \mathcal{L}_1^*\tilde{\rho}_0(\kappa|y)\,T_n^0(y)\,dy + \frac{n}{\nu_0}\int_\alpha^\theta \tilde{\rho}_0(\kappa|y)\,\frac{\widehat{T}_{n-1}^1(y,\kappa)}{\widehat{\varepsilon}(\kappa)}\,dy - \frac{m_n^0}{\nu_0}\,\mathrm{H}^\nu(\kappa), \tag{3.66}$$

which provides a closed expression for the frequency-domain linear response of the ISI moments in terms of the stationary hierarchy, the stationary ISI distribution, and the perturbation operator $\mathcal{L}_1$.

Fig. 3.2 illustrates the linear response properties predicted by the theory. Panel (a) shows the transfer functions of the first two ISI moments in both a drift–dominated and a noise–dominated regime. In each case, the theoretical prediction (solid lines) closely matches the result obtained by numerically integrating the moment hierarchy (3.36). Panel (b) compares the full ISI response function H^ρ (solid lines) with numerical integration of the population equation (3.24) (dashed lines) for a representative modulation frequency ($\omega/2\pi = 0.4/\tau_m$). Both drift and noise modulations are shown. In each case, the linear theory accurately reproduces the phase and amplitude of the ISI perturbation, confirming that the response of the full ISI density is correctly captured.

3.6 Nonlinear dynamics

Finally, we validate our theoretical results through direct numerical simulations, showing that the proposed population equation accurately captures the time evolution of the ISI distribution even when the activity is far from the stationary state. Here, we simulate a coupled population of neurons with deterministic synaptic efficacies J and a distribution of delays in the axonal transmission of spikes $p_\delta(x) = \tau_d^{-1}\,\mathbf{1}(x-\delta)\exp(-(x-\delta)/\tau_d)$, where $\mathbf{1}$ denotes the indicator function. Under mean-field approximation this leads the following infinitesimal mean and variance of the synaptic current

$$\mu(t) = KJ(p_\delta * \nu)(t) + \mu_{\mathrm{ext}} \tag{3.67}$$

$$D(t) = \tfrac{1}{2}KJ^2(p_\delta * \nu)(t) + D_{\mathrm{ext}}\,, \tag{3.68}$$

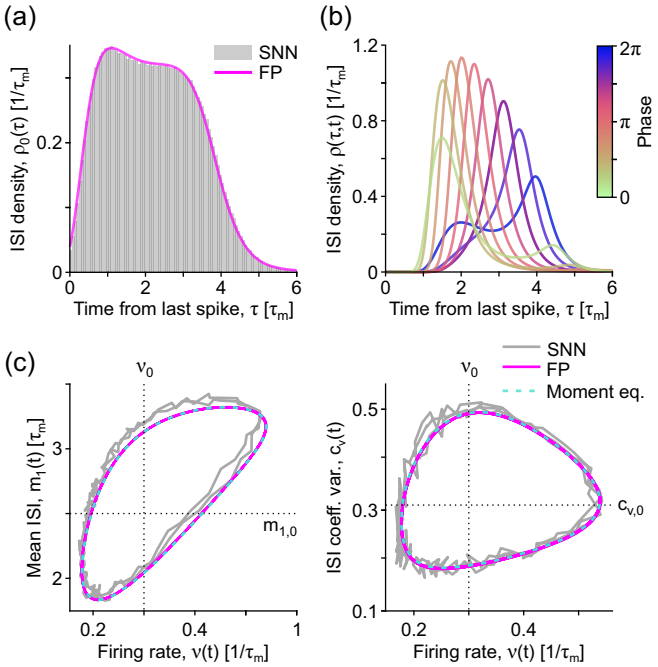

Figure 3.3: Synchronous (limit cycle) regular state in an excitatory population of LIF neurons. (a) ISI distribution $\bar{\rho}(\tau) = \langle \nu_q(\tau, t) \rangle_t / \langle \nu(t) \rangle_t$. Comparison between population dynamics, numerical integration of (3.24) and SNN simulation ($N = 5 \cdot 10^4$). (b) Phase-dependent ISI distribution. (c) Joint trajectory of firing rate, ISI mean (left) and coefficient of variation (right). Comparison between population dynamics (3.24), moment hierarchy from (3.36) and SNN simulation ($N = 5 \cdot 10^4$). Parameters: $J = 0.1\,\mathrm{mV}$, $K = 1000$, $\mu_{\mathrm{ext}} = 21\,\mathrm{mV}/\tau_m$, $D_{\mathrm{ext}} = 1.33\,\mathrm{mV}^2/\tau_m$, $H = 0\,\mathrm{mV}$, $\theta = 20\,\mathrm{mV}$, $\delta = 0.05\,\tau_m$, $\tau_d = 0.1\tau_m$.

where K is the number of presynaptic contacts.

We focus on two representative regimes in which the population converges to a stable limit cycle: an excitatory synchronous–regular oscillation, where all neurons fire on every cycle, and an inhibitory synchronous–irregular oscillation, where only a fraction of the neurons fire on each cycle. In both cases, we compare the predictions of the population density (3.24) and the moment hierarchy (3.36) with large-scale spiking-network (SNN) simulations.

Fig. 3.3 illustrates the excitatory limit cycle. Panel (a) compares the ISI distribution obtained from long-time averages with the prediction of the population equation and with SNN simulations ($N = 5 \cdot 10^4$). The ISI distribution reflects a fully synchronous population in which all neurons fire on every cycle but at different phases. Panel (b) shows the phase-resolved ISI distribution, revealing the temporal organisation of spikes within the cycle. Panel (c) tracks the firing rate, mean ISI, and coefficient of variation over a full period. Both the population equation and the moment hierarchy match the microscopic dynamics, confirming that the ISI evolution is accurately captured even far from stationarity.

The inhibitory limit cycle, shown in Fig. 3.4, displays qualitatively different ISI statistics. Because inhibition prevents a subset of neurons from firing on each cycle, the ISI distribution becomes highly structured and multimodal, with peaks corresponding to different numbers of skipped cycles. Panel (a) shows that the population equation reproduces this complex structure with high accuracy. Panels (b) and (c) track the evolution of the mean ISI and the squared coefficient of variation, again showing excellent agreement between the population equation, the moment hierarchy, and SNN simulations.

Overall, these simulations confirm that the population dynamics equation reliably describes ISI

statistics in regimes where renewal-based approximations break down and the population is far from the stationary state. Moreover, the moment hierarchy offers an efficient complementary description of ISI statistics, accurately capturing moment dynamics at a computational cost comparable to integrating the one-dimensional Fokker–Planck equation.

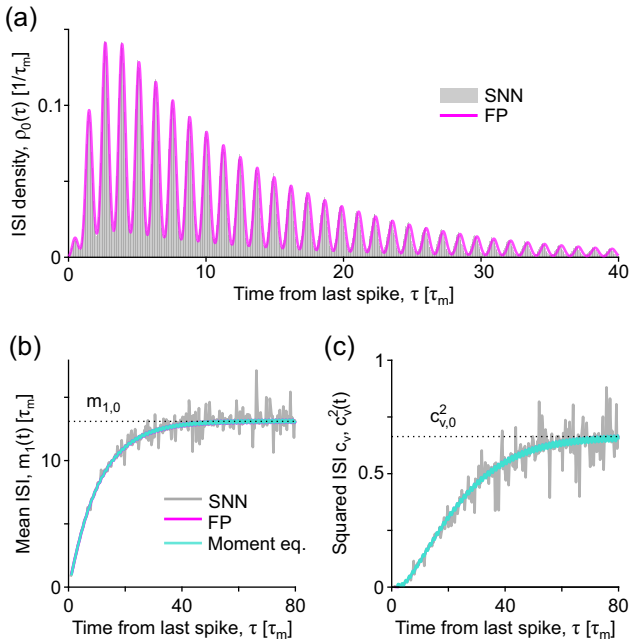

Figure 3.4: Synchronous (limit cycle) irregular state in an inhibitory population of LIF neurons. (a) ISI distribution $\bar{\rho}(\tau) = \langle \nu_q(\tau, t)\rangle_t / \langle \nu(t)\rangle_t$. Comparison between population dynamics, numerical integration of (3.24) and SNN simulation ($N = 5 \cdot 10^4$). (b) Time trajectory of mean ISI. Comparison between population dynamics (3.24), moment hierarchy (3.36) and SNN simulation ($N = 5 \cdot 10^4$). (c) Same as (b) but tracking squared coefficient of variation. Parameters: $J = -0.1\,\mathrm{mV}$, $K = 1000$, $D_{\mathrm{ext}} = 8\,\mathrm{mV}^2/\tau_m$, $\delta = 0.4\,\tau_m$, $\tau_d = 0\,\tau_m$. Other parameters as in Fig. 3.3.

3.7 Discussion

In summary, we developed an analytic framework for characterizing the time evolution of inter-spike interval (ISI) distributions in non-stationary neuronal populations. Starting from the population density equation, we formulated a two-dimensional Fokker–Planck description that jointly tracks membrane voltage and time since the last spike, providing direct access to the ISI distribution under arbitrarily time-dependent inputs.

Numerical integration of this equation shows excellent agreement with large-scale spiking-network simulations, including regimes that strongly depart from stationarity, such as recurrently generated limit cycles. This demonstrates that the framework captures ISI dynamics beyond the reach of classical renewal-based approaches.

Building on this formulation, we derived a one-dimensional hierarchy governing the moments of the ISI distribution and performed a first-order perturbative expansion around the stationary state. This yields a compact analytic expression for the linear response of ISI statistics to weak temporal modulation, offering a reduced yet accurate description of ISI dynamics.

More broadly, the methodology provides a tractable route to analyzing threshold-crossing statistics in driven excitable systems, extending its relevance beyond neuronal models to a wider class of

first-passage problems in statistical physics [123].

Chapter 4

Spectral theory for population dynamics with absolute refractory period

> Questa riconciliazione [...] tradisce la profonda perversione morale che appartiene a un mondo fondato essenzialmente sull'inesistenza del ritorno, perché in un mondo simile tutto è già perdonato e quindi tutto è cinicamente permesso
>
> — Milan Kundera, *L'insostenibile leggerezza dell'essere*

The mathematical analysis of neural population dynamics relies centrally on the Fokker-Planck equation to describe the evolution of the probability density. However, the inclusion of an absolute refractory period τ_0 fundamentally alters the mathematical structure of the problem. Unlike synaptic or axonal transmission delays, which appear as retarded arguments in the interaction terms, the refractory period introduces a delay directly into the boundary conditions of the probability flow: the flux of neurons re-entering the voltage domain at the reset potential depends on the flux that exited at the threshold τ_0 time units earlier. This dependency renders the standard Fokker-Planck equation non-Markovian.

In this chapter, we develop [59] a rigorous spectral theory for the Fokker-Planck operator with delayed reinjection, providing a systematic framework to analyze the stability and response properties of such populations. We begin in Section 4.1 by restoring the Markovian property of the system. By augmenting the state space with a dedicated refractory compartment governed by a transport equation, we map the history-dependent boundary condition onto a local coupling between two partial differential equations. This formulation allows us to define the evolution operator $\mathcal{T}_\gamma$ on a suitable product Banach space.

In Section 4.2, we provide a complete characterization of the spectrum using the rigourous theory of boundary eigenvalue operator functions[131] (or operator pencils). We derive the exact characteristic equation for the eigenvalues and construct the resolvent operator. A key theoretical contribution of this work is the rigorous treatment of *defective eigenvalues* (Section 4.5). We identify these spectral singularities, where the algebraic multiplicity exceeds the geometric multiplicity, as critical transition points in the system's dynamics, necessitating the use of generalized eigenfunctions to construct a complete basis.

To ensure the physical consistency of our framework, in Section 4.6 we prove that the evolution operator generates a strongly continuous contraction semigroup. This is achieved by establishing the dissipativity of the generator in $\mathbb{L}^1$, guaranteeing probability conservation and the existence of unique solutions for the Cauchy problem.
Finally, we apply this spectral machinery to linear response theory. In Section 4.7, we develop a first-order perturbation theory that explicitly accounts for the modulation of boundary conditions by time-dependent external parameters. This yields a generalized transfer function that captures the interplay between refractoriness and noise, and allows us to rigorously analyze stability and bifurcation phenomena, such as the emergence of limit cycles, in coupled networks.

4.1 Augmenting the state space

The primary obstacle to extending the spectral decomposition to networks with a finite absolute refractory period $\tau_0 > 0$ lies in the delay-induced non-Markovianity of the boundary dynamics. As established in the Introduction (see Eq. (4.4d)), the reinjection flux at the reset potential H depends on the firing activity at time $t - \tau_0$. Consequently, the instantaneous probability density $p_t(v)$ alone is insufficient to determine the future evolution of the system; the history of the firing rate must also be tracked.
To recover a Markovian description, we must account for the population of neurons currently in the refractory state. We augment the state space by introducing a *refractory density* function $p_t^r(\tau)$, defined for $\tau \in (0, \tau_0)$, which represents the density of neurons that fired τ time units ago. Since neurons proceed deterministically through the refractory period, this density is simply the transported history of the firing rate:

$$p_t^r(\tau) := \nu(t - \tau), \quad \forall \tau \in (0, \tau_0). \tag{4.1}$$

The dynamical evolution of this variable is governed by a linear transport equation:

$$\partial_t p_t^r(\tau) + \partial_\tau p_t^r(\tau) = 0. \tag{4.2}$$

The full state of the population at time t is thus described by the pair $\mathbf{p}_t = (p_t, p_t^r)$. The evolution of this augmented state is governed by a system of two coupled partial differential equations:

$$\begin{cases} \partial_t p_t(v) = \mathcal{L}_{\boldsymbol{\gamma}(t)} p_t(v) \equiv -\partial_v \mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](v) & v \in (\alpha, H) \cup (H, \theta) \\ \partial_t p_t^r(\tau) = -\partial_\tau p_t^r(\tau) & \tau \in (0, \tau_0) \end{cases} \tag{4.3}$$

where $\mathcal{L}_{\boldsymbol{\gamma}(t)}$ is the standard Fokker-Planck operator defined in Eq. (2.14) with drift $b(v) + \mu(t)$ and diffusion $D(t)$.
The coupling between the membrane potential density p_t and the refractory density p_t^r is mediated entirely through the boundary conditions. Specifically, the flux leaving the potential domain at the threshold enters the refractory domain at $\tau = 0$, and the flux leaving the refractory domain at $\tau = \tau_0$ is reinjected into the potential domain at the reset H. Formally, the boundary conditions

for the augmented system are:

$$p_t(\theta) = 0 \tag{4.4a}$$
$$p_t(H^+) - p_t(H^-) = 0 \tag{4.4b}$$
$$\mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](H^+) - \mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](H^-) = p_t^r(\tau_0) \tag{4.4c}$$
$$p_t^r(0) = \mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](\theta) \tag{4.4d}$$
$$\mathcal{S}_{\boldsymbol{\gamma}(t)}[p_t](\alpha) = 0 \tag{4.4e}$$

Equations (4.4c) and (4.4d) represent the closure of the probability loop: (4.4d) states that the source of refractory neurons is the instantaneous firing rate $\nu(t)$, while (4.4c) states that the source of active neurons at the reset potential is the density of neurons exiting the refractory period $\nu(t-\tau_0)$. Crucially, with this augmentation, the boundary conditions become local in the new state variable $\mathbf{p}_t$, rendering the system mathematically homogeneous and amenable to spectral analysis.

4.1.1 Definition of the evolution operator

The coupled system of partial differential equations derived in Section 4.1 can be formulated abstractly as an evolution equation on a Banach space. The dynamics of the augmented state vector $\mathbf{p}_t = (p_t, p_t^r)$ are governed by:

$$\dot{\mathbf{p}}_t = \mathcal{T}_{\boldsymbol{\gamma}(t)}\mathbf{p}_t, \tag{4.5}$$

where $\mathcal{T}_{\boldsymbol{\gamma}(t)}$ is a linear operator encapsulating both the Fokker-Planck dynamics and the refractory transport.

Functional Setting

We construct our function spaces over the disjoint union of the potential domain and the refractory interval. Due to the singularity at the reset potential H, we partition the voltage domain into sub-intervals (α, H) and (H, θ). The relevant L^p space is defined as:

$$\begin{aligned}\mathbb{L}^p &:= \{\mathbf{p} = (p^-, p^+, p^r) \in \mathrm{L}^p(\alpha, H) \times \mathrm{L}^p(H, \theta) \times \mathrm{L}^p(0, \tau_0)\} \\ &= \{\mathbf{p} = (p, p^r) \in \mathrm{L}^p(\alpha, \theta) \times \mathrm{L}^p(0, \tau_0)\}, \quad p \in [1, \infty)\end{aligned} \tag{4.6}$$

Vectors in this space are denoted $\mathbf{p} = (p^-, p^+, p^r)$. For brevity, when identifying $L^p(\alpha, H) \times L^p(H, \theta)$ with $L^p(\alpha, \theta)$, we write $\mathbf{p} = (p, p^r)$.
The duality pairing $\langle \cdot, \cdot \rangle : \mathbb{L}^q \times \mathbb{L}^p \to \mathbb{C}$ (with $p^{-1} + q^{-1} = 1$) is defined naturally as:

$$\begin{aligned}\langle \mathbf{f}, \mathbf{p} \rangle &:= \int_\alpha^\theta f(v)p(v)dv + \int_0^{\tau_0} f^r(\tau)p^r(\tau)d\tau \\ &= \int_\alpha^{H^-} f^-(v)p^-(v)dv + \int_{H^+}^\theta f^+(v)p^+(v)dv + \int_0^{\tau_0} f^r(\tau)p^r(\tau)d\tau\end{aligned} \tag{4.7}$$

Regularity properties are handled via the Sobolev space $\mathbb{W}^p$:

$$\mathbb{W}^p := \mathrm{W}^{2,p}(\alpha, H) \times \mathrm{W}^{2,p}(H, \theta) \times \mathrm{W}^{1,p}(0, \tau_0). \tag{4.8}$$

The Primal Operator

We are now in a position to formally define the evolution operator.

Definition 4.1.1 (Evolution Operator)**.** *Let $\mu \in \mathbb{R}$ and $D > 0$. The evolution operator $\mathcal{T}_{\boldsymbol{\gamma}}$ is the unbounded linear operator defined by:*

$$\begin{aligned} \mathcal{T}_{\boldsymbol{\gamma}} : {}^{p}\mathbb{D}_{\boldsymbol{\gamma}} \subset \mathbb{L}^p &\to \mathbb{L}^p \\ (p, p^r) &\mapsto (\mathcal{L}_{\boldsymbol{\gamma}} p, -\partial_\tau p^r) \end{aligned} \tag{4.9}$$

where $\mathcal{L}_{\boldsymbol{\gamma}}$ is the differential operator defined in Eq. (2.14). The domain ${}^{p}\mathbb{D}_{\boldsymbol{\gamma}}$ consists of functions in $\mathbb{W}^p$ that satisfy the boundary conditions derived in Eq. (4.4):

$${}^{p}\mathbb{D}_{\boldsymbol{\gamma}} := \{\mathbf{p} \in \mathbb{W}^p \mid \mathcal{B}_{\boldsymbol{\gamma}}\mathbf{p} = 0\} \tag{4.10}$$

where the boundary operator $\mathcal{B}_{\boldsymbol{\gamma}} : \mathbb{W}^p \to \mathbb{C}^5$ is given by:

$$\mathcal{B}_{\boldsymbol{\gamma}}\mathbf{p} = \begin{bmatrix} \mathcal{S}_{\boldsymbol{\gamma}}[p^-](\alpha) \\ p^+(\theta) \\ p^+(H) - p^-(H) \\ \mathcal{S}_{\boldsymbol{\gamma}}[p^+](H) - \mathcal{S}_{\boldsymbol{\gamma}}[p^-](H) - p^r(\tau_0) \\ \mathcal{S}_{\boldsymbol{\gamma}}[p^+](\theta) - p^r(0) \end{bmatrix}. \tag{4.11}$$

Remark 4.1.1. *The operator $\mathcal{T}_{\boldsymbol{\gamma}}$ is closed and densely defined for all $p \in [1, \infty)$.*

Proof. Density follows from the fact that smooth functions with compact support away from the boundaries subset of the domain:

$$\mathrm{C}_c^\infty\big((\alpha, \theta)\big) \times \mathrm{C}_c^\infty(0, \tau_0) \subset {}^{p}\mathbb{D}_{\boldsymbol{\gamma}}.$$

This set is dense in $\mathbb{L}^p$. The closedness of the operator is a direct consequence of the maximal regularity of differential operators with well-posed boundary conditions (see, e.g., [131, Remark 3.4.2]). □

It is useful to explicitly define the "firing rate" functional $\mathcal{N} : \mathbb{W}^p \to \mathbb{C}$ as $\mathcal{N}[\mathbf{p}] = p^r(0)$. By the trace theorem, this operator is continuous from $\mathbb{W}^p$ to $\mathbb{C}$, and thus compact when viewed as an operator on $\mathbb{L}^p$. For any state $\mathbf{p}$ in the domain ${}^{p}\mathbb{D}_{\boldsymbol{\gamma}}$, the boundary condition ensures that $\mathcal{N}[\mathbf{p}] = \mathcal{S}_{\boldsymbol{\gamma}}[p](\theta)$, recovering the standard physical definition of the firing rate as the flux at the threshold.

The Adjoint Operator

The spectral decomposition requires the adjoint basis, necessitating the characterization of the adjoint operator $\mathcal{T}_{\boldsymbol{\gamma}}^*$.

Theorem 4.1.1 (Adjoint Operator)**.** *The adjoint of the evolution operator, $\mathcal{T}_{\boldsymbol{\gamma}}^* : {}^{q}\mathbb{D}_{\boldsymbol{\gamma}}^* \subset \mathbb{L}^q \to \mathbb{L}^q$, is defined by:*

$$\mathcal{T}_{\boldsymbol{\gamma}}^* \mathbf{f} = \big(\mathcal{L}_{\boldsymbol{\gamma}}^* f, \partial_\tau f^r\big) \tag{4.12}$$

where $\mathcal{L}^*_{\boldsymbol{\gamma}} f = [b(v)+\mu]f' + Df''$ is the formal adjoint of the Fokker-Planck operator (the backward Kolmogorov operator). The domain is defined by the adjoint boundary conditions:

$$ {}^{q}\mathbb{D}^*_{\boldsymbol{\gamma}} := \{\mathbf{f} \in \mathbb{W}^q \mid \mathcal{B}^*\mathbf{f} = 0\} \tag{4.13} $$

where the adjoint boundary operator $\mathcal{B}^* : \mathbb{W}^q \to \mathbb{C}^5$ is:

$$ \mathcal{B}^*\mathbf{f} = \begin{bmatrix} f'(\alpha) \\ f'(H^+) - f'(H^-) \\ f(H^+) - f(H^-) \\ f(H) - f^r(\tau_0) \\ f(\theta) - f^r(0) \end{bmatrix}. \tag{4.14} $$

Proof. Consider $\mathbf{f} \in \mathbb{W}^q$ and $\mathbf{p} \in {}^{p}\mathbb{D}_{\boldsymbol{\gamma}}$. Integrating by parts, the duality pairing yields:

$$ \langle \mathbf{f}, \mathcal{T}_{\boldsymbol{\gamma}}\mathbf{p} \rangle = \langle \mathcal{T}^*_{\boldsymbol{\gamma}}\mathbf{f}, \mathbf{p} \rangle + [\mathcal{J}(\mathbf{f},\mathbf{p})]_\partial \tag{4.15} $$

where the boundary term $[\mathcal{J}]_\partial$ collects the surface contributions:

$$ [\mathcal{J}]_\partial = \left[-Df'p - f\mathcal{S}_{\boldsymbol{\gamma}}[p]\right]_\alpha^{H^-} + \left[-Df'p - f\mathcal{S}_{\boldsymbol{\gamma}}[p]\right]_{H^+}^{\theta} - \left[f^r p^r\right]_0^{\tau_0}. \tag{4.16} $$

Substituting the primal boundary conditions $\mathcal{B}_{\boldsymbol{\gamma}}\mathbf{p} = 0$, the surface term vanishes for all $\mathbf{p}$ if and only if $\mathcal{B}^*\mathbf{f} = 0$. □

Several key features of the adjoint problem are worth noting. First, the adjoint boundary conditions are independent of the system parameters $\boldsymbol{\gamma}$, unlike the primal case. Second, the conditions $f(H^+) = f(H^-)$ and $f'(H^+) = f'(H^-)$ imply that any function in the adjoint domain is continuous and has a continuous derivative across the reset potential H. Consequently, functions in ${}^{q}\mathbb{D}^*_{\boldsymbol{\gamma}}$ belong to $\mathrm{W}^{2,q}(\alpha,\theta)$ directly, removing the need for a split domain definition.
Finally, we note a subtle point regarding the domain density. While for $q < \infty$ the domain ${}^{q}\mathbb{D}^*_{\boldsymbol{\gamma}}$ is dense in $\mathbb{L}^q$, this property fails for $q = \infty$. This is a typical feature of adjoints on non-reflexive spaces like L^∞, where the domain of the adjoint is dense in the weak* topology but not necessarily in the norm topology [26, Corollary 2.10].

4.2 Spectrum of the evolution operator

The spectral properties of the evolution operator $\mathcal{T}_{\boldsymbol{\gamma}}$ can be rigorously analyzed using the theory of boundary eigenvalue problems, following the framework established in [131]. Our goal is to determine the values $\lambda \in \mathbb{C}$ for which the operator equation $\mathcal{T}_{\boldsymbol{\gamma}}\mathbf{p} = \lambda\mathbf{p}$ has non-trivial solutions in the domain ${}^{p}\mathbb{D}_{\boldsymbol{\gamma}}$.

4.2.1 Characteristic Matrix Formulation

We begin by defining the boundary eigenvalue operator function $\mathfrak{T}_{\boldsymbol{\gamma}} \in \mathtt{Hol}\left(\mathbb{C}, \mathtt{Lin}\left(\mathbb{W}^p, \mathbb{L}^p \times \mathbb{C}^5\right)\right)$. This operator maps a complex parameter λ to the pair consisting of the differential equation and

the boundary conditions:

$$\begin{aligned} &\mathfrak{T}_{\boldsymbol{\gamma}}(\lambda) : \mathbb{W}^p \to \mathbb{L}^p \times \mathbb{C}^5 \\ &\mathbf{p} = (p, p^r) \mapsto \left(\mathcal{T}_{\boldsymbol{\gamma}}^D(\lambda)\mathbf{p}, \mathcal{B}_{\boldsymbol{\gamma}}\mathbf{p}\right) \end{aligned} \tag{4.17}$$

where $\mathcal{T}_{\boldsymbol{\gamma}}^D(\lambda)\mathbf{p} = (\mathcal{L}_{\boldsymbol{\gamma}}p - \lambda p, -\partial_\tau p^r - \lambda p^r)$. The domain of the operator $\mathcal{T}_{\boldsymbol{\gamma}}$ corresponds to the kernel of the boundary component: ${}^p\mathbb{D}_{\boldsymbol{\gamma}} = \ker \mathcal{B}_{\boldsymbol{\gamma}}$.
The kernel of the differential component $\mathcal{T}_{\boldsymbol{\gamma}}^D(\lambda)$ is finite-dimensional. We parametrize this kernel using the fundamental matrix function $\mathcal{Z}_{\boldsymbol{\gamma}} \in \texttt{Hol}\left(\mathbb{C}, \texttt{Lin}\left(\mathbb{C}^5, \mathbb{W}^p\right)\right)$:

$$\begin{aligned} &\mathcal{Z}_{\boldsymbol{\gamma}}(\lambda) : \mathbb{C}^5 \to \mathbb{W}^p \\ &(a^-, b^-, a^+, b^+, a^r)^\top \mapsto \begin{bmatrix} a^- f_1^-(\cdot, \lambda, \boldsymbol{\gamma}) + b^- f_2^-(\cdot, \lambda, \boldsymbol{\gamma}) \\ a^+ f_1^+(\cdot, \lambda, \boldsymbol{\gamma}) + b^+ f_2^+(\cdot, \lambda, \boldsymbol{\gamma}) \\ a^r \exp(-\lambda \cdot) \end{bmatrix} \end{aligned} \tag{4.18}$$

Here, $f_1(\cdot, \lambda, \boldsymbol{\gamma})$ and $f_2(\cdot, \lambda, \boldsymbol{\gamma})$ are two linearly independent fundamental solutions of the eigenvalue equation $\mathcal{L}_{\boldsymbol{\gamma}} f = \lambda f$ on (α, θ), normalized at the left boundary α:

$$\begin{bmatrix} f_1(\alpha, \lambda, \boldsymbol{\gamma}) & f_2(\alpha, \lambda, \boldsymbol{\gamma}) \\ S_1(\alpha, \lambda, \boldsymbol{\gamma}) & S_2(\alpha, \lambda, \boldsymbol{\gamma}) \end{bmatrix} = \mathbf{Id}_{\mathbb{C}^2} \tag{4.19}$$

where $S_i(\cdot, \lambda, \boldsymbol{\gamma}) := \mathcal{S}_{\boldsymbol{\gamma}}[f_i](\cdot, \lambda, \boldsymbol{\gamma})$ denotes the flux associated with the fundamental solution. Note that this formulation implicitly maps the problem to an equivalent five-dimensional first-order system.
The spectral structure is completely determined by the *characteristic matrix* $M_{\boldsymbol{\gamma}} \in \texttt{Hol}\left(\mathbb{C}, \texttt{Lin}\left(\mathbb{C}^5, \mathbb{C}^5\right)\right)$, obtained by applying the boundary conditions to the fundamental matrix:

$$M_{\boldsymbol{\gamma}}(\lambda) := \mathcal{B}_{\boldsymbol{\gamma}} \mathcal{Z}_{\boldsymbol{\gamma}}(\lambda), \quad \forall \lambda \in \mathbb{C}. \tag{4.20}$$

Explicitly, substituting Eq. (4.18) into Eq. (4.11) yields:

$$M_{\boldsymbol{\gamma}}(\lambda) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & f_1(\theta, \lambda, \boldsymbol{\gamma}) & f_2(\theta, \lambda, \boldsymbol{\gamma}) & 0 \\ -f_1(H, \lambda, \boldsymbol{\gamma}) & -f_2(H, \lambda, \boldsymbol{\gamma}) & f_1(H, \lambda, \boldsymbol{\gamma}) & f_2(H, \lambda, \boldsymbol{\gamma}) & 0 \\ -S_1(H, \lambda, \boldsymbol{\gamma}) & -S_2(H, \lambda, \boldsymbol{\gamma}) & S_1(H, \lambda, \boldsymbol{\gamma}) & S_2(H, \lambda, \boldsymbol{\gamma}) & -\exp(-\lambda \tau_0) \\ 0 & 0 & S_1(\theta, \lambda, \boldsymbol{\gamma}) & S_2(\theta, \lambda, \boldsymbol{\gamma}) & -1 \end{bmatrix} \tag{4.21}$$

4.2.2 Spectral Equivalence and Characterization

The relationship between the operator spectrum and the characteristic matrix is formalized in the following lemma.

Lemma 4.2.1. *The pair $(\mathfrak{T}_{\boldsymbol{\gamma}}, \mathcal{Z}_{\boldsymbol{\gamma}})$ constitutes an abstract boundary eigenvalue problem as defined in [131, pp. 46-50]. specifically, $\mathfrak{T}_{\boldsymbol{\gamma}}$ is a Fredholm operator-valued function. There exist holomorphic operator functions $\mathcal{C}_{\boldsymbol{\gamma}} : \mathbb{C} \to \texttt{Lin}\left(\mathbb{C}^5 \times \mathbb{L}^p, \mathbb{L}^p \times \mathbb{C}^5\right)$ and $\mathcal{D}_{\boldsymbol{\gamma}} : \mathbb{C} \to \texttt{Lin}\left(\mathbb{W}^p, \mathbb{C}^5 \times \mathbb{L}^p\right)$, which are*

boundedly invertible for all $\lambda \in \mathbb{C}$, such that the following factorization holds:

$$\mathfrak{T}_{\boldsymbol{\gamma}}(\lambda) = \mathcal{C}_{\boldsymbol{\gamma}}(\lambda) \begin{bmatrix} M_{\boldsymbol{\gamma}}(\lambda) & \mathbf{0} \\ \mathbf{0} & \mathbf{Id}_{\mathbb{L}^p} \end{bmatrix} \mathcal{D}_{\boldsymbol{\gamma}}(\lambda), \quad \forall \lambda \in \mathbb{C}. \tag{4.22}$$

This implies that $\mathfrak{T}_{\boldsymbol{\gamma}}$ is globally equivalent on $\mathbb{C}$ to the canonical extension of the characteristic matrix $M_{\boldsymbol{\gamma}}$. Consequently, the spectra coincide:

$$\sigma(\mathcal{T}_{\boldsymbol{\gamma}}^*) = \sigma(\mathcal{T}_{\boldsymbol{\gamma}}) = \sigma(M_{\boldsymbol{\gamma}}) =: \sigma(\boldsymbol{\gamma}). \tag{4.23}$$

Moreover, the spectral data of any eigenvalue λ_* (geometric multiplicity, algebraic multiplicity, partial multiplicities) are identical for the operator $\mathcal{T}_{\boldsymbol{\gamma}}$ and the matrix $M_{\boldsymbol{\gamma}}$.

Proof. This result is a direct application of Theorem 1.11.1 and Lemma 1.11.2 in [131] to the present setting. □

Remark 4.2.1. *The differential component $\mathcal{T}_{\boldsymbol{\gamma}}^D(\lambda) : \mathbb{W}^p \to \mathbb{L}^p$ admits a holomorphic right inverse $\mathcal{Q}_{\boldsymbol{\gamma}}(\lambda) : \mathbb{L}^p \to \mathbb{W}^p$, provided by the variation of parameters formula. For any $\mathbf{g} = (g, g^r)^\top \in \mathbb{L}^p$:*

$$\mathcal{Q}_{\boldsymbol{\gamma}}(\lambda)\mathbf{g} = \begin{bmatrix} v \mapsto \int_\alpha^v g(u) \frac{f_1(v,\lambda)f_2(u,\lambda) - f_2(v,\lambda)f_1(u,\lambda)}{w_{\boldsymbol{\gamma}}(u)}\, du \\ \tau \mapsto -e^{-\lambda\tau} \int_0^\tau e^{\lambda z} g^r(z)\, dz \end{bmatrix}$$

where $w_{\boldsymbol{\gamma}}(v)$ is the Wronskian of the fundamental solutions f_1 and f_2. Due to the structure of the Fokker-Planck operator, this Wronskian satisfies (see [198, Section III.B]):

$$\begin{aligned} w_{\boldsymbol{\gamma}}(v) &= D[f_2 f_1' - f_2' f_1](v) \\ &= f_1(v)S_2(v) - f_2(v)S_1(v) = \exp\left(\int_\alpha^v \frac{b(u)+\mu}{D}\, du\right) > 0. \end{aligned} \tag{4.24}$$

4.2.3 Properties of the Spectrum

We can now state the main result regarding the structure of the spectrum.

Theorem 4.2.1. *For any parameter set $\boldsymbol{\gamma} = (\mu, D) \in \mathbb{R} \times \mathbb{R}_{>0}$, the spectrum of the evolution operator $\mathcal{T}_{\boldsymbol{\gamma}}$ consists entirely of a discrete set of eigenvalues with no finite accumulation point. The spectrum $\sigma(\boldsymbol{\gamma})$ is characterized as follows:*

1. ***Characteristic Equation:*** *The eigenvalues are the roots of the scalar characteristic function $\Delta_{\boldsymbol{\gamma}}(\lambda)$:*

$$\sigma(\boldsymbol{\gamma}) = \{\lambda \in \mathbb{C} \mid \Delta_{\boldsymbol{\gamma}}(\lambda) = 0\} \tag{4.25}$$

$$\Delta_{\boldsymbol{\gamma}}(\lambda) := \frac{\det(M_{\boldsymbol{\gamma}}(\lambda))}{w_{\boldsymbol{\gamma}}(H)w_{\boldsymbol{\gamma}}(\theta)} = \frac{f_1(\theta,\lambda,\boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(\theta)} - \frac{f_1(H,\lambda,\boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(H)} e^{-\lambda\tau_0} \tag{4.26}$$

   where f_1 is the fundamental solution satisfying $\mathcal{L}_{\boldsymbol{\gamma}} f_1 = \lambda f_1$ with $f_1(\alpha) = 1, S_1(\alpha) = 0$.
2. ***Existence of Stationary State:*** *$\lambda = 0$ is always an eigenvalue.*
3. ***Simplicity:*** *All eigenvalues have geometric multiplicity equal to 1.*

We denote the spectrum as $\sigma(\boldsymbol{\gamma}) = \{\lambda_n(\boldsymbol{\gamma})\}_{n\in\mathbb{N}}$, ordered by decreasing real part (lexicographically), with each eigenvalue repeated according to its algebraic multiplicity m_n.

Proof. Characteristic Equation: The condition $\lambda \in \sigma(M_{\boldsymbol{\gamma}})$ is equivalent to $\det(M_{\boldsymbol{\gamma}}(\lambda)) = 0$. Computing the determinant of the 5×5 matrix explicitly yields the expression in Eq. (4.25).
Discreteness: The function $\Delta_{\boldsymbol{\gamma}}(\lambda)$ is entire (analytic on $\mathbb{C}$). Moreover $\Delta_{\boldsymbol{\gamma}}$ is not identically zero, since $\Delta'_{\boldsymbol{\gamma}}(0) \neq 0$ ($\lambda = 0$ is a simple eigenvalue, see below); by the global equivalence of Lemma 4.2.1, $\sigma(\boldsymbol{\gamma})$ coincides with the zero set of $\Delta_{\boldsymbol{\gamma}}$, which is therefore a discrete set with no accumulation point in $\mathbb{C}$, each eigenvalue λ_n having finite algebraic multiplicity equal to the order of the corresponding zero.
Stationary State: Consider the case $\lambda = 0$. The fundamental solution $f_1(v, 0, \boldsymbol{\gamma})$ defined by $\mathcal{S}_{\boldsymbol{\gamma}} f_1(\alpha) = 0$ corresponds to the function with constant zero flux. In the absence of boundaries, this is the unnormalized equilibrium density $w_{\boldsymbol{\gamma}}(v) = \exp\left(\int_\alpha^v \frac{A(u)+\mu}{D} du\right)$. Substituting $f_1(v, 0, \boldsymbol{\gamma}) = w_{\boldsymbol{\gamma}}(v)$ into the characteristic equation gives:

$$\Delta_{\boldsymbol{\gamma}}(0) = \frac{w_{\boldsymbol{\gamma}}(\theta)}{w_{\boldsymbol{\gamma}}(\theta)} - \frac{w_{\boldsymbol{\gamma}}(H)}{w_{\boldsymbol{\gamma}}(H)} \cdot e^0 = 1 - 1 = 0. \tag{4.27}$$

Thus, $0 \in \sigma(\boldsymbol{\gamma})$, confirming the existence of a stationary distribution.
Geometric Multiplicity: Assume for contradiction that the geometric multiplicity of some λ_n is > 1. This would imply that the rank of $M_{\boldsymbol{\gamma}}(\lambda_n)$ is at most 3, necessitating that all 4×4 minors vanish. In particular, this requires $f_1(\theta, \lambda_n) = 0$ and $f_1(H, \lambda_n) = 0$. The boundary conditions then force $f_2(\theta) = f_2(H) = 0$. However, this implies $w_{\boldsymbol{\gamma}}(\theta) = f_1(\theta)S_2(\theta) - f_2(\theta)S_1(\theta) = 0$, which contradicts the positivity of the Wronskian ($w_{\boldsymbol{\gamma}} > 0$). Thus, the geometric multiplicity must be 1. □

Theorem 4.2.1 provides a constructive method for finding the eigenvalues. It is worth noting that in the limit of a vanishing refractory period ($\tau_0 \to 0$), the characteristic equation (4.25) reduces to:

$$\frac{f_1(\theta, \lambda)}{w_{\boldsymbol{\gamma}}(\theta)} - \frac{f_1(H, \lambda)}{w_{\boldsymbol{\gamma}}(H)} = 0. \tag{4.28}$$

This recovers the known spectral condition for the standard Fokker-Planck equation derived in previous works (e.g., [126]).
Finally, the characteristic equation can be elegantly reformulated in terms of the adjoint fundamental solutions. Letting $\psi_1(v) = w_{\boldsymbol{\gamma}}(v)^{-1} f_1(v)$, which solves the adjoint equation $\mathcal{L}^*_{\boldsymbol{\gamma}} \psi_1 = \lambda \psi_1$ with boundary condition $\psi'_1(\alpha) = 0$, we have:

$$\Delta_{\boldsymbol{\gamma}}(\lambda) = \psi_1(\theta, \lambda, \boldsymbol{\gamma}) - e^{-\lambda \tau_0} \psi_1(H, \lambda, \boldsymbol{\gamma}) = 0. \tag{4.29}$$

This form highlights the interplay between the value of the adjoint eigenfunction at the threshold and its delayed value at the reset potential.

4.3 Eigenfunctions and associated functions

With the spectral structure characterized, we now turn to the explicit construction of the root functions of the evolution operator $\mathcal{T}_{\boldsymbol{\gamma}}$ and its adjoint $\mathcal{T}^*_{\boldsymbol{\gamma}}$. Given that all eigenvalues have geometric multiplicity one (Theorem 4.2.1), any eigenvalue λ_n with algebraic multiplicity $m_n > 1$ will possess a Jordan chain of length m_n, consisting of one eigenfunction and $m_n - 1$ associated functions.

4.3.1 Adjoint Formulation and Reduced Characteristic Matrix

To study the properties of the spectrum of the evolution operator, it is convenient to consider the adjoint boundary eigenvalue problem, as the domain has no derivative discontinuity at H and the boundary conditions can be reduced. Indeed, inspecting the second and third boundary conditions in Eq. (4.14), specifically $f'(H^+) - f'(H^-) = 0$ and $f(H^+) - f(H^-) = 0$, we observe that $\mathcal{B}^*\mathbf{f} = 0 \Rightarrow f \in \mathrm{W}^{2,q}(\alpha, \theta)$. Thus, functions in the adjoint domain ${}^q\mathbb{D}^*_{\boldsymbol{\gamma}}$ have continuous second derivatives across the reset potential H, and we can naturally define the reduced adjoint space $\mathbb{V}^q := \mathrm{W}^{2,q}(\alpha, \theta) \times \mathrm{W}^{1,q}(0, \tau_0)$.
We can then rewrite the domain of the adjoint operator as

$$\mathbb{D}^* := \{\mathbf{f} \in \mathbb{V}^q \mid \mathcal{B}^{\oplus}\mathbf{f} = 0\}, \tag{4.30}$$

where $\mathcal{B}^{\oplus} : \mathbb{V}^q \to \mathbb{C}^3$ is the reduced adjoint boundary operator, defined as:

$$\mathcal{B}^{\oplus} : \mathbb{V}^q \to \mathbb{C}^3 \tag{4.31}$$

$$\mathbf{f} = (f, f^r) \mapsto \begin{bmatrix} f'(\alpha) \\ f(H) - f^r(\tau_0) \\ f(\theta) - f^r(0) \end{bmatrix}. \tag{4.32}$$

We then define the reduced adjoint boundary eigenvalue operator as the holomorphic function that maps the complex parameter $\lambda \in \mathbb{C}$ to the operator:

$$\mathfrak{T}^{\oplus}_{\boldsymbol{\gamma}}(\lambda) = (\mathfrak{T}^{D\oplus}, \mathfrak{T}^{R\oplus}) : \mathbb{V} \to \mathbb{L}^q \times \mathbb{C}^3 \tag{4.33}$$

$$\mathbf{f} = (f, f^r) \mapsto \left(\left(\mathcal{L}^*_{\boldsymbol{\gamma}} f, \partial_\tau f^r\right) - \lambda \mathbf{f}, \mathcal{B}^{\oplus}\mathbf{f}\right). \tag{4.34}$$

The operator $\mathfrak{T}^{\oplus D}_{\boldsymbol{\gamma}}(\lambda) : \mathbb{W}^p \to \mathbb{L}^p$ has a right inverse $\mathcal{Q}_{\boldsymbol{\gamma}}(\lambda) : \mathbb{L}^p \to \mathbb{W}^p$, holomorphic in λ, given by the variation of parameters formula:

$$\begin{aligned} {}^q\mathbb{D}_{\boldsymbol{\gamma}} \ni \mathbf{g} = \begin{bmatrix} g \\ g^r \end{bmatrix} \mapsto & \mathcal{Q}^{\oplus}_{\boldsymbol{\gamma}}(\lambda)\mathbf{g} = \\ & = \begin{bmatrix} v \mapsto \int_\alpha^v g(u)[\psi_1(v, \boldsymbol{\gamma}, \lambda) f_2(u, \boldsymbol{\gamma}, \lambda) - \psi_2(v, \boldsymbol{\gamma}, \lambda) f_1(u, \boldsymbol{\gamma}, \lambda)]du \\ \tau \mapsto e^{\tau\lambda} \int_0^\tau e^{-z\lambda} g^r(z) dz \end{bmatrix}. \end{aligned} \tag{4.35}$$

As before, we define the reduced adjoint fundamental matrix function as the holomorphic function $\mathcal{Z}^{\oplus}_{\boldsymbol{\gamma}} \in \mathtt{Hol}\left(\mathbb{C}, \mathfrak{L}\left(\mathbb{C}^3, \mathbb{V}\right)\right)$ that parametrizes the kernel of $\mathfrak{T}^{D\oplus}_{\boldsymbol{\gamma}}$:

$$\mathcal{Z}^{\oplus}_{\boldsymbol{\gamma}}(\lambda) : \mathbb{C}^3 \to \mathbb{V} \tag{4.36}$$

$$(a, b, a^r) \mapsto \begin{bmatrix} a\psi_1(\cdot, \lambda, \boldsymbol{\gamma}) + b\psi_2(\cdot, \lambda, \boldsymbol{\gamma}) \\ a^r \exp(\cdot\lambda) \end{bmatrix} \in \mathbb{V}, \tag{4.37}$$

where $\psi_1(\cdot, \lambda, \boldsymbol{\gamma}), \psi_2(\cdot, \lambda, \boldsymbol{\gamma}) \in \mathrm{W}^{2,p}(\alpha, \theta)$ are two independent fundamental solutions of the equation $\mathcal{L}^*_{\boldsymbol{\gamma}}\psi = \lambda\psi$, normalized such that $\psi_1(\alpha) = 1, \psi_1'(\alpha) = 0$ and $\psi_2(\alpha) = 0, \psi_2'(\alpha) = -D^{-1}$.
The spectrum structure can be completely determined by studying the 3×3 reduced adjoint char-

acteristic matrix $M_{\boldsymbol{\gamma}}^{\oplus} \in \mathtt{Hol}\left(\mathbb{C}, \mathtt{Lin}\left(\mathbb{C}^3, \mathbb{C}^3\right)\right)$, defined as:

$$M_{\boldsymbol{\gamma}}^{\oplus}(\lambda) = \mathcal{B}^{\oplus}\mathcal{Z}_{\boldsymbol{\gamma}}^{\oplus}(\lambda) = \begin{bmatrix} 0 & -D^{-1} & 0 \\ \psi_1(H,\lambda,\boldsymbol{\gamma}) & \psi_2(H,\lambda,\boldsymbol{\gamma}) & -\exp(\lambda\tau_0) \\ \psi_1(\theta,\lambda,\boldsymbol{\gamma}) & \psi_2(\theta,\lambda,\boldsymbol{\gamma}) & -1 \end{bmatrix} \quad \forall \lambda \in \mathbb{C}. \tag{4.38}$$

The characteristic equation derived previously is recovered as:

$$0 = \Delta_{\boldsymbol{\gamma}}(\lambda) = D\exp(-\lambda\tau_0)\det\left(M_{\boldsymbol{\gamma}}^{\oplus}(\lambda)\right) = \psi_1(\theta,\boldsymbol{\gamma},\lambda) - \psi_1(H,\boldsymbol{\gamma},\lambda)\exp(-\tau_0\lambda).$$

4.3.2 Eigenfunctions and canonical root functions

The characteristic function $\Delta_{\boldsymbol{\gamma}}$ completely determines the spectral properties of the boundary eigenvalue operator function $\mathcal{T}_{\boldsymbol{\gamma}}$. In particular, one can construct holomorphic *root functions* for the eigenvalue problem associated with $\mathcal{T}_{\boldsymbol{\gamma}}$ and its adjoint.

