A Resolvent Framework for Global and Local Nonlinear Semigroups

Patrick S. Mahon

October 5, 2024

Contents

Symbols

X	Banach space $(X, \|\| \cdot \|\|)$
$w(x, x_t, \theta) : X \times X \times \mathbb{R} \to \mathbb{R}$	Weighting kernel at x_t, parameterized by $\theta \in \mathbb{R}, x_t \in X$

1 Definitions

Definition 1. *X_t is a matrix of a sequence of elements, $x \in X$, indexed by $t \geq 0$*

$$X_t = \begin{bmatrix} x_t \\ x_{t-h} \\ x_{t-2h} \\ \vdots \end{bmatrix}$$

Definition 2. *The t-weighting matrix of X_i is given by*

$$W(X_i, x_t, \theta) = \begin{bmatrix} w(x_j, x_t, \theta) & 0 & 0 & \cdots \\ 0 & w(x_{j-h}, x_t, \theta) & 0 & \cdots \\ 0 & 0 & w(x_{j-2h}, x_t, \theta) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$W_t^\theta X_i$ is short hand for the the product $W(X_i, x_t, \theta)X_i$

$$W_t^\theta X_i := W(X_i, x_t, \theta) \cdot X_i = \begin{bmatrix} w(x_i, x_t, \theta) \cdot x_i \\ w(x_{i-h}, x_t, \theta) \cdot x_{i-h} \\ w(x_{i-2h}, x_t, \theta) \cdot x_{i-2h} \\ \vdots \end{bmatrix}$$

2 Introduction

2.1 C as a globally linear map

Consider the map

$$C(\cdot) : X \to X$$

C may describe a time evolving map

$$C(\cdot) : x_t \mapsto x_{t+h} \qquad\qquad h \geq 0,\ x_i \in X$$

and may be state dependent, $C(x)$, non-autonomous, $C(t)$, or both, $C(t, x)$. If, for the autonomous state-independent case

$$x_{t+h} = Cx_t$$

C is globally linear, then,

$$X_{t+h} = X_t C$$
$$C = X_t^{-1} X_{t+h}$$

provides the solution.

In an applied time series setting, this corresponds to an auto-regressive (AR) model for the process X_t, where future states depend linearly on past states through the operator C. The AR model is a discrete-time example of how a system evolves over time.

2.2 C_t^θ as a locally linear map

If C is not globally linear, it must exhibit either non-autonomous behavior or state dependence. In many empirical settings, implicit non-autonomy arises naturally through state dependence, where $C(x(t))$ depends on the system's evolving state, effectively introducing time dependence via the state trajectory.

To try to account for this dependence we can introduce the t-weighting to the AR solution,

$$\begin{aligned} W_t^\theta X_{t+h} &= W_t^\theta X_t C_t^\theta \\ C_t^\theta &= (W_t^\theta X_t)^{-1} W_t^\theta X_{t+h} \end{aligned}$$

To capture local linearity, the weighting kernel $w(x_i, x_t, \theta)$, parameterized by θ, emphasizes states close to x_t. As $\theta \to \infty$, the local weighting becomes exact, with:

$$\lim_{\theta \to \infty} w(x_i, x_t, \theta) = \delta(x_i - x_t),$$

where the Dirac delta function selects only states arbitrarily close to x_t. This applies to each $x_i \in X_t$, localizing the influence of nearby states.

Typically the kernel is chosen so $w \sim e^{-\theta}$. Taking the limit, W_t^θ reduces to,

$$\lim_{\theta \to \infty} W_t^\theta = \begin{cases} \delta(x_i - x_t) = 1 & \text{if } i = j \wedge x_i = x_t \\ 0 & \text{if } i \neq j \vee x_i \neq x_t \end{cases}$$

which is a diagonal matrix whose only non-zero entries are 1's corresponding to states arbitrarily close to x_t. If X_t never returns to states arbitrarily close to x_t then W_t^θ reduces to the rank-1 matrix with a single non-zero entry,

$$(W_t^\theta)_{1,1} = 1$$

If X_t exhibits periodicity with frequency k, then only every k-th diagonal element of W_t^θ remains non-zero, reflecting the periodic return to the same state:

$$(W_t^\theta)_{i,j} = \begin{cases} 1 & \text{if } i = j = t + nk,\ n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

For ergodic processes, the rank of W_t^θ reflects the frequency and nature of returns to neighborhoods around x_t, with proximity defined by:

$$U = \{x \in X, \delta \in \mathbb{R} : w(x, x_t, \theta) < \delta\}$$

This provides a natural way to define ”closeness” in ergodic settings.