Lemma 4.3.1 (Root functions and principal part of the resolvent)**.** *There exist holomorphic mappings*

$$s \longmapsto \boldsymbol{\varphi}(s,\boldsymbol{\gamma}) \in \mathbb{D}_{\boldsymbol{\gamma}}, \qquad s \longmapsto \boldsymbol{\Psi}(s,\boldsymbol{\gamma}) \in \mathbb{D}^*,$$

given by

$$\boldsymbol{\Psi}(s,\boldsymbol{\gamma}) = \begin{bmatrix} \psi_1(\,\cdot\,,s,\boldsymbol{\gamma}) \\ \psi_1(\theta,s,\boldsymbol{\gamma})\,\exp(\,\cdot\,s) \end{bmatrix}, \tag{4.39}$$

$$\boldsymbol{\varphi}(s,\boldsymbol{\gamma}) = \begin{bmatrix} f_2^-(\,\cdot\,,s,\boldsymbol{\gamma})\,\Delta_{\boldsymbol{\gamma}}(s) - f_1^-(\,\cdot\,,s,\boldsymbol{\gamma})\left(\frac{f_2(\theta,s,\boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(\theta)} - e^{-s\tau_0}\frac{f_2(H,s,\boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(H)}\right) \\ f_2^+(\,\cdot\,,s,\boldsymbol{\gamma})\,\frac{f_1(\theta,s,\boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(\theta)} - f_1^+(\,\cdot\,,s,\boldsymbol{\gamma})\,\frac{f_2(\theta,s,\boldsymbol{\gamma})}{w_{\boldsymbol{\gamma}}(\theta)} \\ \exp(-\,\cdot\,s) \end{bmatrix}. \tag{4.40}$$

For every eigenvalue $\lambda_n \in \sigma(\boldsymbol{\gamma})$ with algebraic multiplicity m_n, the map

$$s \longmapsto (\mathcal{T}_{\boldsymbol{\gamma}} - s)^{-1} - \frac{\boldsymbol{\varphi}(s,\boldsymbol{\gamma}) \otimes \boldsymbol{\Psi}(s,\boldsymbol{\gamma})}{\Delta_{\boldsymbol{\gamma}}(s)} \tag{4.41}$$

is holomorphic in a neighbourhood of λ_n, where $Z_n \in \mathbb{C}\setminus\{0\}$, the leading coefficient of $\Delta_{\boldsymbol{\gamma}}$ at λ_n (used to normalize the root functions below), is given by

$$Z_n = \lim_{s\to\lambda_n} \frac{\Delta_{\boldsymbol{\gamma}}(s)}{(s-\lambda_n)^{m_n}}. \tag{4.42}$$

Proof of Lemma 4.3.1. From Theorem 4.2.1, all eigenvalues have geometric multiplicity one. Then Theorem 1.5.9 and Proposition 3.4.7 in [131] yield the existence of root functions $\boldsymbol{\varphi}(\,\cdot\,,\boldsymbol{\gamma})$ and $\boldsymbol{\Psi}(\,\cdot\,,\boldsymbol{\gamma})$ defined in a neighbourhood of each $\lambda_n \in \sigma(\boldsymbol{\gamma})$ such that (4.41) holds. Moreover, by [131, Lemma 1.11.2], one can explicitly construct a biorthogonal canonical system of root functions (CSRF) from the characteristic matrix $M_{\boldsymbol{\gamma}}$ and its adjoint; equivalently, one may use the reduced adjoint characteristic matrix $M_{\boldsymbol{\gamma}}^{\oplus}$.

We first construct root functions $s \mapsto \mathbf{c}(s,\boldsymbol{\gamma})$ and $s \mapsto \mathbf{d}(s,\boldsymbol{\gamma})$ for $M_{\boldsymbol{\gamma}}^{\oplus}$ and its adjoint:

$$\mathbf{c}(s,\boldsymbol{\gamma}) = \begin{bmatrix} 1 \\ 0 \\ \psi_1(\theta,s,\boldsymbol{\gamma}) \end{bmatrix}, \qquad \mathbf{d}(s,\boldsymbol{\gamma}) = \begin{bmatrix} e^{-\tau_0 s}\psi_2(H,s,\boldsymbol{\gamma}) - \psi_2(\theta,s,\boldsymbol{\gamma}) \\ D^{-1}e^{-\tau_0 s} \\ -D^{-1} \end{bmatrix}. \tag{4.43}$$

Then

$$\boldsymbol{\Psi}(s,\boldsymbol{\gamma}) = \mathcal{Z}_{\boldsymbol{\gamma}}^{\oplus}(s)\mathbf{c}(s,\boldsymbol{\gamma}), \qquad \boldsymbol{\varphi}(s,\boldsymbol{\gamma}) = \big(B^{\oplus}\mathcal{Q}_{\boldsymbol{\gamma}}^{\oplus}(s)\big)^{*}\mathbf{d}(s,\boldsymbol{\gamma}).$$

To determine Z_n we evaluate the pairing on the reduced kernels. A direct computation gives $\mathbf{d}(s,\boldsymbol{\gamma})^* M_{\boldsymbol{\gamma}}^{\oplus}(s)\mathbf{c}(s,\boldsymbol{\gamma}) = -D^{-1}\Delta_{\boldsymbol{\gamma}}(s)$, and, since $\mathcal{Z}_{\boldsymbol{\gamma}}^{\oplus}$ together with $\big(B^{\oplus}\mathcal{Q}_{\boldsymbol{\gamma}}^{\oplus}\big)^*$ contribute a compensating factor $-D$, one has $\langle \boldsymbol{\Psi}(\lambda_n,\boldsymbol{\gamma}), \boldsymbol{\varphi}(\lambda_n,\boldsymbol{\gamma})\rangle = \Delta'_{\boldsymbol{\gamma}}(\lambda_n)$. The eigenvalue λ_n having geometric multiplicity one, the leading element of the chain is normalized by the single constant (4.42), the leading coefficient of $\Delta_{\boldsymbol{\gamma}}$ at λ_n, so that $\langle \boldsymbol{\Psi}_n, \boldsymbol{\varphi}_n\rangle = 1$ when λ_n is simple; the residual scale is fixed by $\mathcal{N}[\boldsymbol{\varphi}] = 1$.

For $m_n \geq 2$ a single constant no longer suffices. First, a defective eigenvalue satisfies $\Delta'_{\boldsymbol{\gamma}}(\lambda_n) = 0$, so $\boldsymbol{\varphi}_n$ is orthogonal to $\boldsymbol{\Psi}_n$ and no rescaling can make that pairing equal to one. Second, the normalization $\mathcal{N}[\boldsymbol{\varphi}] = 1$ fixes only the scale of $\boldsymbol{\varphi}$; the pair $(\boldsymbol{\Psi}, \Delta_{\boldsymbol{\gamma}})$ remains free up to a common holomorphic factor, since (4.41) constrains only the combination $\boldsymbol{\varphi} \otimes \boldsymbol{\Psi}/\Delta_{\boldsymbol{\gamma}}$. The associated vectors must therefore be built from that combination, as in (4.45), which is invariant under the residual freedom, rather than from $\boldsymbol{\Psi}$ and Z_n separately. The chain as a whole is then fixed by the biorthogonality relations of [131, Thm. 1.5.4 and Cor. 1.5.6]. □

Eigenfunctions and normalization. For $n > 0$ we define the eigenfunctions of $\mathcal{T}_{\boldsymbol{\gamma}}$ and $\mathcal{T}_{\boldsymbol{\gamma}}^*$ corresponding to λ_n by

$$\boldsymbol{\varphi}_n(\boldsymbol{\gamma}) := \boldsymbol{\varphi}(\lambda_n,\boldsymbol{\gamma}), \qquad \boldsymbol{\Psi}_n(\boldsymbol{\gamma}) := Z_n^{-1}\boldsymbol{\Psi}(\lambda_n,\boldsymbol{\gamma}).$$

For the stationary eigenvalue $n = 0$ we instead set

$$\boldsymbol{\varphi}_0(\boldsymbol{\gamma}) := Z_0^{-1}\boldsymbol{\varphi}(0,\boldsymbol{\gamma}), \qquad \boldsymbol{\Psi}_0(\boldsymbol{\gamma}) := \boldsymbol{\Psi}(0,\boldsymbol{\gamma}),$$

so that $\|\boldsymbol{\varphi}_0\|_{\mathbb{L}^1} = 1$.

Associated vectors and Laurent expansion. For each eigenvalue λ_n with algebraic multiplicity m_n, define the associated functions by

$$\boldsymbol{\varphi}_n^{(k)}(\boldsymbol{\gamma}) := \frac{1}{k!}\,\partial_s^k \boldsymbol{\varphi}(s,\boldsymbol{\gamma})\big|_{s=\lambda_n} \in \mathbb{D}_{\boldsymbol{\gamma}}, \tag{4.44}$$

$$\boldsymbol{\Psi}_n^{(k)}(\boldsymbol{\gamma}) := \frac{1}{k!}\,\partial_s^k \left[\frac{(s-\lambda_n)^{m_n}\,\boldsymbol{\Psi}(s,\boldsymbol{\gamma})}{\Delta_{\boldsymbol{\gamma}}(s)}\right]_{s=\lambda_n} \in \mathbb{D}^*, \qquad k = 0,1,\dots,m_n-1. \tag{4.45}$$

The bracket is holomorphic at λ_n, since $\Delta_{\boldsymbol{\gamma}}$ has there a zero of order exactly m_n. Writing $\Delta_{\boldsymbol{\gamma}}(s) = (s-\lambda_n)^{m_n} g_n(s)$ with g_n holomorphic and $g_n(\lambda_n) = Z_n \neq 0$, the bracket is $\boldsymbol{\Psi}/g_n$; for $k = 0$ this returns $Z_n^{-1}\boldsymbol{\Psi}(\lambda_n,\boldsymbol{\gamma})$, so $\boldsymbol{\Psi}_n^{(0)} = \boldsymbol{\Psi}_n$ and the simple-eigenvalue case ($m_n = 1$, where g_n contributes only through its value at λ_n) is unchanged. For $m_n \geq 2$, however, g_n is *not* constant (it vanishes

at every other eigenvalue), and its derivatives enter $\boldsymbol{\Psi}_n^{(k)}$ for $k \geq 1$. Replacing g_n by the constant Z_n would therefore misidentify the associated vectors: by [131, Thm. 1.5.4] the left factor in the principal part is a *polynomial* of degree less than m_n, namely the Taylor polynomial of $\boldsymbol{\Psi}/g_n$, which degenerates to the constant $\boldsymbol{\Psi}(\lambda_n)/Z_n$ only when $m_n = 1$.
Taylor expanding $\boldsymbol{\varphi}(\,\cdot\,, \boldsymbol{\gamma})$ and $(\,\cdot - \lambda_n)^{m_n}\boldsymbol{\Psi}(\,\cdot\,, \boldsymbol{\gamma})/\Delta_{\boldsymbol{\gamma}}$ at $s = \lambda_n$ and inserting into (4.41) yields the singular Laurent part of the resolvent at λ_n:

$$\Big[(\mathcal{T}_{\boldsymbol{\gamma}} - s)^{-1}\Big]_{\mathrm{sing},\,\lambda_n} = \sum_{r=1}^{m_n} (s - \lambda_n)^{-r} \sum_{k=0}^{m_n - r} \boldsymbol{\varphi}_n^{(k)} \otimes \boldsymbol{\Psi}_n^{(m_n - r - k)}. \tag{4.46}$$

Biorthogonality of the CSEAV. Let $\lambda_i, \lambda_j \in \sigma(\boldsymbol{\gamma})$ have algebraic multiplicities m_i and m_j. Then the associated vectors form a biorthogonal Canonical System of Eigenvectors and Associated Vectors (CSEAV) in the sense that:

$$\big\langle \boldsymbol{\Psi}_i^{(l)},\, \boldsymbol{\varphi}_j^{(m_j - 1 - k)} \big\rangle = \delta_{ij}\, \delta_{lk}, \qquad 0 \leq l < m_i, \;\; 0 \leq k < m_j. \tag{4.47}$$

In particular, for each fixed eigenvalue λ_n the left chain $\{\boldsymbol{\Psi}_n^{(0)}, \ldots, \boldsymbol{\Psi}_n^{(m_n - 1)}\}$ pairs with the right chain $\{\boldsymbol{\varphi}_n^{(m_n - 1)}, \ldots, \boldsymbol{\varphi}_n^{(0)}\}$ in reversed order.

Flux Properties

Finally, we characterize the firing rate (flux) associated with the spectral modes. This is physically relevant because the total population firing rate is a superposition of these modal fluxes.
Lemma 4.3.2 (Flux normalization)**.** *For any eigenvalue λ_n with $n \geq 1$, the eigenfunction $\boldsymbol{\varphi}_n^{(0)}$ carries unit flux, while all associated functions $\boldsymbol{\varphi}_n^{(k)}$ ($k \geq 1$) carry zero flux.*

Proof. Recall the firing rate operator $\mathcal{N}[\mathbf{p}] = p^r(0)$. For the generating function $\boldsymbol{\varphi}(s, \boldsymbol{\gamma})$, the refractory component equals $\exp(-\tau s)$. Evaluating at $\tau = 0$ gives 1.
Then, for $k = 0$ we obtain

$$\mathcal{N}\big(\boldsymbol{\varphi}_n^{(0)}\big) = 1.$$

For $k \geq 1$ we differentiate the generating function with respect to s:

$$\mathcal{N}\big(\boldsymbol{\varphi}_n^{(k)}\big) = \frac{1}{k!}\, \partial_s^k [\exp(-\tau s)]_{\tau = 0,\, s = \lambda_n} = \frac{1}{k!}\, \Big[(-\tau)^k \exp(-\tau s)\Big]\,\big|_{\tau = 0} = 0.$$

Thus, only the rank-0 eigenfunction contributes to the instantaneous firing rate. □

4.4 Spectrum analysis for LIF neurons

To complement the abstract spectral theory developed thus far, we now turn to a concrete numerical investigation of the spectrum structure.
We numerically solve the characteristic equation $\Delta_{\boldsymbol{\gamma}}(\lambda) = 0$ (Eq. (4.25)) over the domain $[-10, 10] \times [-20i, 20i] \subset \mathbb{C}$ for the LIF neuron ($b(v) = -v$). The diffusion coefficient is fixed at $D = 8\,\mathrm{mV}^2$, and we vary the input mean μ in the range $[-5, 25]\,\mathrm{mV}$ with a step size of 0.05. Figure 4.1 illustrates the evolution of the spectrum as a function of μ for different values of the refractory period τ_0.

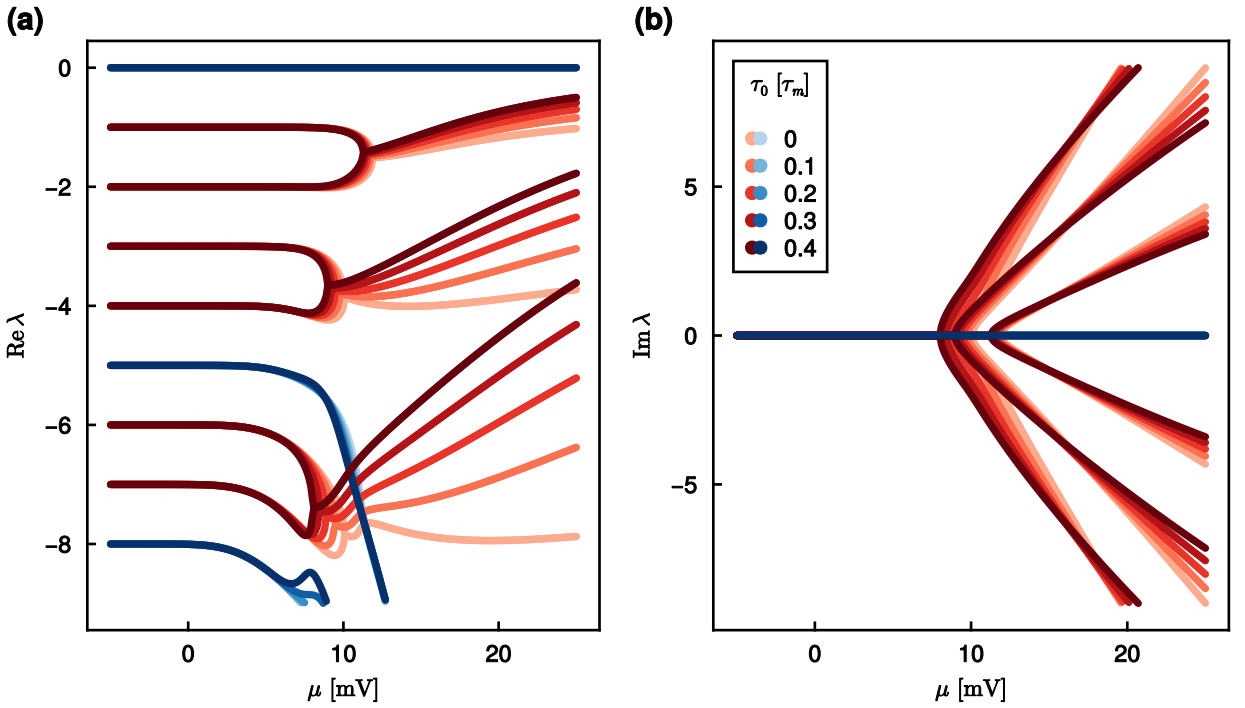

Figure 4.1: Trajectory of the eigenvalues of the evolution operator as a function of the input mean μ, for different values of the refractory period τ_0. (a) Real part of the eigenvalues. (b) Imaginary part. The parameter μ varies from -5 mV to 25 mV. Note the distinct branching structure and the coalescence points marking the transition from real to complex conjugate pairs. Red and blue branches denote mixed and real λ_n, respectively. $D = 8\text{mV}^2$, $H = 0$ mV, $\theta = 20$ mV, $\alpha = -5\theta = -100$ mV, $\tau_m = 1$; shadings represent τ_0 from $0\,\tau_m$ (light) to $0.4\,\tau_m$ (dark); μ is sampled at 0.1 mV steps

Impact of Refractoriness

The introduction of a finite refractory period $\tau_0 > 0$ alters this spectral landscape. First, the onset of the real-to-complex transition shifts to lower values of μ, implying that oscillatory modes emerge more readily. More significantly, for complex eigenvalues, the real part is no longer bounded but increases with μ. This indicates that for strong inputs, the damping rate decreases relative to the oscillation frequency, enhancing the population's propensity for resonant behavior, a phenomenon qualitatively distinct from the $\tau_0 = 0$ case. Furthermore, the growth rate of the imaginary part with respect to μ is reduced as τ_0 increases, resulting in lower oscillation frequencies, a phenomenon related to the reduction of the firing rate ν, as a larger τ_0 leads to longer inter-spike intervals.

4.5 Defective eigenvalues and exceptional points

While the spectral theorem guarantees that each eigenvalue possesses at least one eigenfunction, the algebraic multiplicity m_n may exceed the geometric multiplicity (which is always 1, per Theorem 4.2.1). Such eigenvalues are termed *defective*. They appear precisely at the coalescence points (or exceptional points) of the spectral branches observed in Figure 4.1, where two real eigenvalues merge before becoming complex conjugates.

Mathematically, a defective eigenvalue λ_n corresponds to a root of the characteristic equation $\Delta_{\boldsymbol{\gamma}}(\lambda)$ with multiplicity $m_n \geq 2$. Thus, it must satisfy both the characteristic equation and its derivative:

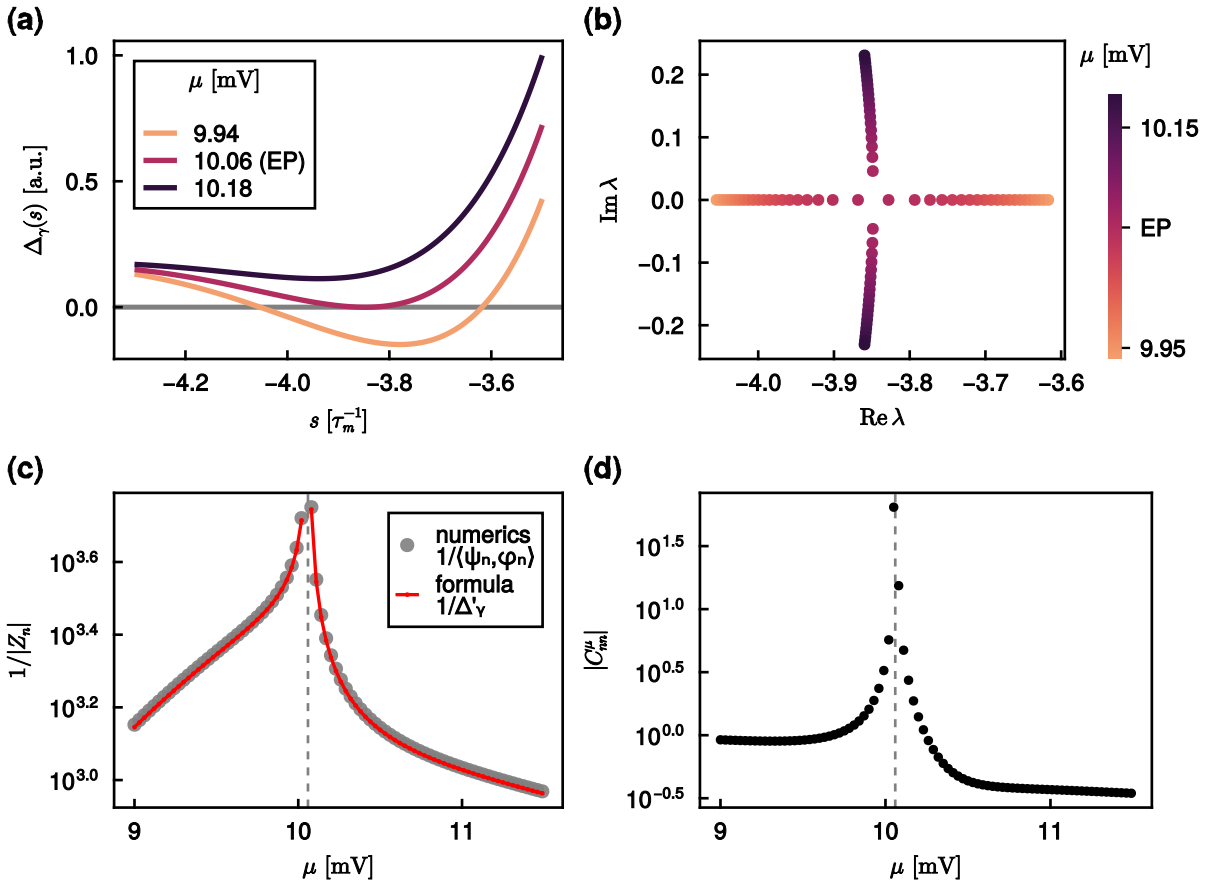

Figure 4.2: (a) Real part of the characteristic function $\Delta_{\boldsymbol{\gamma}}(s)$ for three mean currents around the collision, $\mu = 9.94$, 10.06 (exceptional point) and 10.18 mV; the intersections with the horizontal axis are the real λ_n, and they merge into a second-order zero at the exceptional point. (b) Trajectory of the two eigenvalues in the complex plane colliding at that point (defective eigenvalue), light to dark with increasing μ. (c) Inverse leading coefficient $1/|Z_n|$, from the biorthogonal scalar product $1/\langle\psi_n, \phi_n\rangle$ (grey) and from the closed form $1/\Delta'_{\boldsymbol{\gamma}}(\lambda_n)$ (red); the two agree away from the collision and both diverge at it. (d) Modulus of the self-coupling coefficient $C^{\mu}_{nn} = \langle\partial_\mu\psi_n, \phi_n\rangle$, which diverges at the same point, so that the modal decomposition ceases to be well defined at a defective eigenvalue. The panels show the second mixed branch of Figure 4.1, with $D = 8\,\mathrm{mV}^2$, $H = 0$ mV, $\theta = 20$ mV, $\alpha = -3\theta = -60$ mV.

$$\Delta_{\boldsymbol{\gamma}}(\lambda_n) = \psi_1(\theta, \lambda_n, \boldsymbol{\gamma}) - e^{-\lambda_n \tau_0}\psi_1(H, \lambda_n, \boldsymbol{\gamma}) = 0 \tag{4.48a}$$

$$\partial_\lambda \Delta_{\boldsymbol{\gamma}}(\lambda)\Big|_{\lambda=\lambda_n} = \partial_\lambda\psi_1(\theta) - e^{-\lambda_n\tau_0}\left[\partial_\lambda\psi_1(H) - \tau_0\psi_1(H)\right] = 0. \tag{4.48b}$$

Figure 4.2b visualizes the trajectory of a pair of eigenvalues in the complex plane as they collide. At the collision point, the characteristic function exhibits a second-order zero (Fig. 4.2a), confirming the defective nature of the singularity.

The derivative term $\partial_\lambda\psi_1$ can be computed using variation of parameters on the differentiated adjoint equation $(\mathcal{L}^*_{\boldsymbol{\gamma}} - \lambda)\partial_\lambda\psi_1 = \psi_1$, yielding:

$$\partial_\lambda\psi_1(v, \lambda) = \psi_1(v)\int_\alpha^v f_2(u)\psi_1(u)\,du - \psi_2(v)\int_\alpha^v f_1(u)\psi_1(u)\,du. \tag{4.49}$$

Alternatively, defectiveness can be characterized via the orthogonality of the root functions, which

satisfy

$$\langle \boldsymbol{\Psi}(s,\boldsymbol{\gamma}), \boldsymbol{\varphi}(\lambda_n,\boldsymbol{\gamma})\rangle = \frac{e^{(s-\lambda_n)\tau_0}\,\Delta_{\boldsymbol{\gamma}}(s)}{s-\lambda_n}, \qquad \lambda_n \in \sigma(\boldsymbol{\gamma}). \tag{4.50}$$

Taking $s \to \lambda_n$ yields $\langle \boldsymbol{\Psi}(\lambda_n,\boldsymbol{\gamma}), \boldsymbol{\varphi}(\lambda_n,\boldsymbol{\gamma})\rangle = \Delta'_{\boldsymbol{\gamma}}(\lambda_n)$. An eigenvalue λ_n is therefore defective if and only if the eigenfunction is orthogonal to its corresponding adjoint eigenfunction, that is $\langle \boldsymbol{\Psi}(\lambda_n,\boldsymbol{\gamma}), \boldsymbol{\varphi}(\lambda_n,\boldsymbol{\gamma})\rangle = \Delta'_{\boldsymbol{\gamma}}(\lambda_n) = 0$. Using this criterion, we can rigorously establish the simplicity of the stationary state.

Proposition 4.5.1. *The stationary eigenvalue $\lambda_0 = 0$ is always simple. Its associated eigenfunction and adjoint eigenfunction are:*

$$\boldsymbol{\varphi}_0(\boldsymbol{\gamma}) = (\phi_0(\cdot,\boldsymbol{\gamma}), \nu_0(\boldsymbol{\gamma})), \quad \boldsymbol{\Psi}_0(\boldsymbol{\gamma}) = (1,1) \tag{4.51}$$

where $\phi_0(v,\boldsymbol{\gamma})$ is the stationary density profile:

$$\phi_0(v,\boldsymbol{\gamma}) = Z_0^{-1} D^{-1} w_{\boldsymbol{\gamma}}(v) \int_{v\vee H}^{\theta} w_{\boldsymbol{\gamma}}^{-1}(u)\,du. \tag{4.52}$$

Proof. The form of $\boldsymbol{\varphi}_0$ follows from evaluating the general root function formula (Lemma 4.3.1) at $\lambda = 0$. To prove simplicity, we compute the inner product using the unnormalized basis functions from the characteristic matrix construction:

$$\langle \boldsymbol{\Psi}(0), \boldsymbol{\varphi}(0)\rangle = D^{-1} \int_{\alpha}^{\theta} w_{\boldsymbol{\gamma}}(v) \left(\int_{v\vee H}^{\theta} w_{\boldsymbol{\gamma}}^{-1}(u)\,du \right) dv + \tau_0. \tag{4.53}$$

Since the integrand and τ_0 are strictly positive, the inner product is non-zero, proving simplicity. Moreover, by (4.42) this quantity is exactly Z_0, so the same computation gives $\|\boldsymbol{\varphi}_0\|_{\mathbb{L}^1} = 1$: the normalization of the stationary density follows from the definition of Z_n rather than being imposed as a separate convention. □

4.5.1 Implications for Spectral Expansions

Previous studies on the spectral decomposition of neuronal population dynamics [52, 101, 126, 154] have implicitly assumed that all eigenvalues are simple. This assumption is mathematically convenient because simple eigenvalues depend smoothly on the system parameters $\boldsymbol{\gamma}$.

Remark 4.5.1. *If $\lambda^{(0)} \in \sigma(\boldsymbol{\gamma}_\star)$ is a simple eigenvalue, then by the Implicit Function Theorem applied to the characteristic equation $\Delta_{\boldsymbol{\gamma}}(\lambda) = 0$, there exists a neighborhood of $\boldsymbol{\gamma}_\star$ and a continuously differentiable function $\lambda(\boldsymbol{\gamma})$ tracking the eigenvalue. The sensitivity of the eigenvalue to parameter variations is explicitly given by:*

$$\partial_{\boldsymbol{\gamma}}\lambda = -\frac{\partial_{\boldsymbol{\gamma}}\Delta_{\boldsymbol{\gamma}}(\lambda)}{\partial_\lambda \Delta_{\boldsymbol{\gamma}}(\lambda)}. \tag{4.54}$$

Critically, the definition of a defective eigenvalue (algebraic multiplicity $m > 1$) implies that λ is a multiple root of the characteristic equation, hence $\partial_\lambda \Delta_{\boldsymbol{\gamma}}(\lambda) = 0$. Consequently, at these exceptional points, the denominator in (4.54) vanishes, causing the sensitivity $\partial_{\boldsymbol{\gamma}}\lambda$ to diverge.

This singularity indicates that the smooth dependence breaks down; the eigenvalue $\lambda(\boldsymbol{\gamma})$ exhibits branch-point behavior and is non-differentiable at the exceptional point. Since the eigenfunctions

$\boldsymbol{\Psi}_n(\lambda(\boldsymbol{\gamma}), \boldsymbol{\gamma})$ depend on the parameters through λ, the loss of differentiability of λ implies that the eigenfunctions themselves are not differentiable with respect to $\boldsymbol{\gamma}$. Thus, the coupling coefficients

$$\mathbf{C}^{\boldsymbol{\gamma}}_{nn} = \langle \partial_{\boldsymbol{\gamma}} \boldsymbol{\Psi}_n, \boldsymbol{\varphi}_n \rangle. \tag{4.55}$$

diverge, causing numerical instabilities often reported in the literature but rarely explained. Our analysis clarifies that these are not numerical artifacts but structural features of the operator. To correctly describe the dynamics near these points, one must abandon the assumption of a diagonal basis and include the associated functions $\boldsymbol{\varphi}_n^{(k)}$ to form a complete biorthogonal system (CSEAV). The regularized description of Section 5.4 builds on exactly these associated functions: by transporting the invariant spectral subspace rather than the individual eigenvectors, it keeps the coupling coefficients bounded across the exceptional point (Fig. 5.2).

4.6 Semigroup

The formal study of solutions to the evolution equation (4.5) can be approached via the theory of Markov semigroups associated with the stochastic process governing the membrane potential. In [111], the authors established a formal link, valid for stationary currents ($\dot{\boldsymbol{\gamma}} = 0$) and in the absence of an absolute refractory period, between the integrate-and-fire process on the half-line and its corresponding Fokker–Planck PDE.

Related results were obtained in [77], where Brownian motion on an interval with resetting at the origin was analyzed. The Markov semigroup for the process is defined on continuous functions over a "figure-eight" topology, where the endpoints are identified with the origin. These results were later extended to compact domains in $\mathbb{R}^n$ [19], where the ergodicity of the semigroup and its convergence to equilibrium were studied through the spectral properties of its generator. Further extensions include processes with "holding and jumping boundaries" [152], where trajectories are held at the boundary for an exponentially distributed time before being reset. The spectral properties of such processes, in the driftless Brownian case, were analyzed in [107], where analytical bounds on the spectral gap were established.

While our setting shares similarities with these models, it differs in two biologically motivated aspects: the presence of a reflecting barrier at one endpoint and a deterministic holding time τ_0. The techniques developed in previous works could, in principle, be adapted to account for the reflecting barrier and analyze the evolution equation (4.5) for stationary external currents. In this setting, the semigroup generated by the evolution operator uniquely determines the solution to the abstract Cauchy problem associated with the equation (4.5) in the case of an uncoupled population with a stationary external current. Here, we will briefly establish a generation theorem via dissipative operator theory, we refer to Appendix A.1 for the complete proofs of the statements.

We begin by showing that $\mathcal{T}_{\boldsymbol{\gamma}}$ is a dissipative operator in $\mathbb{L}^1$.

Lemma 4.6.1. *The operator $\mathcal{T}_{\boldsymbol{\gamma}} : {}^1\mathbb{D}_{\boldsymbol{\gamma}} \subset \mathbb{L}^1 \to \mathbb{L}^1$ is dissipative and the following inequality holds:*

$$\forall \mathbf{p} \in {}^1\mathbb{D}_{\boldsymbol{\gamma}} \tag{4.56}$$

$$Re\ \langle \mathrm{sign}(\mathbf{p}), \mathcal{T}_{\boldsymbol{\gamma}} \mathbf{p} \rangle = Re \left(\int_\alpha^\theta \mathrm{sign}\left(p(v)\right) \mathcal{L}_{\boldsymbol{\gamma}} p(v) dv - \int_0^{\tau_0} \mathrm{sign}\left(p^r(\tau)\right) \partial_\tau p^r(\tau) d\tau \right) \leq 0$$

We can then prove the following:

Proposition 4.6.1. $\forall p \in [1,\infty)$ *The spectrum of the operator* $\mathcal{T}_{\boldsymbol{\gamma}}$ *lies closed left-half plane:*

$$\sigma(\boldsymbol{\gamma}) \subset \{\, s \in \mathbb{C} \mid Re(s) \leq 0 \,\} \tag{4.57}$$

and $0 = \lambda_0 \in \sigma(\boldsymbol{\gamma})$ *is the only eigenvalue with non-negative real part.*

By Lumer-Philips Theorem [55, Therem 3.15] we can immediately deduce

Theorem 4.6.1. *The evolution operator* $\mathcal{T}_{\boldsymbol{\gamma}} : {}^1\mathbb{D}_{\boldsymbol{\gamma}} \subset \mathbb{L}^1 \to \mathbb{L}^1$ *is the generator of a positive strongly continuous contraction semigroup* $\left(\mathcal{F}^1_{\boldsymbol{\gamma}}(t)\right)_{t\geq 0}$ *on* $\mathbb{L}^1$.

Given a real positive initial datum $0 < \mathbf{p}_0 \in \mathbb{L}^1$ the probability density $0 < \mathbf{p}_t = \mathcal{F}_1(t)\mathbf{p}_0$ is the unique positive solution of the abstract differential equation (4.5) for stationary current $\boldsymbol{\gamma}(t) \equiv \boldsymbol{\gamma}$. By duality, the adjoint $\mathcal{T}^*_{\boldsymbol{\gamma}} : {}^\infty\mathbb{D} \subset \mathbb{L}^\infty \to \mathbb{L}^\infty$ generates a strongly continuous Markov semigroup when restricted on the closure of ${}^\infty\mathbb{D}_{\boldsymbol{\gamma}} \subset \mathbb{L}^\infty$ [55, p. 62-3] (See also [77, 152]). In our case $\overline{{}^\infty\mathbb{D}_{\boldsymbol{\gamma}}} = \mathrm{C}(P)$ is given by the continuous functions on a "P" shaped domain, union of the two distinct intervals $[\alpha,\theta]$, $[0,\tau_0]$ where we identify H (resp θ) in the first domain with τ_0 (resp. 0) in the second. The adjoint semigroup characterizes the relaxation of observables for uncoupled networks with a stationary external current. Extending this framework to coupled systems with time-dependent input is naturally done within a spectral setting, which underscores the need for a rigorous characterization of the spectrum and eigenfunctions. In the next section, we show how the spectral theory developed here can be used, through a perturbative approach, to derive the linear response and analyze the stability of fixed points in coupled populations with $\tau_0 \geq 0$.

4.7 First order Perturbation theory

In this section, we extend our analysis beyond the stationary case, where $\boldsymbol{\gamma}$ is constant, by studying the linearized dynamics around the stationary solution using first-order perturbation theory. We start by considering an uncoupled population in which the moments of the external current are modulated by a small time-dependent factor, given by $\boldsymbol{\gamma}(t) = \boldsymbol{\gamma}_0 + \boldsymbol{\gamma}_1\varepsilon(t)$, with $\boldsymbol{\gamma}_1 = (\mu_1, D_1)$. We can then write

$$\mathcal{T}_{\boldsymbol{\gamma}(t)} = \mathcal{T}_0 + \mathcal{T}_1\varepsilon(t) \quad \mathcal{B}_{\boldsymbol{\gamma}(t)} = \mathcal{B}_0 + \mathcal{B}_1\varepsilon(t) \tag{4.58}$$

Where we defined $\mathcal{T}_0 := \mathcal{T}_{\boldsymbol{\gamma}_0}$, $\mathcal{B}_0 := \mathcal{B}_{\boldsymbol{\gamma}_0}$ and $\mathcal{T}_1 := \nabla_{\boldsymbol{\gamma}}\mathcal{T}_{\boldsymbol{\gamma}}|_{\boldsymbol{\gamma}_0} \cdot \boldsymbol{\gamma}_1 =: (\mathcal{L}_1, 0)$ and $\mathcal{B}_1 := \nabla_{\boldsymbol{\gamma}}\mathcal{B}_{\boldsymbol{\gamma}}|_{\boldsymbol{\gamma}_0} \cdot \boldsymbol{\gamma}_1$, such that:

$$\begin{aligned}
\mathcal{T}_1\boldsymbol{\varphi}_0 &= (\mathcal{L}_1\phi_0, 0) = \left(-\mu_1\phi_0'(v) + D_1\phi_0''(v), 0\right) \\
\mathcal{B}_1\boldsymbol{\varphi}_0 &= \begin{bmatrix} \mu_1\phi_0(\alpha) - D_1\phi_0'(\alpha) \\ 0 \\ 0 \\ -D_1\left(\phi_0'(H^+) - \phi_0'(H^-)\right) \\ -D_1\phi_0'(\theta, \boldsymbol{\gamma}) \end{bmatrix} = \begin{bmatrix} \mu_1\phi_0(\alpha) - D_1\phi_0'(\alpha) \\ 0 \\ 0 \\ \nu_0 D_1/D_0 \\ \nu_0 D_1/D_0 \end{bmatrix}
\end{aligned} \tag{4.59}$$

To avoid unnecessary clutter, in the rest of the section, we remove the dependence on $\boldsymbol{\gamma}$ on the symbols, which are evaluated as a fixed value $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0$. Expanding the solution as $\mathbf{p}_t = \boldsymbol{\varphi}_0 + \mathbf{p}^1_t + o(\varepsilon(t))$ and neglecting higher-order terms, the first-order perturbation $\mathbf{p}^1_t$ satisfies the following

evolution equation:

$$\begin{cases} \partial_t \mathbf{p}_t^1 - \mathcal{T}_{\boldsymbol{\gamma}} \mathbf{p}_t^1 & = \mathcal{T}_1 \boldsymbol{\varphi}_0 \varepsilon(t) \\ -\mathcal{B}_0 \mathbf{p}_t^1 & = \mathcal{B}_1 \boldsymbol{\varphi}_0 \varepsilon(t) \end{cases} \tag{4.60}$$

Equation (4.60) describes a linear system with an external forcing. As we are interested in the long time behaviour, for simplicity we will always assume the initial condition $\mathbf{p}_0^1 = \mathbf{0}$. The main quantity of interest, in this context, is the frequency domain response of the population to the stimulus, given by the Laplace transform of the firing rate perturbation $\nu_1(t) = \mathcal{N}[\mathbf{p}_t^1]$. This quantity is known in the literature as the "transfer function" [1]:

$$\mathrm{H}(s) := \frac{\widehat{\nu}_1(s)}{\widehat{\varepsilon}(s)} = \frac{\mathcal{N}\left[\widehat{\mathbf{p}}^1(s)\right]}{\widehat{\varepsilon}(s)} \tag{4.61}$$

The Laplace transform of the solution to the forced problem (4.60) is given by the resolvent of the boundary eigenvalue problem, as:

$$\begin{cases} (\mathcal{T}_0 - s)\, \widehat{\mathbf{p}}^1(s) & = -\mathcal{T}_1 \boldsymbol{\varphi}_0 \hat{\varepsilon}(s) \\ \mathcal{B}_0 \widehat{\mathbf{p}}^1(s) & = -\mathcal{B}_1 \boldsymbol{\varphi}_0 \hat{\varepsilon}(s) \end{cases} \tag{4.62}$$

Such that:

$$\widehat{\mathbf{p}}^1(s) = -\left(\mathcal{T}_0 - s\right)^{-1} \mathcal{T}_1 \boldsymbol{\varphi}_0 \hat{\varepsilon}(s) - \mathcal{Z}(s) M^{-1}(s) \mathcal{B}_1 \boldsymbol{\varphi}_0 \hat{\varepsilon}(s) \quad s \in \mathbb{C} \setminus \sigma(\boldsymbol{\gamma}) \tag{4.63}$$

Lemma 4.7.1. *Let $\widehat{\mathbf{p}}^1(s)$ solution of the system (4.62) we have that The transfer function is given by:*

$$\begin{aligned} \mathrm{H}(s) =& \frac{1}{\Delta(s)} \left(\frac{\{f_1, Q\}_{\boldsymbol{\gamma}}(\theta, s)}{w_{\boldsymbol{\gamma}}(\theta)} - \frac{\{f_1, [\![Q]\!]\}_{\boldsymbol{\gamma}}(H, s)}{w_{\boldsymbol{\gamma}}(H)} \right) + \\ &+ \frac{\nu_0 D_1}{D_0} \frac{\psi_1(\theta, s) - \psi_1(H, s)}{\Delta(s)} - \frac{\mu_1 \phi_0(\alpha) - D_1 \phi_0'(\alpha)}{\Delta(s)} . \end{aligned} \tag{4.64}$$

Where $Q(\cdot, s)$, is any particular solution to the problem $(s - \mathcal{L}_0) Q(\cdot, s) = \mathcal{L}_1 \phi_0$. $\{f, h\}_{\boldsymbol{\gamma}} = f \mathcal{S}_{\boldsymbol{\gamma}} h - h \mathcal{S}_{\boldsymbol{\gamma}} f$, and $[\![Q]\!](v, s) = \lim_{u \to 0} Q(v + |u|, s) - Q(v - |u|, s)$.

Proof. We begin by observing that $Q(\cdot, s)$, is any particular solution to the problem $(s - \mathcal{L}_0) Q(\cdot, s) = \mathcal{L}_1 \phi_0$ we have that $\widetilde{\mathbf{Q}}(s) = [Q(\cdot, s), 0] \in \mathbb{W}^p$ is a particular solution to the equation $(\mathcal{T}_0 - s)\, \widehat{\mathbf{p}}^1(s) = -\mathcal{T}_1 \boldsymbol{\varphi}_0 \hat{\varepsilon}(s)$. We then search for a solution of the form $\widehat{\mathbf{p}}^1(s) = \mathcal{Z}(s)\mathbf{g} + \widetilde{\mathbf{Q}}(s)$ where $\mathbf{g} \in \mathbb{C}^5$ is determined by the boundary conditions:

$$B_0 \left(\mathcal{Z}(s)\mathbf{g} + \widetilde{\mathbf{Q}}(s) \right) = -\mathcal{B}_1 \boldsymbol{\varphi}_0 \Rightarrow \tag{4.65}$$

$$\mathbf{g} = -M_0^{-1} B_0 \widetilde{\mathbf{Q}}(s) - M_0^{-1} \mathcal{B}_1 \boldsymbol{\varphi}_0 \tag{4.66}$$

Considering that $\mathcal{N}[\widehat{\mathbf{p}}^1(s)] = \mathcal{N}[\mathcal{Z}(s)\mathbf{g}] + \mathcal{N}[\widetilde{\mathbf{Q}}(s)] = g_5$, the fifth entry of $\mathbf{g}$. Solving the linear system (4.66) after some computation, we arrive at equation (4.64). □

[1] We indicate with $\widehat{\cdot}$ the Laplace transform of a time-dependent function.

Remark 4.7.1. *If we use the particular solution given by the variation of parameters formula (see 4.2.1), we obtain an expression dependent on the eigenfunctions:*

$$\frac{1}{\Delta(s)}\left(\frac{\{f_1,Q\}_{\boldsymbol{\gamma}}(\theta,s)}{w_{\boldsymbol{\gamma}}(\theta)}-\frac{\{f_1,[\![Q]\!]\}_{\boldsymbol{\gamma}}(H,s)}{w_{\boldsymbol{\gamma}}(H)}\right)=\frac{\int_\alpha^\theta \psi_1(v,s)\mathcal{L}_1\phi_0(v)dv}{\Delta(s)}=\frac{\langle \boldsymbol{\Psi}(s),\mathcal{T}_1\boldsymbol{\varphi}_0\rangle}{\Delta(s)} \tag{4.67}$$

We refer to [198, Section III.B] for the details in the derivation.
This expression generalizes the known form of the transfer function for integrate-and-fire neurons without an absolute refractory period. The term in (4.67) corresponds to the standard contribution derived in [198, Eq. 49], with the important distinction that the characteristic equation in the denominator, $\Delta_{\boldsymbol{\gamma}}(s)$ defined in (4.25), now depends explicitly on the absolute refractory period. In addition to this term, two further contributions appear due to the explicit dependence of the boundary conditions on $\boldsymbol{\gamma}$, which was not fully accounted for in previous literature [31, 108, 164, 198]. The first additional term arises from the modulation of the noise intensity. In the limit $\tau_0 \to 0$, it reduces to $\nu_0 D_1/D_0$. We expect similar corrections to appear in more general integrate-and-fire models with state-dependent noise, such as conductance-based neurons. The second term reflects the presence of a reflecting barrier at α. In models with a confining force field $A(v)$, this term vanishes as $\alpha \to -\infty$, but it could play a significant role in linear integrate-and-fire models, where $A(v)$ is constant. The investigation of these effects will be the subject of future work. Here, we will focus on analyzing the effect of the absolute refractory period on the transfer function for the LIF model.

Transfer function of LIF neurons with refractory period

In the case of the LIF neuron, the particular solution $Q(v,s)$ can be expressed in terms of derivatives of the stationary state $\phi_0(v)$ [31]. This allows for a further simplification of the term (4.67), leading to a non-integral expression for the transfer function [198]. By isolating the contributions due to the modulation of drift and noise, the transfer function can be written as: $\mathrm{H} = \mu_1 \mathrm{H}_\mu + D_1 \mathrm{H}_D$. This yields the following expressions for the frequency response of a population of LIF neurons with an absolute refractory period:

$$\begin{aligned}
\mathrm{H}_\mu(s) &= \frac{\nu_0}{s+1}\,\frac{\partial_v\psi_1(\theta,s)-\partial_v\psi_1(H,s)}{\psi_1(\theta,s)-e^{-s\tau_0}\psi_1(H,s)} - \frac{s}{s+1}\,\frac{\phi_0(\alpha)}{\psi_1(\theta,s)-e^{-s\tau_0}\psi_1(H,s)}\\
\mathrm{H}_D(s) &= \frac{\nu_0\big(s[\psi_1(\theta,s)-\psi_1(H,s)]+(\theta-\mu_0)\partial_v\psi_1(\theta,s)-(H-\mu_0)\partial_v\psi_1(H,s)\big)}{D_0(s+2)\left(\psi_1(\theta,s)-e^{-s\tau_0}\psi_1(H,s)\right)}\\
&\quad + \frac{s}{s+2}\,\frac{\phi_0'(\alpha)}{\psi_1(\theta,s)-e^{-s\tau_0}\psi_1(H,s)}
\end{aligned}$$

We refer to [198] for details in the derivation. Figure 4.3 shows the effect of the absolute refractory period on the transfer function. As τ_0 increases, the drift-modulated response peaks become more pronounced at frequencies $\omega = \mathrm{Im}(\lambda_n) \simeq 2\pi n\,\nu_0$; this follows from the inter-spike interval distribution becoming relatively less variable as its mean increases, leading to more regular firing and a smaller coefficient of variation. As we will see, this amplification influences the emergence of limit cycles in coupled populations.

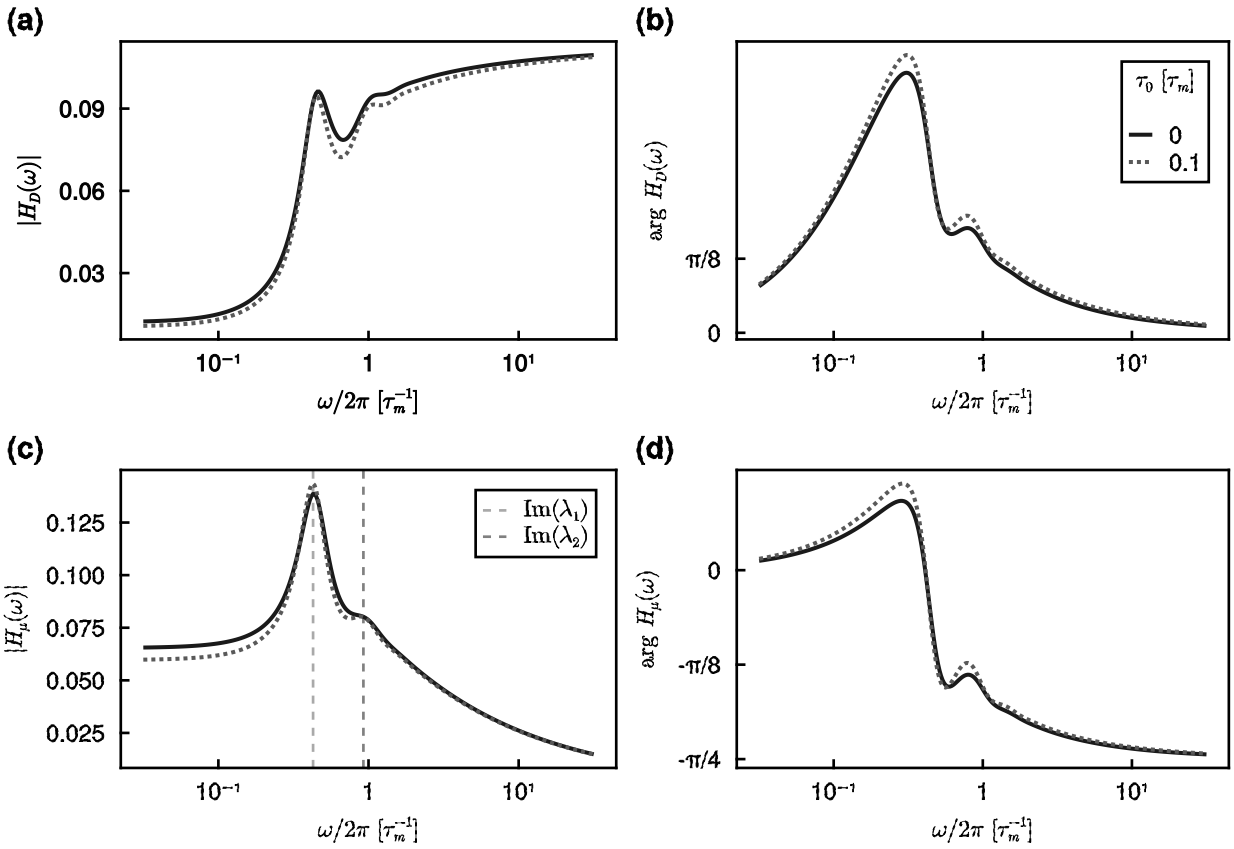

Figure 4.3: Transfer function of a LIF population for varying refractory period τ_0. Module (left) and phase (right) of the response to a noise (H_D, top) and drift (H_μ, bottom) sinusoidal perturbation, with (dotted, $\tau_0 = 0.1\,\tau_m$) or without (solid, $\tau_0 = 0$) absolute refractory period; the vertical dashed lines mark Im $\lambda_n \simeq 2\pi n\,\nu_0$; the abscissa is the ordinary frequency $\omega/2\pi$ in units of τ_m^{-1}. Note the sharpening of the resonance peaks as τ_0 increases. Parameters: $\mu = 21$ mV, $D = 3.54\text{mV}^2$, $\theta = 20$ mV, $H = 0$ mV, $\alpha = -3\theta = -60$ mV, $\nu_0 = 0.4$.

4.7.1 Linear stability for coupled populations

We now consider a coupled neuronal population. Under the 'extended' mean-field approximation [6, 30, 68, 99, 101, 126, 145], the drift and the diffusion coefficients of the Fokker-Planck equation depend self-consistently on past values of the firing rate, making the PDE nonlinear:

$$\boldsymbol{\gamma}(t) = \boldsymbol{\gamma} \cdot \nu(t-\delta) + \boldsymbol{\gamma}_0 \tag{4.68}$$

where $\delta \geq 0$ is a deterministic delay term and $\boldsymbol{\gamma}_0$ represents a stationary external current. Fixed points of the dynamics are given by solutions of the form $\mathbf{p}_t = \boldsymbol{\varphi}_0(\boldsymbol{\gamma})$, where the moments of the current satisfy the self-consistency equation $\boldsymbol{\gamma} = N(\boldsymbol{\varphi}_0(\boldsymbol{\gamma})) + \boldsymbol{\gamma}_0$. Notice that since now the evolution equation is nonlinear the system can display multistability. We can use the linearized first-order dynamics to study the stability of these fixed points. Expanding the equa of the first-order perturbation is given by Equation (4.60) where now $\varepsilon(t) = \nu_1(t-\delta) = \mathcal{N}(\mathbf{p}^1_{t-\delta})$ depends self-consistently on the firing rate perturbation. The delay δ turns the linearized dynamics into a delay differential equation, whose stability, in the framework of delay semigroups [18, 55], is governed by the eigenvalues of the perturbed boundary eigenvalue problem $\mathfrak{T}_0 + \mathcal{K}_1 \in \mathtt{Hol}\left(\mathbb{C}, \mathtt{Lin}\left(\mathbb{W}^p, \mathbb{L}^p \times \mathbb{C}^5\right)\right)$, where $\mathfrak{T}_0 = \mathfrak{T}_{\boldsymbol{\gamma}_0}$ and $\mathcal{K}_1 = \mathcal{K}_1 = \left(\mathcal{K}_1^D, \mathcal{K}_1^R\right)$ is a rank-1 perturbation given by:

$$\mathcal{K}_1^{\mathcal{T}}(s) = \mathcal{T}_1\boldsymbol{\varphi}_0 \cdot e^{-s\delta}\mathcal{N} \quad \mathcal{K}_1^{\mathcal{B}}(s) = \mathcal{B}_1\boldsymbol{\varphi}_0 \cdot e^{-s\delta}\mathcal{N} \tag{4.69}$$

Theorem 4.7.1. *$\mathfrak{T}_0 + \mathcal{K}_1$ is a Fredholm operator valued function such that:*

$$\mathfrak{T}_0 + \mathcal{K}_1 = \mathcal{C}(s)\begin{bmatrix} M(s) + M_1(s) & \mathbf{0} \\ \mathcal{K}_1^D(s)\mathcal{Z}(s) & \mathbf{Id}_{\mathbb{L}^p} \end{bmatrix}\mathcal{D}(s) \tag{4.70}$$

Where $\mathcal{D} = \mathcal{D}_{\boldsymbol{\gamma}_0}$, $\mathcal{C} = \mathcal{C}_{\boldsymbol{\gamma}_0}$, $\mathcal{Z} = \mathcal{Z}_{\boldsymbol{\gamma}_0}$ are the same as Lemma 4.2.1 and $M_1(s)$ is given by:

$$M_1(s) = \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & (\mathcal{B}_1\boldsymbol{\varphi}_0 - \mathcal{B}_0 Q(s)) \cdot e^{-\delta s} \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \end{array}\right] \tag{4.71}$$

where $Q(s) = \mathcal{Q}(s)\mathcal{T}_1\boldsymbol{\varphi}_0$ is the particular solution of the equation $\mathfrak{T}_0^D(s)Q = \mathcal{T}_1\boldsymbol{\varphi}_0$ given by separation of variables (see Remark 4.2.1). Then the spectrum of $\mathfrak{T}_0 + \mathcal{K}_1$ is given by the solution to the characteristic equation:

$$\sigma\left(\mathfrak{T}_0 + \mathcal{K}_1\right) = \{s \in \mathbb{C} | \mathrm{H}(s) = e^{\delta s}\} \tag{4.72}$$

Proof. By Theorem 1.11 in [131] we have that:

$$\mathcal{C}(s)^{-1} = \begin{bmatrix} -B_0\mathcal{Q}(s) & \mathbf{Id}_{\mathbb{C}^5} \\ \mathbf{Id}_{\mathbb{L}^p} & \mathbf{0} \end{bmatrix} \quad \mathcal{D}(s)^{-1} = (\mathcal{Q}(s), \mathcal{Z}(s)) \tag{4.73}$$

Then:

$$\mathfrak{T}_0 + \mathcal{K}_1 = \mathcal{C} \begin{bmatrix} M - B_0 \mathcal{Q} \mathcal{K}_1^D \mathcal{Z} + \mathcal{K}_1^R \mathcal{Z} & -B_0 \mathcal{Q}(s) \mathcal{K}_1^D \mathcal{Q} + \mathcal{K}_1^R \mathcal{Q} \\ \mathcal{K}_1^D \mathcal{Z} & \mathcal{K}_1^D \mathcal{Q} + \mathbf{Id}_{\mathbb{L}^p} \end{bmatrix} \mathcal{D} \tag{4.74}$$

We observe that, as the firing rate of the particular solution is $\mathcal{N}\mathcal{Q}(s) = 0$, we have $\mathcal{K}_1^D \mathcal{Q} = 0$ and $\mathcal{K}_1^R \mathcal{Q} = 0$. From the definition $B_0 \mathcal{Q} \mathcal{K}_1^D(s) \mathcal{Z}(s) = B_0 Q(s) \cdot e^{-\delta s} \mathcal{N} \mathcal{Z}(s)$ and $\mathcal{K}_1^R \mathcal{Z} = B_1 \boldsymbol{\varphi}_0 \cdot e^{-\delta s} \mathcal{N} \mathcal{Z}(s)$ such that $-B_0 \mathcal{Q} \mathcal{K}_1^D \mathcal{Z} + \mathcal{K}_1^R \mathcal{Z} = M_1$. The spectrum of the perturbed boundary value problem $\mathfrak{T}_0 + \mathcal{K}_1$, is then given by the characteristic equation $0 = \det(M(s) + M_1(s))$, computing this matrix determinant we arrive at equation (4.72). □

4.7.2 Emergence of limit cycles in excitatory populations

We apply this stability analysis to a recurrently coupled excitatory population with delay. We consider a fixed point that is stable in the absence of refractoriness ($\tau_0 = 0$) and numerically compute the spectrum of the perturbed boundary eigenvalue problem as τ_0 increases.
Following [126], the spectrum typically contains two types of poles: *diffusion poles*, which originate from the intrinsic Fokker-Planck spectrum, and *transmission poles*, which arise from the delay-induced feedback. Figure 4.4 demonstrates that as the refractory period τ_0 increases, a pair of complex conjugate diffusion poles moves towards the right half-plane. At a critical value of τ_0, these poles cross the imaginary axis (Hopf bifurcation), destabilizing the fixed point and leading to the emergence of a stable limit cycle. We note that the spectral analysis proves only the loss of linear stability of the fixed point; the emergence of a stable limit cycle is verified numerically and not proven here, as this would require a nonlinear bifurcation analysis beyond the present scope.
This highlights the possibility of inducing stable oscillations in the positive drift regime by destabilizing diffusion poles rather than transmission poles. Whereas transmission delays act on the transmissio poles, the refractory period introduces a delay in the "mass" rejection that specifically influences the diffusion poles associated with the single-neuron dynamics. This confirms that the refractory period acts as a singular perturbation that can fundamentally alter the macroscopic network dynamics from a stable focus to a limit cycle oscillator.

4.8 Discussion

In this chapter, we have presented a rigorous mathematical framework for the spectral analysis of neuronal population dynamics with an absolute refractory period. To the best of our knowledge, this is the first work to systematically apply the theory of boundary eigenvalue problems [131] to the Fokker-Planck operator governing population density dynamics, moving beyond the heuristic treatments often found in the literature.
Our central contribution is the analytical characterization of the spectrum, the resolvent, and the eigenfunctions of the evolution operator. By extending the state space to include the refractory history, we recovered a Markovian description amenable to rigorous operator theoretic analysis. We proved the dissipativity of the generator (Section4.6), a fundamental property ensuring physical stability and the existence of a contraction semigroup. This formally demonstrates that a decoupled population always relaxes to its unique stationary state.

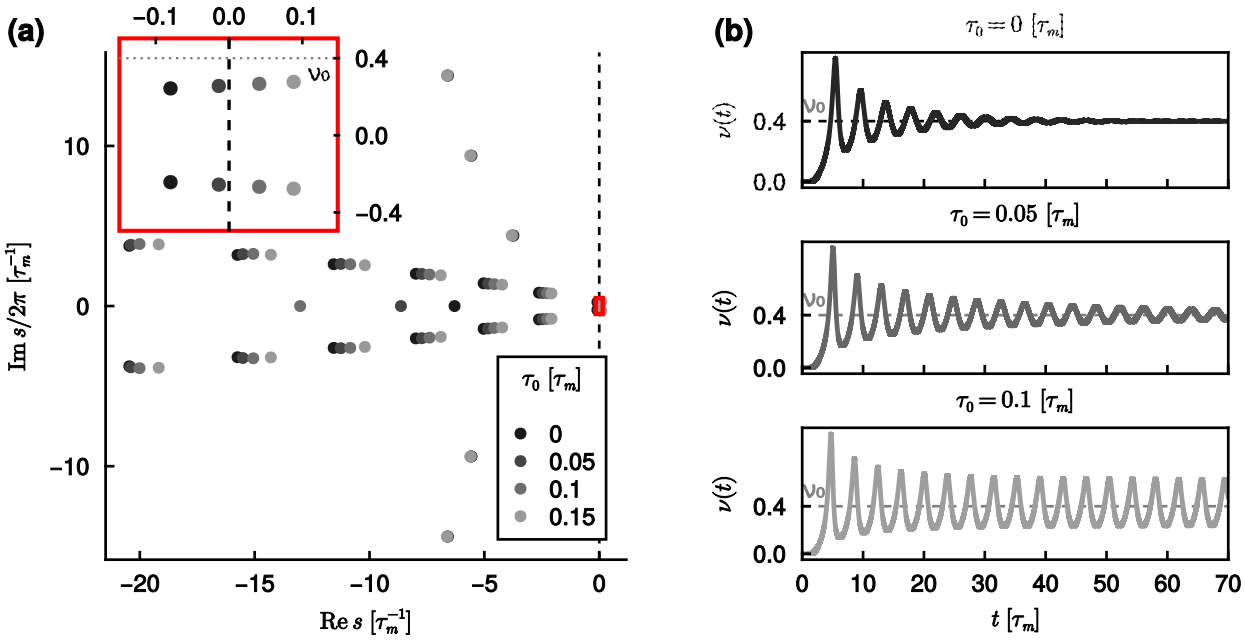

Figure 4.4: (a) Locus of the leading eigenvalues (poles) of the coupled system as the refractory period τ_0 is varied; the inset magnifies the pair closest to the imaginary axis. The crossing of the imaginary axis by the diffusion poles indicates a Hopf bifurcation and the onset of global oscillations. (b) Population firing rate $\nu(t)$ from the corresponding Fokker–Planck simulations at $\tau_0 = 0$, 0.05 and $0.1\,\tau_m$ (top to bottom): the damped relaxation towards ν_0 (dashed) at $\tau_0 = 0$ gives way to a sustained limit cycle once the diffusion pair has crossed. Same single-neuron parameters as Figure 4.3, with $\nu_0 = 0.4$, $K = 1000$, $J = 0.012$ ($KJ = 12$), transmission delay $\tau_D = 0.2\,\tau_m$, and τ_0 ranging over $\{0, 0.05, 0.1, 0.15\}\,\tau_m$

A key insight from our analysis is the identification and characterization of defective eigenvalues (Section 4.5). These spectral singularities, often overlooked or treated as numerical instabilities, represent structural features of the operator where the algebraic multiplicity exceeds the geometric one. We have shown that they correspond to exceptional points in the parameter space where oscillatory modes emerge from the coalescence of relaxational modes. Properly accounting for the associated generalized eigenfunctions is essential for a complete description of the system's transient dynamics near these critical points.

Furthermore, our rigorous treatment of the boundary conditions has refined the linear response theory for neuronal populations. We derived an exact expression for the transfer function (Section 4.7) that explicitly accounts for boundary conditions modulated by external parameters. This correction resolves ambiguities in previous derivations and reveals new contributions to the population response, specifically terms related to noise modulation at the threshold, that are likely to be significant in conductance-based or state-dependent noise models [108, 164].

We also addressed the role of the lower boundary α. While often sent to $-\infty$ in theoretical studies of confining potentials (e.g., Ornstein-Uhlenbeck processes), a finite reflecting barrier is biologically plausible for neurons with reversal potentials bounded by ion channel physics. Our results show that this boundary introduces specific spectral signatures, although these vanish in the limit of strong confinement.

This chapter lays the foundation for a deeper mathematical exploration of population density equations. Two major open questions remain: the rigorous proof of the completeness of the eigenbasis and the detailed regularity properties of the generated semigroup (e.g., analyticity). Establish-

ing completeness is the necessary next step to fully validate the spectral expansion methods used throughout computational neuroscience. In the next chapter, we will tackle this challenge directly by establishing the Birkhoff regularity of the boundary value problem and proving the sectoriality of the generator.

Chapter 5

Spectral decomposition of the dynamics

> Self-respect is a question of recognizing that anything worth having has a price
>
> — Joan Didion, *On Self Respect*

In this chapter, we study the spectral decomposition of the population dynamics generated by the operator associated with the evolution equation. Throughout the remainder of this thesis, we assume $\tau_0 = 0$, which corresponds to the absence of an absolute refractory period. Consequently, the auxiliary compartment $(0, \tau_0)$ is discarded, and the dynamics are described by a probability density on the union of two intervals.

5.1 Eigenfunction Expansion

Following standard practice in the literature (e.g., [131]), we parameterize the complex spectral parameter by setting $\lambda = -z^2$. We consider the boundary eigenvalue problem associated with the operator $\mathcal{L}_{\boldsymbol{\gamma}}$, defined by the pencil

$$\mathfrak{L}_{\boldsymbol{\gamma}}(z) = (\mathfrak{L}^D, \mathfrak{L}^R) : \mathbb{W}^p \to \mathbb{L}^p \times \mathbb{C}^4, \tag{5.1}$$

$$f \mapsto \left(\mathcal{L}_{\boldsymbol{\gamma}} f + z^2 f,\ \mathcal{B}_{\boldsymbol{\gamma}} f\right). \tag{5.2}$$

This pencil corresponds to $\mathfrak{T}_{\boldsymbol{\gamma}}(-z^2)$ from the previous chapter under the restriction $\tau_0 = 0$. Upon discarding the compartment $(0, \tau_0)$, the function spaces reduce to

$$\mathbb{L}^p = \mathrm{L}^p(\alpha, H) \times \mathrm{L}^p(H, \theta), \tag{5.3}$$

$$\mathbb{W}^p = \mathrm{W}^{2,p}(\alpha, H) \times \mathrm{W}^{2,p}(H, \theta), \tag{5.4}$$

and the boundary operator $\mathcal{B}_{\boldsymbol{\gamma}} : \mathbb{W}^p \to \mathbb{C}^4$ takes the simplified form

$$\mathcal{B}_{\boldsymbol{\gamma}}\mathbf{p} = \begin{bmatrix} \big(b(v)+\mu\big)p^-(\alpha) - D\partial_v p^-(\alpha) \\ p^-(\theta) \\ p^+(H) - p^-(H) \\ \partial_v p^+(\theta) - \partial_v p^+(H) + \partial_v p^-(H) \end{bmatrix}, \tag{5.5}$$

where $\mathbf{p} = (p^-, p^+)$ with $p^- \in \mathtt{W}^{2,p}(\alpha, H)$ and $p^+ \in \mathtt{W}^{2,p}(H, \theta)$.
Since $\mathfrak{L}_{\boldsymbol{\gamma}}(z) = \mathfrak{T}_{\boldsymbol{\gamma}}(-z^2)$ with $\tau_0 = 0$, the spectral properties established in the previous chapter carry over directly. In particular, the spectrum $\sigma(\mathcal{L}_{\boldsymbol{\gamma}})$ consists of a countable set of eigenvalues $\{\lambda_n\}_{n\in\mathbb{Z}}$ with finite algebraic multiplicities, accumulating only at infinity.
For the spectral decomposition of the adjoint dynamics, we consider the adjoint operator $\mathcal{L}^*_{\boldsymbol{\gamma}}$ with domain

$$\mathbb{D}^* := \left\{ f \in \mathbb{V}^q \,\middle|\, \mathcal{B}^{\oplus} f = 0 \right\}, \tag{5.6}$$

where $\mathbb{V}^q := \mathrm{W}^{2,q}(\alpha, \theta)$ and $\mathcal{B}^{\oplus} : \mathbb{V}^q \to \mathbb{C}^2$ denotes the reduced adjoint boundary operator defined by

$$\mathcal{B}^{\oplus} : \mathbb{V}^q \to \mathbb{C}^2, \tag{5.7}$$

$$f \mapsto \begin{pmatrix} f'(\alpha) \\ f(\theta) - f(H) \end{pmatrix}. \tag{5.8}$$

A spectral decomposition of the dynamics is achievable if observables can be expanded in terms of the eigenfunctions of $\mathcal{L}^*_{\boldsymbol{\gamma}}$. In the next section, we establish this property by proving that the associated boundary eigenvalue problem is *Birkhoff regular* in the sense of Mennicken and Möller [131]. This regularity ensures the completeness of the corresponding system of root functions and provides convergence results for eigenfunction expansions. The adjoint boundary eigenvalue problem is defined by the pencil

$$\mathfrak{L}^+_{\boldsymbol{\gamma}}(z) = (\mathfrak{L}^{D+}, \mathfrak{L}^{R+}) : \mathbb{V}^q \to \mathbb{L}^q \times \mathbb{C}^2, \tag{5.9}$$

$$f \mapsto \left(\mathcal{L}^*_{\boldsymbol{\gamma}} f + z^2 f,\ \mathcal{B}^{\oplus} f\right). \tag{5.10}$$

5.1.1 Birkhoff regularity of the adjoint

To establish Birkhoff regularity of the boundary eigenvalue problem $\mathfrak{L}^+_{\boldsymbol{\gamma}}(z)$, we work with the associated first-order system. Introducing the vector $\tilde{\mathbf{y}} = (f, f')^\top \in (\mathrm{L}^q(\alpha, \theta))^2$, the eigenvalue equation $\mathcal{L}^*_{\boldsymbol{\gamma}} f + z^2 f = 0$ together with the boundary conditions $\mathcal{B}^{\oplus} f = 0$ can be written as

$$\tilde{\mathbf{y}}'(v) - \begin{pmatrix} 0 & 1 \\ -\dfrac{z^2}{D} & -\dfrac{b(v)+\mu}{D} \end{pmatrix} \tilde{\mathbf{y}}(v) = \mathbf{0}, \tag{5.11}$$

with boundary conditions

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tilde{\mathbf{y}}(\alpha) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tilde{\mathbf{y}}(\theta) + \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \tilde{\mathbf{y}}(H) = 0. \tag{5.12}$$

To bring this system into the standard form required by the theory of Birkhoff regular boundary eigenvalue problems [131], we apply the z-dependent transformation

$$\tilde{\mathbf{y}} = C(z)\mathbf{y}, \qquad C(z) = \begin{pmatrix} 1 & 1 \\ \dfrac{zi}{\sqrt{D}} & -\dfrac{zi}{\sqrt{D}} \end{pmatrix}. \tag{5.13}$$

This transformation diagonalises the principal part with respect to the spectral parameter z and yields the standard form:

$$\mathbf{y}'(v) - \big(zA_1 + A_0(v)\big)\mathbf{y}(v) = 0, \tag{5.14}$$

$$W^{(0)}\mathbf{y}(\alpha) + W^{(1)}\mathbf{y}(\theta) + W^{(2)}\mathbf{y}(H) = 0, \tag{5.15}$$

where:

$$A_1 = \begin{pmatrix} \dfrac{i}{\sqrt{D}} & 0 \\ 0 & -\dfrac{i}{\sqrt{D}} \end{pmatrix}, \qquad A_0(v) = -\frac{\mu + b(v)}{2D} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{5.16}$$

and the boundary matrices are given by

$$W^{(0)} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad W^{(1)} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad W^{(2)} = \begin{pmatrix} 0 & 0 \\ -1 & -1 \end{pmatrix}. \tag{5.17}$$

The boundary functionals here are normalised so that the $W^{(j)}$ are constant in z: under $\tilde{\mathbf{y}} = C(z)\mathbf{y}$ the condition at α acquires a factor $zi/\sqrt{D}$, removed by left-multiplying that row by $\sqrt{D}/(zi)$ (admissible for $z \neq 0$). Note that A_1 is a diagonal matrix with nonzero entries of opposite imaginary parts. This places the problem precisely in the framework of 2×2 systems treated in Chapter IV of [131].

Proposition 5.1.1. *The boundary eigenvalue problem* (5.14)–(5.15) *is Birkhoff regular in the sense of Definition 4.1.2 in [131].*

Proof. Birkhoff regularity requires that the matrix

$$B(z) = W^{(0)}\big(I - \triangle(z)\big) + W^{(1)} \triangle(z) \tag{5.18}$$

is invertible for all $z \in \mathbb{C} \setminus \{0\}$. Here $\triangle(z)$ is the Birkhoff projector (Definition 4.1.22 in [131]), which for the present 2×2 system with diagonal leading matrix A_1 having entries $\pm i/\sqrt{D}$ is given

by

$$\triangle(z) = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \text{if } \operatorname{Im}(z) < 0 \text{ or } \big(\operatorname{Im}(z) = 0 \wedge \operatorname{Re}(z) > 0\big), \\ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } \operatorname{Im}(z) > 0 \text{ or } \big(\operatorname{Im}(z) = 0 \wedge \operatorname{Re}(z) < 0\big). \end{cases} \tag{5.19}$$

From equations (5.19) and (5.17), a direct computation yields the Birkhoff matrix in each region:

- If $\operatorname{Im}(z) < 0$ or $(\operatorname{Im}(z) = 0 \wedge \operatorname{Re}(z) > 0)$, then $\triangle(z) = \operatorname{diag}(1, 0)$ and

$$B(z) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{5.20}$$

  which has determinant 1.
- If $\operatorname{Im}(z) > 0$ or $(\operatorname{Im}(z) = 0 \wedge \operatorname{Re}(z) < 0)$, then $\triangle(z) = \operatorname{diag}(0, 1)$ and

$$B(z) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5.21}$$

  which is also invertible.

On the real axis both admissible values of $\triangle(z)$ give an invertible $B(z)$, so the boundary assignment there is immaterial. Hence $B(z)$ is invertible for all $z \in \mathbb{C}\backslash\{0\}$, and the first-order problem is Birkhoff regular. □

Remark 5.1.1. *By Definition 7.3.1 in [131], Birkhoff regularity of the first-order system* (5.14)–(5.15) *implies Birkhoff regularity of the original boundary eigenvalue problem* $\mathfrak{L}^+_{\boldsymbol{\gamma}}(z)$*, since the transformation from the second-order scalar equation to the first-order system is of the type treated in Theorem 7.2.4 of [131].*

5.1.2 Spectral projections and expansion theorem

As a consequence of Birkhoff regularity, Theorem 4.3.9 in [131] guarantees the existence of a sequence of circles $\Gamma_k = \{z \in \mathbb{C} : |z| = r_k\}$ with $r_k \to \infty$ such that the characteristic determinant of $\mathfrak{L}^+_{\boldsymbol{\gamma}}(z)$ is bounded away from zero on $\bigcup_k \Gamma_k$. In particular, the operator $\mathfrak{L}^+_{\boldsymbol{\gamma}}(z)$ is invertible for all z with $|z| = r_k$.

This allows us to define spectral projections via contour integration. With the canonical embedding $\mathbb{W}^q \hookrightarrow \mathbb{L}^q$, we define

$$\mathcal{P}^*_{\Gamma_k}(\boldsymbol{\gamma}) : \mathbb{L}^q \to \mathbb{L}^q, \qquad \mathcal{P}^*_{\Gamma_k}(\boldsymbol{\gamma}) f = +\frac{1}{2\pi i} \oint_{|z|=r_k} z\, \mathfrak{L}^+_{\boldsymbol{\gamma}}(z)^{-1}(f, 0)\, dz. \tag{5.22}$$

Using the substitution $\lambda = -z^2$, the resolvent $(\mathcal{L}^*_{\boldsymbol{\gamma}} - \lambda)^{-1}$ can be expressed in terms of $\mathfrak{L}^+_{\boldsymbol{\gamma}}(z)^{-1}$. The factor z is the Jacobian of the substitution $\lambda = -z^2$, equivalently the frame factor of the first-order reduction (cf. Proposition 7.4.1 in [131]), and the circle $\{|z| = r_k\}$ covers $\{|\lambda| = r_k^2\}$ twice. A change

of variables then yields the equivalent representation

$$\mathcal{P}^*_{\Gamma_k}(\boldsymbol{\gamma})f = -\frac{1}{2\pi i}\oint_{|\lambda|=r_k^2}(\mathcal{L}^*_{\boldsymbol{\gamma}}-\lambda)^{-1}f\,d\lambda. \tag{5.23}$$

By the residue theorem and the general theory of operator functions with Fredholm resolvent, this projection can be expanded in terms of the eigenvalues λ_n of $\mathcal{L}^*_{\boldsymbol{\gamma}}$ inside the disc $|\lambda| \leq r_k^2$. If λ_n has algebraic multiplicity m_n, and $\{\boldsymbol{\varphi}_n^{(0)},\dots,\boldsymbol{\varphi}_n^{(m_n-1)}\}$ and $\{\boldsymbol{\Psi}_n^{(0)},\dots,\boldsymbol{\Psi}_n^{(m_n-1)}\}$ are biorthogonal systems of eigenfunctions and associated functions (CSEAV) of $\mathcal{L}_{\boldsymbol{\gamma}}$ and $\mathcal{L}^*_{\boldsymbol{\gamma}}$ at λ_n, respectively, then

$$\mathcal{P}^*_{\Gamma_k}(\boldsymbol{\gamma})f = \sum_{\substack{\lambda_n\in\sigma(\boldsymbol{\gamma}):\\ |\lambda_n|\leq r_k^2}}\sum_{\ell=0}^{m_n-1}\langle f,\boldsymbol{\varphi}_n^{(\ell)}\rangle\,\boldsymbol{\Psi}_n^{(m_n-1-\ell)}. \tag{5.24}$$

The main expansion theorems for Birkhoff regular boundary eigenvalue problems are Theorems 7.4.3 and 7.4.4 in [131]. Adapted to the present setting, they yield:

Theorem 5.1.1 (Eigenfunction expansion)**.** *Let $k \in \mathbb{N}$ and let $\mathcal{P}^*_{\Gamma_k}(\boldsymbol{\gamma})$ be the spectral projection defined by* (5.24)

1. *Let $q \in (1,\infty)$ and $f \in \mathbb{L}^q$. Then*

$$\mathcal{P}^*_{\Gamma_k}(\boldsymbol{\gamma})f \xrightarrow[k\to\infty]{\mathbb{L}^q} f. \tag{5.25}$$

   In particular, the system of root functions of $\mathcal{L}^*_{\boldsymbol{\gamma}}$ is complete in $\mathbb{L}^q$.

2. *Let $q=\infty$ and $f \in \mathtt{C}[\alpha,\theta]\cap\mathtt{BV}[\alpha,\theta]$. If f satisfies the condition*

$$f(\theta)-f(H)=0, \tag{5.26}$$

   then

$$\mathcal{P}^*_{\Gamma_k}(\boldsymbol{\gamma})f \xrightarrow[k\to\infty]{\mathbb{L}^\infty} f. \tag{5.27}$$

Proof. The first assertion follows directly from Theorem 7.4.3 in [131], which establishes the $\mathbb{L}^q$-convergence of the spectral projections for any Birkhoff regular boundary eigenvalue problem when $1 < q < \infty$.

For the second assertion, we use Theorem 7.4.4 in [131]. This theorem guarantees uniform convergence for functions in $\mathtt{C}[\alpha,\theta]\cap\mathtt{BV}[\alpha,\theta]$ that lie in the kernel of specific limiting boundary functionals. For convergence in $\mathbb{L}^\infty$, the only relevant BC are those of order zero (i.e., involving function evaluations but not derivatives). □

5.1.3 Birkhoff regularity of the Fokker-Planck operator

Since the characteristic determinants of the operator $\mathcal{L}_{\boldsymbol{\gamma}}$ and its adjoint $\mathcal{L}^*_{\boldsymbol{\gamma}}$ differ only by a non-vanishing constant factor $D\,w_{\boldsymbol{\gamma}}(H)w_{\boldsymbol{\gamma}}(\theta) > 0$ which is independent of the spectral parameter, the lower bound on the characteristic determinant established for the adjoint operator remains valid for $\mathcal{L}_{\boldsymbol{\gamma}}$ on the same sequence of regularity circles Γ_k. We may therefore define the spectral projections

for the original dynamics analogously:

$$\mathcal{P}_{\Gamma_k}(\boldsymbol{\gamma}) : \mathbb{L}^p \to \mathbb{L}^p, \tag{5.28}$$

$$f \mapsto -\frac{1}{2\pi i} \oint_{|s|=r_k^2} (\mathcal{L}_{\boldsymbol{\gamma}} - s)^{-1} f \, ds = \sum_{\substack{\lambda_n \in \sigma(\mathcal{L}_{\boldsymbol{\gamma}}): \\ |\lambda_n| \le r_k^2}} \sum_{\ell=0}^{m_n-1} \langle \boldsymbol{\Psi}_n^{(\ell)}, f \rangle \boldsymbol{\varphi}_n^{(m_n-1-\ell)}. \tag{5.29}$$

This common bound ensures the completeness of the system of root functions $\{\boldsymbol{\varphi}_n^{(\ell)}\}_{n,\ell}$ in the same sense as for the adjoint.

For completeness, we provide a direct proof of the Birkhoff regularity for the boundary eigenvalue problem associated with $\mathcal{L}_{\boldsymbol{\gamma}}$. This problem is defined by the operator pencil

$$\mathfrak{L}_{\boldsymbol{\gamma}}(z) = (\mathfrak{L}^D, \mathfrak{L}^R) : \mathbb{W}^p \to \mathbb{L}^p \times \mathbb{C}^4, \tag{5.30}$$

$$f \mapsto \left(\mathcal{L}_{\boldsymbol{\gamma}} f + z^2 f, \ \mathcal{B}_{\boldsymbol{\gamma}} f \right), \tag{5.31}$$

which corresponds to the pencil $\mathfrak{T}(-z^2)$ studied previously (with $\tau_0 = 0$). Recall that in the absence of the refractory period, the state space is the union of two intervals (α, H) and (H, θ). To apply the standard theory for first-order systems on a single interval, we map both compartments to the reference interval $(0, 1)$ via the affine transformation

$$v(x) = \begin{cases} v_-(x) = (H - \alpha)x + \alpha, & x \in (0,1) \quad \text{compartment } (\alpha, H), \\ v_+(x) = (\theta - H)x + H, & x \in (0,1) \quad \text{compartment } (H, \theta). \end{cases} \tag{5.32}$$

The associated differential operators transform according to the chain rule:

$$\partial_v f^- = \frac{1}{H - \alpha} \partial_x f^-, \qquad \partial_v^2 f^- = \frac{1}{(H-\alpha)^2} \partial_x^2 f^-, \tag{5.33}$$

$$\partial_v f^+ = \frac{1}{\theta - H} \partial_x f^+, \qquad \partial_v^2 f^+ = \frac{1}{(\theta - H)^2} \partial_x^2 f^+. \tag{5.34}$$

Substituting these into the boundary conditions $\mathcal{B}_{\boldsymbol{\gamma}} \mathbf{p} = 0$ yields the transformed conditions on $(0, 1)$:

$$\mathcal{B}_{\boldsymbol{\gamma}} \mathbf{p} = \begin{bmatrix} (b(\alpha) + \mu) p^-(0) - \frac{D}{H-\alpha} \partial_x p^-(0) \\ p^+(1) \\ p^+(0) - p^-(1) \\ \frac{1}{\theta - H} \partial_x p^+(1) - \frac{1}{\theta - H} \partial_x p^+(0) + \frac{1}{H-\alpha} \partial_x p^-(1) \end{bmatrix}. \tag{5.35}$$

We now cast the eigenvalue problem $\mathcal{L}_{\boldsymbol{\gamma}} f + z^2 f = 0$, as a first-order system for the vector-valued function $\tilde{\mathbf{y}} = (f^-, \partial_x f^-, f^+, \partial_x f^+)^\top \in (\mathrm{L}^p(0,1))^4$. The resulting system is:

$$\tilde{\mathbf{y}}'(x) - \mathcal{A}(x, z) \tilde{\mathbf{y}}(x) = 0, \tag{5.36}$$

where the matrix $\mathcal{A}(x,z)$ contains the coefficients of the differential equation; explicitly,

$$\mathcal{A}(x,z) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{(H-\alpha)^2}{D}\big(z^2 - b'(v_-(x))\big) & +\frac{H-\alpha}{D}\big(b(v_-(x)) + \mu\big) & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{(\theta-H)^2}{D}\big(z^2 - b'(v_+(x))\big) & +\frac{\theta-H}{D}\big(b(v_+(x)) + \mu\big) \end{pmatrix}.$$

To bring this system into the canonical form required for Birkhoff regularity analysis, we apply the z-dependent transformation

$$\tilde{\mathbf{y}} = C(z)\mathbf{y}, \quad C(z) = \begin{pmatrix} C_-(z) & 0 \\ 0 & C_+(z) \end{pmatrix}, \tag{5.37}$$

where the blocks $C_\pm(z)$ are defined by

$$C_-(z) = \begin{pmatrix} 1 & 1 \\ \frac{zi(H-\alpha)}{\sqrt{D}} & -\frac{zi(H-\alpha)}{\sqrt{D}} \end{pmatrix}, \quad C_+(z) = \begin{pmatrix} 1 & 1 \\ \frac{zi(\theta-H)}{\sqrt{D}} & -\frac{zi(\theta-H)}{\sqrt{D}} \end{pmatrix}. \tag{5.38}$$

This transformation diagonalizes the principal part of the system with respect to the spectral parameter z, leading to the standard form:

$$\mathbf{y}'(x) - \big(zA_1 + A_0(x) + z^{-1}A_{-1}(x)\big)\mathbf{y}(x) = 0, \tag{5.39}$$

subject to the boundary conditions

$$W^{(0)}(z)\mathbf{y}(0) + W^{(1)}(z)\mathbf{y}(1) = 0. \tag{5.40}$$

Here, the leading coefficient matrix A_1 is (explicitly) given by

$$A_1 = \begin{pmatrix} \frac{i(H-\alpha)}{\sqrt{D}} & 0 & 0 & 0 \\ 0 & -\frac{i(H-\alpha)}{\sqrt{D}} & 0 & 0 \\ 0 & 0 & \frac{i(\theta-H)}{\sqrt{D}} & 0 \\ 0 & 0 & 0 & -\frac{i(\theta-H)}{\sqrt{D}} \end{pmatrix}. \tag{5.41}$$

The boundary matrices $W^{(0)}(z)$ and $W^{(1)}(z)$ are derived by substituting $\tilde{\mathbf{y}} = C(z)\mathbf{y}$ into the transformed boundary conditions:

$$W^{(0)}(z) = \begin{pmatrix} 1 + \frac{b(\alpha)+\mu}{\sqrt{D}z}i & -1 + \frac{b(\alpha)+\mu}{\sqrt{D}z}i & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \tag{5.42}$$

$$W^{(1)}(z) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & 0 \\ 1 & -1 & 1 & -1 \end{pmatrix}. \tag{5.43}$$

To check for Birkhoff regularity, we consider the asymptotic limit of these matrices as $|z| \to \infty$:

$$W_0^{(0)} := \lim_{z\to\infty} W^{(0)}(z) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \tag{5.44}$$

$$W_0^{(1)} := \lim_{z\to\infty} W^{(1)}(z) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & 0 \\ 1 & -1 & 1 & -1 \end{pmatrix}. \tag{5.45}$$

Proposition 5.1.2. *The boundary eigenvalue problem defined by* (5.39)–(5.40) *is Birkhoff regular in the sense of Definition 4.1.2 in [131].*

Proof. Birkhoff regularity requires the invertibility of the characteristic boundary matrix $B(z)$ for all $z \in \mathbb{C} \setminus \{0\}$. This matrix is constructed from the asymptotic boundary matrices $W_0^{(0)}, W_0^{(1)}$ and the Birkhoff projector $\triangle(z)$. Since A_1 is diagonal with purely imaginary entries, $\triangle(z)$ is the projection onto the subspaces corresponding to the exponentially decaying modes. Explicitly,

$$\triangle(z) = \begin{cases} \operatorname{diag}(1,0,1,0) & \text{if } \operatorname{Im} z < 0 \vee (\operatorname{Im} z = 0 \wedge \operatorname{Re} z > 0), \\ \operatorname{diag}(0,1,0,1) & \text{if } \operatorname{Im} z > 0 \vee (\operatorname{Im} z = 0 \wedge \operatorname{Re} z < 0). \end{cases} \tag{5.46}$$

The Birkhoff matrix is defined as

$$B(z) = W_0^{(0)}(I - \triangle(z)) + W_0^{(1)} \triangle(z). \tag{5.47}$$

Evaluating the determinant of $B(z)$ in the two complementary regions of the complex plane yields $|\det B(z)| = 2$ in both cases. Since the determinant is non-zero, $B(z)$ is invertible for all $z \neq 0$, which proves that the problem is Birkhoff regular. □

Theorem 5.1.2 (Eigenfunction expansion for $\mathcal{L}_{\boldsymbol{\gamma}}$)**.** *Let $k \in \mathbb{N}$ and let $\mathcal{P}_{\Gamma_k}(\boldsymbol{\gamma})$ be the spectral projection defined above.*

1. *If $p \in (1,\infty)$ and $f \in \mathbb{L}^p$, then*

$$\mathcal{P}_{\Gamma_k}(\boldsymbol{\gamma}) f \xrightarrow[k\to\infty]{\mathbb{L}^p} f. \tag{5.48}$$

2. *If $p = \infty$ and $f \in \mathtt{C}[\alpha,\theta] \cap \mathtt{BV}[\alpha,\theta]$ satisfies the essential boundary condition $f(\theta) = 0$, then*

$$\mathcal{P}_{\Gamma_k}(\boldsymbol{\gamma}) f \xrightarrow[k\to\infty]{\mathbb{L}^\infty} f. \tag{5.49}$$

Proof. The first part follows from Theorem 7.4.3 in [131]. The second part follows from Theorem 7.4.4 in [131], noting that for the Fokker-Planck operator, the order-zero boundary condition relevant for uniform convergence is the absorption condition at the threshold, $f(\theta) = 0$. □

5.2 Eigenvalue Asymptotics

In this section, we derive the asymptotic distribution of the eigenvalues λ_n for large modulus. We utilize the asymptotic analysis of the fundamental system of the differential equation as developed in [131, Chapter 4].

Asymptotic Fundamental Matrix

We consider the first-order system formulation of the adjoint boundary value problem associated with the operator $\mathfrak{L}^+(z)$, where the spectral parameter is $\lambda = -z^2$.
According to [131, Theorem 2.8.2], for sufficiently large $|z|$, there exists a fundamental matrix solution $Y(v,z)$ which admits the asymptotic representation:

$$Y(v,z) = \left(P^{[0]}(v) + O(z^{-1})\right) E(v,z), \tag{5.50}$$

where $E(v,z)$ is the diagonal exponential matrix capturing the leading oscillations, and $P^{[0]}(v)$ is the transformation matrix diagonalising the leading coefficient. In our specific case, these matrices are given by:

$$E(v,z) = \begin{pmatrix} e^{iz\ell(v)} & 0 \\ 0 & e^{-iz\ell(v)} \end{pmatrix}, \tag{5.51}$$

$$P^{[0]}(v) = \begin{pmatrix} e^{-U(v)/2} & 0 \\ 0 & e^{-U(v)/2} \end{pmatrix}, \quad \text{where } U(v) = \int_\alpha^v \frac{b(u)+\mu}{D}\, du. \tag{5.52}$$

We denote $\ell(v) = (v-\alpha)/\sqrt{D}$ for brevity.
Let $M^{\oplus}(-z^2)$ denote the characteristic matrix of the adjoint boundary eigenvalue problem constructed using the fundamental solution $Y(v,z)$. The determinant $\det M^{\oplus}(-z^2)$ is an entire function whose zeros correspond to the eigenvalues. However, this determinant grows exponentially as $|\mathrm{Im}(z)| \to \infty$. To apply the spectral distribution results of [131], we must introduce a normalized characteristic matrix function bounded in the complex plane.
Following the construction in [131, Eq. 4.3.4], we define the normalized matrix $\tilde{M}(z)$ by strictly factoring out the exponential growth. The sectorial projection $\triangle(z)$ selects the growing exponential modes. We then define:

$$\tilde{M}(z) := M^{\oplus}(-z^2)\left[(I - \triangle(z)) + E(\theta,z)^{-1}\triangle(z)\right]. \tag{5.53}$$

The linear transformation on the right-hand side acts as a renormalization factor. Specifically, it relates the modulus of the raw determinant to that of the normalized determinant via the factor:

$$\left|\det M^{\oplus}(-z^2)\right| = e^{\ell(\theta)|\mathrm{Im}(z)|}\left|\det \tilde{M}(z)\right|. \tag{5.54}$$

According to [131, Corollary 4.3.4], the matrix function $\tilde{M}(z)$ is uniformly bounded in $\mathbb{C}\setminus\{0\}$. Crucially, since the exponential factor $e^{\ell(\theta)|\mathrm{Im}(z)|}$ never vanishes, the zeros of $\tilde{M}(z)$ coincide exactly with the zeros of $M^{\oplus}(-z^2)$.

The Asymptotic Limit Matrix M_1

To locate the zeros, we approximate $\tilde{M}(z)$ by its asymptotic limit for large $|z|$. We define the asymptotic characteristic matrix $M_1(z)$ by replacing $Y(v,z)$ with its leading term $P^{[0]}(v)E(v,z)$ in the boundary conditions. Based on the boundary matrices W_0, W_1, W_2 corresponding to points α, θ, H, the asymptotic matrix is defined as:

$$\begin{aligned} M_1(z) := W_0 P^{[0]}(\alpha) E(\alpha, z) \mathcal{N}(z) \\ + W_1 P^{[0]}(\theta) E(\theta, z) \mathcal{N}(z) \\ + W_2 P^{[0]}(H) E(H, z) \mathcal{N}(z), \end{aligned} \tag{5.55}$$

where $\mathcal{N}(z) = (I - \triangle(z)) + E(\theta, z)^{-1} \triangle(z)$ is the normalization factor. From the asymptotic estimates in [131, Proposition 4.3.6], we have the convergence:

$$|M_1(z) - \tilde{M}(z)| = o(1) \quad \text{as } |z| \to \infty. \tag{5.56}$$

Evaluating the matrix product explicitly yields:

$$M_1(z) = \begin{pmatrix} e^{-U(\alpha)/2} & -e^{iz\ell(\theta) - U(\alpha)/2} \\ -e^{iz\ell(H) - U(H)/2} + e^{iz\ell(\theta) - U(\theta)/2} & -e^{iz(\ell(\theta) - \ell(H)) - U(H)/2} + e^{-U(\theta)/2} \end{pmatrix}. \tag{5.57}$$

Through $\triangle(z)$, the normalization $\mathcal{N}(z)$ depends on $\operatorname{sign}(\operatorname{Im} z)$; the matrix above is the $\operatorname{Im} z > 0$ representative, and in the complementary region $\det M_1(z)$ differs only by a non-vanishing factor, so the characteristic equation is unchanged. Computing the determinant of $M_1(z)$, we find:

$$\begin{aligned} \det M_1(z) = e^{-U(\alpha)/2} \left[-e^{iz(\ell(\theta) - \ell(H)) - U(H)/2} + e^{-U(\theta)/2} \right] \\ - \left(-e^{iz\ell(\theta) - U(\alpha)/2} \right) \left(-e^{iz\ell(H) - U(H)/2} + e^{iz\ell(\theta) - U(\theta)/2} \right). \end{aligned} \tag{5.58}$$

Factoring this expression reveals the characteristic equation governing the eigenvalues. We obtain:

$$\det M_1(z) = 2e^{-U(\alpha)/2} e^{iz\ell(\theta)} \left[e^{-U(\theta)/2} \cosh\left(iz\ell(\theta)\right) - e^{-U(H)/2} \cosh\left(iz\ell(H)\right) \right]. \tag{5.59}$$

Since $e^{-U(\alpha)/2} e^{iz\ell(\theta)}$ is strictly non-zero (in fact $U(\alpha) = 0$), the asymptotic spectrum is determined by the vanishing of the term in the brackets. Thus, $\det M_1(z) = 0$ if and only if:

$$\cosh\left(\frac{iz(H - \alpha)}{\sqrt{D}} \right) = e^{-\frac{U(\theta) - U(H)}{2}} \cosh\left(\frac{iz(\theta - \alpha)}{\sqrt{D}} \right). \tag{5.60}$$

This transcendental equation defines the asymptotic location of the eigenvalues $\lambda_n = -z_n^2$.

5.2.1 The asymptotic characteristic equation

The asymptotic characteristic equation

$$\cosh\big(iz\,\ell(H)\big) = \exp\Big(-\tfrac{U(\theta)-U(H)}{2}\Big)\,\cosh\big(iz\,\ell(\theta)\big), \tag{5.61}$$

depends on the specific neuron model only through the scalar factor

$$\kappa := \exp\Big(-\tfrac{U(\theta)-U(H)}{2}\Big). \tag{5.62}$$

All remaining parameters enter only via the geometric lengths $\ell(H) = (H-\alpha)/\sqrt{D}$ and $\ell(\theta) = (\theta-\alpha)/\sqrt{D}$.
To analyze (5.61), we introduce the exponential variable

$$X := e^{iz}. \tag{5.63}$$

Using $\cosh(w) = \frac{1}{2}(e^w + e^{-w})$, equation (5.61) is equivalent to

$$X^{\ell(H)} + X^{-\ell(H)} = \kappa\Big(X^{\ell(\theta)} + X^{-\ell(\theta)}\Big). \tag{5.64}$$

Which gives a polynomial equation (with possibly irrational exponents)
If $z \in \mathbb{R}$, then $|X| = |e^{iz}| = 1$, so real roots z correspond to solutions X lying on the unit circle $|X| = 1$. Furthermore if $\ell(\theta)$ and $\ell(H)$ are rationally dependent then there exists $L > 0$ and integers $m, n \in \mathbb{N}$ such that $\ell(\theta) = mL, \qquad \ell(H) = nL$. Setting $Y := X^L$, equation (5.64) becomes (after multiplying by Y^m) a proper polynomial equation

$$\kappa\, Y^{2m} - Y^{m+n} - Y^{m-n} + \kappa = 0. \tag{5.65}$$

In particular, the periodicity of the asymptotic solutions in the z-plane is governed by the commensurability scale L, and hence by the fraction $\ell(\theta)/\ell(H) = m/n$.

Example: $\ell(\theta) = 2\ell(H)$

As an illustration, consider the commensurate case $\ell(\theta) = 2\ell(H)$. Choosing $\theta = 1$, $H = 0$, $\alpha = -1$, $D = 1$ we have the quartic equation:

$$\kappa\, X^4 - X^3 - X + \kappa = 0, \tag{5.66}$$

which can be solved explicitly. In Fig. 5.1 we compare the analytical asymptotic prediction obtained from (5.66) against the numerical eigenvalues of the BPIF boundary eigenvalue problem ($\kappa = e^{-\xi}$ with $\xi = \mu/(2D)$, i.e. $\kappa = e^{-\mu/(2D)}$ since $b \equiv 0$).
The transition of a pair of eigenvalue branches from the real axis into a complex-conjugate pair is the exceptional point $\kappa = 1$ of the asymptotic characteristic equation (5.61). By (5.62),

$$\kappa = 1 \iff U(\theta) - U(H) = \int_H^\theta \frac{b(v)+\mu}{D}\,dv = 0,$$

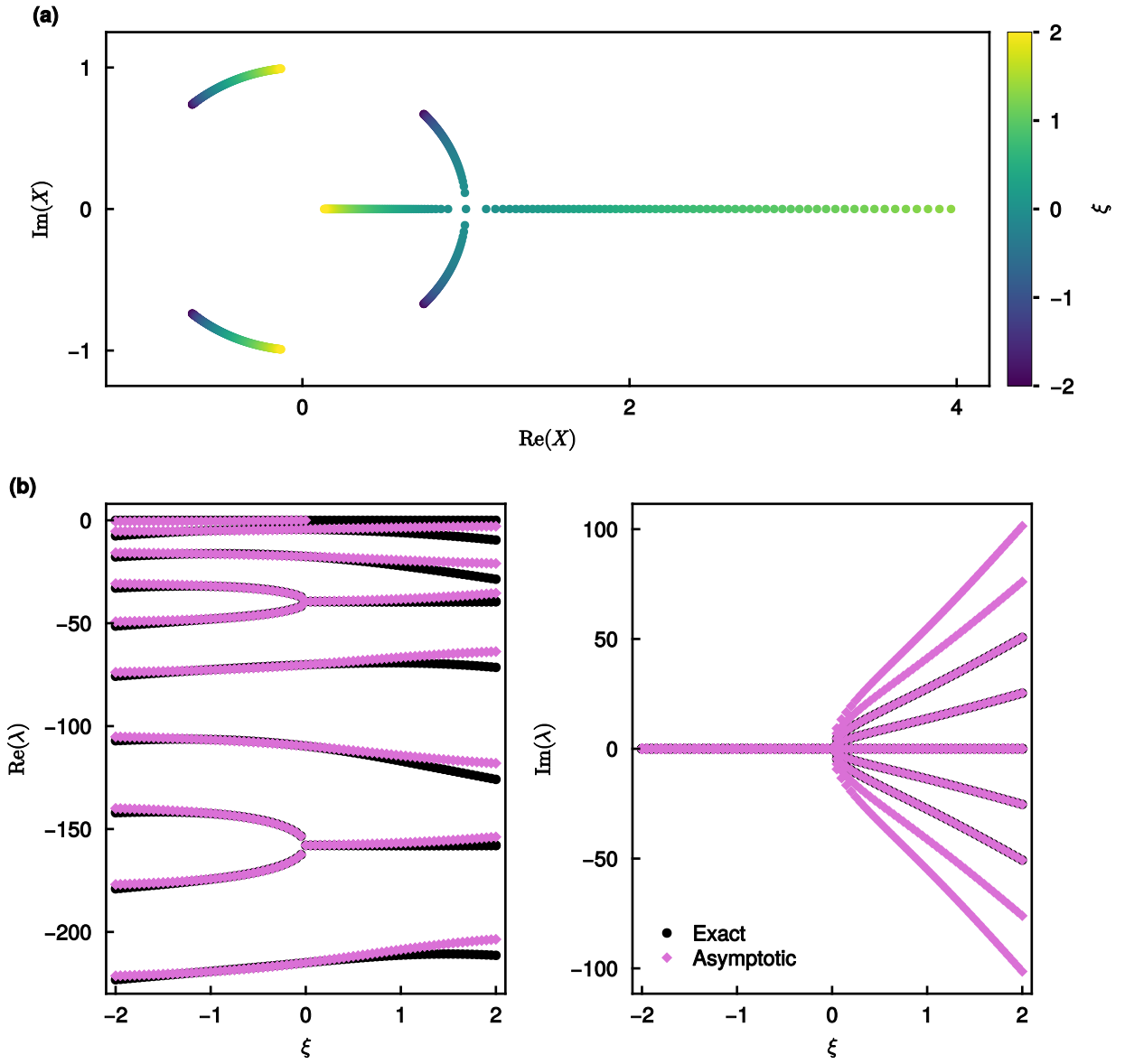

Figure 5.1: Asymptotic eigenvalue distribution for the BPIF neuron. **(a)** Solutions X of the quartic equation (5.66) in the complex plane as a function of $\xi = \frac{\mu}{2D}$. Real λ-eigenvalues correspond to solutions on the unit circle $|X| = 1$. At $\xi = 0$, two solutions coalesce at an exceptional point, and for $\xi > 0$ they split into a complex-conjugate pair in the λ-plane. **(b)** Real (left) and imaginary (right) parts of the eigenvalues $\lambda = (\log X)^2$, comparing the analytical asymptotic prediction with the numerically computed spectrum.

so the bifurcation occurs at a value of the input current μ that depends on the neuron model. For the driftless BPIF ($b \equiv 0$) it is $\mu = 0$, whereas for the LIF neuron ($b(v) = -v$) it follows from

$$\int_H^\theta (\mu - v)\, dv = (\theta - H)\left(\mu - \tfrac{H+\theta}{2}\right) = 0 \tag{5.67}$$

that $\mu = \frac{H+\theta}{2}$. The asymptotic analysis thus recovers, from first principles, the merging threshold $\mu > (H + \theta)/2$ of the mixed eigenvalue branches observed for the LIF model in Chapter 2.

5.2.2 Spectral Localization

Proposition 5.2.1 (Spectral Localization). *Let $\tilde{M}(z)$ be the normalized characteristic matrix function defined in* (5.53). *There exists a constant $y_M > 0$ and a uniform lower bound $C_M > 0$ such that:*

$$|\det \tilde{M}(z)| \geq C_M \quad \textit{for all } z \in \mathbb{C} \textit{ satisfying } |\mathrm{Im}(z)| \geq y_M. \tag{5.68}$$

Consequently, the roots $\{z_n\}_{n\in\mathbb{Z}}$ of the characteristic equation lie within the horizontal strip:

$$\mathcal{S} = \{z \in \mathbb{C} : |\mathrm{Im}(z)| < y_M\}. \tag{5.69}$$

Furthermore, the eigenvalues $\lambda_n = -z_n^2 \in \sigma(\mathcal{L})$ are contained within the parabolic region $\mathcal{P}_{y_M}$ defined by:

$$\mathcal{P}_{y_M} = \left\{\lambda \in \mathbb{C} : \mathrm{Re}(\lambda) \leq y_M^2 - \frac{\mathrm{Im}(\lambda)^2}{4y_M^2}\right\}. \tag{5.70}$$

Proof. The bound (5.68) is a direct consequence of [131, Proposition 4.3.7]. Indeed, A_1 has eigenvalues $\pm i/\sqrt{D}$, so the associated Birkhoff sector angles are $\phi_\pm = \pm\pi/2$ and $|\mathrm{Re}(z\, e^{i\phi_\pm})| = |\mathrm{Im}(z)|$; the hypothesis of that proposition then reduces to $|\mathrm{Im}(z)| \geq y_M$. This implies that all zeros z_n of $\det \tilde{M}(z)$, which are identical to the roots of the characteristic equation in the z-plane, must satisfy the condition $|\mathrm{Im}(z_n)| < y_M$.
Finally, we map this strip to the λ-plane via the transformation $\lambda = -z^2$. Let $z = x + iy$ with $|y| < y_M$. Then:

$$\lambda = -(x + iy)^2 = (y^2 - x^2) - 2ixy. \tag{5.71}$$

Letting $u = \mathrm{Re}(\lambda)$ and $v = \mathrm{Im}(\lambda)$, we have $u = y^2 - x^2$ and $v = -2xy$. Substituting $x = -v/(2y)$ into the equation for u, we obtain:

$$u = y^2 - \frac{v^2}{4y^2}. \tag{5.72}$$

Since $y^2 < y_M^2$, it follows that $u \leq y_M^2 - \frac{v^2}{4y_M^2}$. □

5.3 Spectral decomposition for the relaxation dynamics

5.3.1 Green's Function with Reinjection

We now derive the explicit form of the resolvent kernel (Green's function) for the operator with reinjection, relating it to the resolvent kernel $G^A(v, v_0; s)$ of the problem without reijnection.

Theorem 5.3.1 (Resolvent Identity for Reinjection). *The resolvent kernel $G(v, v_0; s)$ of the Fokker–Planck operator with reinjection $\mathcal{L}_\gamma$ satisfies the identity*

$$G(v, v_0; s) = G^A(v, v_0; s) + G^A(v, H; s)\frac{\tilde{\rho}_0(s|v_0)}{1 - \tilde{\rho}_0(s|H)} \quad \forall s \in \mathbb{C} \setminus \sigma(\boldsymbol{\gamma}). \tag{5.73}$$

Proof. The proof relies on a renewal decomposition in the domain of absolute convergence, followed by analytic continuation.
Consider first $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 0$. We decompose the transition density $p(v,t|v_0)$ based on the number of reinjections n:

$$p(v,t|v_0) = \sum_{n=0}^{\infty} p_n(v,t|v_0), \tag{5.74}$$

where $p_0(v,t|v_0) = p^A(v,t|v_0)$. For $n \geq 1$, the terms satisfy the renewal recursion:

$$p_n(v,t|v_0) = \int_0^t p^A(v,t-\tau|H)\rho_{n-1}(\tau|v_0)\,d\tau, \tag{5.75}$$

where ρ_{n-1} is the flux density of the n-th threshold crossing. In the Laplace domain, this convolution becomes

$$\tilde{p}_n(v,s|v_0) = G^A(v,H;s)\tilde{\rho}_{n-1}(s|v_0). \tag{5.76}$$

Iterating the corresponding recursion for the fluxes, $\tilde{\rho}_n(s|v_0) = \tilde{\rho}_0(s|H)\tilde{\rho}_{n-1}(s|v_0)$, yields

$$\tilde{p}_n(v,s|v_0) = G^A(v,H;s)\tilde{\rho}_0(s|v_0)[\tilde{\rho}_0(s|H)]^{n-1}, \quad n \geq 1. \tag{5.77}$$

Summing the series for the total Green's function $G(v,v_0;s) = \sum_{n=0}^{\infty}\tilde{p}_n(v,s|v_0)$:

$$G(v,v_0;s) = G^A(v,v_0;s) + G^A(v,H;s)\tilde{\rho}_0(s|v_0)\sum_{k=0}^{\infty}[\tilde{\rho}_0(s|H)]^k. \tag{5.78}$$

Since $\rho_0(t|H)$ is a probability density on $(0,\infty)$, its Laplace transform satisfies $|\tilde{\rho}_0(s|H)| < 1$ for $\mathrm{Re}(s) > 0$. Thus, the geometric series converges absolutely to $(1-\tilde{\rho}_0(s|H))^{-1}$, establishing (5.73) for $\mathrm{Re}(s) > 0$.
Both sides of (5.73) define meromorphic functions of the spectral parameter s on $\mathbb{C}\setminus\sigma(\boldsymbol{\gamma})$.
The right-hand side is *a priori* also singular at the poles μ_i of $G^A(\cdot,\cdot;s)$ (the absorbing eigenvalues, equivalently the roots of $\psi_1(\theta;s)=0$), but a direct residue computation yields

$$\begin{aligned}\operatorname*{Res}_{s=\mu_i}\left(G^A(v,v_0;s) + G^A(v,H;s)\frac{\tilde{\rho}_0(s|v_0)}{1-\tilde{\rho}_0(s|H)}\right) = \\ = f_1(v,\mu_i)\left[\psi_1(v_0,\mu_i) + \psi_1(H,\mu_i)\,\frac{\psi_1(v_0;\mu_i)}{\psi_1(\theta;\mu_i)-\psi_1(H;\mu_i)}\right] = 0.\end{aligned} \tag{5.79}$$

so each μ_i is a removable singularity of the right-hand side and the latter extends holomorphically across $\{\mu_i\}$.
Consequently, the right-hand side is meromorphic on $\mathbb{C}\setminus\sigma(\boldsymbol{\gamma})$ with poles only at $\sigma(\boldsymbol{\gamma})$ (i.e. at the solutions of $1-\tilde{\rho}_0(s|H)=0$). Since $G(v,v_0;s)$ has the same pole set and the two meromorphic functions coincide for $\mathrm{Re}(s)>0$, the identity theorem implies that they coincide on $\mathbb{C}\setminus\sigma(\boldsymbol{\gamma})$.

□

Remark 5.3.1. *Finally, applying the boundary flux functional $\mathcal{N}$ to the Green's function with*

respect to the initial condition H yields the Laplace transform of the population firing rate:

$$\tilde{\nu}(s) = \mathcal{N}[G(\cdot, H; s)] = \frac{\tilde{\rho}_0(s|H)}{1 - \tilde{\rho}_0(s|H)}. \tag{5.80}$$

This expression recovers the classical renewal result relating the population activity to the single-neuron first-passage time density, corresponding to Equation (33) in [198].

5.3.2 Sectoriality of the Operator with Reinjection

We now address the functional analytic properties of the operator with reinjection. While the operator without reinjection (absorbing case) is self-adjoint in the weighted space, the non-local reinjection condition destroys self-adjointness. However, we prove that the operator generates an analytic semigroup by showing it is sectorial.

Theorem 5.3.2 (Sectoriality of the Reinjection Operator)**.** *The Fokker–Planck operator with reinjection $\mathcal{L}_\gamma$, defined on the domain incorporating the reinjection boundary condition, is a sectorial operator in $L^2(\alpha, \theta)$. Consequently, it generates a strongly continuous analytic semigroup.*

Proof. We first establish the bound for the simplified *driftless* Bayesian Point Integrate-and-Fire (BPIF) model. Without loss of generality, let $\mu = 0$, $D = 1$, and the domain be $(\alpha, \theta) = (0, \pi)$. The operator reduces to the Laplacian ∂_v^2 with a reflective boundary at $v = 0$ and absorbing condition at $v = \pi$.

The homogeneous solution ψ_1 of the associated second order differential equation ($\partial_v^2 \psi_1 = s\psi_1$ with $\psi_1'(0, z) = 0$, $\psi_1(0, z) = 1$) is given by $\psi_1(v, z) = \cos(vz)$, where $s = -z^2$ for $z \in \mathbb{C}$. The term $\tilde{\rho}_0(s|x)$ corresponds to the ratio of the eigenfunctions:

$$\tilde{\rho}_0(s|x) = \frac{\psi_1(x, z)}{\psi_1(\theta, z)} = \frac{\cos(zx)}{\cos(z\pi)}. \tag{5.81}$$

The characteristic equation for the eigenvalues then becomes:

$$1 - \tilde{\rho}_0(s|H) = 1 - \frac{\cos(zH)}{\cos(z\pi)} = \frac{\cos(z\pi) - \cos(zH)}{\cos(z\pi)}. \tag{5.82}$$

The zeros of this function, which determine the spectrum $\sigma(\mathcal{L}^*_{\text{BPIF}})$, are the real roots of $\cos(zH) = \cos(z\pi)$. Thus, the spectrum lies on the real axis and hence $\sigma(\mathcal{L}^*_{\text{BPIF}}) \subset (-\infty, 0]$ in the s–plane (recall $s = -z^2$).

Using the decomposition

$$G^*(x, y; -z^2) = G^A(y, x; -z^2) + G^A(y, H; -z^2)\, \frac{\cos(zx)}{\cos(z\pi) - \cos(zH)}, \qquad x, y \in (0, \pi),$$

and the pointwise bound on the absorbing kernel (3.15), it remains to control the reinjection factor. For $-z^2 \in \Sigma_{\pi-\delta}$ we have $|\operatorname{Im}(z)| \geq c_\delta |z|$, and the standard asymptotic estimates yield

$$|\cos(\pi z)| \simeq \tfrac{1}{2} e^{\pi |\operatorname{Im}(z)|}, \qquad |\cos(Hz)| \simeq \tfrac{1}{2} e^{H |\operatorname{Im}(z)|}, \qquad |\cos(zx)| \simeq \tfrac{1}{2} e^{x |\operatorname{Im}(z)|}.$$

Since $H < \pi$, it follows that $|\cos(\pi z) - \cos(Hz)| \simeq |\cos(\pi z)|$, and hence

$$\left|\frac{\cos(zx)}{\cos(z\pi) - \cos(zH)}\right| \leq C'_\delta e^{-(\pi - x)|\operatorname{Im}(z)|} \quad - z^2 \in \Sigma_{\pi-\delta}.$$

We estimate the two terms of the decomposition above by Schur's test, as in Section 3.1.1, in the regime of large $|\operatorname{Im}(z)|$ in which (3.15) and the reinjection-factor bound above hold; on $\Sigma_{\pi-\delta}$ this is the regime $|\lambda| \to \infty$, since $|\operatorname{Im}(z)| \geq c_\delta |z|$ and $|z|^2 = |\lambda|$. The absorbing term $G^A(y, x; -z^2)$ obeys (3.15), so both its Schur integrals are $\leq 2C_\delta/(|z|\,|\operatorname{Im}(z)|)$, and it thus defines an integral operator of norm $\leq 2C_\delta/(|z|\,|\operatorname{Im}(z)|)$. For the reinjection term, (3.15) at the pair (y, H) and the factor bound give the separable estimate

$$\left|G^A(y, H; -z^2)\,\frac{\cos(zx)}{\cos(z\pi) - \cos(zH)}\right| \leq \frac{C_\delta C'_\delta}{|z|}\, e^{-|\operatorname{Im}(z)|\,|y-H|}\, e^{-|\operatorname{Im}(z)|\,(\pi - x)},$$

whose Schur integrals, using $\int_0^\pi e^{-|\operatorname{Im}(z)|\,|y-H|}\,dy \leq 2/|\operatorname{Im}(z)|$ and $\int_0^\pi e^{-|\operatorname{Im}(z)|\,(\pi-x)}\,dx \leq 1/|\operatorname{Im}(z)|$, are $\leq 2C_\delta C'_\delta/(|z|\,|\operatorname{Im}(z)|)$ and $\leq C_\delta C'_\delta/(|z|\,|\operatorname{Im}(z)|)$; the reinjection term thus defines an integral operator of norm $\leq \sqrt{2}\,C_\delta C'_\delta/(|z|\,|\operatorname{Im}(z)|)$. Adding the two, for $\lambda = -z^2 \in \Sigma_{\pi-\delta}$ with $|\lambda|$ large,

$$\left\|(\lambda - \mathcal{L}^*_{\boldsymbol{\gamma}})^{-1}\right\|_{L^2} \leq \frac{2C_\delta + \sqrt{2}\,C_\delta C'_\delta}{|z|\,|\operatorname{Im}(z)|} \leq \frac{M_\delta}{|\lambda|}.$$

Since $\sigma(\mathcal{L}^*_{\boldsymbol{\gamma}}) \subset (-\infty, 0]$ and $\lambda = 0$ is a simple eigenvalue, the resolvent is meromorphic with a simple pole at $\lambda = 0$, so the bound $M_\delta/|\lambda|$ holds throughout $\Sigma_{\pi-\delta}$. Hence $\mathcal{L}^*_{\boldsymbol{\gamma}}$ is sectorial and generates an analytic semigroup.

For the general Integrate-and-Fire neuron with drift $\mu \neq 0$ and diffusion $D > 0$, the operator $\mathcal{L}^*_{\boldsymbol{\gamma}}$ comprises the second-order diffusion term (equivalent to the driftless operator) and a first-order drift perturbation. The drift term $\partial_v[(b(v) + \mu)\cdot\,]$ is relatively bounded with respect to the second-derivative operator ∂_v^2 with relative bound zero (since higher-order Sobolev norms control lower-order ones). By the standard perturbation theory for sectorial operators (see [55, Thm. III.2.10]), the addition of a relatively bounded perturbation with sufficiently small bound preserves sectoriality. Therefore, $\mathcal{L}^*_{\boldsymbol{\gamma}}$ is sectorial in $\mathrm{L}^2(\alpha, \theta)$. By duality, the original forward operator $\mathcal{L}_{\boldsymbol{\gamma}}$ is sectorial. □

5.3.3 Relaxation dynamics

Having established the Birkhoff regularity and the sectoriality of $\mathcal{L}_{\boldsymbol{\gamma}}$ in the Hilbert space $\mathbb{L}^2 = \mathrm{L}^2(\alpha, \theta)$, the spectral decomposition of the neural population dynamics is well-posed. As a consequence of sectoriality, $\mathcal{L}_{\boldsymbol{\gamma}}$ generates a strongly continuous analytic semigroup

$$\mathcal{F}_2(t) : \mathbb{L}^2 \to \mathbb{L}^2, \qquad t \geq 0. \tag{5.83}$$

For any initial condition $p_* \in \mathbb{L}^2$, the solution is

$$p(\cdot, t) = \mathcal{F}_2(t) p_*. \tag{5.84}$$

Moreover, analytic smoothing implies that for every $t > 0$ one has $p(\cdot, t) \in \mathbb{D}_{\boldsymbol{\gamma}}$.

Since $\mathcal{L}_{\boldsymbol{\gamma}}$ is Birkhoff regular, its system of root functions is complete in $\mathbb{L}^2$, and we

can expand the solution in the biorthogonal canonical system of eigenfunctions and associated functions (CSEAV) $\{\boldsymbol{\varphi}_n^{(k)}, \boldsymbol{\Psi}_n^{(k)}\}$ associated with the eigenvalues $\{\lambda_n\}_{n\geq 0}$, where $k = 0, \dots, m_n - 1$ indexes the associated vectors in the Jordan chain of length m_n:

$$p(\cdot, t) = \sum_{n=0}^{\infty} \sum_{k=0}^{m_n-1} a_n^{(k)}(t)\, \boldsymbol{\varphi}_n^{(k)}. \tag{5.85}$$

The modal coefficients are obtained by projecting onto the dual vector of complementary index:

$$a_n^{(k)}(t) := \left\langle \boldsymbol{\Psi}_n^{(m_n-1-k)},\, p(\cdot, t) \right\rangle_{\mathbb{L}^2}, \qquad k = 0, 1, \dots, m_n - 1. \tag{5.86}$$

Dynamics of the modal coefficients. The right CSEAV satisfies the Jordan chain relations

$$(\mathcal{L}_{\boldsymbol{\gamma}} - \lambda_n)\boldsymbol{\varphi}_n^{(0)} = 0, \qquad (\mathcal{L}_{\boldsymbol{\gamma}} - \lambda_n)\boldsymbol{\varphi}_n^{(k)} = \boldsymbol{\varphi}_n^{(k-1)}, \qquad 1 \leq k \leq m_n - 1. \tag{5.87}$$

Projecting $\partial_t p = \mathcal{L}_{\boldsymbol{\gamma}} p$ using (5.86), the Jordan relations (5.87), and the biorthogonality $\langle \boldsymbol{\Psi}_i^{(l)}, \boldsymbol{\varphi}_j^{(m_j-1-k)} \rangle_{\mathbb{L}^2} = \delta_{ij}\delta_{lk}$, yields the coupled dynamics:

$$\dot{a}_n^{(k)}(t) = \lambda_n a_n^{(k)}(t) + a_n^{(k+1)}(t), \qquad k = 0, 1, \dots, m_n - 2, \qquad \dot{a}_n^{(m_n-1)}(t) = \lambda_n a_n^{(m_n-1)}(t). \tag{5.88}$$

Integrating (5.88) gives the explicit evolution

$$a_n^{(k)}(t) = e^{\lambda_n t} \sum_{j=0}^{m_n-1-k} a_n^{(k+j)}(0)\, \frac{t^j}{j!}, \qquad k = 0, 1, \dots, m_n - 1, \tag{5.89}$$

and in particular

$$a_n^{(0)}(t) = e^{\lambda_n t} \sum_{k=0}^{m_n-1} a_n^{(k)}(0)\, \frac{t^k}{k!}. \tag{5.90}$$

For simple eigenvalues ($m_n = 1$), (5.89) reduces to $a_n^{(0)}(t) = e^{\lambda_n t} a_n^{(0)}(0)$.

Flux observable and relaxation of the firing rate. Define the population firing rate via the boundary flux functional $\mathcal{N} : \mathbb{D}_{\boldsymbol{\gamma}} \to \mathbb{R}$,

$$\nu(t) = \mathcal{N}[p(\cdot, t)]. \tag{5.91}$$

For $t > 0$ the analytic semigroup implies $p(\cdot, t) \in \mathbb{D}_{\boldsymbol{\gamma}}$, and the series (5.85) converges in the graph norm. Since $\mathcal{N}$ is continuous with respect to this topology, we may interchange $\mathcal{N}$ and the spectral summation:

$$\nu(t) = \sum_{n=0}^{\infty} \sum_{k=0}^{m_n-1} a_n^{(k)}(t)\, \mathcal{N}[\boldsymbol{\varphi}_n^{(k)}]. \tag{5.92}$$

By the flux normalization property (Lemma 4.3.2),

$$\mathcal{N}[\boldsymbol{\varphi}_n^{(0)}] = 1, \qquad \mathcal{N}[\boldsymbol{\varphi}_n^{(k)}] = 0 \ (k \geq 1), \qquad n \geq 1. \tag{5.93}$$

Thus, for $n \geq 1$ only the eigenfunctions $\boldsymbol{\varphi}_n^{(0)}$ contribute to the firing rate, and using (5.90) we obtain:

$$\nu(t) = \Phi(\boldsymbol{\gamma}) + \sum_{n=1}^{\infty} e^{\lambda_n t} \left(\sum_{k=0}^{m_n-1} a_n^{(k)}(0) \, \frac{t^k}{k!} \right), \tag{5.94}$$

where $\Phi(\boldsymbol{\gamma})$ denotes the stationary firing rate associated with the simple eigenvalue $\lambda_0 = 0$, whose coefficient $a_0^{(0)}$ is constant by conservation of probability.

Block-vector formulation. For each spectral block n define the coefficient vector

$$\mathbf{a}_n(t) := \big(a_n^{(0)}(t), a_n^{(1)}(t), \dots, a_n^{(m_n-1)}(t)\big)^\top \in \mathbb{C}^{m_n}, \tag{5.95}$$

and introduce the Jordan block matrix

$$\mathbf{J}_n := \lambda_n \mathbf{I}_{m_n} + \mathbf{S}_{m_n}, \qquad (\mathbf{S}_{m_n})_{k,k+1} = 1 \ (k = 0, \dots, m_n - 2), \qquad (\mathbf{S}_{m_n})_{ij} = 0 \text{ otherwise.} \tag{5.96}$$

Then (5.88) can be written compactly as

$$\dot{\mathbf{a}}_n(t) = \mathbf{J}_n \, \mathbf{a}_n(t), \qquad \mathbf{a}_n(t) = e^{t\mathbf{J}_n} \mathbf{a}_n(0). \tag{5.97}$$

5.4 Regularized coupled dynamics

We finally shift our focus to the evolution problem driven by a *time-dependent* (parameter-dependent) generator $t \mapsto \mathcal{L}_{\boldsymbol{\gamma}(t)}$, where $\boldsymbol{\gamma}(\cdot)$ is a prescribed parameter path. At the formal level, the dynamics reads

$$\partial_t p(\cdot, t) = \mathcal{L}_{\boldsymbol{\gamma}(t)} \, p(\cdot, t), \qquad p(\cdot, t) \in \mathbb{D}_{\boldsymbol{\gamma}(t)}. \tag{5.98}$$

When $\boldsymbol{\gamma}$ varies in time the problem becomes *non-autonomous*: in general, one should expect a *two-parameter* evolution family described by the propagator $\mathcal{U}(t, s)$ satisfying $\mathcal{U}(s, s) = \mathbf{I}$ and the composition rule $\mathcal{U}(t, r)\mathcal{U}(r, s) = \mathcal{U}(t, s)$ for $s \leq r \leq t$. In principle, well-posedness for these problems can be established by verifying the uniform sectoriality of the generator [4]. While we do not derive the resolvent bounds required by the non-autonomous well-posedness theory here, they are expected to follow from the Green's function asymptotics implied by Birkhoff regularity.

For coupled neural populations, $\boldsymbol{\gamma}(t)$ itself depends on the state $p(\cdot, t)$, leading to a nonlinear self-consistency constraint given by the mean-field closure. Establishing well-posedness in that setting is substantially more delicate, and we do not attempt it here. Instead, we adopt a formal perspective: we treat $\boldsymbol{\gamma}(t)$ as given, assume the existence of a sufficiently regular solution $p(\cdot, t) \in \mathbb{D}_{\boldsymbol{\gamma}(t)}$, and analyze the dynamics after projection onto instantaneous spectral structures.

A natural route is expansion on the instantaneous eigenmodes, which yields the non-adiabatic system (2.53). However, as discussed in Section 4.5, this approach breaks down at *exceptional points* (EPs), that is, parameter values $\boldsymbol{\gamma}_*$ where eigenvalues coalesce, and the instantaneous eigenbasis becomes defective; correspondingly, the non-adiabatic coefficients (in particular the diagonal terms $\mathbf{C}_{nn}$) become singular.

To obtain a description that remains well-defined across EP crossings, we avoid transporting individual eigenvectors and instead transport *invariant spectral subspaces*. Concretely, we select a

contour Γ enclosing the relevant spectral cluster and define the spectral projector by the Riesz formula:

$$\mathbf{P}_\Gamma(\boldsymbol{\gamma}(t)) = \frac{1}{2\pi i}\oint_\Gamma (z - \mathcal{L}_{\boldsymbol{\gamma}(t)})^{-1}\, dz,$$

which varies smoothly as long as Γ separates the cluster from the rest of the spectrum. This strategy of subspace transport remains well-defined even when the instantaneous eigenbasis becomes singular [89]. To implement this, we construct a single-valued moving basis for the subspace $\operatorname{Ran}\mathbf{P}_\Gamma(\boldsymbol{\gamma}(t))$ by adapting the method of *contour moments* (originally developed for static eigensolvers [21, 169]) to the time-dependent setting. This yields a "regularized" projected dynamics whose coefficients remain bounded at EPs, thereby providing a non-adiabatic modal description compatible with eigenvalue coalescence.

5.4.1 Dynamics in the global moment basis

Fix $K \geq 1$. We assume that the (simple) stationary eigenvalue $\lambda_0(\boldsymbol{\gamma}) = 0$ remains spectrally isolated along the parameter path.

For each $\boldsymbol{\gamma}$ we choose a positively oriented simple closed contour $\Gamma_K(\boldsymbol{\gamma}) \subset \mathbb{C}$ with the separation property that $\Gamma_K(\boldsymbol{\gamma}) \subset \rho(\mathcal{L}_{\boldsymbol{\gamma}})$, $\sigma(\mathcal{L}_{\boldsymbol{\gamma}}) \cap \Gamma_K(\boldsymbol{\gamma}) = \varnothing$, and the only spectral points of $\mathcal{L}_{\boldsymbol{\gamma}}$ contained in the bounded component $\operatorname{int}\Gamma_K(\boldsymbol{\gamma})$ are the K eigenvalues $\lambda_1(\boldsymbol{\gamma}), \ldots, \lambda_K(\boldsymbol{\gamma})$ counted with algebraic multiplicity. Moreover, $0 \notin \operatorname{int}\Gamma_K(\boldsymbol{\gamma})$ and there exists a second positively oriented contour $\Gamma_0(\boldsymbol{\gamma}) \subset \rho(\mathcal{L}_{\boldsymbol{\gamma}})$, disjoint from $\Gamma_K(\boldsymbol{\gamma})$, such that $\sigma(\mathcal{L}_{\boldsymbol{\gamma}}) \cap \operatorname{int}\Gamma_0(\boldsymbol{\gamma}) = \{0\}$.

Such a contour defines the corresponding spectral projector:

$$\mathbf{P}_K(\boldsymbol{\gamma}) := -\frac{1}{2\pi i}\oint_{\Gamma_K(\boldsymbol{\gamma})} (\mathcal{L}_{\boldsymbol{\gamma}} - s)^{-1}\, ds, \qquad \mathbb{E}_K(\boldsymbol{\gamma}) := \operatorname{Ran}\mathbf{P}_K(\boldsymbol{\gamma}). \tag{5.99}$$

The adjoint (dual) projector is defined by

$$\mathbf{P}_K^*(\boldsymbol{\gamma}) := (\mathbf{P}_K(\boldsymbol{\gamma}))^* = -\frac{1}{2\pi i}\oint_{\Gamma_K(\boldsymbol{\gamma})} \left((\mathcal{L}_{\boldsymbol{\gamma}} - s)^{-1}\right)^* ds = -\frac{1}{2\pi i}\oint_{\Gamma_K(\boldsymbol{\gamma})} (\mathcal{L}_{\boldsymbol{\gamma}}^* - s)^{-1}\, ds, \tag{5.100}$$

and the corresponding adjoint invariant subspace by

$$\mathbb{E}_K^*(\boldsymbol{\gamma}) := \operatorname{Ran}\mathbf{P}_K^*(\boldsymbol{\gamma}). \tag{5.101}$$

We further define the monic polynomial whose roots are the enclosed eigenvalues:

$$\pi_K(s, \boldsymbol{\gamma}) := \prod_{k=1}^{K} (s - \lambda_k(\boldsymbol{\gamma})) = s^K + \beta_{K-1}(\boldsymbol{\gamma}) s^{K-1} + \cdots + \beta_0(\boldsymbol{\gamma}). \tag{5.102}$$

Away from exceptional points, let $\{(\lambda_k(\boldsymbol{\gamma}), \boldsymbol{\varphi}_k(\boldsymbol{\gamma}), \boldsymbol{\Psi}_k(\boldsymbol{\gamma}))\}_{k=1}^K$ be the normalized right/adjoint eigenpairs enclosed by $\Gamma_K(\boldsymbol{\gamma})$.

Global moment basis and Vandermonde structure

Adjoint moments Define, for $p = 0, \dots, K-1$,

$$\boldsymbol{\eta}_p(\boldsymbol{\gamma}) := \frac{1}{2\pi i} \oint_{\Gamma_K(\boldsymbol{\gamma})} s^p \, \frac{\boldsymbol{\Psi}(s, \boldsymbol{\gamma})}{\Delta_{\boldsymbol{\gamma}}(s)} \, ds \;\in\; \mathbb{E}_K^*(\boldsymbol{\gamma}). \tag{5.103}$$

If the enclosed eigenvalues are simple, then

$$\boldsymbol{\eta}_p(\boldsymbol{\gamma}) = \sum_{k=1}^{K} \lambda_k(\boldsymbol{\gamma})^p \, \boldsymbol{\Psi}_k(\boldsymbol{\gamma}), \qquad p = 0, \dots, K-1, \tag{5.104}$$

so the coefficients form a Vandermonde matrix $(V^\top)_{pk} = \lambda_k^p$.

Right basis Define $\{\mathbf{u}_q(\boldsymbol{\gamma})\}_{q=0}^{K-1} \subset \mathbb{E}_K(\boldsymbol{\gamma})$ as the unique family satisfying

$$\langle \boldsymbol{\eta}_p(\boldsymbol{\gamma}), \mathbf{u}_q(\boldsymbol{\gamma}) \rangle = \delta_{pq}, \qquad p, q = 0, \dots, K-1. \tag{5.105}$$

Equivalently, one may write

$$\mathbf{u}_q(\boldsymbol{\gamma}) = \frac{1}{2\pi i} \oint_{\Gamma_K(\boldsymbol{\gamma})} \frac{\chi_q(s, \boldsymbol{\gamma})}{\pi_K(s, \boldsymbol{\gamma})} \, \boldsymbol{\varphi}(s, \boldsymbol{\gamma}) \, ds, \qquad q = 0, \dots, K-1, \tag{5.106}$$

where $\chi_q(\cdot, \boldsymbol{\gamma})$ is the unique polynomial of degree $\leq K-1$ such that

$$\frac{1}{2\pi i} \oint_{\Gamma_K(\boldsymbol{\gamma})} s^p \, \chi_q(s, \boldsymbol{\gamma}) \, \frac{ds}{\pi_K(s, \boldsymbol{\gamma})} = \delta_{pq}, \qquad p = 0, \dots, K-1. \tag{5.107}$$

Explicitly, with $\pi_K(s, \boldsymbol{\gamma}) = s^K + \beta_{K-1} s^{K-1} + \cdots + \beta_0$ and $\beta_K := 1$,

$$\chi_q(s, \boldsymbol{\gamma}) = \sum_{r=0}^{K-1-q} \beta_{r+1+q}(\boldsymbol{\gamma}) \, s^r. \tag{5.108}$$

This definition remains meaningful when eigenvalues coalesce. If the enclosed eigenvalues are simple then (5.106) reduces to a residue sum

$$\mathbf{u}_q(\boldsymbol{\gamma}) = \sum_{k=1}^{K} U_{kq}(\boldsymbol{\gamma}) \, \boldsymbol{\varphi}_k(\boldsymbol{\gamma}), \qquad \mathbf{U}(\boldsymbol{\gamma}) = \Big(\mathbf{V}^\top(\boldsymbol{\gamma})\Big)^{-1}, \tag{5.109}$$

and the entries of $\Big(\mathbf{V}^\top(\boldsymbol{\gamma})\Big)^{-1}$ can be expressed in terms of Lagrange polynomials.

Companion form on $\mathbb{E}_K$

The matrix representing $\mathcal{L}_{\boldsymbol{\gamma}}|_{\mathbb{E}_K}$ in the basis $\{\mathbf{u}_0, \dots, \mathbf{u}_{K-1}\}$ has companion form:

$$\boldsymbol{\Omega}_K(\boldsymbol{\gamma}) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -\beta_0(\boldsymbol{\gamma}) & -\beta_1(\boldsymbol{\gamma}) & \cdots & -\beta_{K-2}(\boldsymbol{\gamma}) & -\beta_{K-1}(\boldsymbol{\gamma}) \end{pmatrix}, \tag{5.110}$$

We substitute $\boldsymbol{\gamma} = \boldsymbol{\gamma}(t)$ and expand

$$p(\cdot, t) = \boldsymbol{\varphi}_0(\boldsymbol{\gamma}(t)) + \sum_{q=0}^{K-1} b_q(t)\, \mathbf{u}_q(\boldsymbol{\gamma}(t)) + p_\perp(\cdot, t), \qquad b_q(t) := \big\langle \boldsymbol{\eta}_q(\boldsymbol{\gamma}(t)), p(\cdot, t) \big\rangle, \tag{5.111}$$

where $p_\perp := \big(\mathbf{I} - \mathbf{P}_0(\boldsymbol{\gamma}(t)) - \mathbf{P}_K(\boldsymbol{\gamma}(t))\big)p$ is the component outside $\operatorname{span}\{\boldsymbol{\varphi}_0\} \oplus \mathbb{E}_K$. Projecting and *discarding* $p_\perp$ yields

$$\dot{\mathbf{b}} = \boldsymbol{\Omega}_K(\boldsymbol{\gamma}(t))\, \mathbf{b} + \dot{\boldsymbol{\gamma}}(t) \cdot \boldsymbol{\Upsilon}_K^{\boldsymbol{\gamma}}(\boldsymbol{\gamma}(t))\, \mathbf{b} + \dot{\boldsymbol{\gamma}}(t) \cdot \tilde{\mathbf{c}}_K^{\boldsymbol{\gamma}}(\boldsymbol{\gamma}(t)), \tag{5.112}$$

where $\mathbf{b} = (b_0, \dots, b_{K-1})^\top$. The matrices $\boldsymbol{\Upsilon}_K^{\boldsymbol{\gamma}}$ and $\tilde{\mathbf{c}}_K^{\boldsymbol{\gamma}}$ are the moment-basis analogues of the non-adiabatic coupling coefficients $\mathbf{C}^{\boldsymbol{\gamma}}$, $\mathbf{c}^{\boldsymbol{\gamma}}$ of (2.53), to which they reduce when $K = 1$. The neglected contribution is $\dot{\boldsymbol{\gamma}}(t) \cdot \langle \partial_{\boldsymbol{\gamma}} \boldsymbol{\eta}_p, p_\perp \rangle$. Note that the $\boldsymbol{\Omega}$ term is exact irrespective of $p_\perp$, since $\mathbb{E}_K^*$ is invariant; the truncation affects only the non-adiabatic term, because a moving $\boldsymbol{\gamma}$ transfers amplitude across Γ_K. The closure is therefore exact when $\dot{\boldsymbol{\gamma}} = 0$ or $p_\perp = 0$, and is otherwise a Galerkin approximation. Here

$$(\boldsymbol{\Upsilon}_K^{\boldsymbol{\gamma}}(\boldsymbol{\gamma}))_{pq} := \big\langle \partial_{\boldsymbol{\gamma}} \boldsymbol{\eta}_p(\boldsymbol{\gamma}),\, \mathbf{u}_q(\boldsymbol{\gamma}) \big\rangle, \tag{5.113}$$

$$\tilde{\mathbf{c}}_{K,p}^{\boldsymbol{\gamma}}(\boldsymbol{\gamma}) := -\big\langle \boldsymbol{\eta}_p(\boldsymbol{\gamma}),\, \partial_{\boldsymbol{\gamma}} \boldsymbol{\varphi}_0(\boldsymbol{\gamma}) \big\rangle. \tag{5.114}$$

Firing Rate Let $\nu(t) = \mathcal{N}[p(\cdot, t)]$ be the population firing rate. By the flux normalization of the right root functions, the global basis satisfies

$$\mathcal{N}[\mathbf{u}_0(\boldsymbol{\gamma})] = 1, \qquad \mathcal{N}[\mathbf{u}_q(\boldsymbol{\gamma})] = 0 \quad (q = 1, \dots, K-1), \tag{5.115}$$

hence

$$\nu(t) = \Phi(\boldsymbol{\gamma}(t)) + b_0(t), \tag{5.116}$$

where $\Phi(\boldsymbol{\gamma}) = \mathcal{N}[\boldsymbol{\varphi}_0(\boldsymbol{\gamma})]$ is the stationary firing rate.

5.4.2 Complete block expansion in $\mathbb{L}^2$

As $\mathcal{L}_{\boldsymbol{\gamma}}$ is Birkhoff regular, its root functions yield complete expansions in $\mathbb{L}^2$ [131]. Accordingly, we now introduce a block decomposition of the full spectrum.

Spectral clusters and fixed contours

Assume that for each $\boldsymbol{\gamma}$ the spectrum splits into disjoint clusters $\sigma(\mathcal{L}_{\boldsymbol{\gamma}}) = \bigcup_{n\geq 0} \sigma_n(\boldsymbol{\gamma})$ and that for each n there exists a positively oriented contour $\Gamma_n(\boldsymbol{\gamma}) \subset \rho(\mathcal{L}_{\boldsymbol{\gamma}})$ such that $\sigma_n(\boldsymbol{\gamma}) \subset \operatorname{int}\Gamma_n(\boldsymbol{\gamma})$, and the spectrum stays uniformly away from $\Gamma_n(\boldsymbol{\gamma})$ along the path (no eigenvalue ever crosses Γ_n).
Define the block projectors and subspaces

$$\mathbf{P}_n(\boldsymbol{\gamma}) := -\frac{1}{2\pi i}\oint_{\Gamma_n(\boldsymbol{\gamma})} (\mathcal{L}_{\boldsymbol{\gamma}} - s)^{-1}\, ds, \qquad \mathbb{E}_n(\boldsymbol{\gamma}) := \operatorname{Ran}\mathbf{P}_n(\boldsymbol{\gamma}), \qquad \mathbb{E}_n^*(\boldsymbol{\gamma}) := \operatorname{Ran}\mathbf{P}_n(\boldsymbol{\gamma})^*.$$

Shifted block basis

Fix a block $n \geq 1$ with algebraic multiplicity m_n, and define the shifted quantities

$$\hat{\lambda}_n(\boldsymbol{\gamma}) := \frac{1}{m_n}\sum_{k=0}^{m_n-1} \lambda_{n,k}(\boldsymbol{\gamma}), \qquad \delta\hat{\lambda}_{n,k}(\boldsymbol{\gamma}) := \lambda_{n,k}(\boldsymbol{\gamma}) - \hat{\lambda}_n(\boldsymbol{\gamma})$$

where $\{\lambda_{n,k}(\boldsymbol{\gamma})\}_{k=0}^{m_n-1}$ enumerates $\sigma_n(\boldsymbol{\gamma})$ repeated according to algebraic multiplicity. We denote by π_n the block polynomial and by $\hat{\pi}_n^{\mathrm{sh}}$ its mean-centered (shifted) form:

$$\pi_n(s,\boldsymbol{\gamma}) := \prod_{k=0}^{m_n-1}\big(s - \lambda_{n,k}(\boldsymbol{\gamma})\big) = \hat{\pi}_n^{\mathrm{sh}}\big(s - \hat{\lambda}_n(\boldsymbol{\gamma}), \boldsymbol{\gamma}\big), \tag{5.117}$$

$$\hat{\pi}_n^{\mathrm{sh}}(z,\boldsymbol{\gamma}) := \prod_{k=0}^{m_n-1}\big(z - \delta\hat{\lambda}_{n,k}(\boldsymbol{\gamma})\big) = z^{m_n} + \hat{\beta}_{m_n-1}(\boldsymbol{\gamma})\, z^{m_n-1} + \cdots + \hat{\beta}_0(\boldsymbol{\gamma}).$$

Its coefficients are, up to sign, the elementary symmetric polynomials of the shifted eigenvalues, $\hat{\beta}_{m_n-i} = (-1)^i \hat{e}_i$ with $\hat{\beta}_{m_n} := 1$; in particular $\hat{\beta}_{m_n-1} = 0$ since the shift has zero mean. Although the individual branches $\lambda_{n,k}(\boldsymbol{\gamma})$ are in general only continuous and multivalued near an exceptional point, every *symmetric* function of them is single-valued and analytic in $\boldsymbol{\gamma}$: by the argument principle the power sums $\mathfrak{p}_r(\boldsymbol{\gamma}) = \frac{1}{2\pi i}\oint_{\Gamma_n} s^r\, \partial_s\Delta_{\boldsymbol{\gamma}}(s)/\Delta_{\boldsymbol{\gamma}}(s)\, ds$ depend analytically on $\boldsymbol{\gamma}$ as long as no eigenvalue crosses Γ_n, and by Newton's identities so do $\hat{\lambda}_n$, the coefficients $\hat{\beta}_j$, and hence $\pi_n(\cdot,\boldsymbol{\gamma})$. This is precisely what makes (5.118)–(5.119) single-valued and smooth across $\boldsymbol{\gamma}_*$, whereas the individual eigenvectors are not.
Define the right/left shifted symmetric basis by

$$\hat{\mathbf{u}}_{n,j}(\boldsymbol{\gamma}) := \frac{1}{2\pi i}\oint_{\Gamma_n(\boldsymbol{\gamma})} \frac{\hat{\chi}_j(s,\boldsymbol{\gamma})}{\pi_n(s,\boldsymbol{\gamma})}\, \boldsymbol{\varphi}(s,\boldsymbol{\gamma})\, ds, \qquad j = 0,1,\ldots,m_n-1, \tag{5.118}$$

$$\hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma}) := \frac{1}{2\pi i}\oint_{\Gamma_n(\boldsymbol{\gamma})} (s - \hat{\lambda}_n(\boldsymbol{\gamma}))^p\, \frac{\boldsymbol{\Psi}(s,\boldsymbol{\gamma})}{\Delta_{\boldsymbol{\gamma}}(s)}\, ds, \qquad p = 0,1,\ldots,m_n-1. \tag{5.119}$$

Here $\hat{\chi}_j(\cdot,\boldsymbol{\gamma})$ is the unique polynomial of degree $\leq m_n - 1$ dual to the moments of (5.119), i.e. such that

$$\frac{1}{2\pi i}\oint_{\Gamma_n(\boldsymbol{\gamma})} (s - \hat{\lambda}_n(\boldsymbol{\gamma}))^p\, \hat{\chi}_j(s,\boldsymbol{\gamma})\, \frac{ds}{\pi_n(s,\boldsymbol{\gamma})} = \delta_{pj}, \qquad p = 0,\ldots,m_n-1, \tag{5.120}$$

which mirrors (5.107) in the shifted variable. Explicitly, with the $\hat{\beta}_j$ of (5.117) and the convention

$\hat{\beta}_{m_n} := 1$,

$$\hat{\chi}_j(s,\boldsymbol{\gamma}) = \sum_{r=0}^{m_n-1-j} \hat{\beta}_{r+1+j}(\boldsymbol{\gamma})\,\big(s-\hat{\lambda}_n(\boldsymbol{\gamma})\big)^r, \tag{5.121}$$

so $\hat{\chi}_{m_n-1} = 1$ and $\hat{\chi}_{m_n-2}(s) = s - \hat{\lambda}_n$.
For simple eigenvalues in the block, these reduce to residue sums; at EPs they reduce to confluent residues and recover the Jordan chain attached to λ_n [131].

Exceptional-point limit (generalized eigenfunctions). At an exceptional point $\boldsymbol{\gamma} = \boldsymbol{\gamma}_*$ of order m_n,

$$\pi_n(s,\boldsymbol{\gamma}) \longrightarrow (s-\lambda_n)^{m_n}, \qquad \Delta_{\boldsymbol{\gamma}}(s) = (s-\lambda_n)^{m_n}\, g_n(s,\boldsymbol{\gamma}), \quad g_n(\lambda_n,\boldsymbol{\gamma}_*) = Z_n(\boldsymbol{\gamma}_*) \neq 0,$$

hence (5.118)–(5.119) reduce to confluent residues and recover the generalized eigenfunctions (the canonical system of eigenvectors and associated vectors, CSEAV) [131]

$$\hat{\mathbf{u}}_{n,j}(\boldsymbol{\gamma}_*) = \frac{1}{2\pi i}\oint_{\Gamma_n(\boldsymbol{\gamma}_*)} \frac{\boldsymbol{\varphi}(s,\boldsymbol{\gamma}_*)}{(s-\lambda_n)^{j+1}}\,ds = \frac{1}{j!}\,\partial_s^j[\boldsymbol{\varphi}(s,\boldsymbol{\gamma}_*)]_{s=\lambda_n} \equiv \boldsymbol{\varphi}_n^{(j)}(\boldsymbol{\gamma}_*), \tag{5.122}$$

$$\begin{aligned}\hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma}_*) &= \frac{1}{2\pi i}\oint_{\Gamma_n(\boldsymbol{\gamma}_*)} \frac{\boldsymbol{\Psi}(s,\boldsymbol{\gamma}_*)}{g_n(s,\boldsymbol{\gamma}_*)\,(s-\lambda_n)^{m_n-p}}\,ds \\ &= \frac{1}{(m_n-1-p)!}\,\partial_s^{m_n-1-p}\left[\frac{(s-\lambda_n)^{m_n}\,\boldsymbol{\Psi}(s,\boldsymbol{\gamma}_*)}{\Delta_{\boldsymbol{\gamma}_*}(s)}\right]_{s=\lambda_n} \equiv \boldsymbol{\Psi}_n^{(m_n-1-p)}(\boldsymbol{\gamma}_*).\end{aligned} \tag{5.123}$$

Biorthogonality (moving-basis form). With the CSEAV normalization of (4.47), the families $\{\hat{\mathbf{u}}_{n,j}(\boldsymbol{\gamma})\}_{j=0}^{m_n-1}$ and $\{\hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma})\}_{p=0}^{m_n-1}$ satisfy

$$\big\langle \hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma}), \hat{\mathbf{u}}_{l,q}(\boldsymbol{\gamma})\big\rangle = \delta_{nl}\,\delta_{pq}, \qquad \big\langle \hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma}), \boldsymbol{\varphi}_0(\boldsymbol{\gamma})\big\rangle = 0, \tag{5.124}$$

for $n,l \geq 1$, $p = 0,\dots,m_n-1$ and $q = 0,\dots,m_l-1$. The cross-block relations are immediate from $\hat{\boldsymbol{\eta}}_{n,p} \in \mathbb{E}_n^*$, $\hat{\mathbf{u}}_{l,q} \in \mathbb{E}_l$ and the disjointness of the contours; they are needed because (5.125) sums over all blocks.

Block dynamics and flux in the truncated model

Using completeness we expand (formally)

$$p(\cdot,t) = \boldsymbol{\varphi}_0(\boldsymbol{\gamma}(t)) + \sum_{n\geq 1}\sum_{j=0}^{m_n-1} b_{n,j}(t)\,\hat{\mathbf{u}}_{n,j}(\boldsymbol{\gamma}(t)), \qquad b_{n,j}(t) := \big\langle \hat{\boldsymbol{\eta}}_{n,j}(\boldsymbol{\gamma}(t)), p(\cdot,t)\big\rangle. \tag{5.125}$$

Projecting gives

$$\dot{\mathbf{b}}_n = \hat{\boldsymbol{\Omega}}_n(\boldsymbol{\gamma}(t))\,\mathbf{b}_n + \dot{\boldsymbol{\gamma}}(t)\cdot\sum_{l\geq 1}\hat{\boldsymbol{\Upsilon}}_{nl}^{\boldsymbol{\gamma}}(\boldsymbol{\gamma}(t))\,\mathbf{b}_l + \dot{\boldsymbol{\gamma}}(t)\cdot\hat{\mathbf{c}}_n^{\boldsymbol{\gamma}}(\boldsymbol{\gamma}(t)). \tag{5.126}$$

By the flux normalization, the shifted basis satisfies

$$\mathcal{N}[\hat{\mathbf{u}}_{n,0}(\boldsymbol{\gamma})] = 1, \qquad \mathcal{N}[\hat{\mathbf{u}}_{n,j}(\boldsymbol{\gamma})] = 0 \;\; (j = 1,\dots,m_n-1),$$

so the firing rate reads

$$\nu(t) = \Phi(\boldsymbol{\gamma}(t)) + \sum_{n\geq 1} b_{n,0}(t).$$

Companion form of $\hat{\boldsymbol{\Omega}}_n$ and explicit $\hat{\boldsymbol{\Upsilon}}^{\boldsymbol{\gamma}}$

Fix a block n of algebraic multiplicity m_n and recall the hatted (shifted) eigenvalues

$$\delta\hat{\lambda}_{n,k}(\boldsymbol{\gamma}) := \lambda_{n,k}(\boldsymbol{\gamma}) - \hat{\lambda}_n(\boldsymbol{\gamma}), \qquad \hat{\lambda}_n(\boldsymbol{\gamma}) := \frac{1}{m_n}\sum_{k=0}^{m_n-1} \lambda_{n,k}(\boldsymbol{\gamma}). \tag{5.127}$$

and the shifted characteristic polynomial $\hat{\pi}_n^{\mathrm{sh}}(z,\boldsymbol{\gamma})$ of (5.117), with coefficients $\hat{\beta}_j(\boldsymbol{\gamma})$. Recall that $\hat{\beta}_{m_n-1} = 0$ because the shift has zero mean; this is what makes the diagonal of (5.128) exactly $\hat{\lambda}_n$.

Internal block generator (companion matrix). In the shifted symmetric basis $\{\hat{\mathbf{u}}_{n,j}(\boldsymbol{\gamma})\}_{j=0}^{m_n-1}$ (ordered as in (5.125)), the restriction of $\mathcal{L}_{\boldsymbol{\gamma}}$ to the block subspace is represented by a (Frobenius) companion matrix associated with $\hat{\pi}_n^{\mathrm{sh}}$:

$$\hat{\boldsymbol{\Omega}}_n(\boldsymbol{\gamma}) = \begin{pmatrix} \hat{\lambda}_n(\boldsymbol{\gamma}) & 1 & 0 & \cdots & 0 \\ 0 & \hat{\lambda}_n(\boldsymbol{\gamma}) & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\lambda}_n(\boldsymbol{\gamma}) & 1 \\ -\hat{\beta}_0(\boldsymbol{\gamma}) & -\hat{\beta}_1(\boldsymbol{\gamma}) & \cdots & -\hat{\beta}_{m_n-2}(\boldsymbol{\gamma}) & \hat{\lambda}_n(\boldsymbol{\gamma}) \end{pmatrix}. \tag{5.128}$$

Exceptional-point limit (Jordan block). As $\boldsymbol{\gamma} \to \boldsymbol{\gamma}_*$ one has $\delta\hat{\lambda}_{n,k}(\boldsymbol{\gamma}) \to 0$ for all k, hence $\hat{\beta}_j(\boldsymbol{\gamma}) \to 0$ and

$$\hat{\boldsymbol{\Omega}}_n(\boldsymbol{\gamma}) \longrightarrow \hat{\lambda}_n(\boldsymbol{\gamma}_*)\mathbf{I}_{m_n} + \mathbf{S}_{m_n}, \qquad (\mathbf{S}_{m_n})_{q,q+1} = 1, \tag{5.129}$$

i.e. $\hat{\boldsymbol{\Omega}}_n$ converges to a Jordan block.

Regularized coupling blocks. The non-adiabatic coupling matrices are defined by

$$(\hat{\boldsymbol{\Upsilon}}^{\boldsymbol{\gamma}}_{nl}(\boldsymbol{\gamma}))_{pq} := \big\langle \partial_{\boldsymbol{\gamma}}\hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma}),\, \hat{\mathbf{u}}_{l,q}(\boldsymbol{\gamma})\big\rangle, \qquad p = 0,\dots,m_n-1,\;\; q = 0,\dots,m_l-1. \tag{5.130}$$

Projecting $\partial_t\boldsymbol{\varphi}_0(\boldsymbol{\gamma}(t)) = \dot{\boldsymbol{\gamma}}(t)\cdot\partial_{\boldsymbol{\gamma}}\boldsymbol{\varphi}_0(\boldsymbol{\gamma}(t))$ onto the moving dual block basis yields

$$\hat{\mathbf{c}}^{\boldsymbol{\gamma}}_{n,p}(\boldsymbol{\gamma}) := -\big\langle \hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma}),\, \partial_{\boldsymbol{\gamma}}\boldsymbol{\varphi}_0(\boldsymbol{\gamma})\big\rangle, \qquad p = 0,1,\dots,m_n-1. \tag{5.131}$$

For $\boldsymbol{\gamma} \neq \boldsymbol{\gamma}_*$ we have:

$$\hat{\mathbf{c}}^{\boldsymbol{\gamma}}_{n,p}(\boldsymbol{\gamma}) = \sum_{k=0}^{m_n-1} \big(\delta\hat{\lambda}_{n,k}(\boldsymbol{\gamma})\big)^p\, \mathbf{c}^{\boldsymbol{\gamma}}_{n,k}(\boldsymbol{\gamma}), \qquad \mathbf{c}^{\boldsymbol{\gamma}}_{n,k}(\boldsymbol{\gamma}) := -\big\langle \boldsymbol{\Psi}_{n,k}(\boldsymbol{\gamma}), \partial_{\boldsymbol{\gamma}}\boldsymbol{\varphi}_0(\boldsymbol{\gamma})\big\rangle. \tag{5.132}$$

We now establish the property anticipated in the introduction: unlike the instantaneous eigenbasis, the regularized coupling coefficients stay finite, indeed analytic, through an exceptional point. Two

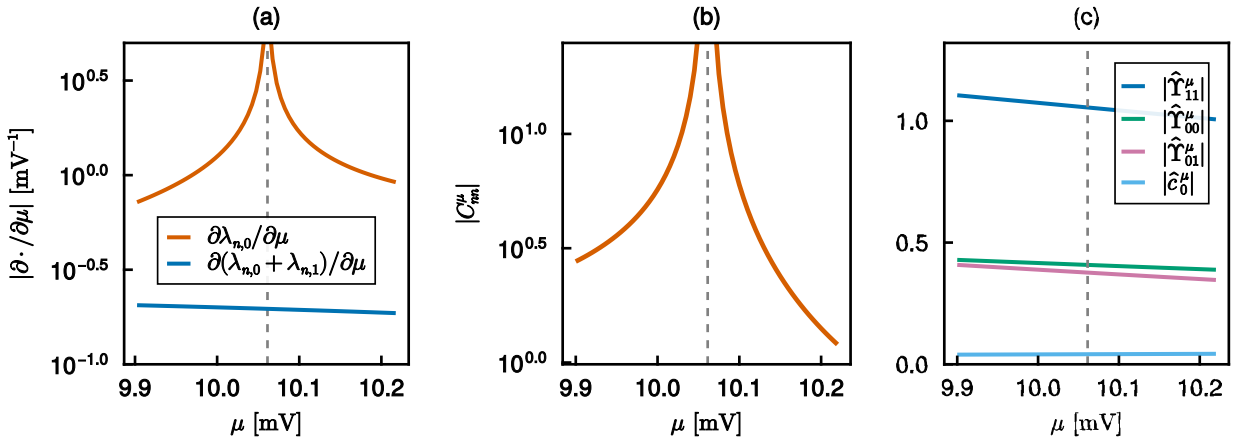

Figure 5.2: Regularization of the coupled dynamics across an exceptional point, for the collision on the second mixed branch of Fig. 4.2 ($D = 8$ mV^2, $H = 0$ mV, $\theta = 20$ mV, $\alpha = -3\theta$, $\tau_0 = 0$), as the mean input μ is swept across the exceptional point $\boldsymbol{\gamma}_*$ (dashed line). (a) The sensitivity $\partial\lambda_{n,0}/\partial\mu$ of a single block eigenvalue diverges at $\boldsymbol{\gamma}_*$, whereas the sensitivity of the eigenvalue sum $\lambda_{n,0} + \lambda_{n,1}$, a symmetric function of the block, stays finite. (b) The standard modal coupling C^{μ}_{nn} diverges accordingly. (c) The regularized coefficients $\hat{\Upsilon}^{\mu}_{pq}$ and $\hat{c}^{\mu}_{p}$ of the moving block basis remain bounded and smooth, as established in Lemma 5.4.1.

structural facts are used. First, near $\boldsymbol{\gamma}_*$ the contour Γ_n may be chosen *independent of* $\boldsymbol{\gamma}$: since the block stays clustered through the collision, a single curve encloses its m_n eigenvalues, and no others, for all $\boldsymbol{\gamma}$ in a neighbourhood Q of $\boldsymbol{\gamma}_*$. Second, $\boldsymbol{\varphi}(s, \boldsymbol{\gamma})$, $\boldsymbol{\Psi}(s, \boldsymbol{\gamma})$ and $\Delta_{\boldsymbol{\gamma}}(s)$ are jointly holomorphic in $(s, \boldsymbol{\gamma})$, as solutions of an ordinary differential equation whose coefficients are analytic in $\boldsymbol{\gamma} = (\mu, D)$.

Lemma 5.4.1 (Regularity at exceptional points)**.** *Under these two conditions the maps* $\boldsymbol{\gamma} \mapsto \hat{\mathbf{u}}_{n,j}(\boldsymbol{\gamma})$ *and* $\boldsymbol{\gamma} \mapsto \hat{\boldsymbol{\eta}}_{n,p}(\boldsymbol{\gamma})$ *are holomorphic on* Q*. Consequently the coupling coefficients* (5.130)–(5.131) *are analytic on* Q*; in particular they are bounded, and their limits as* $\boldsymbol{\gamma} \to \boldsymbol{\gamma}_*$ *exist and are finite.*

Proof. Denominators bounded below. On the fixed contour Γ_n both $\Delta_{\boldsymbol{\gamma}}$ and $\pi_n(\cdot, \boldsymbol{\gamma})$ vanish only at the enclosed eigenvalues, which lie strictly inside. Being continuous and nonzero on the compact set $\Gamma_n \times Q$, they are bounded below there: $|\Delta_{\boldsymbol{\gamma}}(s)| \geq c$ and $|\pi_n(s, \boldsymbol{\gamma})| \geq c$ for some $c > 0$.
Numerators analytic. The shift $\hat{\lambda}_n(\boldsymbol{\gamma})$ and coefficients $\hat{\beta}_j(\boldsymbol{\gamma})$ are power sums, respectively elementary symmetric functions, of the enclosed eigenvalues, given by the fixed-contour integrals

$$\mathfrak{p}_r(\boldsymbol{\gamma}) = \frac{1}{2\pi i} \oint_{\Gamma_n} s^r \, \frac{\partial_s \Delta_{\boldsymbol{\gamma}}(s)}{\Delta_{\boldsymbol{\gamma}}(s)} \, ds, \tag{5.133}$$

holomorphic in $\boldsymbol{\gamma}$ by the bound above and the joint analyticity of $\Delta_{\boldsymbol{\gamma}}$. Hence the numerators $\hat{\chi}_j(s, \boldsymbol{\gamma})$ and $(s - \hat{\lambda}_n(\boldsymbol{\gamma}))^p \boldsymbol{\Psi}(s, \boldsymbol{\gamma})$ are jointly analytic in $(s, \boldsymbol{\gamma})$.
Basis functions and coefficients. The integrands of (5.118) and (5.119) are then jointly analytic on $\Gamma_n \times Q$, an analytic numerator over a denominator bounded below. Integrating over the fixed contour and differentiating under the integral sign shows that $\hat{\mathbf{u}}_{n,j}$ and $\hat{\boldsymbol{\eta}}_{n,p}$, together with $\partial_{\boldsymbol{\gamma}} \hat{\boldsymbol{\eta}}_{n,p}$, are holomorphic maps $Q \to \mathbb{L}^2$. The coefficients (5.130) and (5.131) are inner products of these maps with one another and with $\partial_{\boldsymbol{\gamma}} \boldsymbol{\varphi}_0$, itself analytic since $\lambda_0 = 0$ stays simple and isolated. They are therefore analytic scalars, and analyticity on the compact Q gives boundedness with finite limits at $\boldsymbol{\gamma}_*$. □

The contrast with the instantaneous eigenbasis is sharp. The standard coefficient $\mathbf{C}^{\boldsymbol{\gamma}}_{nn} = \langle \partial_{\boldsymbol{\gamma}} \boldsymbol{\Psi}_n, \boldsymbol{\varphi}_n \rangle$

diverges at $\boldsymbol{\gamma}_*$ because $\boldsymbol{\Psi}_n(\boldsymbol{\gamma}) = Z_n(\boldsymbol{\gamma})^{-1}\boldsymbol{\Psi}(\lambda_n(\boldsymbol{\gamma}), \boldsymbol{\gamma})$ depends on $\boldsymbol{\gamma}$ through $\lambda_n(\boldsymbol{\gamma})$, which has a square-root branch point there (Section 4.5): both $\partial_{\boldsymbol{\gamma}}\lambda_n$ and Z_n^{-1} blow up like $|\boldsymbol{\gamma} - \boldsymbol{\gamma}_*|^{-1/2}$. The regularized construction never evaluates a root function at λ_n and never divides by $Z_n = \Delta'_{\boldsymbol{\gamma}}(\lambda_n)$; it divides by $\Delta_{\boldsymbol{\gamma}}$ on a contour where that quantity is bounded below. This is the precise sense in which transporting the invariant subspace, rather than the individual eigenvectors, removes the singularity. Figure 5.2 shows this numerically at the exceptional point of Fig. 4.2: the sensitivity of a single eigenvalue and the standard coupling $\mathbf{C}^{\boldsymbol{\gamma}}_{nn}$ both diverge, while the regularized coefficients stay bounded and smooth.

Chapter 6

Multi-Task Reservoir Computing

Rose is a rose is a rose is a rose

— Gertrude Stein, *Sacred Emily*

6.1 From Population Density to Rate Dynamics

Having established the spectral description of a single population density, we now shift our perspective from the mesoscopic to the macroscopic scale. Our goal is to understand how networks formed by multiple interacting populations can perform computation. We have seen that the PDE dynamics can be decomposed into a system of ordinary differential equations for the eigenmode projections. To simulate a large network of interacting populations, we can further simplify the dynamics by exploiting the timescale separation inherent in the spectral decomposition. In many biological regimes, the higher-order modes ($K \gg 1$) correspond to fast relaxation timescales that vanish almost instantaneously compared to the relevant network dynamics. We can therefore effectively reduce the system dimensionality by truncating the expansion to the first K non-stationary modes.
Following the standard heuristic approach of [124], we now show how to describe the dynamics of a neural population by a single K-th order ODE on the firing rate. In the present setting, this reduction follows directly from the adiabatic approximation applied to the regularized coupled dynamics developed above. The classical result of [124] thus emerges as the adiabatic limit of our dynamics in the global moment basis.

Adiabatic truncation and K-th order firing-rate ODE. Starting from the truncated regularised dynamics (5.112), we assume slow parameter variation $\dot{\boldsymbol{\gamma}}(t) \approx 0$ and drop the non-adiabatic coupling terms $\dot{\boldsymbol{\gamma}} \cdot \boldsymbol{\Upsilon}_K^{\boldsymbol{\gamma}}$, $\dot{\boldsymbol{\gamma}} \cdot \tilde{\mathbf{c}}_K^{\boldsymbol{\gamma}}$, obtaining

$$\dot{\mathbf{b}}(t) = \boldsymbol{\Omega}_K(\boldsymbol{\gamma})\, \mathbf{b}(t), \qquad \mathbf{b} = (b_0, \dots, b_{K-1})^\top. \tag{6.1}$$

Recalling that $\boldsymbol{\Omega}_K$ has the Frobenius companion form given by Equation (5.110), it follows immediately that the transient rate $b_0(t) = \nu(t) - \Phi(\boldsymbol{\gamma})$ satisfies a scalar K-th order ODE, with coefficients

(β_j) (see [191, Section 3.3]). We can then write the following:

$$\nu^{(K)}(t) + \beta_{K-1}(\boldsymbol{\gamma})\,\nu^{(K-1)}(t) + \cdots + \beta_1(\boldsymbol{\gamma})\,\nu'(t) + \beta_0(\boldsymbol{\gamma})\,\big(\nu(t) - \Phi(\boldsymbol{\gamma})\big) = 0, \tag{6.2}$$

exactly recovering the result of [124].

6.1.1 Balanced Dynamics and the First-Order Limit

The general K-th order ODE allows for complex transient behaviors, including oscillations. However, we are specifically interested in the dynamical regime of *balanced networks*.
In a balanced Excitation-Inhibition (EI) network, neurons are driven by large currents that fluctuate significantly. This corresponds to the noise-dominated regime. Physically, the strong diffusion suppresses coherent oscillations in the probability density, resulting in an overdamped relaxation.
Mathematically, this regime implies that real eigenvalues dominate the spectrum of the FP operator $\mathcal{L}$. Furthermore, the first eigenvalue λ_1 is typically separated from the subsequent modes by a large spectral gap ($|\lambda_1| \ll |\lambda_2| < \dots$). Consequently, the macroscopic dynamics are dominated by the slowest decay mode.
Truncating the expansion to $K = 1$, the dynamics collapse to a first-order relaxation process:

$$-\frac{1}{\lambda_1}\dot{\nu}(t) = \Phi(\boldsymbol{\gamma}(t)) - \nu(t) \tag{6.3}$$

Defining the effective population time constant as $\tau = -1/\lambda_1$, we arrive at the Wilson-Cowan-like rate equation:

$$\tau\dot{\nu}(t) = -\nu(t) + \Phi(\mu(t), D(t)) \tag{6.4}$$

6.1.2 Network of Rate Equations

We can now extend this description to a macroscopic network of N interacting populations. We identify the instantaneous firing rate of the i-th population, $\nu_i(t)$, with the state variable $h_i(t)$.
The input statistics $\mu_i(t)$ and $D_i(t)$ for population i are determined by the sum of recurrent inputs from other populations and external drives. Assuming the variance D is effectively constant in the balanced state, the mean drive μ_i becomes the primary time-dependent variable:

$$\mu_i(t) = \sum_{j=1}^{N} J_{ij} h_j(t) + u_i(t) \tag{6.5}$$

where $u_i(t)$ is the external current. Substituting this into the single-population dynamics, we obtain the system of coupled differential equations that defines the rate-based Recurrent Neural Network (RNN) model:

$$\tau\dot{h}_i(t) = -h_i(t) + \Phi\left(\sum_{j=1}^{N} J_{ij} h_j(t) + u_i(t)\right) \tag{6.6}$$

where $\Phi(\cdot)$ is the transfer function derived from the stationary Fokker-Planck solution Φ. This derivation provides a rigorous biophysical justification for the use of Eq. (6.6) in modeling the

dynamics of interacting neural populations in the limit of high-noise, balanced activity where the spectral expansion is dominated by the fundamental decay mode.

6.2 Recurrent Neural Networks: A Theoretical Framework

Recurrent Neural Networks occupy a unique interface between computational neuroscience, machine learning, and statistical physics. As we have seen, the rate-based dynamics derived in Eq. (6.6) constitute an effective macroscopic description of interacting neuronal populations. Simultaneously, these architectures are heavily utilized in machine learning for temporal processing and learning dynamical systems from data. In the limit of large network size N, random RNNs behave as disordered physical systems that can be studied through mean-field analysis.

6.2.1 RNNs in Machine Learning

In **Machine Learning**, RNNs are heavily utilized for temporal processing and modeling dynamical systems from data. Their ubiquity stems from their remarkable theoretical expressivity: they constitute universal approximators of dynamical systems. Rigorously, the rate-based hidden dynamics can approximate any finite-time trajectory of an open dynamical system with arbitrary precision [66]. More broadly, RNNs with rational weights and unbounded computation time are Turing-complete [181].

Despite this theoretical universality, RNN-based architectures face profound practical challenges in optimization and interpretability. Training these networks involves shaping the system's state-space trajectory through optimization. The standard approach, *Backpropagation Through Time* (BPTT), typically treats the dynamics as a discrete process: it unfolds the recurrent activity across temporal steps and manages this structure as a deep feedforward network with weight sharing, enabling gradient-based learning with respect to the connectivity matrix J [205].

However, BPTT suffers from pathologies rooted in how gradients propagate through the network's nonlinear dynamics, most notably the vanishing or exploding gradient problem[87, 149]. To overcome this, one must specifically constrain the architecture, for example by introducing gating mechanisms as found in Long Short-Term Memory (LSTM) networks. Beyond optimization, BPTT typically yields solutions that are theoretically universal but practically opaque. The learned connectivity matrices J are high-dimensional, full-rank objects that successfully solve the task but offer limited insight into the underlying algorithmic mechanism.

6.2.2 Computation Through Dynamics

Computational Neuroscience has moved beyond these black-box descriptions by treating the RNN as a normative model for cortical dynamics. This paradigm, termed "computation through dynamics," posits that cognitive functions emerge from the collective trajectories of neural populations rather than from single-neuron selectivity [201].

A foundational example is the generation of movement. Churchland and collaborators demonstrated that motor cortex activity during reaching is best understood not as a representation of muscle parameters, but as a dynamical system generating oscillatory patterns. The preparatory state serves as an initial condition that seeds specific rotational trajectories, a mechanism naturally implemented

by strong recurrent dynamics [45].
This perspective clarifies higher cognitive functions as well. Mante ad collaborators revealed that prefrontal cortex solves context-dependent decision tasks not by rewiring connections, but by utilizing a fixed connectivity that supports flexible dynamics. The "context" input acts as a control parameter, shifting the system's flow field to direct trajectories toward task-relevant attractors [118]. Recent theoretical work has formalized this, showing how linear approximations of these manifolds can disentangle the specific mechanisms that circuits use to implement flexible computations [185]. "Computation" thus corresponds to neural dynamic and learing becomes equivalent to the geometric shaping of the system's phase space.

6.2.3 Statistical Mechanics of RNNs

To understand how such collective properties emerge from connectivity, we turn to **Statistical Physics**. The entry point is the Amari-Hopfield model [5, 90], which assumes symmetric connectivity ($J_{ij} = J_{ji}$). This symmetry guarantees the existence of an energy function that the dynamics minimizes, relaxing into stable fixed points representing stored memories
Because an energy function exists, these systems can be analyzed using equilibrium statistical mechanics. Amit, Gutfreund, and Sompolinsky applied the *replica method* from spin-glass theory to rigorously derive storage capacity and phase transitions in the thermodynamic limit [7, 8]. Their analysis revealed critical boundaries between retrieval phases and spin-glass phases, dominated by spurious metastable states. While mathematically elegant, the symmetry constraint is restrictive: biological circuits are inherently asymmetric, and functional behaviors like oscillation or transient amplification require non-equilibrium dynamics.

Asymmetry, Chaos, and Non-Equilibrium Dynamics

Relaxing the symmetry constraint removes the energy landscape, allowing for more complex phenomena to arise. The rigorous study of **asymmetric random networks** was pioneered by Sompolinsky, Crisanti, and Sommers [186], who analyzed networks with weights drawn from independent Gaussian distributions. They identified a sharp phase transition driven by the connectivity variance g^2: as g crosses unity, the stable origin destabilizes into *high-dimensional chaos*.
To solve this, they developed *Dynamical Mean-Field Theory* (DMFT). DMFT replaces the deterministic interaction term with a time-dependent stochastic process whose autocorrelation is determined self-consistently. These results have been placed on rigorous ground via *Large Deviation Principles*. Moynot and Samuelides proved [136] that the empirical measure of trajectories in asymmetric random networks converges exponentially fast to the unique solution of the mean-field equations. Crucially, recent work has extended this to **spatially structured networks**, proving that mean-field descriptions remain asymptotically exact even when position-dependent coupling introduces strong spatial correlations [32].

Structured Connectivity

While random networks capture generic dynamical features, functional computation requires structure. A powerful theoretical framework for understanding this involves decomposing the connectivity matrix into a structured low-rank component and a random bulk term.

In the purely low-rank limit, the high-dimensional neural activity is confined to a finite-dimensional subspace determined by the rank of the structure. This reduction allows the dynamics to be described by a small set of ODEs, making it possible to explicitly embed attractors by designing the constituent latent vectors [122].
When a random component is added, this low-dimensional manifold interacts with the high-dimensional disorder. As shown by Schuessler et al. [173], while uncorrelated noise simply suppresses structured dynamics, correlations between the structured and random components, typical of trained networks, can fundamentally reshape the eigenspectrum, allowing robust computation to emerge from the disordered substrate.

6.3 Reservoir Computing

Reservoir Computing (RC) is a machine learning paradigm grounded in the theory of nonlinear dynamical systems and statistical physics. Originally developed independently as *Echo State Networks* (ESN) by Jaeger [92] and *Liquid State Machines* by Maass [114], RC provides a computationally efficient method for modeling complex temporal dynamics. Unlike fully trainable Recurrent Neural Networks (RNNs), which require expensive and often unstable gradient-based optimization of internal weights, RC relies on a fixed, random dynamical substrate (the reservoir) to project inputs into a high-dimensional feature space. This approach separates the generation of dynamics from the learning process, which is restricted to a linear readout layer.
The core of the RC framework is a fixed, random RNN that acts as a non-linear temporal filter. We consider a continuous-time formulation where the reservoir is driven by a time-dependent input signal $\mathbf{u}(t) \in \mathbb{R}^K$. The evolution of the N hidden units follows the RNN dynamics (6.6). Written in matrix form, it reads:

$$\tau \dot{\mathbf{h}}(t) = -\mathbf{h}(t) + \Phi\left(J\mathbf{h}(t) + J^{\text{in}}\mathbf{u}(t) + \mathbf{b}\right) \tag{6.7}$$

The connectivity matrix $J \in \mathbb{R}^{N\times N}$ and the input projection matrix $J^{\text{in}} \in \mathbb{R}^{N\times K}$ are fixed random matrices drawn from specific distributions (typically Gaussian or sparse). The vector $\mathbf{b}$ represents a bias term, which shifts the operating point of the non-linearity; in practice, this term is often absorbed into the definition of the transfer function $\Phi(\cdot)$ or treated as an additional constant input.
The reservoir transforms the input history into a high-dimensional state representation. To utilize these features, we a linear readout is trained to approximate a target signal $\mathbf{y}_{\text{target}}(t) \in \mathbb{R}^M$. The network output $\mathbf{y}(t)$ is given by:

$$\mathbf{y}(t) = \Theta\mathbf{h}(t) \tag{6.8}$$

The readout weights $\Theta \in \mathbb{R}^{M\times N}$ are optimized to minimize the squared error loss $\mathcal{E}(\Theta) = \langle|\mathbf{y}_{\text{target}}(t) - \Theta\mathbf{h}(t)|^2\rangle_t$, typically via ridge regression. Mathematically, the reservoir implements a functional mapping from the input history to the output:

$$\mathbf{y}(t) = \mathrm{RNN}_{\Theta}(\{\mathbf{u}\}_t) \tag{6.9}$$

where $\{\mathbf{u}\}_t = \{\mathbf{u}(s) : s \leq t\}$ represents the input trajectory up to time t. For this mapping to be well-defined and independent of the initial conditions at $t \to -\infty$, the network must satisfy the *Echo State Property* (ESP) [212]. This property guarantees that the state $\mathbf{h}(t)$ is asymptotically determined solely by the driving signal.

Dynamical System Modeling and Autonomous Mode

A primary application of RC is the data-driven modeling of dynamical systems. In this context, the task is to predict the future evolution of the input itself, i.e., $\mathbf{y}_{\text{target}}(t) = \mathbf{u}(t + dt)$. Once trained, the network can be switched to *autonomous mode* by closing the loop, feeding the prediction back as the input:

$$\tau\dot{\mathbf{h}}(t) = -\mathbf{h}(t) + \Phi\left((J + J^{\text{in}}\Theta)\mathbf{h}(t) + \mathbf{b}\right) \tag{6.10}$$

This results in a new autonomous dynamical system where the effective connectivity $J_{\text{eff}} = J + J^{\text{in}}\Theta$ consists of a fixed random component and a trained low-rank perturbation, a structure that has been shown to support rich dynamical behaviors [173].
The capability of reservoir computing to reconstruct unobserved dynamics relies on embedding theory. Generalized Synchronization guarantees that under mild conditions, the coupling between a driving system and a driven response system (the reservoir) induces a continuous map between their state spaces [83].
Hart et al. [83] provided a rigorous foundation for this phenomenon, proving that a reservoir effectively constructs an embedding of the driving system's attractor. This can be viewed as a neural implementation of the Takens embedding theorem, which states that the state of a dynamical system can be reconstructed from delay coordinates of a single observable. Specifically, Hart et al. demonstrate that for a sufficiently large reservoir, the mapping from the reservoir state $\mathbf{h}(t)$ to the next input value is a smooth function that can be uniformly approximated by the linear readout Θ. Consequently, the reservoir does not merely memorize patterns but learns the topological structure of the underlying attractor.

6.4 Readout as a Latent Variable Model

We propose a probabilistic framework that explicitly models the shared structure across a family of dynamical tasks. Let $\mathcal{A}$ denote a family of related dynamical systems or tasks. For each task $\alpha \in \mathcal{A}$, we observe a state trajectory $\mathbf{x}_\alpha(t)$ and use a fixed reservoir to extract a corresponding feature trajectory $\mathbf{h}_\alpha(t)$.
As in standard Reservoir Computing, the features are obtained by driving the reservoir with the

input sequence:

$$\mathbf{h}_\alpha(t) = \mathrm{RNN}(\{\mathbf{x}_\alpha\}_t) \tag{6.11}$$

However, instead of learning an independent full-rank readout matrix Θ_α for each task, we parameterize the readout as a linear combination of a shared basis set. Specifically, we define a set of K global readout components $\Xi = \{\Theta^{(k)}\}_{k=1}^K$, where $\Theta^{(k)} \in \mathbb{R}^{M\times N}$, and a task-specific latent vector $\mathbf{g}_\alpha \in \mathbb{R}^K$. The effective readout for task $\boldsymbol{\alpha}$ is given by:

$$\Theta_\alpha = \sum_{k=1}^{K} g_\alpha^{(k)} \Theta^{(k)} \tag{6.12}$$

where $g_\alpha^{(k)}$ denotes the k-th component of the vector $\mathbf{g}_\alpha$. This formulation imposes a low-rank structure on the family of readouts, effectively treating the task identity as a continuous latent variable embedded in a K-dimensional space.

6.4.1 Probabilistic Formulation and Variational Inference

We frame the learning problem within the formalism of Variational Inference (VI). We assume a generative model where the observations $\{\mathbf{x}_\alpha(t)\}$ are produced from the reservoir states - which themselves encode the history of past observations - via the latent readout with additive Gaussian noise:

$$\mathbf{x}_\alpha(t) = \left(\sum_{k=1}^{K} g_\alpha^{(k)} \Theta^{(k)}\right) \mathbf{h}_\alpha(t) + \boldsymbol{\epsilon}(t), \quad \boldsymbol{\epsilon}(t) \sim \mathrm{Gaussian}(0, \upsilon I) \tag{6.13}$$

We place zero-mean Gaussian priors on both the latent task variables and the basis matrices to enforce regularization:

$$p(\mathbf{g}_\alpha) = \mathrm{Gaussian}(0, \gamma_g^{-1} I) \tag{6.14}$$

$$p(\Xi) = \prod_{k=1}^{K} \mathrm{Gaussian}(\Theta^{(k)} | 0, \gamma_\Theta^{-1} I) \tag{6.15}$$

where γ_g and γ_Θ are precision hyperparameters controlling the regularization strength.
Our goal is to minimize the marginal negative log-likelihood of the data. Although the integral over the latent variables is analytically defined for linear-Gaussian models, evaluating and optimizing the resulting marginal likelihood is computationally prohibitive due to the high dimensionality of the data and the bilinear coupling between the basis Θ and the latent codes $\mathbf{g}$. To circumvent this, we construct the Evidence Lower Bound (ELBO), denoted here as the free energy functional $\mathcal{E}$:

$$-\log p(\{\mathbf{x}_\alpha\}) \leq \mathcal{E}(q) = \mathbb{E}_q\left[\log p(\{\mathbf{x}_\alpha\}, \{\mathbf{g}_\alpha\}, \Xi)\right] - \mathbb{E}_q\left[\log q(\{\mathbf{g}_\alpha\})\right] - \mathbb{E}_{\tilde{q}}\left[\log \tilde{q}(\Xi)\right] \tag{6.16}$$

where $q(\{\mathbf{g}_\alpha\})$ is the variational posterior over the latent codes and $q(\Xi)$ is the variational posterior over the basis. This formulation replaces the computationally prohibitive integral over the latent space with a tractable optimization problem, searching for the distribution q within a restricted

family that best approximates the true posterior .

6.4.2 Inference via Alternating Least Squares

We adopt a computationally efficient *deterministic approximation* by restricting both variational posteriors to Dirac delta distributions:

$$q(\{\mathbf{g}_\alpha\}) = \prod_\alpha \delta(\mathbf{g}_\alpha - \hat{\mathbf{g}}_\alpha) \tag{6.17}$$

$$\tilde{q}(\Xi) = \prod_k \delta(\Theta^{(k)} - \hat{\Theta}^{(k)}) \tag{6.18}$$

The energy functional then simplifies to[1]:

$$\mathcal{E}(\{\hat{\mathbf{g}}_\alpha\}, \hat{\Xi}) = \sum_\alpha \left(\frac{1}{2} \int_0^T \left\| \mathbf{x}_\alpha(t) - \sum_{k=1}^K \hat{g}_\alpha^{(k)} \hat{\Theta}^{(k)} \mathbf{h}_\alpha(t) \right\|^2 dt + \frac{\gamma_g}{2} \|\hat{\mathbf{g}}_\alpha\|^2 \right) + \frac{\gamma_\Theta}{2} \sum_k \|\hat{\Theta}^{(k)}\|^2 \tag{6.19}$$

We maximize this functional using a coordinate ascent scheme corresponding to a MAP-EM algorithm [22]. For implementation, we discretize the integral as a sum over time steps. The algorithm alternates between two convex sub-problems (*Alternating Least Squares*, ALS) [104]:

E-Step (Latent Code Inference): Fixing $\hat{\Xi}$, we update $\hat{\mathbf{g}}_\alpha$ via ridge regression:

$$\hat{\mathbf{g}}_\alpha \leftarrow \arg\min_{\mathbf{g}} \sum_t \|\mathbf{x}_\alpha(t) - \mathbf{F}_\alpha(t)\mathbf{g}\|^2 + \gamma_g \|\mathbf{g}\|^2 \tag{6.20}$$

where $\mathbf{F}_\alpha(t) \in \mathbb{R}^{M \times K}$ has columns $\hat{\Theta}^{(k)} \mathbf{h}_\alpha(t)$.

M-Step (Basis Learning): Fixing $\{\hat{\mathbf{g}}_\alpha\}$, we update the shared basis $\hat{\Xi}$ by regularized least squares over all tasks and times. Because the reconstruction $\sum_k \hat{g}_\alpha^{(k)} \Theta^{(k)} \mathbf{h}_\alpha$ couples all K basis matrices, they cannot be fitted independently and must be solved jointly. Identifying Ξ with its stacked matrix form $\Xi = [\,\Theta^{(1)} \;\cdots\; \Theta^{(K)}\,] \in \mathbb{R}^{M \times KN}$ and defining the lifted feature $\tilde{\mathbf{h}}_\alpha(t) := \mathbf{g}_\alpha \otimes \mathbf{h}_\alpha(t) \in \mathbb{R}^{KN}$, for which $\Xi \tilde{\mathbf{h}}_\alpha = \sum_k g_\alpha^{(k)} \Theta^{(k)} \mathbf{h}_\alpha$, the update becomes a single ridge regression

$$\hat{\Xi} \leftarrow \arg\min_{\Xi} \sum_{\alpha,t} \left\| \mathbf{x}_\alpha(t) - \Xi\, \tilde{\mathbf{h}}_\alpha(t) \right\|^2 + \gamma_\Theta \|\Xi\|_F^2, \tag{6.21}$$

with the closed-form solution

$$\hat{\Xi} = \Big(\sum_{\alpha,t} \mathbf{x}_\alpha(t)\, \tilde{\mathbf{h}}_\alpha(t)^\top \Big) \Big(\sum_{\alpha,t} \tilde{\mathbf{h}}_\alpha(t)\, \tilde{\mathbf{h}}_\alpha(t)^\top + \gamma_\Theta \mathbf{I}_{KN} \Big)^{-1}. \tag{6.22}$$

The diagonal blocks of the Gram matrix carry the correct quadratic weights $(\hat{g}_\alpha^{(k)})^2$, while the off-diagonal blocks $\hat{g}_\alpha^{(j)} \hat{g}_\alpha^{(k)}$ encode the coupling between basis components. Equivalently, the basis may

[1] As v is a proportionality constant irrelevant for the optimization, we set $v = 1$

be refined one component at a time by a Gauss–Seidel sweep, updating each $\hat{\Theta}^{(j)}$ against the partial residual $\mathbf{x}_\alpha - \sum_{k\neq j} \hat{g}_\alpha^{(k)} \hat{\Theta}^{(k)} \mathbf{h}_\alpha$; both schemes converge to the same stationary point.
This deterministic framework provides a scalable foundation for learning shared dynamics, though it is important to note that the resulting optimization objective is non-convex. Consequently, the coordinate ascent scheme is guaranteed to converge only to a local minimum, making the solution sensitive to initialization [22].
Future extensions could address this limitation by restoring the full Bayesian formalism, replacing the Dirac delta approximations with Gaussian variational posteriors. This would not only enable the quantification of uncertainty in the inferred latent trajectories but also smooth the optimization landscape, potentially helping to escape poor local minima. Furthermore, a fully Bayesian treatment allows for automatic relevance determination (ARD), a mechanism that naturally prunes unnecessary latent dimensions and effectively determines the optimal rank of the shared basis without cross-validation [115, 140].

6.4.3 Experiment: Generalization of the Lorenz Return Map

We validate the framework on the Lorenz system, a canonical model of chaotic dynamics governed by the equations:

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= x(\rho - z) - y \\ \dot{z} &= xy - \beta z \end{aligned} \tag{6.23}$$

A defining signature of the Lorenz attractor is the *return map* of the z-coordinate. If we define the sequence $\{z_n\}$ as the values of successive local maxima of the variable $z(t)$, the dynamics follow an approximate 1D difference equation:

$$z_{n+1} \approx F(z_n) \tag{6.24}$$

This function F takes the form of a *tent map*. The slope and peak height of this map are determined by β, providing a rigorous geometric metric for the fidelity of the learned dynamics.

Setup and Training. We generated a dataset of Lorenz trajectories by varying β while keeping $\sigma = 10$ and $\rho = 70$ fixed. The model was trained using trajectories from only three discrete parameter values: $\beta_{\text{train}} \in \{3.0, 4.0, 5.0\}$. During this phase, the algorithm jointly optimized the shared basis Ξ, with $K = 2$ and the specific latent codes $\mathbf{g}_\alpha$ for these three tasks.

Inference and Extrapolation. To test generalization, we presented the model with trajectories from unseen values of β, including interpolation ($\beta = 3.5, 4.5$) and extrapolation ($\beta = 2.5, 5.5$) points. For these test cases, the global basis Ξ was frozen. The task-specific latent coordinates $\hat{\mathbf{g}}_\beta$ were inferred by performing a single **E-step** (minimizing Eq. 6.19 with respect to $\hat{\mathbf{g}}_\beta$), effectively locating the new dynamical system within the learned readout manifold.
Once the code $\hat{\mathbf{g}}_\beta$ was inferred, we constructed the effective readout $\Theta_\beta = \sum_k \hat{g}_\beta^{(k)} \Theta^{(k)}$ and ran the reservoir in *autonomous closed-loop mode*. Figure 6.1 (Left) compares the return maps derived from

these autonomously generated trajectories against the ground truth.

The results show that the model successfully extrapolates the vector field structure. As shown in Fig. 6.1 (Right), the latent space spontaneously organizes into a 1D manifold ordered by β. This geometric order allows the model to synthesize the correct vector field for $\beta = 2.5$ or $\beta = 5.5$ simply by linear combination of the basis fields, recovering the correct tent map profile even for parameters outside the training distribution.

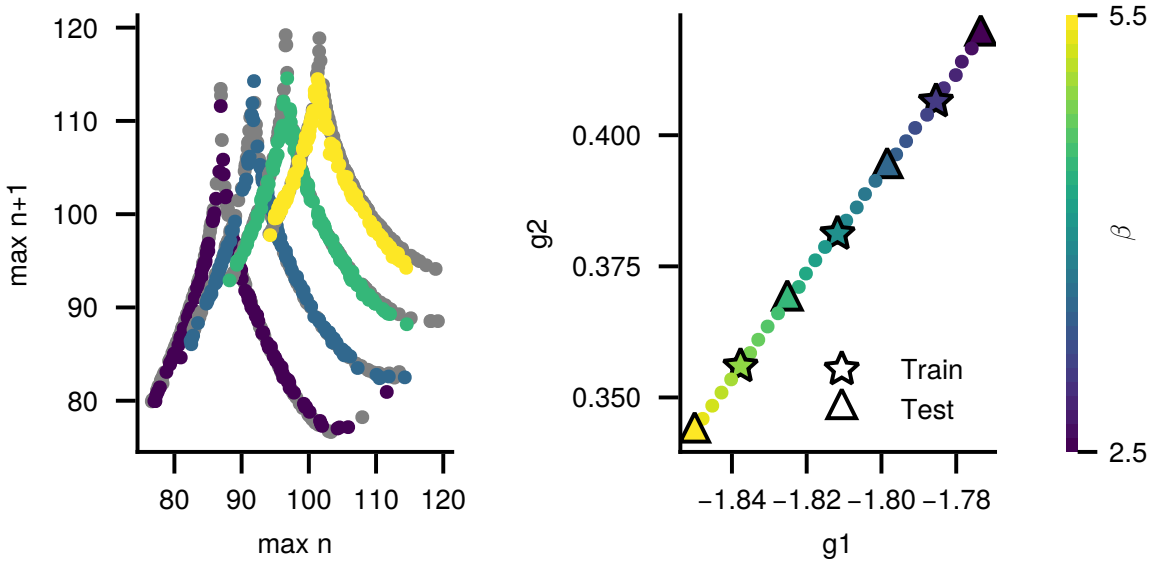

Figure 6.1: Latent extrapolation of the Lorenz return map. Left: Closed-loop predictions of the return map z_{n+1} vs z_n (colored dots) compared to the ground truth data (gray dots) for test set values $\beta \in \{2.5, 3.5, 4.5, 5.5\}$. The model accurately captures the varying cusp shape of the map, including for values outside the training range. **Right:** The inferred latent space. The model was trained using only $\beta \in \{3, 4, 5\}$ (black stars). For the test trajectories, the latent coordinates g_α were inferred via a single E-step (colored dots). The resulting latent space spontaneously organizes with respect to the physical parameter β, enabling the successful interpolation and extrapolation seen in the left panel. Network parameters: $N = 300$, $\Phi = \tanh$, $\tau = 0.05$.

Chapter 7

Adaptive Behaviour with Gain-modulated Networks

Colmo di fiori è il pesco,
Non tutti diventeranno frutto,
Splendono limpidi come schiuma di rose
Per l'azzurra fuga delle nubi.

Come fiori sbocciano i pensieri,
Cento al giorno –
Lasciali fiorire! lascia alle cose il loro corso!
Non domandare del raccolto!

Occorrono anche giuoco e innocenza
E fiori in abbondanza,
Altrimenti il mondo ci sarebbe angusto
E la vita priva di piacere.

— Hermann Hesse, *Colmo di fiori*

In Section 6.4, we established a probabilistic framework where task-specific dynamics emerge from a bilinear interaction between the reservoir state $\mathbf{h}$ and a latent modulatory variable $\mathbf{g}$. Within the associated Free Energy functional $\mathcal{E}(\Xi, \mathbf{g})$, $\mathbf{g}$ acts as a fast variable that adapts online to the current context, distinct from the slow evolution of the stable synaptic basis Ξ. This interaction mathematically mirrors the biological mechanism of gain modulation, often implemented via input segregation in pyramidal neurons, where context signals modulate afferent drives without requiring synaptic modification.

This perspective offers a computational explanation for a key biological phenomenon: behavioral changes in animals and humans, triggered by errors or verbal instructions, can occur extremely rapidly. While learning theories typically attribute improvements in performance to synaptic plasticity, recent findings suggest that such fast adaptations may instead result from dynamic reconfiguration of the networks involved without changes to synaptic weights.

Recently, similar capabilities have been observed in Transformers, foundational architectures in machine learning capable of *in-context learning*: the ability to adapt and acquire new information dynamically within the context of the task they are currently engaged in, without changing their

parameters.
We argue that this property may stem fundamentally from gain modulation. In this chapter, we propose a constructive approach to induce in-context learning in an architecture composed of recurrent networks with gain modulation, demonstrating abilities inaccessible to standard networks. In particular, we show that such an architecture can dynamically implement standard gradient-based optimization by encoding weight changes in the activity of another network. We argue that, while these algorithms are traditionally associated with synaptic plasticity, their reliance on non-local terms suggests that they may be more naturally realized in the brain at the level of neural circuits.
We demonstrate that we can extend our approach to temporal tasks and reinforcement learning. We further validate our approach in a MuJoCo ant navigation task, showcasing a neuromorphic control paradigm via real-time network reconfiguration. This framework, developed in collaboration with Cristiano Capone, was published in [36].

7.1 Introduction

Adaptive Behavior in Biological Agents The study of adaptive behavior in biological agents, including humans and animals, has been a longstanding subject of research [106]. Observations of rapid behavioral shifts in response to environmental cues have raised questions about whether traditional mechanisms, such as synaptic plasticity, can fully account for such flexibility.
A broad range of neural and behavioral data indicates the existence of distinct systems for behavioral choice [48, 130]. The conventional view posits that subcortical areas support habitual, stimulus-based responses, enabling reflex-based and rapid reactions. In contrast, higher cortical areas, such as the prefrontal cortex, are associated with reflective, goal-directed behavior, which relies on cognitive action planning and deliberate evaluation. Critically, goal-directed behavior allows for the flexible construction of value at each decision point through online planning procedures. This mechanism makes agents immediately sensitive to changes in outcome values, enabling them to dynamically adapt their actions to novel or changing circumstances [142]. Studies in network neuroscience have identified dynamic network reconfiguration, characterized by the flexible recruitment and integration of neural circuits, as a fundamental mechanism supporting this adaptability across cognitive[25] and motor domains[187].
Computational models within the predictive coding framework have recently provided insights into the neural mechanisms underlying dynamic adaptation [134]. These models suggest that adaptation can be understood as a form of Bayesian inference, where the brain continuously infers latent contextual states to guide behavior [86]. A prominent example is the COIN model[84, 85], which posits that adaptation arises through two complementary processes: proper learning, involving the creation and updating of memories, and apparent learning, which dynamically adjusts the expression of pre-existing memories based on inferred context. While these frameworks capture key phenomena of adaptive behavior, such as context-dependent single-trial learning and spontaneous recovery, a significant gap remains between these theoretical models and the neural substrates that implement these computations in the brain.

In-context learning as an emerging property in AI In machine learning, the capacity of a model to adapt its behavior, without weight updates, is referred to as in-context learning (ICL). Initially, ICL was observed in architectures tailored for few-shot learning [199] or even zero-shot learning [166]. However, the game changed when it was observed that ICL emerges naturally in large-scale transformers [29, 183]. These architectures leverage self-attention, a mechanism that dynamically changes the way past inputs are attended, in order to shape the current response. Their exceptional capability to adapt to contextual information allowed transformer-based architectures to achieve state-of-the-art performances in many domains, such as natural language processing and image analysis[3, 158, 196]. While there are several intriguing explanations [144, 159] for the emergence of ICL in transformers, a more formal understanding was pioneered by Von Oswald and collaborators [200], who demonstrated that a linear self-attention layer can implement gradient descent steps within its forward pass. This intuition was empirically validated by analyzing trained transformers and was formally proven [213].
Despite their success, transformer architectures face notable shortcomings, particularly in their memory requirements, which scale quadratically with sequence length. These requirements make those architectures largely biologically implausible. Moreover, memory demands constrain their scalability and restrict their applicability to tasks involving long sequences, such as full-length document analysis or video understanding. To overcome these challenges, several efficient deep linear RNN architectures have been proposed [78, 147], demonstrating substantial performance gains over transformers, especially in long-sequence tasks.
However, these deep learning architectures rely on several mechanisms that do not have an obvious correspondence in biological neural systems (such as the attention mechanism itself), leaving an unresolved gap in relating them to the in-context learning abilities exhibited by biological networks. This work bridges this gap by introducing a constructive method to induce in-context learning in biologically plausible neural networks, demonstrating its applicability in a wide range of scenarios.

Gain-Modulation for dynamic behavioral adaptation There is strong experimental and theoretical support for gain modulation as a mechanism for rapid adaptation, independent of synaptic rewiring. Dendritic amplification, for instance, shows that top-down inputs to apical dendrites of layer 5 pyramidal neurons can instantly increase neuronal gain and alter firing patterns, enabling rapid association of inputs without synaptic changes [79, 105]. Computational models further demonstrate that gain changes alone can rapidly shift network dynamics and behavior without requiring synaptic modifications [103], a possibility supported by experimental findings involving optogenetic disinhibition and dendritic amplification [39, 61]. In the realm of sensorimotor adaptation, human motor learning experiments reveal that feedback gains are rapidly tuned to new environments or tasks within a single session, allowing for flexible control without relying on synaptic plasticity [84]. More recently, in [203] authors provided direct evidence that phasic neuromodulatory bursts dynamically modulate neural gain to mediate rapid perceptual switches; through pupillometry, fMRI, and recurrent neural network modeling, they show that heightened gain transiently destabilizes neural dynamics during periods of uncertainty, thereby accelerating perceptual updating. In [188], the authors also discuss how the cerebellum modulates motor cortex activity in response to error signals. Taken together, these findings offer strong biological grounding for gain

modulation as a fast and flexible learning mechanism, supporting the framework presented in this manuscript.

Biological support for in-context learning Building on this premise, our investigation explores how in-context learning, as demonstrated in deep learning architectures, might also be realized in biological recurrent neural networks thanks to gain modulation. Are there unique features in transformers that are also present in biological networks?
We propose that in-context learning can be facilitated by input segregation and dendritic amplification, two features widely observed in biological networks. We further argue that these mechanisms provide a biologically plausible foundation for implementing a process analogous to the attention mechanism in transformers. Recent findings on dendritic computational properties [155] and on the complexity of pyramidal neurons dynamics [105] motivated the study of multi-compartment neuron models in the development of new biologically plausible learning rules [80, 151, 168, 194]. It has been proposed that segregation of dendritic input [80] (i. e., neurons receive sensory information and higher-order feedback in segregated compartments) and generation of high-frequency bursts of spikes [151] would support backpropagation in biological neurons.
Recently, it has been suggested [37] that this neuronal architecture naturally allows for orchestrating "hierarchical imitation learning", enabling the decomposition of challenging long-horizon decision-making tasks into simpler subtasks. They show a possible implementation of this in a two-level network, where the high-network produces the contextual signal for the low-network. Here, we propose an architecture composed of gain-modulated recurrent networks that demonstrate remarkable in-context learning capabilities, which we refer to as 'dynamical adaptation'. Specifically, we illustrate that our biologically plausible architecture can dynamically adapt its behavior in response to feedback from the environment without altering its synaptic weights. We present results for supervised learning of temporal trajectories and reinforcement learning, involving non trivial input-output temporal relations. This novel architecture aims to bridge the gap between biological-inspired in-context learning and the capabilities of artificial neural networks, offering a promising avenue for advancing our understanding of adaptive behavior in both natural and artificial intelligence domains.

7.2 From Latent Readouts to Gain Modulation

The latent variable readout model detailed in Section 6.4 provides a natural starting point for understanding how context-dependent computation can be implemented without synaptic plasticity. Recall that the effective readout for task α was given by a bilinear form:

$$\Theta_\alpha = \sum_{k=1}^{K} g_\alpha^{(k)} \Theta^{(k)} \tag{7.1}$$

where the task-specific latent code $\mathbf{g}_\alpha$ linearly combines a set of fixed basis matrices, adapting the readout to the current task α. When interpreted dynamically through the lens of predictive coding, the dynamics of the gain variables correspond to inference of the latent causes, minimizing the free energy functional $\mathcal{E}$.

7.2.1 Framework Formalization

Consider a general learning scenario where an agent engages in an environment $\boldsymbol{\alpha}$ from a family of tasks $\mathcal{A}$. The agent receives temporal input $\mathbf{x}(t)$, computes a response $y(t)$, and receives feedback $f(t)$ from the environment. A low-level network (the reservoir) extracts features $\mathbf{h}(t)$ relevant to the task family. The agent adapts its response to the specific environment by learning readout parameters $\mathbf{w}$, which determine how features are recombined to produce the correct response:

$$y(t) = Y(\mathbf{h}(t), \mathbf{w}) \tag{7.2}$$

where Y is a generic function that depends on the parameters $\mathbf{w}$. These parameters are adjusted at each time step based on an internally computed error signal e. The overall learning process can be written as the following dynamical system:

$$\begin{cases} \boldsymbol{\tau}_e \, \dot{\mathbf{e}} & = E(\mathbf{e}, \mathbf{y}, \mathbf{f}), \\ \boldsymbol{\tau}_y \, \dot{\mathbf{y}} & = Y(\mathbf{h}, \mathbf{w}), \\ \boldsymbol{\tau}_w \, \dot{\mathbf{w}} & = W(\mathbf{w}, \mathbf{h}, \mathbf{e}). \end{cases} \tag{7.3}$$

The specific instantiation of the functions W and E depends on the learning algorithm employed and can encompass both supervised and reinforcement learning settings (see Methods for precise definitions).

Classically, the parameters $\mathbf{w}$ are interpreted as synaptic weights, and the functions E and W describe synaptic plasticity mechanisms. However, this interpretation imposes stringent locality requirements: all information must be present at the synapse for learning to occur.

Rather than interpreting $\mathbf{w}$ as synaptic weights, we propose viewing them as dynamic activity variables of an auxiliary network that modulates the response Y in real time based on feedback and context. We refer to these dynamically adapting state variables as *virtual weights* because:

- **Functional role**: They parametrize the mapping from features $\mathbf{h}(t)$ to outputs, playing the same computational role as physical synaptic weights.
- **Neural substrate**: They are encoded in the activity of neurons in a separate network (W-network) rather than in synaptic weights, with their activity provided as input currents to the output network (Y-network).

In this sense, the virtual weights $\mathbf{w}$ play a role analogous to the latent task codes $\mathbf{g}$ from Section 6.4: both adapt the effective response of the network to the current context, enabling task-specific computation while maintaining fixed underlying structure.

This dynamic adaptation paradigm offers key advantages over synaptic plasticity:

- **Fast timescales**: It naturally operates on faster timescales, enabling rapid behavioral changes compatible with observations of learning occurring within a single trial.
- **Non-locality**: Dynamic adaptation is not constrained by the locality requirements of synaptic plasticity, allowing computation at the circuit level rather than the synapse level.

In this framework, the functions E and W in Eq. (7.3) arise from the dynamics of the auxiliary network of virtual weights. We show that such networks can dynamically implement standard gradient-based algorithms widely used in machine learning. While these algorithms are traditionally associated with synaptic plasticity, their reliance on non-local terms suggests that in the brain they

may be more naturally realized at the level of neural circuits.

7.2.2 The Role of Gain Modulation in Adaptive Behavior

Recent research has highlighted a growing interest in understanding modulatory mechanisms and their role in neural computation. It has been observed that gain modulation allows neurons to adjust their response to inputs based on context [105]. This dynamic regulation of neuronal sensitivity plays a critical role in adaptive behavior and cognitive flexibility. Moreover, it has been implicated in various functions, including feedback control [62, 132] and attention, where it allows for the selective amplification of relevant information. A recent study [96] identified gain modulation as a key neural mechanism underlying hierarchical spatiotemporal prediction and sequence learning, demonstrating that such models can replicate various cortical processing features, including temporal abstraction and space-time receptive fields in the primary visual cortex. Furthermore, gain modulation is closely tied to meta-learning and multitask learning [153, 209], enabling neural networks to efficiently generalize across diverse tasks by leveraging contextual cues.

Within biological recurrent networks, dendritic segregation provides a structural basis for gain modulation. By compartmentalizing synaptic inputs, neurons can effectively implement nonlinear interactions between different input streams. This is reminiscent of in-context learning (ICL) in machine learning, where multiplicative interactions are crucial for enabling models to leverage previously learned information in new contexts. Although the focus here is on gain modulation mediated by apical dendrites, other mechanisms are also known. For instance, neuro-glial interactions at the synapse have been proposed as a substrate for working memory, with astrocytes playing a crucial role in modulating neural activity and maintaining memory through glia–synapse interactions [50]. This gain modulation mechanism could complement other adaptive processes, such as subtle, correlated changes in connectivity [62].

The present work explores the role of gain modulation in recurrent networks, employing a simplified model where apical currents act multiplicatively on basal currents within a transfer function. A key parameter, γ, controls the strength of these nonlinear interactions (see Methods for further details). By modulating the gain of neuronal responses, these networks exhibit enhanced adaptability and in-context learning capabilities. The bilinear structure, where one signal multiplicatively gates another, emerges naturally from this dendritic geometry. In the framework presented, $\mathbf{x}^{\mathrm{ap}}(t)$ corresponds to contextual or error signals arriving at apical dendrites, while $\mathbf{x}(t)$ represents basal sensory input, creating a circuit-level implementation of nonlinear learning rules.

7.2.3 Gain-Modulated Reservoir Computing

The bilinear structure of gain-modulated networks, where the readout is linear in both the reservoir state $\mathbf{h}$ and the context variable $\mathbf{g}$, naturally implements task adaptation provided a rule exists to update $\mathbf{g}$ dynamically. Building upon this insight, a gain-modulated reservoir network (GM-RC) is introduced. This architecture draws inspiration from the morphology and function of pyramidal neurons in the cortex, which nonlinearly integrate inputs from basal and apical dendrites [105, 179]. As shown below, such architectures possess the natural structure to implement gradient-based learning rules.

Here, we consider an RNN with an additional input source $\mathbf{x}^{\mathrm{ap}}$ randomly projected into the apical

dendrite of each neuron by the matrix R^{ap}. Consistent with experimental observations in L5 pyramidal neurons, we allow the apical inputs to modulate the gain of the activation function, thereby altering its slope. Consequently, the resulting RNN equation is formulated as:

$$\tau_h \dot{\mathbf{h}} = \phi\left((\mathbf{b}^{\text{ap}} + \gamma \cdot R^{\text{ap}}\mathbf{x}^{\text{ap}}) \odot (J\mathbf{h} + \beta \cdot R^{\text{ap}}\mathbf{x}^{\text{ap}} + R\mathbf{x})\right) - \mathbf{h} \tag{7.4}$$

Where $\mathbf{b}^{\text{ap}}$ is a constant bias vector. The hyperparameters β, γ modulate the effect of the gain modulation of the apical inputs. Specifically, when $\gamma = 0$, we obtain a network in which $\mathbf{x}^{\text{ap}}$ does not affect the gain modulation.
Similarly, as before the gain-modulated RNN (GMRNN) extracts features

$$\mathbf{h}(t) = \text{GMRNN}^{\gamma}(\{\mathbf{x}|\mathbf{x}^{\text{ap}}\}_t) \tag{7.5}$$

such that the expression $\mathbf{y}(t) = \text{GMRNN}^{\gamma}_{\Theta}(\{\mathbf{x}|\mathbf{x}^{\text{ap}}\}_t)$ will denote the input $\rightarrow$ output mapping implemented by a gain-modulated reservoir with readout Θ.
Removing time dependencies [1], we will have an instantaneous function

$$\mathbf{y}(t) = \text{GMNN}^{\gamma}_{\Theta}(\mathbf{x}|\mathbf{x}^{\text{ap}}) = \Theta\phi\left((\mathbf{b}^{\text{ap}} + \gamma \cdot R^{\text{ap}}\mathbf{x}^{\text{ap}}(t)) \odot (R\mathbf{x})(t)\right) \tag{7.6}$$

of the inputs $\mathbf{x}^{\text{ap}}$ and $\mathbf{x}$. We explicitly maintain the dependence on γ because, in our experiments, we use this parameter to regulate the gain modulation effect of the apical inputs on the network. The choice for the name $\mathbf{x}^{\text{ap}}$ is inspired by the current received in the apical dendrites of L5 pyramidal neurons, that are believed to carry contextual/high-level information [105].
Experimentally, networks with strong gain modulation ($\gamma > 0$) generalize to out-of-distribution tasks, while networks without gain modulation ($\gamma = 0$) fail to adapt, demonstrating the criticality of this parameter.

7.2.4 The Bilinear Readout Layer and Function Parametrization

To better understand the expressive power of gain modulated networks we study the approximation properties of feedforward networks with linear transfer functions ($\phi = Id$). In this case, we have a bilinear layer, such that it naturally parameterizes a family of linear functions, indexed by the virtual weights. Consider the output given by:

$$Y(\{\mathbf{x}|\mathbf{w}\}) = \Theta\left((R^{\text{ap}}\mathbf{x}^{\text{ap}}) \odot (R\mathbf{x})\right) \tag{7.7}$$

where $\Theta \in \mathbb{R}^{1\times N_h}$ is the readout matrix, and the random matrices R^{ap}, R encode the apical and basal contributions.
The following proposition establishes that such bilinear networks can express arbitrary linear functions of the virtual weights:

Proposition 7.2.1. *Let $Y(\{\mathbf{x}|\mathbf{w}\}) = \Theta\left((R^{ap}\mathbf{x}^{ap}) \odot (R\mathbf{x})\right)$ where Θ, R^{ap}, and R are random matrices of dimensions $1 \times N_h$, $N_h \times N_{in}$, and $N_h \times N_{in}$, respectively, with i.i.d. entries drawn from any distribution that is absolutely continuous with respect to the Lebesgue measure.*[2] *Then almost*

[1] Consiering the limit $J \to 0$, $\tau \to 0$
[2] Here for simplicity we assumed $\gamma = 1$, $\beta = 0$, $\mathbf{b}^{\text{ap}} = 0$.

surely, if $N_h \geq N_{in}$ for every linear function, $f : \mathbb{R}^{N_{in}} \to \mathbb{R}$ there exists a set of virtual parameters $\mathbf{w}_f \in \mathbb{R}^{N_{in}}$ such that $Y(\{\cdot|\mathbf{w}_f\}) = f(\cdot)$.

Proof. We first observe that the output can be written as a bilinear form:

$$Y(\{\mathbf{x}|\mathbf{w}\}) = \mathbf{w}^\top A\mathbf{x} \tag{7.8}$$

where the matrix $A \in \mathbb{R}^{N_{in}\times N_{in}}$ has entries:

$$A_{ij} = \sum_{k=1}^{N_h} \Theta_{1k} R^{\text{ap}}_{ki} R_{kj} \tag{7.9}$$

That is, A is a sum of N_h rank-one matrices of the form $\Theta_{1k}\mathbf{R}^{\text{ap}}_k \mathbf{R}_k^\top$, where $\mathbf{R}^{\text{ap}}_k$ and $\mathbf{R}_k$ are the k-th rows of R^{ap} and R, respectively.
Since the entries of Θ, R^{ap}, and R are i.i.d. and drawn from a distribution absolutely continuous with respect to the Lebesgue measure, the collection of vectors $\{\mathbf{R}^{\text{ap}}_k\}_{k=1}^{N_h}$ and $\{\mathbf{R}_k\}_{k=1}^{N_h}$ almost surely span a subspace of dimension $\min(N_h, N_{in})$. Consequently, the matrix A has rank $\text{rank}(A) = \min(N_h, N_{in})$ almost surely. Thus, if $N_h \geq N_{in}$, then A is almost surely invertible. Now, given any linear function $f(\mathbf{x}) = \mathbf{w}_*^\top \mathbf{x}$ with $\mathbf{w}_* \in \mathbb{R}^{N_{in}}$, define $\mathbf{w}_f := (A^\top)^{-1}\mathbf{w}_*$. Then, for all $\mathbf{x}$, $Y(\{\mathbf{x}|\mathbf{w}_f\}) = \mathbf{w}_*^\top \mathbf{x} = f(\mathbf{x})$ as required. □

This proposition demonstrates that the bilinear architecture is sufficiently expressive to implement any linear readout, with the effective weight vector encoded in the apical input.

7.2.5 Expressing Gradient Descent Dynamics

The central claim is that gradient-based learning can be implemented dynamically via the activity of a second gain-modulated network. Consider a gain-modulated bilinear network that dynamically approximates a continuous in-context target signal $\hat{y}(t)$ by adjusting the activity of its virtual weights to minimize the mean squared error $(Y(\{\mathbf{x}|\mathbf{w}\}) - \hat{y}(t))^2$.
The following proposition shows that a separate gain-modulated bilinear network can express the corresponding online gradient update using a number of units that scales linearly with the number of features.

Proposition 7.2.2. *Let $Y(\{\mathbf{z}|\mathbf{w}\})$ defined as in Proposition 7.2.1, and $W_{\Theta^w}(\{\mathbf{z}|e\}) = \Theta^w\left((R^{ap}e) \odot (R\mathbf{z})\right)$ another gain-modulated bilinear network with R^{ap}, R are random matrices of dimensions $1 \times N_h$, and $N_h \times N_{in}$, respectively, with i.i.d. entries drawn from any distribution that is absolutely continuous with respect to the Lebesgue measure. Then there exists a readout matrix Θ^w such that $W_{\Theta^w}(\{\mathbf{z}|e\}) = e\mathbf{z} = e\nabla_{\mathbf{w}} Y(\{\mathbf{z}|\mathbf{w}\})$.*

Proof. We can express that the output can be written as $W_{\Theta^w}(\{\mathbf{z}|e\}) = e\Theta^w C\mathbf{z}$ where $C_{ij} = R^{\text{ap}}_{1j} R_{ij}$. Since the entries of R^{ap}, and R are i.i.d. and drawn from a distribution absolutely continuous with respect to the Lebesgue measure, when $N_h \geq N_{in}$ the matrix C is almost surely injective. Then there exists a left inverse Θ_* such that $\Theta_* C = Id$. □

Combining these results, we obtain a two-network architecture where the synaptic weights Θ^w and Θ remain fixed during deployment; only the activity (the virtual weights and apical modulation signals) change in real time.

7.2.6 Two-Phase Implementation: Development and Dynamical Learning

The construction of this architecture occurs in two distinct phases:

Phase 1: Development of Network Structure. Using an algorithmic distillation protocol, we collect learning trajectories $\{(\mathbf{h}_\alpha(t), \mathbf{w}_\alpha(t), e_\alpha(t))_{t=0}^T \mid \alpha \in \mathcal{A}_{ID}\}$ by running the learning dynamics in Eq. (7.3) on a subset of tasks $\mathcal{A}_{ID}$ (in-distribution). We then train the static readout weights Θ^y and Θ^w of two random networks to replicate the functions Y and W, respectively:

$$L_Y(\Theta^y) = \sum_{\alpha,t} \left| Y_{\Theta^y}(\mathbf{h}_\alpha(t), \mathbf{w}_\alpha(t)) - Y(\mathbf{h}_\alpha(t), \mathbf{w}_\alpha(t)) \right|^2 \tag{7.10}$$

$$L_W(\Theta^w) = \sum_{\alpha,t} \left| W_{\Theta^w}(\mathbf{h}_\alpha(t), e_\alpha(t)) - W(\mathbf{h}_\alpha(t), e_\alpha(t)) \right|^2 \tag{7.11}$$

Phase 2: Dynamical Learning on Novel Tasks. During deployment on an out-of-distribution task $\alpha \in \mathcal{A}_{OOD} = \mathcal{A} \setminus \mathcal{A}_{ID}$, all synaptic weights remain fixed. Instead, virtual weights evolve dynamically:

$$\dot{\mathbf{w}}(t) = W_{\Theta^w}(\mathbf{h}(t), e(t)) \tag{7.12}$$

and the output is computed via:

$$y(t) = Y_{\Theta^y}(\mathbf{h}(t), \mathbf{w}(t)) \tag{7.13}$$

Critically, $\mathbf{w}(t)$ is no longer a synaptic weight but a contextual signal encoded in neural activity. This architecture shifts the locus of learning from synapses to circuits, enabling fast, context-dependent computation compatible with observations of single-trial learning in biology.

7.3 Results: Dynamical Learning Across Tasks

7.3.1 Dynamical Adaptation for Temporal Trajectories

We consider the task of autonomously predicting a temporal trajectory $\{f(t) = y^{\text{targ}}(t)\}_t$. Following the reservoir computing paradigm, we employ a random RNN whose dynamics follow Equation (6.7) to extract temporal features $\mathbf{h}(t)$. During training, the reservoir receives the target dynamics $y^{\text{targ}}(t)$ as input and is tasked with predicting the subsequent step of the trajectory via a linear readout of its activity. This setup can be reformulated as a dynamical supervised learning problem. The readout weights can then be dynamically adjusted, minimizing the error between the target and the current prediction $y(t)$.
This results in the following dynamics:

$$\begin{cases} e & = (y^{\text{targ}} - y) \\ y & = Y_{\Theta^y}(\{\mathbf{h}, \mathbf{w}\}) \simeq \mathbf{h} \cdot \mathbf{w} \\ \tau_w \, \dot{\mathbf{w}} & = W_{\Theta^w}(\mathbf{h}, e) \simeq \mathbf{h}\, e \end{cases} \tag{7.14}$$

Here, the network $W_{\Theta^w}(\{\mathbf{h}, e\})$ is dedicated to estimating the gradient of virtual weights, while

$Y_{\Theta^y}(\{\mathbf{h}, \mathbf{w}\})$ is tasked with estimating the predicted $y(t)$ as a function of the new virtual weights. These gain-modulated architectures are pre-trained to replicate, respectively, gradient descent updates and scalar products obtained on a set of N_{train} training sequences. More specifically, the target sequences are 5 sinusoidal functions with different frequencies.
In the closed-loop phase, the features $\mathbf{h}$ are obtained by directly feeding the network estimation $y(t)$ as the input to the RNN. This results in an autonomous dynamical system that reproduces the target trajectory. Errors, represented by the blue trajectory, are fed back to the initial network to assess necessary updates to the virtual weights. These updates are then transmitted to the second network, modifying the decoding of reservoir dynamics. Initially, the reservoir receives the target trajectory as input in an open-loop fashion, which is later replaced by the estimated trajectory itself in a closed-loop configuration.
The networks were pre-trained on five target frequencies and then tested across a range of frequencies. Evaluation was performed by assessing the mean squared error (MSE) between the target and estimated trajectories in closed-loop scenarios. The network generalizes smoothly across tested frequencies. Individual learning curves show that the distilled networks approximate the original gradient descent dynamics closely, tracking the same MSE evolution as the source algorithm.

7.3.2 Reinforcement Learning and the Role of Gain Modulation: Multi-Armed Bandits

Dynamical adaptation is now investigated within the framework of reinforcement learning. We provide robust evidence supporting the hypothesis that gain modulation represents a fundamental component in implementing in-context learning in biological agents.
We first explore stateless environments represented by Bernoulli K-armed bandits [20]. Within each environment $\boldsymbol{\alpha} \in \mathcal{A}^{\text{bandits}}$, there exists a subset $P_{\boldsymbol{\alpha}} \subset [K]$ of arms that yield a reward with high probability ($p = 0.95$). During the training phase, the in-distribution environments give a high reward probability to even-numbered arms, whereas during testing, the out-of-distribution environments assign a high reward probability to odd-numbered arms.
In this simplified bandit scenario where state information is absent, the virtual weights $\mathbf{w}$ directly parameterize the policy probabilities: $\boldsymbol{\pi} = \text{softmax}(\mathbf{w})$. Consequently, our focus lies primarily on analyzing the behavior of the network $W^{\gamma}_{\Theta^w}(\cdot)$ acting on the virtual parameters. This serves as an ideal test bed for evaluating the network's ability to learn the policy gradient update rule and generalize it to out-of-distribution scenarios.
The network $W^{\gamma}_{\Theta^w}(\boldsymbol{\pi}, a, r)$ has a gain-modulated architecture and is a function of the reward r and the action a. Here γ indicates the dependence on the gain-modulation strength. We train the network to approximate the policy gradient update rule using policy gradient estimates from a single learning trajectory (1000 rounds, learning rate $lr = 0.1$) in an in-distribution environment. We then test the distilled policy gradient networks with a higher learning rate ($lr = 1.0$) in out-of-distribution (OOD) environments.
To systematically investigate the impact of gain modulation ($\gamma = 1$) on out-of-distribution performance, we compare it with networks where the reward does not modulate the network gain ($\gamma = 0$). For each case, we select the optimal hyperparameters through a grid-based search.

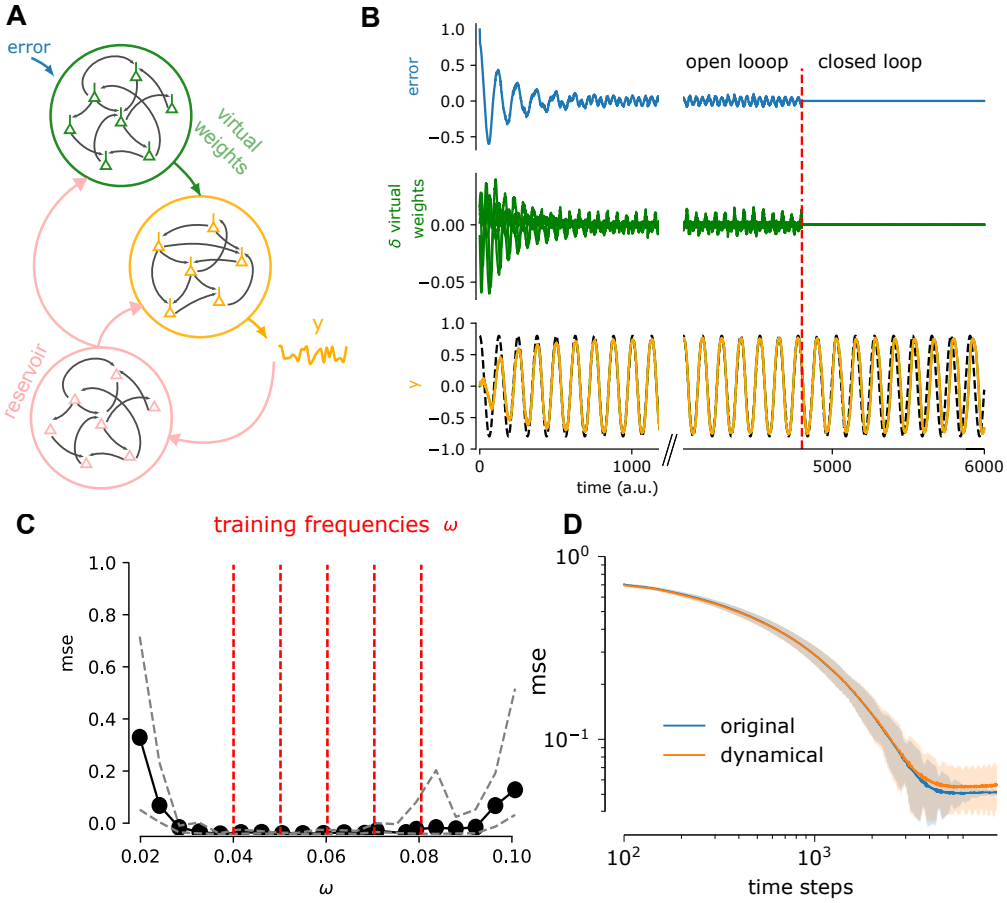

Figure 7.1: Dynamical adaptation of temporal trajectories: A. Overview of the network architecture employed. A recurrent network composed of $N = 20$ units (described by the vector $\mathbf{h}(t)$, depicted in pink, which follows (6.7)) is utilized for learning a periodic trajectory, following the prescription of reservoir computing. One recurrent network is tasked with estimating the gradient of virtual weights, while another is dedicated to estimating the behavioral reconfiguration (illustrated in green and orange, respectively) resulting from these updated virtual weights. **B.** Illustration of our architecture dynamically adjusting (without synaptic alterations) to adhere to the desired dynamics. Errors (represented by the blue trajectory) are fed back to the initial network to assess necessary updates to the virtual weights δw, a N-dimensional vector (shown in the green trajectory). Subsequently, these updates are transmitted to the second network, modifying the decoding of reservoir dynamics (indicated by the orange lines). Initially, the reservoir receives the target trajectory as input (in open loop, before the red vertical line), which is later replaced by the estimated trajectory itself (in closed loop, after the red vertical line). **C.** Our networks were pre-trained on five target frequencies (marked by red vertical lines) and tested across a range of frequencies, evaluating the mean squared error (MSE) between the target and estimated trajectories in closed-loop scenarios (solid: median, dashed: 20th-80th percentile (statistics evaluated over 10 realizations). **D.** Comparison of the learning curve, mse in function of the weights update, for the original learning algorithm, and the one evaluated by the two networks, in blue and orange respectively.

The policy gradient update rule for the bandit task is given by:

$$\begin{cases} \mathbf{e} & = r\,(\mathbf{1}_a - \boldsymbol{\pi}) \\ \mathbf{y} & = \mathbf{w} \cdot \mathbf{h} \\ \tau_w\,\dot{\mathbf{w}} & = \mathbf{e} \odot \mathbf{h} \end{cases} \tag{7.15}$$

where $\mathbf{1}_{a(t)}$ represents the one-hot encoded action at time t. It is a D-element vector where the $a(t)$-th element is one, and all other elements are zero.

As before, we model that the policy is modulated by an auxiliary population of virtual weights $\mathbf{w}$, determined by the activity of an additional GMRNN W_{Θ^w}. This auxiliary network adjusts its internal activity based on the rewards received, effectively implementing policy gradient updates of the virtual weights and thereby modulating the agent behavior in real-time. Then, the above dynamics can be approximated by two networks, one for estimating the gradients, and one for the scalar product:

$$\begin{cases} \tau_w \, \dot{\mathbf{w}} & = W_{\Theta^w}(\{\boldsymbol{\pi}|a, r\}) \simeq r \, (\mathbf{1}_a - \boldsymbol{\pi}) \\ \boldsymbol{\pi} & = \text{softmax}(Y_{\Theta^y}(\{\mathbf{h}|\mathbf{w}\})) \end{cases} \tag{7.16}$$

Our experimental findings show that models with gain modulation require significantly fewer neurons to achieve maximum scores in OOD environments. A gain-modulated network achieves a regret distribution comparable to the policy gradient in both ID and OOD environments, with a lower median. In contrast, networks without gain modulation show significantly higher regret.

Analyzing the regret curves in the OOD setting, we observe that a model with gain modulation often learns a more data-efficient algorithm than its source, even in OOD environments. Conversely, a model without gain modulation fails to generalize to OOD environments and does not converge to zero regret.

In summary, gain modulation enables the network to consistently and efficiently distill the correct gradient update rule and generalize it to unseen environments, predicting the correct virtual weight update in regions of the input space far from the training data.

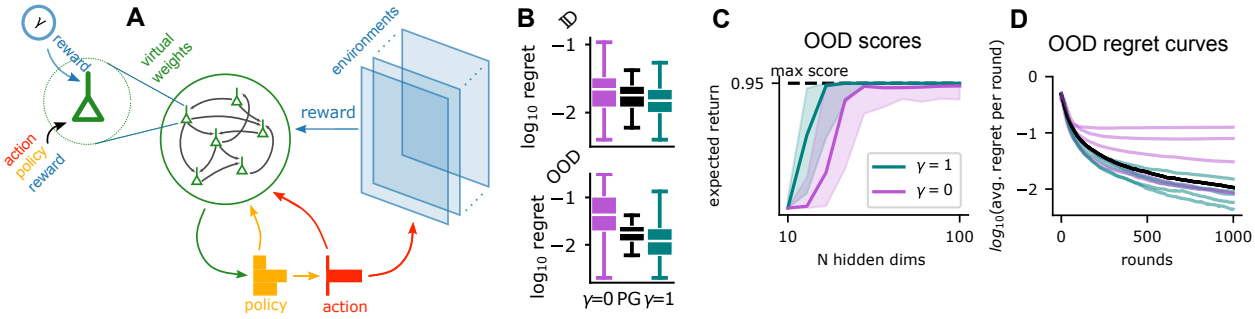

Figure 7.2: Dynamical reinforcement learning: Multi Armed Bandits. **A.** Schematic of network architecture and task. Virtual weights parameterize the agent's policy (orange). At each round, an action (red) is sampled from the policy and played. A reward is then sampled from the current environment. A GM-network (green) then predicts the virtual parameter update. **B.** Regret comparison between the distilled and the original policy gradient algorithm. We report log regret per round distribution achieved at the 100-th round for 100 independently trained models in ID and OOD settings. We compare a gain-modulated network ($\gamma = 1$), and a network without gain modulation ($\gamma = 0$) with the policy gradient training source (PG). **C.** OOD model performance varying number of hidden dimensions. Solid lines indicate the median score (expected reward), computed over 100 independently trained models. The filled area indicates 20-80% confidence interval. **D.** We report average regret per round curves in OOD setting. We compare a gain-modulated network ($\gamma = 1$, teal), and a network without gain modulation ($\gamma = 0$, orchid) with the policy gradient training source (in black). Each curve represents the average regret per round for one fixed trained model, computed by averaging over 100 independent simulations.

7.3.3 Sensorimotor Control: Dynamical Reinforcement Learning in a Reaching Task

We examine a reaching task, known as the dark room task (a simple instance of the water maze task [135]), set in a 2D maze within the domain $(-1,1)\times(-1,1)$ with a grid size of 0.1. Within this grid-like environment, the agent navigates by selecting one of four actions: up, down, left, or right, thereby determining its subsequent position. The primary goal is to locate a concealed object within the maze, with the agent having sole awareness of its own position. Feedback is provided via rewards, where the agent receives a reward of 10 if the distance to the object is 0.1, 15 if the distance is 0, and 0 otherwise. Through iterative exploration, the agent develops a strategy to efficiently traverse the maze and pinpoint the object despite the limited visibility.
The position of the agent is encoded separately using 25 input units $\mathbf{x}$ each, employing Gaussian activation functions distributed on a 5x5 grid in the maze and with a width of 0.2.
To accomplish this task we consider a network architecture composed of two networks. One GM-network, $W^{\gamma}_{\Theta^w}(\{\mathbf{h}, r, a, \boldsymbol{\pi}\})$ is responsible for estimating the necessary update of virtual weights through policy gradient, while another GM-network $Y^{\Gamma}_{\Theta^y}(\{\mathbf{h}, \mathbf{w}\})$ estimates policy reconfiguration resulting from changes in virtual weights. Here, the parameters γ, Γ define the strength of the gain modulation. Their dynamics can be described by the following equation:

$$\begin{cases} \tau_w \, \dot{\mathbf{w}} & = W_{\Theta^w}(\{\mathbf{h}|r, a, \boldsymbol{\pi}\}) \simeq W(\mathbf{h}, r, a, \boldsymbol{\pi}) = r\,(\mathbf{1}_a - \boldsymbol{\pi}(\mathbf{y})) \odot \mathbf{h} \\ \mathbf{y} & = Y_{\Theta^y}(\{\mathbf{h}|\mathbf{w}\}) \simeq Y(\mathbf{h}, \mathbf{w}) = \mathbf{h} \cdot \mathbf{w} \end{cases} \tag{7.17}$$

Indeed, we compare the case with (gain modulation, $\gamma = 1$) and without (no gain modulation, $\gamma = 0$) gain modulation.
Firstly, we assessed the performance of the policy gradient algorithm using a set of 8 food locations, where the total reward averaged over 200 trials was observed against the number of trials. After multiple trials, the agent successfully achieved precise targeting of the circles. Data collected from these experiments were utilized to train our networks to estimate gradients and scalar products, as defined in Eq. (7.17).
To validate that our trained model is capable of dynamically implementing policy gradient itself, we tested it on both the training set locations (ID) and new test positions (OOD). For each food position, we illustrate a sample trajectory executed by the agent to reach the target at the end of the training. The agent's precision closely matches that of the plastic policy gradient learning rule. We present the reward plotted against the number of trials for both training and test food locations.
We compared these performances against an architecture lacking gain modulation ($\gamma = 0$), observing worst performances. Notably, while performances for ID food locations are acceptable, those for OOD cases are extremely poor. This observation, coupled with the results from the preceding section, suggests that gain modulation is a crucial component in facilitating the generalization of adaptive behavioral capabilities in recurrent networks.

Temporal Credit Assignment and Future Reward We demonstrate that our architecture is capable to perform the temporal computation required to learn delayed action-reward temporal relations, requiring evaluating temporal credit assignment. Indeed, in reinforcement Learning (RL)

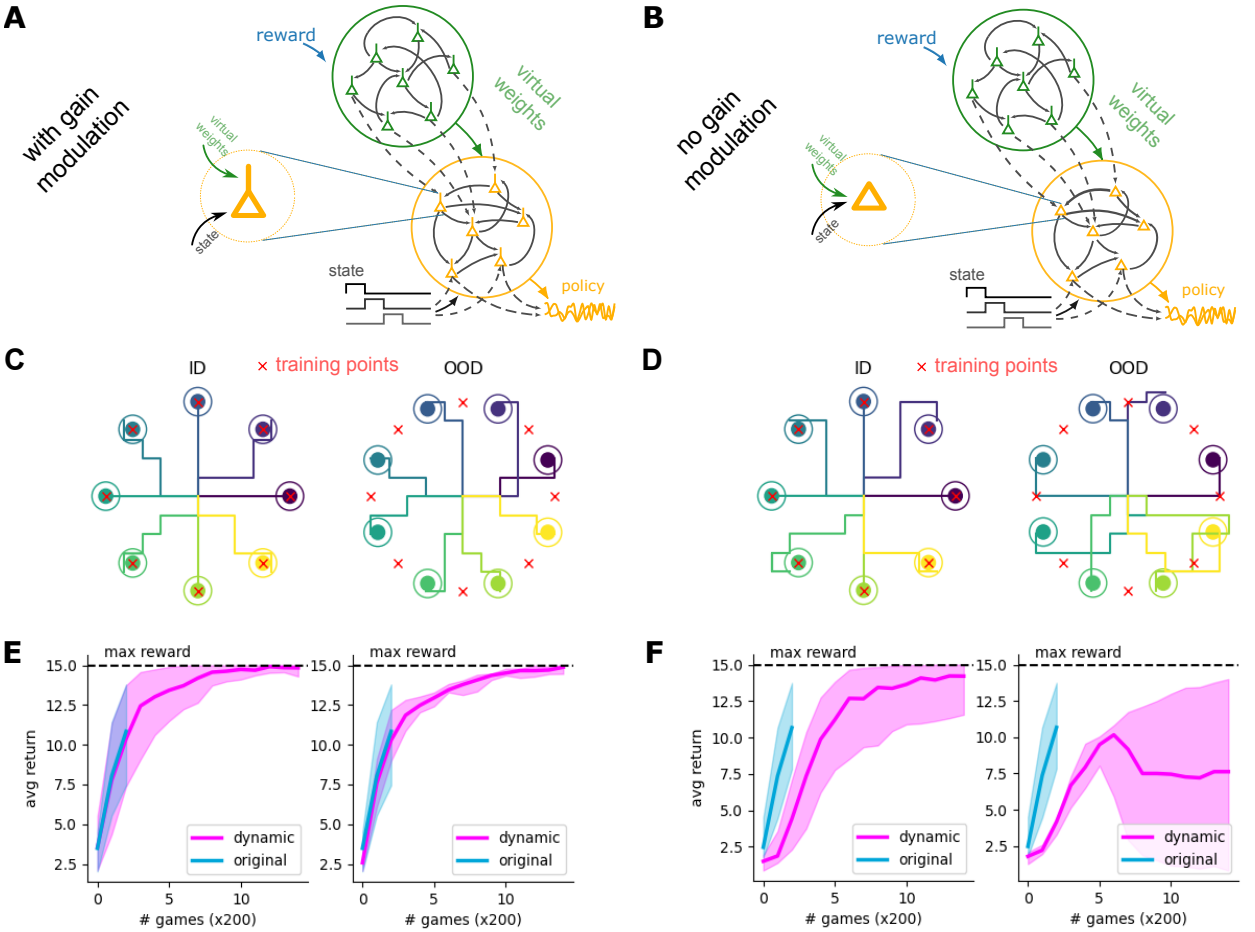

Figure 7.3: Dynamical Reinforcement Learning Dark Room A. Overview of the network architecture employed. One GM-network (depicted in green) is responsible for estimating the necessary update of virtual weights through policy gradient, while another GM-reservoir is focused on estimating policy reconfiguration (shown in orange) resulting from changes in virtual weights. **B.** Illustration of the task: the agent begins from the center and learns to reach not observable objects positioned at various locations (represented by colored circles). After several trials, the agent achieves precise targeting of the circles. **C.** Reward as a function of number of trials (or games), blue: policy gradient, left panel pink: dynamical learning for training positions (ID), right panel pink: dynamical learning for new positions (OOD). Line: median, shading: 20-th/80-th percentile range. **D, E, F.** Similar to **A, B, C**, respectively, but without gain modulation.

an agent learns to maximize cumulative rewards. The discount factor, usually denoted by γ_d (where $0 \leq \gamma_d \leq 1$), is crucial in RL as it determines the importance of future rewards. A higher γ_d values future rewards more significantly, encouraging the agent to consider the long-term consequences of its actions. Conversely, a lower γ_d makes the agent prioritize immediate rewards. The cumulative reward R_t at time step t is given by:

$$\overline{r}_t = \sum_{s \geq t} r_s \gamma_d^{(s-t)} \tag{7.18}$$

To account for this, the term $\mathbf{e}$ in Eq. (7.17) should be replaced by its exponential temporal filtering, with a timescale $\tau_e = -\frac{1}{\log(\gamma_d)}$. In this way actions in the past, can be potentiated when delayed rewards are received.

If the discount factor is zero, the agent's policy would change only when it is very close to the moment the reward is received. In a reaching task, the agent will only change its policy when it is very close to the reward, completely ignoring the need for long-term planning and making it unlikely

to ever reach the goal from distant starting points. As a result, the policy points towards the target position only when nearby the target itself, resulting in failure when the agent randomly moves in the wrong direction at the beginning of the task.
On the other hand, recurrent connections enable a network to maintain a memory of past states and actions, effectively allowing it to use information from previous time steps to inform current decisions, achieving optimal planning and decisions even when far from the target location (and hence, from the reward).

7.3.4 Application to a Robotic Benchmark

We test our framework to a more complex and realistic setting by applying it to a MuJoCo ant environment, demonstrating its scalability to continuous control tasks. The agent, modeled as a quadruped, operates in a high-dimensional control space with 27 environmental state variables and 8 control torques (2 per leg), requiring coordinated control of multiple degrees of freedom.
To manage low-level motor execution, we first trained a network to perform four fundamental motor primitives: moving left, right, forward, and backward. This pretraining step allows higher-level adaptation without direct joint-level optimization.
Building on this, we use our gain-modulated network as a meta-policy that dynamically selects motor primitives based on the agent's position. The meta-policy network receives positional inputs and determines the appropriate primitive, while a secondary network modulates its behavior through virtual weights, mirroring the architecture used in the grid-based task.
The high-level policy is defined as a discrete policy over four options, corresponding to the four sub-policies. It is computed using a softmax over a linear readout of the observation vector:

$$\boldsymbol{\pi}_{\text{high}}(a_{\text{high}} \mid s) = \text{softmax}(Ws) \tag{7.19}$$

where s is the current state of the environment, and $a_{\text{high}} \in \{\text{left}, \text{right}, \text{forward}, \text{backward}\}$. The selected action determines which sub-policy to invoke at each high-level decision point. As in the previous experiment, W and $\boldsymbol{\pi}_{\text{high}}$ are evaluated respectively by the networks $W^{\gamma}_{\Theta^w}$ and $Y^{\Gamma}_{\Theta^y}$.
This architecture enables flexible and efficient adaptation: the agent successfully generalizes to target locations not included in the training set, adjusting its movement strategies dynamically without modifying its underlying synaptic structure. This capability is particularly significant for neuromorphic computing and robotics, where adaptive behavior is often constrained by the need for on-chip plasticity. Our results demonstrate that a fixed neuromorphic chip, programmed with a gain-modulated network, can dynamically adjust to new goals without requiring structural modifications. This approach reduces computational overhead and power consumption while enabling efficient, low-power autonomous systems that adapt in real time. By eliminating the need for weight updates at the hardware level, this architecture provides a scalable and robust foundation for adaptive physical agents, paving the way for next-generation neuromorphic robotics.

7.4 Discussion

The brain's ability to flexibly and rapidly adapt to novel tasks remains only partially understood. Different studies have highlighted that this adaptability relies on mechanisms such as dynamic

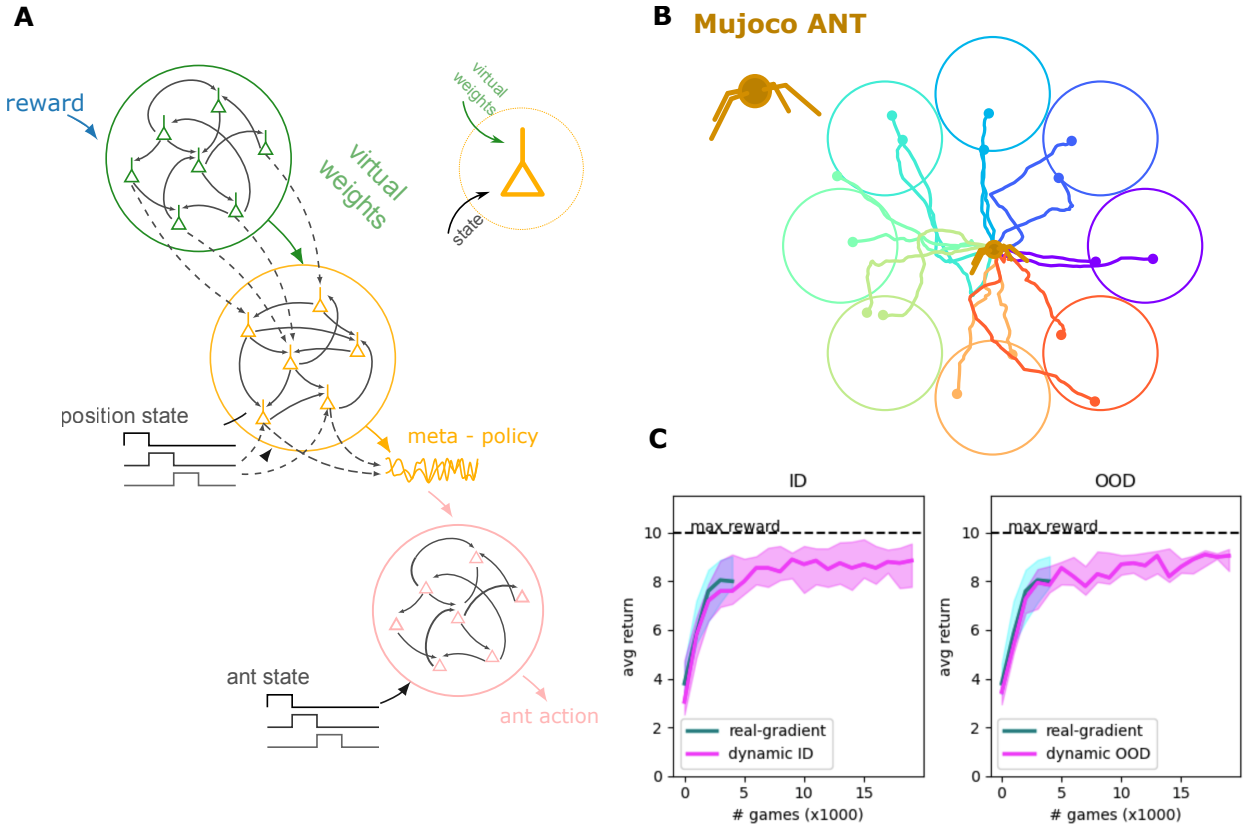

Figure 7.4: Dynamical Reinforcement Learning: MuJoCo Ant Navigation A. Overview of the network architecture employed. One GM-network (depicted in green) is responsible for estimating the necessary update of virtual weights through policy gradient, while another GM-network is focused on estimating policy reconfiguration (shown in orange) resulting from changes in virtual weights. **B.** Illustration of the task and of the MuJoCo environment: the agent begins from the center and learns to reach objects positioned at various locations (represented by colored circles). After several trials, the agent achieves precise targeting of the circles. Our networks are pre-trained on a set of target points (red crosses) and then tested on the same (displayed in the left panel) and new positions (shown in the right panel). **C.** Reward as a function of number of trials (or games), blue: policy gradient, left panel pink: dynamical learning for training positions (ID), right panel pink: dynamical learning for new positions (OOD). Line: median, shading: 20-th/80-th percentile range.

network reconfiguration and context-dependent modulation, allowing agents to respond flexibly to environmental changes. However, a significant gap remains in understanding how these processes map onto specific neural substrates.

While synaptic plasticity has been extensively studied as a mechanism underlying learning and memory [2], emerging evidence suggests that it alone may not account for the full spectrum of adaptive behavior. Abraham et al. critically evaluate the synaptic plasticity and memory (SPM) hypothesis, concluding that non-synaptic mechanisms, such as intrinsic plasticity, structural neural changes, epigenetic regulation, glial contributions, network reorganization, and protein turnover, likely contribute significantly to memory storage and retrieval, thus broadening the framework for understanding dynamic adaptation. This perspective questions long-standing assumptions about neural circuits, highlighting that we may still lack full understanding of many processes we consider well-established.

Inspired by the brain's adaptability, artificial neural network architectures have been challenged to replicate similar capabilities. Fast synaptic plasticity, hypothesized to orchestrate flexible cogni-

tive functions, inspired several artificial architectures, such as fast-weight networks [91, 172] and hyper-networks that predict main network weights [12]. However, these approaches often rely on biologically unrealistic mechanisms, such as requiring neurons to directly act on the synapses of a second network or implementing "inner loops" in network dynamics [12, 40]. More biologically plausible alternatives, like coupled neural-synaptic dynamics [133], come closer to natural systems but still depend on complex credit assignment mechanisms.

Authors in [134, 176] address the problem of dynamic adaptation by proposing a temporal predictive coding model that uses recurrent connections to predict its own future neural responses and updates its parameters online through local plasticity. This allows the network to dynamically infer and track the state of continuously changing stimuli in real time.

Complementary to those approaches, the framework presented here does not presuppose a specific learning paradigm or fixed synaptic plasticity rule. Instead, it demonstrates how a neural system can dynamically reproduce a target learning rule through activity-based mechanisms.

Transformers, known for their remarkable in-context learning (ICL) capabilities, excel at tasks requiring adaptability, such as natural language and image processing [3, 158, 196]. Mechanisms such as attention, associative memory [159], and induction-based copying by attention heads [144] have been proposed to explain ICL in transformers, with theoretical work showing that attention mechanisms can implement gradient updates [200]. However, the emergence of these capabilities remains poorly understood, particularly when comparing them to the in-context learning observed in biological systems.

A constructive method is proposed to induce in-context learning in biologically plausible neural networks in a broad variety of scenarios. An alternative formulation is offered in which dynamic adaptation is modulated by dendritic input segregation and amplification. Virtual (fast) weights are encoded by the activity of a second network that modulates the transfer function of neurons in the base network. As a result, virtual weights are not synaptic weights, but their values and updates are evaluated and encoded by the activity of other neurons, which were pre-trained to perform gradient updates. Another pre-trained network receives those weights signals as an input, along with the current environment features, and computes an adapted in-context response.

This mechanism overcomes the strict locality constraints of synaptic plasticity, which require all elements involved in evaluating weight updates to be present at the synaptic site. Consequently, biologically plausible synaptic updates cannot explicitly rely on information from other weights, a requirement often seen in backpropagation methods. In contrast, virtual weights are encoded by neuronal activity and can propagate across the neural circuit, bypassing these limitations.

This allows virtual weight updates to depend on other virtual weights, enabling a more natural and precise evaluation of credit assignment.

The work generalizes and provides a biologically plausible implementation for deep gated RNNs performing ICL [215], maintaining a similar number of trainable parameters. By dividing the architecture into two components and introducing an additional objective that forces one network to approximate the gradient of the other, robust dynamic adaptation is achieved in tasks requiring temporal learning. When tasked to learn temporal trajectories, through error feedback and virtual weight updates, the network achieves successful dynamical adaptation, without synaptic plasticity. Similarly, in reinforcement learning scenarios, the policy in response to the environment state is dynamically modulated by virtual weights, which are updated in function of the reward.

Networks with gain modulation exhibit improved performance and robustness compared to those without gain modulation. Moreover, these gain-modulated networks demonstrate more data-efficient learning algorithms, outperforming non-gain-modulated counterparts in both in-distribution and out-of-distribution environments, showing their capability to face novel scenarios.
Adaptive network architectures can emerge by selecting appropriate readout weights in initially random networks, without requiring explicit backpropagation. These weights can be shaped through biologically plausible mechanisms such as evolutionary processes, slow reinforcement learning, or local plasticity rules relying on feedback alignment. Unlike standard deep-learning models, such as transformers, gated RNNs, and hyper-networks, that depend on non-local credit assignment mechanisms like backpropagation through time, this approach leverages local, dynamically reconfigurable processes, making it more compatible with biological learning principles.
In line with this, the model suggests that dynamic reconfiguration provides a biologically plausible mechanism for rapid adaptation, complementing traditional fast synaptic plasticity. Different neural circuits in the brain may rely on varying balances of dynamic reconfiguration and synaptic plasticity, shaped by their functional roles and evolutionary pressures. This distinction supports the potential for experimental validation.
In conclusion, the approach presented here provides an explicit framework to induce in-context learning (ICL) in biologically plausible networks, offering a constructive pathway toward understanding ICL in biological systems and bridging the gap between neuroscience and machine learning.

Chapter 8

Conclusion

Puoi disperdere
tutti i suoi petali
al vento
che tu lo voglia o no
un fiore avrà sempre la forza
di rinascere.

— Carmelina Rotundo, *Un fiore*

This thesis set out from the tension highlighted in the Introduction: bottom-up approaches grounded in statistical physics offer mechanistic realism but face severe analytical barriers at the level of complex circuit behavior, while top-down machine-learning models achieve striking algorithmic performance but often lack transparent links to biological substrate. The work developed mathematical instruments that operate in the space between these perspectives, refining our understanding on both sides of the divide. The main contributions are naturally organized around three pillars.

Rigorous Spectral Theory for Population Density Approach. A first contribution is a rigorous functional-analytic foundation for the spectral decomposition methods used in the population density approach. Motivated by the long-standing absence of a completeness proof for the biorthogonal eigenfunction expansions employed in practice, the thesis formulates the spectral problem in a boundary eigenvalue framework and establishes conditions under which the spectral expansion is mathematically well-posed. In particular, it resolves the completeness problem by proving Birkhoff regularity and develops the associated machinery needed to place spectral projections and semigroup evolution on solid ground. The analysis also characterizes *exceptional points* in the spectrum as defective eigenvalues, clarifying their role as singularities where eigenbranches coalesce and where generalized root functions (Jordan-chain structure) are required for completeness. These results close a foundational gap in the theory of non-self-adjoint Fokker–Planck operators with fire-and-reset boundary conditions, i.e. in a genuinely out-of-equilibrium setting.

Non-Stationary ISI Statistics. A second contribution addresses the mismatch between voltage-based population equations (which provide access to firing rates but lose explicit spike-time information) and age-structured descriptions (which track spike timing but typically rely on quasi-stationary

approximations for voltage statistics). The thesis bridges these viewpoints by formulating the dynamics on an *augmented state space* and deriving a two-dimensional Fokker–Planck equation for the joint density of membrane potential and time since last spike. This framework yields exact inter-spike-interval statistics far from stationarity, including regimes driven by time-dependent stimuli or internally generated oscillations. From this two-dimensional formulation, the thesis derives a tractable hierarchy of one-dimensional moment equations and develops a theory of the linear response for higher-order temporal statistics. The framework is validated through numerical simulations of large-scale spiking-neuron networks in strongly non-stationary regimes.

In-Context Learning via Gain Modulation. A third contribution shifts to a top-down computational lens and asks how rapid adaptation—of the kind seen in modern machine learning models performing in-context learning—could be implemented in biologically plausible recurrent circuits. Building on the latent-variable view of readouts for families of dynamical systems, the thesis proposes gain-modulated recurrent architectures in which adaptation is implemented through fast changes in neural activity ("virtual weights") rather than through immediate synaptic rewiring. In this setting, a fixed random recurrent substrate supports a rich dynamical feature space, while gain modulation enables context-dependent reconfiguration of the effective readout map. The thesis demonstrates constructively that such architectures can implement standard gradient-based optimization dynamics in real time, providing a concrete hypothesis for how a circuit-level mechanism (consistent with dendritic gain modulation) can realize gradient descent-like updates without changing synaptic weights during deployment. This offers a mechanistic bridge between biological plausibility constraints and algorithmic adaptation principles central to contemporary machine learning.

Future directions. While this thesis works with state-independent noise, most of the results can be extended to incorporate state-dependent diffusion $D = D(v)$, as naturally arises in conductance-based (COBA) synapses where fluctuations depend on the instantaneous membrane state [165]. In particular, our rigorous treatment of *parameter-dependent boundary conditions* in the derivation of transfer functions should become even more relevant in this setting.

A much more difficult task is to extend our analysis to neurons whose intrinsic state space has dimension > 1, such as models with adaptation currents, slow conductance dynamics, or additional gating variables. A possible starting point in this direction is the perturbative approach developed in [125]. Conceptually, we expect the properties established here to hold in broad generality, but proving them becomes substantially more challenging, as we can no longer rely on the theory of boundary eigenvalue problems. Discreteness of the spectrum comes from compactness of the resolvent, which stems from boundedness of the domain or from the presence of confining force fields. Dissipativity can likely be established by adapting the proofs in [19]. Sectoriality of the generator depends on strong ellipticity [28].

A particularly interesting route for studying multidimensional models would be to generalize Section 5.3.2, expressing the Green's function of the operator with reinjection as a correction of the (typically more tractable) operator without reinjection.

A further direction concerns the rigorous PDE study of *coupled populations*, where the self-consistent boundary flux can trigger finite-time blow-up. A promising avenue here relies on *time-dilation*

techniques [53, 81, 207]. In [53] the FP dynamics is reformulated in a rate-dependent dilated timescale, transforming it into a uniformly parabolic equation with bounded feedback, ensuring global well-posedness even through blow-up events (which map to instantaneous jumps in the original time). This perspective suggests that combining non-autonomous evolution-family theory with fixed-point arguments for the boundary flux could provide a systematic framework for treating existence, uniqueness, and long-time behaviour in coupled networks.
Another important direction is to quantify the approximation error introduced by projecting the infinite-dimensional dynamics onto the first K eigenmodes. While spectral truncations are widely used for dimensionality reduction, rigorous error bounds are lacking, particularly in non-normal regimes where transient growth can amplify neglected modes. Establishing such bounds would require estimating the magnitude of the *non-adiabatic coupling coefficients* $\boldsymbol{\Upsilon}^{\gamma}$ (which capture how time-varying inputs drive transitions between modes). In the introduced regularized framework, these coefficients remain bounded even near exceptional points, offering a promising path to derive uniform error estimates for low-dimensional reductions of the population dynamics.

Integration of Physics and Machine Learning in Cortical Theory The history of neural computation reveals a productive tension between two perspectives rooted in distinct disciplinary traditions. In the earliest formulations, the relationship between biophysics and computation was direct and unambiguous. Warren McCulloch and Walter Pitts, in their seminal 1943 work, explicitly modeled the fundamental units of computation on the physical behavior of biological neurons. By idealizing the neuron's "all-or-none" firing spike as a simple binary switch, they established a paradigm where the logical operations of a machine were derived directly from the physiology of the brain [128].
This bottom-up, physics-first approach reached its maturity with the Amari-Hopfield network in the early 1980s. While Amari was influenced by the associative memory mechanisms of Nakano's Associatron [139], the resulting framework emerged largely from statistical physics [5, 90], Energy minimization, attractor dynamics, and collective phenomena arising from interactions between microscopic components became the language through which neural computation was understood. The earliest associative memories, the first rigorous treatments of attractor-based computation, and the first explicit recognition that neural networks could be understood via energy functions all arose from this perspective.
In recent decades, this historical relationship has been substantially inverted. The most influential contemporary models of brain function now function as top-down normative hypotheses that specify *what* the brain ought to compute for optimal inference or information processing, then interrogate the neural substrate for evidence of those computations. This approach has proven remarkably generative, linking algorithmic principles to neural measurements and generating testable predictions about cortical circuitry.
The "Bayesian Brain" hypothesis [102] exemplifies this shift, positing that neural circuits are evolutionarily optimized to represent probability distributions, where biophysical noise serves as a resource for computation. Probabilistic Population Codes demonstrate how Poisson-like variability allows cortical networks to implement optimal Bayesian inference via simple linear operations [113], while sampling-based theories show that variability can directly represent uncertainty [146]. In sensory processing, "goal-driven" deep learning models have successfully predicted cortical responses

by optimizing for functional performance such as object recognition, showing that complex representations often emerge from the normative demands of the task itself [210].
At the macroscopic scale, predictive coding and the Free Energy Principle have emerged as unifying frameworks [64, 65]. The Free Energy Principle proposes that all self-organizing biological dynamics can be understood as minimizing a variational bound on surprise, offering a thermodynamic perspective on neural computation. Yet, these normative frameworks often treat biophysical constraints as implementation details rather than as fundamental features that actively shape what computations are possible and efficient.
However, a growing body of evidence suggests that the physical structure of biological neural networks is not a neutral substrate for arbitrary computation, but an active participant in defining the space of computations and their efficiency. Biological networks operate far from equilibrium, maintained in regimes of *driven stochasticity* and irreversible dynamics [34]. Rather than noise degrading an ideal function, empirical and theoretical work on perceptual decision-making suggests that cortical dynamics can be understood in terms of *noisy attractors* and noise-driven transitions among metastable states [24, 27]. In multistable perception, this same attractor picture makes explicit how fluctuations can both destabilize an ongoing state and enable flexible switching [24].
More specifically, work on binocular rivalry indicates that human observers operate in a narrow dynamical regime that balances two competing objectives: *stability* (maintaining a cognitive state to guide behavior) and *sensitivity* (responding quickly to new evidence) [150]. This regime has been characterized as lying near the boundary between *bistable* and *oscillatory* dynamics, where intrinsic fluctuations are neither fully damped nor fully chaotic, but tuned to maximize the system's susceptibility to input [150].
More broadly, noise in the nervous system is not a unitary nuisance but a ubiquitous feature spanning molecular to behavioral scales [57]. Far from merely degrading performance, this variability is often actively exploited: in nonlinear systems, *stochastic facilitation* mechanisms allow intermediate noise levels to enhance signal detection and information transmission, with *stochastic resonance* being the most canonical example [129]. From a dynamical systems perspective, intrinsic fluctuations do not just perturb trajectories but determine the transitions between metastable states, thereby shaping probabilistic perception and decision outcomes [51].
The synthesis required is not a simple interpretation of bottom-up emergent phenomena through top-down perspectives. Rather, it relies on a reciprocal feedback loop. Normative models specify the functional desiderata (what computation is required for survival) while physics-based analysis reveals how the biological substrate navigates these goals within strict material limits. Crucially, this is not a one-way implementation of abstract logic; the very constraints of the hardware (stochasticity, delays, metabolic costs) actively reshape the definition of optimality itself. By forcing the system to exploit, rather than suppress, its own physical instabilities, evolution has unlocked computational capabilities, such as noise-enhanced sensitivity and low-energy state switching, that pure algorithmic analysis would deem inefficient. Thus, the physics does not merely serve the computation; it expands the very horizon of what is computationally possible.

Learning and Synaptic Plasticity. The study of learning represents perhaps the highest-tension point between biological and machine learning perspectives. On one side, deep learning and modern optimization theory have achieved remarkable empirical success by treating learning as a determin-

istic optimization process: minimize a well-defined objective function via gradient descent, backpropagate error signals through multiple layers, and update weights iteratively. This framework is elegant, mathematically tractable, and computationally powerful. Yet, on the other side, biological neural systems must learn under stringent local and causal constraints that this framework largely ignores. These constraints reflect fundamental properties of biological tissue: a synapse has no direct access to the loss function of the entire network. Furthermore, biological learning happens at multiple timescales and operates continuously [143], learning and adapting while performing real-time computation, continuously responding to changing environments while maintaining functional performance—not in discrete offline phases.

Modern "machine-learning driven" approaches fundamentally attempt to work around these constraints [16, 137]. Approaches such as local learning rules with random feedback projections attempt to approximate backpropagation while respecting locality. Glia-inspired plasticity mechanisms and neuromodulatory signals have been proposed to bridge the credit assignment problem. Yet, most advances remain fundamentally phenomenological, as they attempt to explain what biological mechanism could plausibly approximate a global learning algorithm rather than investigating what principles of learning emerge naturally from the physics of biological circuits.

The insight offered here is a complementary perspective: that nonlocal learning rules—or more precisely, the effects of global optimization—can often be efficiently implemented *via* neural dynamics and rapid adaptation without requiring explicit weight modification. In this framework, a random recurrent reservoir, shaped by its connectivity and initial conditions, provides a high-dimensional substrate for computation. Rapid, non-synaptic adaptation mechanisms (gain modulation, facilitation, dendritic integration) can sculpt the effective input–output mapping of this substrate on the timescale of individual trials or even single stimuli. Local synaptic plasticity, operating at different timescales [67], then consolidates these dynamical adaptations into a persistent network structure. This framework, exemplified by random RNNs with fixed readout weights tuned via gain modulation, raises a profound open question: what local synaptic learning rules, starting from random initial connectivity, sculpt the network to support the gain-modulated rapid adaptation observed in practice?

Switching perspectives again: deriving a rigorous mean-field theory of plasticity and learning from first principles is extraordinarily difficult. Mean-field theory relies on homogeneity and independence assumptions, but these methods break down when plasticity enhances correlations and introduces symmetry breaking—precisely the regime where learning shapes the network structure. Recent work [46] shows promise in addressing this challenge, but the general theory remains underdeveloped.

A related and equally fundamental gap concerns the link between spike-based and rate-based plasticity rules. Classical treatments of synaptic plasticity distinguish sharply between spike-timing-dependent learning rules [38, 120] and rate-based Hebbian rules that depend only on the firing rates of pre- and postsynaptic populations. Yet much of theoretical neuroscience continues to employ rate-based descriptions despite growing evidence that spike timing matters critically[63, 76]. The two approaches have been shown to be consistent under certain assumptions (such as Poisson firing), but a unifying framework that allows for the study of spike-timing-dependent learning rules at the global population level remains elusive.

The framework presented in Chapter 3 represents a fundamental contribution to this objective, as it

enables rigorous characterization of single-neuron spiking statistics at the population level. Applied to pairs of neurons coupled via synapse dynamics, the same techniques could derive the population-level double dynamics that emerge from single-neuron circuits and synaptic interactions—a crucial step toward a unified theory of dynamics and learning in cortical neural networks.

Bibliography

[1] L. F. Abbott and C. van Vreeswijk. Asynchronous states in networks of pulse-coupled oscillators. *Physical Review E*, 48(2):1483–1490, 1993.

[2] W. C. Abraham, O. D. Jones, and D. L. Glanzman. Is plasticity of synapses the mechanism of long-term memory storage? *NPJ science of learning*, 4(1):9, 2019.

[3] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[4] P. Acquistapace and B. Terreni. A unified approach to abstract linear nonautonomous parabolic equations. *Rendiconti del seminario matematico della Università di Padova*, 78:47–107, 1987.

[5] S.-i. Amari. Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Transactions on Computers*, C-21(11):1197–1206, 1972.

[6] D. J. Amit and N. Brunel. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex*, 7(3):237–52, 1997.

[7] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters*, 55(14):1530, 1985.

[8] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Statistical mechanics of neural networks near saturation. *Annals of Physics*, 173(1):30–67, 1987.

[9] D. J. Amit and M. Tsodyks. Quantitative study of attractor neural network retrieving at low spike rates. i. substrate-spikes, rates and neuronal gain. *Network: Computation in neural systems*, 2(3):259, 1991.

[10] A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, 273(5283):1868–71, 1996.

[11] M. Augustin, J. Ladenbauer, F. Baumann, and K. Obermayer. Low-dimensional spike rate models derived from networks of adaptive integrate-and-fire neurons: comparison and implementation. *PLoS computational biology*, 13(6):e1005545, 2017.

[12] J. Ba, G. E. Hinton, V. Mnih, J. Z. Leibo, and C. Ionescu. Using fast weights to attend to the recent past. *Advances in neural information processing systems*, 29, 2016.

[13] L. Badel, S. Lefort, R. Brette, C. C. H. Petersen, W. Gerstner, and M. J. E. Richardson. Dynamic I-V curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. *Journal of Neurophysiology*, 99(2):656–66, 2008.

[14] W. Bair and C. Koch. Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comput.*, 8(6):1185–1202, 1996.

[15] W. Bair, C. Koch, W. T. Newsome, and K. Britten. Power spectrum analysis of bursting cells in area MT in the behaving monkey. *J. Neurosci.*, 14:2870–92, 1994.

[16] A. Balwani, A. Q. Wang, F. Najafi, and H. Choi. Constructing biologically constrained rnns via dale's backpropagation and topologically informed pruning. *Science Advances*, 11(50):eadw4970, 2025.

[17] P. Bashivan, K. Kar, and J. J. DiCarlo. Neural population control via deep image synthesis. *Science*, 364(6439):eaav9436, 2019.

[18] A. Bátkai and S. Piazzera. *Semigroups for Delay Equations*, volume 10 of *Research Notes in Mathematics*. A K Peters, Wellesley, MA, 2005.

[19] I. Ben-Ari and R. G. Pinsky. Ergodic behavior of diffusions with random jumps from the boundary. *Stochastic processes and their applications*, 119(3):864–881, 2009.

[20] D. A. Berry and B. Fristedt. Bandit problems: sequential allocation of experiments (monographs on statistics and applied probability). *London: Chapman and Hall*, 5(71-87):7–7, 1985.

[21] W.-J. Beyn. An integral method for solving nonlinear eigenvalue problems. *Linear Algebra and its Applications*, 436(10):3839–3863, 2012.

[22] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[23] V. Braitenberg and A. Schüz. *Anatomy of the cortex—Statistics and geometry*. Springer-Verlag, Berlin, 1991.

[24] J. Braun and M. Mattia. Attractors and noise: twin drivers of decisions and multistability. *Neuroimage*, 52(3):740–751, 2010.

[25] U. Braun, A. Schäfer, H. Walter, S. Erk, N. Romanczuk-Seiferth, L. Haddad, J. I. Schweiger, O. Grimm, A. Heinz, H. Tost, et al. Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proceedings of the National Academy of Sciences*, 112(37):11678–11683, 2015.

[26] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.

[27] B. A. W. Brinkman, H. Yan, A. Maffei, I. M. Park, A. Fontanini, J. Wang, and G. La Camera. Metastable dynamics of neural circuits and networks. *Applied Physics Reviews*, 9(1):011313, mar 2022.

[28] F. BROWDER. On the spectcral theory of elliptic differential operators. i. *Mathematische Annalen*, 142:22–130, 1961.

[29] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[30] N. Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208, 2000.

[31] N. Brunel and V. Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural computation*, 11(7):1621–1671, 1999.

[32] T. Cabana and J. Touboul. Large deviations for randomly connected neural networks: I. spatially extended systems. *Advances in Applied Probability*, 50(3):944–982, 2018.

[33] M. J. Cáceres, J. A. Carrillo, and B. Perthame. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *The Journal of Mathematical Neuroscience*, 1:1–33, 2011.

[34] R. Cao, A. Pastukhov, S. Aleshin, M. Mattia, and J. Braun. Binocular rivalry reveals an out-of-equilibrium neural dynamics suited for decision-making. *elife*, 10:e61581, 2021.

[35] R. M. Capocelli and L. M. Ricciardi. Diffusion approximation and first passage time problem for a model neuron. *Kybernetik*, 8(6):214–223, 1971.

[36] C. Capone and L. Falorsi. Adaptive behavior with stable synapses. *Neural Networks*, page 108082, 2025.

[37] C. Capone, C. Lupo, P. Muratore, and P. S. Paolucci. Beyond spiking networks: The computational advantages of dendritic amplification and input segregation. *Proceedings of the National Academy of Sciences*, 120(49):e2220743120, 2023.

[38] N. Caporale and Y. Dan. Spike timing–dependent plasticity: A Hebbian learning rule. *Annual Review of Neuroscience*, 31:25–46, 2008.

[39] J. A. Cardin, L. A. Palmer, and D. Contreras. Cellular mechanisms underlying stimulus-dependent gain modulation in primary visual cortex neurons in vivo. *Neuron*, 59(1):150–160, 2008.

[40] M. Chalvidal, T. Serre, and R. VanRullen. Meta-reinforcement learning with self-modifying networks. *Advances in Neural Information Processing Systems*, 35:7838–7851, 2022.

[41] J. Chevallier, M. J. Cáceres, M. Doumic, and P. Reynaud-Bouret. Microscopic approach of a time elapsed neural model. *Mathematical Models and Methods in Applied Sciences*, 25(14):2669–2719, 2015.

[42] A. V. Chizhov and L. J. Graham. Population model of hippocampal pyramidal neurons, linking a refractory density approach to conductance-based neurons. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 75(1):011924, 2007.

[43] A. V. Chizhov and L. J. Graham. Efficient evaluation of neuron populations receiving colored-noise current based on a refractory density method. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 77(1):011910, 2008.

[44] A. V. Chizhov, L. J. Graham, and A. A. Turbin. Simulation of neural population dynamics with a refractory density approach and a conductance-based threshold neuron model. *Neurocomputing*, 70(1-3):252–262, 2006.

[45] M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy. Neural population dynamics during reaching. *Nature*, 487(7405):51–56, 2012.

[46] D. Clark and L. Abbott. Theory of coupled neuronal–synaptic dynamics. *Physical Review X*, 14:021001, 2024.

[47] A. Compte, C. Constantinidis, J. Tegner, S. Raghavachari, M. V. Chafee, P. S. Goldman-Rakic, and X.-j. Wang. Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J. Neurophysiol.*, 90:3441–3454, 2003.

[48] N. D. Daw, Y. Niv, and P. Dayan. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*, 8(12):1704–1711, 2005.

[49] C. P. J. de Kock and B. Sakmann. High frequency action potential bursts (>or= 100 Hz) in L2/3 and L5B thick tufted neurons in anaesthetized and awake rat primary somatosensory cortex. *J. Physiol.*, 586(14):3353–64, 2008.

[50] M. De Pittà and N. Brunel. Multiple forms of working memory emerge from synapse–astrocyte interactions in a neuron–glia network model. *Proceedings of the National Academy of Sciences*, 119(43):e2207912119, 2022.

[51] G. Deco, E. T. Rolls, and R. Romo. Stochastic dynamics as a principle of brain function. *Progress in neurobiology*, 88(1):1–16, 2009.

[52] T. Deniz and S. Rotter. Solving the two-dimensional fokker-planck equation for strongly correlated neurons. *Physical Review E*, 95(1):012412, 2017.

[53] X. Dou and Z. Zhou. Dilating blow-up time: A generalized solution of the nnlif neuron model and its global well-posedness. *Communications in Mathematical Sciences*, 23(4):1081–1138, 2025.

[54] A. Duarte and S. Ostojic. Spectral dimensions of large-scale spiking neural networks. *arXiv preprint arXiv:1612.03723*, 2016.

[55] K.-J. Engel and R. Nagel. One-parameter semigroups for linear evolution equations. *Semigroup Forum*, 63:278–280, 1999.

[56] T. A. Engel, L. Schimansky-Geier, A. V. M. Herz, S. Schreiber, and I. Erchova. Subthreshold membrane-potential resonances shape spike-train patterns in the entorhinal cortex. *J. Neurophysiol.*, 100(3):1576–1589, 2008.

[57] A. A. Faisal, L. P. J. Selen, and D. M. Wolpert. Noise in the nervous system. *Nat. Rev. Neurosci.*, 9(4):292–303, 2008.

[58] L. Falorsi, G. V. Vinci, and M. Mattia. Non-stationary dynamics of interspike intervals in neuronal populations. *arXiv preprint arXiv:2512.23922*, 2025.

[59] L. Falorsi, G. V. Vinci, and M. Mattia. Population density dynamics of interacting spiking neurons with absolute refractory period. *In preparation*, 2025.

[60] W. Feller. The parabolic differential equations and the associated semi-groups of transformations. *Annals of Mathematics*, pages 468–519, 1952.

[61] K. A. Ferguson and J. A. Cardin. Mechanisms underlying gain modulation in the cortex. *Nature Reviews Neuroscience*, 21(2):80–92, 2020.

[62] B. Feulner, M. G. Perich, L. E. Miller, C. Clopath, and J. A. Gallego. Feedback-based motor control can guide plasticity and drive rapid learning. *bioRxiv*, pages 2022–10, 2022.

[63] N. Frémaux and W. Gerstner. Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Frontiers in Neural Circuits*, 9(JAN2016):85, jan 2015.

[64] K. Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138, 2010.

[65] K. Friston, J. Kilner, and L. Harrison. A free energy principle for the brain. *Journal of physiology-Paris*, 100(1-3):70–87, 2006.

[66] K.-i. Funahashi and Y. Nakamura. Approximation of dynamical systems by continuous time recurrent neural networks. *Neural networks*, 6(6):801–806, 1993.

[67] S. Fusi, P. J. Drew, and L. F. Abbott. Cascade models of synaptically stored memories. *Neuron*, 45(4):599–611, feb 2005.

[68] S. Fusi and M. Mattia. Collective behavior of networks with linear (vlsi) integrate-and-fire neurons. *Neural Computation*, 11(3):633–652, 1999.

[69] A. P. Georgopoulos, J. T. Lurito, M. Petrides, A. B. Schwartz, and J. T. Massey. Mental rotation of the neuronal population vector. *Science*, 243(4888):234–236, 1989.

[70] A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner. Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419, 1986.

[71] G. L. Gerstein and B. Mandelbrot. Random walk models for the spike activity of a single neuron. *Biophysical Journal*, 4(1):41–68, 1964.

[72] W. Gerstner. Time structure of the activity in neural network models. *Phys. Rev. E*, 51(1):738–758, 1995.

[73] W. Gerstner. Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural computation*, 12(1):43–89, 2000.

[74] W. Gerstner and J. L. van Hemmen. Associative memory in a network of 'spiking' neurons. *Network: Computation in Neural Systems*, 3(2):139–164, 1992.

[75] A. Goldstein, Z. Zada, E. Buchnik, M. Schain, A. Price, B. Aubrey, S. A. Nastase, A. Feder, D. Emanuel, A. Cohen, A. Jansen, H. Gazula, G. Choe, A. Rao, C. Kim, C. Casto, L. Fanda, W. Doyle, D. Friedman, P. Dugan, L. Melloni, R. Reichart, S. Devore, A. Flinker, L. Hasenfratz, O. Levy, A. Hassidim, M. Brenner, Y. Matias, K. A. Norman, O. Devinsky, and U. Hasson. Shared computational principles for language processing in humans and deep language models. *Nature Neuroscience*, 25(3):369–380, 2022.

[76] M. Graupner and N. Brunel. Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proceedings of the National Academy of Sciences of the USA*, 109(10):3991–6, 2012.

[77] I. Grigorescu and M. Kang. Brownian motion on the figure eight. *Journal of Theoretical Probability*, 15:817–844, 2002.

[78] A. Gu, K. Goel, and C. Re. Efficiently modeling long sequences with structured state spaces. In *International Conference on Learning Representations*, 2021.

[79] D. Guarino, I. Carannante, and A. Destexhe. Neuromodulators control neuronal dynamics through feature space reshaping. *bioRxiv*, pages 2025–05, 2025.

[80] J. Guerguiev, T. P. Lillicrap, and B. A. Richards. Towards deep learning with segregated dendrites. *eLife*, 6:e22901, 2017.

[81] R. Gütig and H. Sompolinsky. Time-warp-invariant neuronal processing. *PLoS Biology*, 7(7), 2009.

[82] M. Haase. *The Functional Calculus for Sectorial Operators*, volume 169 of *Operator Theory: Advances and Applications*. Birkhäuser Basel, 2006.

[83] A. Hart, J. Hook, and J. Dawes. Embedding and approximation theorems for echo state networks. *Neural Networks*, 128:234–247, 2020.

[84] J. B. Heald, M. Lengyel, and D. M. Wolpert. Contextual inference underlies the learning of sensorimotor repertoires. *Nature*, 600(7889):489–493, 2021.

[85] J. B. Heald, M. Lengyel, and D. M. Wolpert. Contextual inference in learning and memory. *Trends in cognitive sciences*, 27(1):43–64, 2023.

[86] J. B. Heald, D. M. Wolpert, and M. Lengyel. The computational and neural bases of context-dependent learning. *Annual Review of Neuroscience*, 46(Volume 46, 2023):233–258, 2023.

[87] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, and C. Elvezia. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies.

[88] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544, 1952.

[89] J. Höller, N. Read, and J. G. E. Harris. Non-hermitian adiabatic transport in spaces of exceptional points. *Physical Review A*, 102(3):032216, 2020.

[90] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.

[91] K. Irie, I. Schlag, R. Csordás, and J. Schmidhuber. Going beyond linear transformers with recurrent fast weight programmers. *Advances in neural information processing systems*, 34:7703–7717, 2021.

[92] H. Jaeger. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *GMD Report*, 148(34):13, 2001.

[93] H. Jaeger. Controlling recurrent neural networks by conceptors. *arXiv preprint arXiv:1403.3369*, 2014.

[94] H. Jaeger. Using conceptors to manage neural long-term memories for temporal patterns. *Journal of Machine Learning Research*, 18(13):1–43, 2017.

[95] H. Jaeger and H. Haas. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *science*, 304(5667):78–80, 2004.

[96] L. P. Jiang and R. P. Rao. Dynamic predictive coding: A model of hierarchical sequence learning and prediction in the neocortex. *PLoS Computational Biology*, 20(2):e1011801, 2024.

[97] P. Johannesma. Diffusion models for the stochastic activity of neurons. In E. Caianiello, editor, *Neural Networks*, pages 116–144. Springer, Berlin, 1968.

[98] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth. *Principles of Neural Science*. McGraw-Hill Companies, New York, fifth edit edition, 2013.

[99] B. W. Knight. Dynamics of encoding in neuron populations: some general mathematical features. *Neural Comput.*, 12(3):473–518, 2000.

[100] B. W. Knight, D. Manin, and L. Sirovich. Dynamical models of interacting neuron populations. In E. Gerf, editor, *Symposium on Robotics and Cybernetics: Computational Engineering in Systems Applications*, pages 9–12, Lille, France, 1996.

[101] B. W. Knight, D. Manin, and L. Sirovich. Dynamical models of interacting neuron populations in visual cortex. In E. Gerf, editor, *Symposium on Robotics and Cybernetics: Computational Engineering in Systems Applications*, pages 1–5, Cite Scientifique, Lille, France, 1996. Cite Scientifique.

[102] D. C. Knill and A. Pouget. The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719, 2004.

[103] E. Köksal-Ersöz, P. Chossat, and F. Lavigne. Gain modulation of actions selection without synaptic relearning. *arXiv preprint arXiv:2501.01964*, 2024.

[104] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[105] M. Larkum. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends in Neurosciences*, 36(3):141 – 151, 2013.

[106] I. Lee and C.-H. Lee. Contextual behavior and neural circuits. *Frontiers in neural circuits*, 7:84, 2013.

[107] Y.-J. Leung. One dimensional brownian motion with holding and jumping boundary. *Mathematical Methods in the Applied Sciences*, 47(17):13051–13062, 2022.

[108] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier. Effects of noise in excitable systems. *Physics Reports*, 392(6):321–424, 2004.

[109] B. Lindner, A. Longtin, and A. Bulsara. Analytic expressions for rate and CV of a type I neuron driven by white gaussian noise. *Neural Comput.*, 15(8):1761–1788, 2003.

[110] A. Litwin-Kumar and B. Doiron. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat. Neurosci.*, 15(11):1498–1505, 2012.

[111] J.-G. Liu, Z. Wang, Y. Zhang, and Z. Zhou. Rigorous justification of the fokker–planck equations of neural networks based on an iteration perspective. *SIAM Journal on Mathematical Analysis*, 54(1):1270–1312, 2022.

[112] A. Longtin, A. R. Bulsara, and F. Moss. Time-interval sequences in bistable systems and the noise-induced transmission of information by sensory neurons. *Phys. Rev. Lett.*, 67(5):656–659, 1991.

[113] W. J. Ma, J. M. Beck, P. E. Latham, and A. Pouget. Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11):1432–1438, 2006.

[114] W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531–2560, 2002.

[115] D. J. C. MacKay. Bayesian interpolation. *Neural Computation*, 4(3):415–447, 1992.

[116] G. Maimon and J. A. Assad. Beyond Poisson: increased spike-time regularity across primate parietal cortex. *Neuron*, 62(3):426–40, 2009.

[117] P. Mandl. *Analytical treatment of one-dimensional Markov processes*. Springer, 1968.

[118] V. Mante, D. Sussillo, K. V. Shenoy, and W. T. Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474):78–84, 2013.

[119] N. T. Markov, M. Ercsey-Ravasz, D. C. Van Essen, K. Knoblauch, Z. Toroczkai, and H. Kennedy. Cortical high-density counterstream architectures. *Science*, 342(6158):1238406, nov 2013.

[120] H. Markram, W. Gerstner, and P. Sjöström. Spike-timing-dependent plasticity: A comprehensive overview. *Frontiers in Synaptic Neuroscience*, 4:2, 2012.

[121] D. Marr and T. Poggio. From understanding computation to understanding neural circuitry. 1976.

[122] F. Mastrogiuseppe and S. Ostojic. Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron*, 99(3):609–623, 2018.

[123] M. V. Matthews, W. L. Ellsworth, and P. A. Reasenberg. A Brownian Model for Recurrent Earthquakes. *Bulletin of the Seismological Society of America*, 92(6):2233–2250, 2002.

[124] M. Mattia. Low-dimensional firing rate dynamics of spiking neuron networks. *arXiv preprint arXiv:1609.08855*, 2016.

[125] M. Mattia, M. Biggio, A. Galluzzi, and M. Storace. Dimensional reduction in networks of non-Markovian spiking neurons: Equivalence of synaptic filtering and heterogeneous propagation delays. *PLoS Computational Biology*, 15(10):e1007404, oct 2019.

[126] M. Mattia and P. Del Giudice. Population dynamics of interacting spiking neurons. *Physical Review E*, 66(5):051917, 2002.

[127] M. Mattia and G. V. Vinci. Low dimensional dynamics of spiking neuron networks. *Zenodo*, page 5518215, 2021.

[128] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4):115–133, 1943.

[129] M. D. McDonnell and L. M. Ward. The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience*, 12(7):415–425, 2011.

[130] S. D. McDougle, K. M. Bond, and J. A. Taylor. Explicit and implicit processes constitute the fast and slow processes of sensorimotor learning. *Journal of Neuroscience*, 35(26):9568–9579, 2015.

[131] R. Mennicken and M. Möller. *Non-self-adjoint boundary eigenvalue problems*, volume 192. Gulf Professional Publishing, 2003.

[132] A. Meulemans, N. Zucchet, S. Kobayashi, J. Von Oswald, and J. Sacramento. The least-control principle for local learning at equilibrium. *Advances in Neural Information Processing Systems*, 35:33603–33617, 2022.

[133] T. Miconi, K. Stanley, and J. Clune. Differentiable plasticity: training plastic neural networks with backpropagation. In *International Conference on Machine Learning*, pages 3559–3568. PMLR, 2018.

[134] B. Millidge, M. Tang, M. Osanlouy, N. S. Harper, and R. Bogacz. Predictive coding networks for temporal prediction. *PLOS Computational Biology*, 20(4):e1011183, 2024.

[135] R. G. Morris. Spatial localization does not require the presence of local cues. *Learning and motivation*, 12(2):239–260, 1981.

[136] O. Moynot and M. Samuelides. Large deviations and mean-field theory for asymmetric random recurrent neural networks. *Probability Theory and Related Fields*, 123(1):41–75, 2002.

[137] J. Murray. Local online learning in recurrent networks with random feedback. *eLife*, 8:e43299, 2019.

[138] M. Musila and P. Lánský. On the interspike intervals calculated from diffusion approximations of Stein's neuronal model with reversal potentials. *J. Theor. Biol.*, 171(2):225–232, 1994.

[139] K. Nakano. Associatron-A Model of Associative Memory. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-2(3):380–388, jul 1972.

[140] R. M. Neal. *Bayesian Learning for Neural Networks*, volume 118 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1996.

[141] D. Q. Nykamp and D. Tranchina. A population density approach that facilitates large-scale modeling of neural networks: analysis and an application to orientation tuning. *J. Comput. Neurosci.*, 8(1):19–50, 2000.

[142] J. P. O'Doherty, J. Cockburn, and W. M. Pauli. Learning, reward, and decision making. *Annual review of psychology*, 68(1):73–100, 2017.

[143] C. O'Donnell. Nonlinear slow-timescale mechanisms in synaptic plasticity. *Current Opinion in Neurobiology*, 79:102671, 2023.

[144] C. Olsson, N. Elhage, N. Nanda, N. Joseph, N. DasSarma, T. Henighan, B. Mann, A. Askell, Y. Bai, A. Chen, et al. In-context learning and induction heads. *arXiv preprint arXiv:2209.11895*, 2022.

[145] A. Omurtag, B. W. Knight, and L. Sirovich. Dynamics of neuronal populations: The equilibrium solution. *SIAM Journal on Applied Mathematics*, 60(6):2009–2028, 2000.

[146] G. Orbán, P. Berkes, J. Fiser, and M. Lengyel. Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron*, 92(2):530–543, 2016.

[147] A. Orvieto, S. L. Smith, A. Gu, A. Fernando, C. Gulcehre, R. Pascanu, and S. De. Resurrecting recurrent neural networks for long sequences. In *International Conference on Machine Learning*, pages 26670–26698. PMLR, 2023.

[148] S. Ostojic. Interspike interval distributions of spiking neurons driven by fluctuating inputs. *J. Neurophysiol.*, 106(1):361–73, 2011.

[149] R. Pascanu, T. Mikolov, and Y. Bengio. On the difficulty of training recurrent neural networks. In *International conference on machine learning*, pages 1310–1318. Pmlr, 2013.

[150] A. Pastukhov, P. E. García-Rodríguez, J. Haenicke, A. Guillamon, G. Deco, and J. Braun. Multi-stable perception balances stability and sensitivity. *Frontiers in computational neuroscience*, 7:17, 2013.

[151] A. Payeur, J. Guerguiev, F. Zenke, B. A. Richards, and R. Naud. Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nature neuroscience*, 24(7):1010–1019, 2021.

[152] J. Peng and W. V. Li. Diffusions with holding and jumping boundary. *Science China Mathematics*, 56:161 – 176, 2012.

[153] E. Perez, F. Strub, H. De Vries, V. Dumoulin, and A. Courville. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[154] B. Pietras, N. Gallice, and T. Schwalger. Low-dimensional firing-rate dynamics for populations of renewal-type spiking neurons. *Phys. Rev. E*, 102(2):022407, 2020.

[155] P. Poirazi and A. Papoutsi. Illuminating dendritic function with computational models. *Nature Reviews Neuroscience*, 21(6):303–321, 2020.

[156] C. Pozzorini, S. Mensi, O. Hagens, R. Naud, C. Koch, and W. Gerstner. Automated high-throughput characterization of single neurons by means of simplified spiking models. *PLoS Computational Biology*, 11(6):1–29, 2015.

[157] R. Pyle and R. Rosenbaum. Spatiotemporal dynamics and reliable computations in recurrent spiking neural networks. *Physical Review Letters*, 118(1):018103, 2017.

[158] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever. Zero-shot text-to-image generation. In *International conference on machine learning*, pages 8821–8831. Pmlr, 2021.

[159] H. Ramsauer, B. Schäfl, J. Lehner, P. Seidl, M. Widrich, T. Adler, L. Gruber, M. Holzleitner, M. Pavlović, G. K. Sandve, et al. Hopfield networks is all you need. *arXiv preprint arXiv:2008.02217*, 2020.

[160] A. Rauch, G. La Camera, H.-R. Lüscher, W. Senn, and S. Fusi. Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo-like input currents. *Journal of Neurophysiology*, 90(3):1598–612, 2003.

[161] L. M. Ricciardi. Diffusion approximations and computational problems for single neuron's activity. *Competition and Cooperation in Neural Nets*, pages 300–319, 1976.

[162] L. M. Ricciardi and L. Sacerdote. The Ornstein-Uhlenbeck process as a model for neuronal activity. I. Mean and variance of the firing time. *Biol. Cybern.*, 35(1):1–9, 1979.

[163] L. M. Ricciardi and S. Sato. First-passage-time density and moments of the Ornstein-Uhlenbeck process. *J. Appl. Prob.*, 25(1):43–57, 1988.

[164] M. J. E. Richardson. The effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Physical Review E*, 69(4):044405, 2004.

[165] M. J. E. Richardson. Effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Physical Review E*, 69:051918, 2004.

[166] B. Romera-Paredes and P. Torr. An embarrassingly simple approach to zero-shot learning. In *International conference on machine learning*, pages 2152–2161. PMLR, 2015.

[167] R. Rosenbaum and B. Doiron. Balanced networks of spiking neurons with spatially dependent recurrent connections. *Physical Review X*, 4(2):021039, 2014.

[168] J. Sacramento, R. Ponte Costa, Y. Bengio, and W. Senn. Dendritic cortical microcircuits approximate the backpropagation algorithm. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 8721–8732. Curran Associates, Inc., 2018.

[169] T. Sakurai and H. Sugiura. A projection method for generalized eigenvalue problems using numerical integration. *Journal of computational and applied mathematics*, 159(1):119–128, 2003.

[170] M. V. Sanchez-Vives, M. Massimini, and M. Mattia. Shaping the default activity pattern of the cortical network. *Neuron*, 94(5):993–1001, 2017.

[171] A. Sanzeni, M. H. Histed, and N. Brunel. Emergence of irregular activity in networks of strongly coupled conductance-based neurons. *Phys. Rev. X*, 12(1):11044, 2022.

[172] J. Schmidhuber. Learning to control fast-weight memories: An alternative to dynamic recurrent networks. *Neural Computation*, 4(1):131–139, 1992.

[173] F. Schuessler, F. Mastrogiuseppe, A. Dubreuil, S. Ostojic, and O. Barak. Dynamics of random recurrent networks with correlated low-rank structure. *Physical Review Research*, 2(1):013111, 2020.

[174] T. Schwalger. Mapping input noise to escape noise in integrate-and-fire neurons: a level-crossing approach. *Biological Cybernetics*, 115(5):539–562, 2021.

[175] T. Schwalger and A. V. Chizhov. Mind the last spike — firing rate models for mesoscopic populations of spiking neurons. *Curr. Opin. Neurobiol.*, 58:155–166, 2019.

[176] W. Senn, D. Dold, A. F. Kungl, B. Ellenberger, J. Jordan, Y. Bengio, J. Sacramento, and M. A. Petrovici. A neuronal least-action principle for real-time learning in cortical circuits. *ELife*, 12:RP89674, 2024.

[177] M. N. Shadlen and W. T. Newsome. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.*, 4(4):569–579, 1994.

[178] M. N. Shadlen and W. T. Newsome. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.*, 18(10):3870–96, 1998.

[179] A. S. Shai, C. A. Anastassiou, M. E. Larkum, and C. Koch. Physiology of layer 5 pyramidal neurons in mouse primary visual cortex: coincidence detection through bursting. *PLoS computational biology*, 11(3):e1004090, 2015.

[180] S. Shinomoto, H. Kim, T. Shimokawa, N. Matsuno, S. Funahashi, K. Shima, I. Fujita, H. Tamura, T. Doi, K. Kawano, N. Inaba, K. Fukushima, S. Kurkin, K. Kurata, M. Taira, K.-I. Tsutsui, H. Komatsu, T. Ogawa, K. Koida, J. Tanji, and K. Toyama. Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Comput. Biol.*, 5(7):e1000433, 2009.

[181] H. T. Siegelmann and E. D. Sontag. Turing computability with neural nets. *Applied Mathematics Letters*, 4(6):77–80, 1991.

[182] A. Siegert. On the first passage time probability problem. *Phys. Rev.*, 81(4):617–623, 1951.

[183] A. Singh, S. Chan, T. Moskovitz, E. Grant, A. Saxe, and F. Hill. The transient nature of emergent in-context learning in transformers. *Advances in Neural Information Processing Systems*, 36, 2024.

[184] W. R. Softky and C. Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.*, 13(1):334–50, 1993.

[185] J. Soldado-Magraner, V. Mante, and M. Sahani. Inferring context-dependent computations through linear approximations of prefrontal cortex dynamics. *Science Advances*, 10(51):eadl4743, 2024.

[186] H. Sompolinsky, A. Crisanti, and H.-J. Sommers. Chaos in random neural networks. *Physical Review Letters*, 61(3):259, 1988.

[187] D. I. Standage, C. N. Areshenkoff, D. J. Gale, J. Y. Nashed, J. R. Flanagan, and J. P. Gallivan. Whole-brain dynamics of human sensorimotor adaptation. *Cerebral Cortex*, 33(8):4761–4778, 2023.

[188] J. P. Stroud, M. A. Porter, G. Hennequin, and T. P. Vogels. Motor primitives in space and time via targeted gain modulation in cortical networks. *Nature neuroscience*, 21(12):1774–1783, 2018.

[189] N. V. Swindale, M. A. Spacek, M. R. Krause, and C. Mitelut. Spontaneous activity in cortical neurons is stereotyped and non-Poisson. *Cereb. Cortex*, 33, 2023.

[190] C. Teeter, R. Iyer, V. Menon, N. Gouwens, D. Feng, J. Berg, A. Szafer, N. Cain, H. Zeng, M. Hawrylycz, C. Koch, and S. Mihalas. Generalized leaky integrate-and-fire models classify multiple neuron types. *Nature Communications*, 9(1):709, 2018.

[191] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. Section 3.1, Eq. (3.5).

[192] H. C. Tuckwell. *Introduction to theoretical neurobiology: linear cable theory and dendritic structure*, volume 1. Cambridge University Press, 1988.

[193] H. C. Tuckwell. *Introduction to theoretical neurobiology: volume 2, nonlinear and stochastic theories*. Cambridge University Press, Cambridge, 1988.

[194] R. Urbanczik and W. Senn. Learning by the dendritic prediction of somatic spiking. *Neuron*, 81(3):521–528, 2014.

[195] C. van Vreeswijk and H. Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724–1726, 1996.

[196] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[197] R. D. Vilela and B. Lindner. Comparative study of different integrate-and-fire neurons: Spontaneous activity, dynamical response, and stimulus-induced correlation. *Phys. Rev. E*, 80(3):1–12, 2009.

[198] G. V. Vinci and M. Mattia. Rosetta stone for the population dynamics of spiking neuron networks. *Physical Review E*, 110(3):034303, 2024.

[199] O. Vinyals, C. Blundell, T. Lillicrap, D. Wierstra, et al. Matching networks for one shot learning. *Advances in neural information processing systems*, 29, 2016.

[200] J. Von Oswald, E. Niklasson, E. Randazzo, J. Sacramento, A. Mordvintsev, A. Zhmoginov, and M. Vladymyrov. Transformers learn in-context by gradient descent. In *International Conference on Machine Learning*, pages 35151–35174. PMLR, 2023.

[201] S. Vyas, M. D. Golub, D. Sussillo, and K. V. Shenoy. Computation through Neural Population Dynamics. *Annual Review of Neuroscience*, 43:249–275, 2020.

[202] V. V. Vyazovskiy, U. Olcese, Y. M. Lazimy, U. Faraguna, S. K. Esser, J. C. Williams, C. Cirelli, and G. Tononi. Cortical firing and sleep homeostasis. *Neuron*, 63(6):865–878, 2009.

[203] G. Wainstein, C. J. Whyte, K. A. E. Martens, E. J. Müller, V. Medel, B. Anderson, E. Stöttinger, J. Danckert, B. R. Munn, and J. M. Shine. Evidence from pupillometry, fmri, and rnn modelling shows that gain neuromodulation mediates task-relevant perceptual switches. *eLife*, 13:RP93191, 2025.

[204] X.-J. Wang. Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nature Reviews Neuroscience*, 21(3):169–178, mar 2020.

[205] P. J. Werbos. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560, 2002.

[206] W. J. Wilbur and J. Rinzel. A theoretical basis for large coefficient of variation and bimodality in neuronal interspike interval distributions. *J. Theor. Biol.*, 105(2):345–368, 1983.

[207] A. H. Williams, B. Poole, N. Maheswaranathan, A. K. Dhawale, T. Fisher, C. D. Wilson, D. H. Brann, E. M. Trautmann, S. Ryu, R. Shusterman, D. Rinberg, B. P. Ölveczky, K. V. Shenoy, and S. Ganguli. Discovering precise temporal patterns in large-scale neural recordings through robust and interpretable time warping. *Neuron*, 105(2):246–259.e8, 2020.

[208] M. A. Wilson and B. L. McNaughton. Dynamics of the hippocampal ensemble code for space. *Science*, 261(5124):1055–1058, 1993.

[209] W. A. Wybo, M. C. Tsai, V. A. K. Tran, B. Illing, J. Jordan, A. Morrison, and W. Senn. Nmda-driven dendritic modulation enables multitask representation learning in hierarchical sensory processing pathways. *Proceedings of the National Academy of Sciences*, 120(32):e2300558120, 2023.

[210] D. L. K. Yamins and J. J. DiCarlo. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 19(3):356–365, 2016.

[211] G. R. Yang and X.-J. Wang. Artificial neural networks for neuroscientists: A primer. *Neuron*, 107(6):1048–1070, 2020.

[212] I. B. Yildiz, H. Jaeger, and S. J. Kiebel. Re-visiting the echo state property. *Neural networks*, 35:1–9, 2012.

[213] R. Zhang, S. Frei, and P. L. Bartlett. Trained transformers learn linear models in-context. *Journal of Machine Learning Research*, 25(49):1–55, 2024.

[214] D. Zipser, B. Kehoe, G. Littlewort, and J. M. Fuster. A spiking network model of short-term active memory. *J. Neurosci.*, 13(8):3406–20, 1993.

[215] N. Zucchet, S. Kobayashi, Y. Akram, J. Von Oswald, M. Larcher, A. Steger, and J. Sacramento. Gated recurrent neural networks discover attention. *arXiv preprint arXiv:2309.01775*, 2023.

Appendix A

Appendix

A.1 Further details on Dissipativity

Given $\mathbf{p} = (p, p^r) \in \mathbb{L}^p \subseteq \mathbb{L}^1, p \in [1, \infty]$ we define the sign of $\mathbf{p}$ as $\text{sign}(\mathbf{p}) = (\text{sign}(p), \text{sign}(p^r)) \in \mathbb{L}^\infty$ where:

$$\text{sign}(p)(v) = \begin{cases} \frac{\overline{p(v)}}{|p(v)|} & p(v) \neq 0 \\ 0 & p(v) = 0 \end{cases} \qquad \text{sign}(p^r)(\tau) = \begin{cases} \frac{\overline{p^r(\tau)}}{|p^r(\tau)|} & p^r(\tau) \neq 0 \quad v \in (\alpha, \theta) \\ 0 & p^r(\tau) = 0 \quad \tau \in (0, \tau_0) \end{cases}$$

such that $\langle \text{sign}(\mathbf{p}), \mathbf{p} \rangle = \|\mathbf{p}\|_{\mathbb{L}^1}$. Then $\|\mathbf{p}\|_{\mathbb{L}^1} \text{sign}\mathbf{p} \in \mathbb{L}^\infty$ is a element of the duality set of $\mathbf{p} \in \mathbb{L}^1$:

$$\|\mathbf{p}\|_{\mathbb{L}^1} \text{ sign}\mathbf{p} \in J(p) = \left\{ \mathbf{f} \in \mathbb{L}^\infty \;\middle|\; \langle \mathbf{f}, \mathbf{p} \rangle = \|\mathbf{p}\|^2_{\mathbb{L}^1} \right\} \tag{A.1}$$

If we rewrite $\mathbf{p} = (p, p^r)$ is polar cordinates:

$$p(v) = h(v)e^{i\beta(v)} \quad p^r(\tau) = h^r(\tau)e^{i\beta^r(\tau)} \tag{A.2}$$

Where $\mathbb{R} \ni h(v), h^r(\tau) \geq 0$ and $\beta(v), \beta^r(\tau) \in \mathbb{R}$. The sign is given by:

$$\text{sign}(p)(v) = \begin{cases} e^{-i\beta(v)} & h(v) > 0 \\ 0 & h(v) = 0 \end{cases} \quad \text{sign}(p^r)(\tau) = \begin{cases} e^{-i\beta^r(\tau)} & h^r(\tau) > 0 \quad v \in (\alpha, \theta) \\ 0 & h(\tau) = 0 \quad \tau \in (0, \tau_0) \end{cases}$$

Proof of Lemma 4.6.1. As $\mathbf{p} \in {}^1\mathbb{D}_{\boldsymbol{\gamma}}$ is a continuous function the set $\{v \in (\alpha, \theta) : h(v) > 0\}$ is open and can be written as a (at most) countable union of intervals $\{v \in (\alpha, \theta) : h(v) > 0\} = \cup_n (a_n, b_n)$. Integrating by parts and imposing boundary conditions we have:

$$\begin{aligned} &Re\left(\int_\alpha^\theta \text{sign}\left(p(v)\right) \mathcal{L}_{\boldsymbol{\gamma}} p(v) dv - \int_0^{\tau_0} \text{sign}\left(p^r(\tau)\right) \partial_\tau p^r(\tau) d\tau \right) = \\ &= -D \sum_n \int_{a_n}^{b_n} h(v)(\beta'(v))^2 dv + D \sum_n \left(h'(b_n^-) - h'(a_n^+) \right) + \\ &\quad + h^r(0)(\cos(\beta(\theta) - \beta^r(0)) - 1) + h^r(\tau_0)(\cos(\beta(H) - \beta^r(\tau_0)) - 1) \leq 0 \end{aligned} \tag{A.3}$$

notice that since $h(v) \geq 0 \; \forall v \in (\alpha, \theta)$ and $h(a_n) = h(b_n) = 0$, we have $h'(b_n^-) \leq 0, h'(a_n^+) \geq 0$. Dissipativity follows from [55, Proposition 3.23] □

and $0 = \lambda_0 \in \sigma(\boldsymbol{\gamma})$ is the only eigenvalue with non-negative real part.

Proof of Proposition 4.6.1. Let $\lambda \in \sigma(\boldsymbol{\gamma})$, with $\boldsymbol{\varphi}(\lambda, \boldsymbol{\gamma}) \in {}^{p}\mathbb{D}_{\boldsymbol{\gamma}} \subset {}^{1}\mathbb{D}_{\boldsymbol{\gamma}}$, then by (4.56):

$$Re(\lambda)\|\boldsymbol{\varphi}(\lambda, \boldsymbol{\gamma})\|_{\mathbb{L}^1} \leq 0 \tag{A.4}$$

This proves (4.57). Suppose now that there exists $i\omega \in \sigma(\boldsymbol{\gamma})$ with $\omega \in \mathbb{R}$. Up to a complex scalar factor, the corresponding eigenfunction will have the form $\boldsymbol{\varphi}(i\omega, \boldsymbol{\gamma}) = (h_\omega(v)exp(i\beta_\omega(v)), \exp(i\tau\omega))$. We then have:

$$Re\langle \text{sign}\boldsymbol{\varphi}(i\omega, \boldsymbol{\gamma}), \mathcal{T}_{\boldsymbol{\gamma}}\boldsymbol{\varphi}(i\omega, \boldsymbol{\gamma})\rangle = Re(i\omega\|\boldsymbol{\varphi}(i\omega, \boldsymbol{\gamma})\|_{\mathbb{L}^1}) = 0 \tag{A.5}$$

This holds if and only if all the negative terms in (A.3) vanish. The reader can verify that this occurs precisely when $\beta_w(v) \equiv 0 \wedge h_w(v) > 0,\ \forall v \in (\alpha, \theta)$ and ω is of the form $\omega = n2\pi/\tau_0$ for some $n \in \mathbb{Z}$. Then as $0 \in \sigma(\boldsymbol{\gamma}) \Rightarrow \boldsymbol{\Psi}(0, \boldsymbol{\gamma}) \in \mathbb{D}^*$ we have:

$$0 = \langle \boldsymbol{\Psi}(0, \boldsymbol{\gamma}), \mathcal{T}_{\boldsymbol{\gamma}}\boldsymbol{\varphi}(iw, \boldsymbol{\gamma})\rangle = in\frac{2\pi}{\tau_0}\left(\int_\alpha^\theta h_\omega(v)dv + \int_0^{\tau_0} cos\left(n\frac{2\pi\tau}{\tau_0}\right)d\tau\right) \tag{A.6}$$

this implies $n = 0 \Rightarrow \omega = 0$. □

Acknowledgements

Grazie alla mia Famiglia, Mamma, Papà e Giulio per aver sempre creduto in me e per avermi sostenuto e supportato in tutto il mio percorso di studi e soprattutto in questi ultimi anni. Grazie a chi c'è da sempre: Dragos, Luca, Marco, Davide e Filippo, compagni di una vita, per esserci stati in ogni occasione, per me il legame che condividiamo è veramente speciale. Grazie Brainjuice, per tutto quello che abbiamo vissuto e costruito insieme e per avermi insegnato cosa significa essere Comunità Scientifica. Grazie a Gabriele per le interminabili chiacchierate e per aver ascoltato i miei sproloqui sulla completezza davanti ad una Peroni in piazzetta.
Grazie a chi c'è da poco e con cui ho condiviso questa fine che è stata un inizio: grazie a Luca, Maite e Matteo, e grazie al Coro, per esistere e per essere stupend3.
Grazie a Fede e Camilla: siete la mia famiglia e il dottorato rimarrà sempre legato alle giornate passate in allegria e spensieratezza insieme nel giardino di Via Arati.
Anche se non posso ringraziare chi fugge, questa tesi è dedicata a te, Giorgia, che resti un mistero più grande di queste equazioni. La maggior parte di queste righe le ho scritte con te accanto e sognando una vita insieme, grazie a te sono una persona migliore.
Grazie a chi ha condiviso con me e dato luce a questi anni di dottorato in ISS: grazie Sofia, Gabriele e Lorenzo. Grazie per gli spunti, il confronto e il reciproco conforto e tutti i momenti anche semplici ma mai noiosi passati insieme. Ma anche a Cristiano, per le interminabili discussioni sull'adattamento dinamico, e anche ad Andrea per la sua saggezza. Grazie a Gianni per tutto quello che mi hai insegnato sulla Fokker-Planck e il suo spettro, per me tu sei il gigante sulle cui spalle mi ergo.
Grazie ad Elena per la sua Gentilezza, per i saggi consigli, la guida e per avermi aiutato a navigare questi anni di dottorato in acque sicure.
Grazie a Maurizio, grazie per la libertà e la fiducia che mi hai dato, il tuo prezioso tempo e le lunghe discussioni sempre stimolanti. Mai avrei detto di trovare un maestro a 27 anni e dopo due magistrali... mi ricorderò per sempre la prima visita in ISS con le mascherine, gli autovalori sul monitor e tu e Gianni che mi spiegavate come degli operatori possono descrivere la dinamica neuronale. Ho capito subito di aver trovato la mia casa.

Luca Falorsi, Roma, 24 luglio 2026