3 One Parameter Semigroups

Semigroup theory provides a powerful framework for understanding the evolution of systems, especially when dealing with local linearity and the influence of higher-order terms. By introducing one-parameter semigroups, we gain tools to connect the operators that generate the semigroup with the semigroup's overall dynamics. This helps us explore the relationship between locality (controlled by θ) and the time evolution of the elements in X_t.

3.1 $C(h)$ as a one parameter semigroup

We define $C(h)$

$$C(h) : \mathbb{R}_{\geq 0} \times X \to X$$

$\{C(h)\}_{h\geq 0}$ forms a one parameter semigroup if the the following requirements are met:

Associativity:

$$\forall h, s \geq 0 : C(h+s) = C(h)C(s) \tag{1}$$

Identity:

$$C(0) = I \tag{2}$$

In the context of global and local linear maps, the parameter t referred to specific time points in the system's evolution, and h represented the time step or interval over which the system evolved.

In this context, h describes the continuous evolution of the operator family $\{C(h)\}_{h\geq 0}$, functioning similarly to the time step h in the discrete-time models. However, here h represents the time duration over which the system evolves, extending the discrete case into a continuous framework.

3.1.1 C_0-semigroup

Continuity for one-parameter semigroups is defined by the following properties:

Uniform Continuity: A one-parameter semigroup $\{C(h)\}_{h\geq 0}$ is uniformly continuous if:

$$\lim_{h\to 0^+} ||C(h)x - x|| = 0, \quad \forall x \in X \tag{3}$$

Uniform continuity implies that the operator family $C(h)$ behaves "smoothly" as $h \to 0^+$, ensuring that the system's evolution is well-behaved and bounded.

The linear operator A is the infinitesimal generator of a uniformly continuous semigroup if and only if A is a *bounded linear operator*. In this case, the infinitesimal generator A of the semigroup $C(h)$ must satisfy:

$$C(h) = e^{Ah} := \sum_{k=0}^{\infty} \frac{A^k}{k!} h^k \tag{4}$$

This exponential form provides a strong connection to the classical matrix exponential, commonly seen in linear systems.

Strong Continuity: If the semigroup is not uniformly continuous but still exhibits strong continuity, it is called a C_0*-semigroup*. This requires that for any $x \in X$:

$$\lim_{h\to 0^+} ||C(h)x - x|| = 0, \quad \forall x \in X \tag{5}$$

This property ensures that the semigroup $C(h)$ is *strongly continuous*, even if the system's evolution is not uniformly bounded.

For a strongly continuous semigroup, we define the *infinitesimal generator* A, acting on the domain $D(A)$, as:

$$Ax = \lim_{h \to 0^+} \frac{1}{h}(C(h) - I)x \tag{6}$$

The limit must exist for $x \in D(A)$, but unlike the uniformly continuous case, the generator A need not be bounded.

In this more general case, the semigroup $C(h)$ is still governed by the operator A, but the series e^{Ah} need not converge if A is unbounded.

The C_0-semigroup provides a strong foundation for understanding time evolution in terms of linear operators, but it does not fully capture the role of locality or the higher-order terms in systems like X_t. To extend this, we will need to explore additional terms and relationships in the generator.

3.2 C_n-semigroup

The C_0-semigroup can be viewed as a special case of a more general continuous one-parameter semigroup, the nonlinear C_n-semigroup. This semigroup's infinitesimal generator expresses higher-order effects through the limit of ever smaller finite differences.

We guide and justify the definition of such a semigroup based on the fact that the exponential of the infinitesimal generator yields a well-known generalization of the Taylor series—using the infinitesimal limit of finite differences—known as the Hille series.

The key advantage of the C_n-semigroup and its infinitesimal generator is that they carry intuition regarding continuity and nonlinearity into a more pragmatic operator-theoretic framework. This makes the C_n-semigroup particularly well-suited for investigating local dynamics and the role of locality in nonlinear systems.

3.2.1 The Hille Operator for Nonlinear Semigroups

We can relate the forward differences to the semigroup where:

$$\Delta_h^1 x(t) = x(t+h) - x(t) = C_1(h)x(t) - x(t).$$

For higher-order differences, the difference operator series is:

$$\begin{aligned} \Delta_h^n x(t) &= \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} x(t+kh) \\ &= \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} C_k(h)x(t) \end{aligned} \tag{7}$$

If we define the n-order generator of a strongly continuous nonlinear semigroup:

$$A^n x = \lim_{h \to 0^+} \frac{1}{h^n}(C_n(h) - I)x,$$

then the higher-order generators, $\{A^n\}_{n\geq 2}$, describe higher-order terms of the semigroup, capturing corrections to the first-order dynamics A^1.

We can express local continuity of the nonlinear C_n-semigroup as the matrix exponential of the n-order generators,

$$\begin{aligned} e^{Ah} &:= \sum_{k=0}^{n} \frac{h^k}{k!} A^k \\ &:= \lim_{h\to 0^+} \sum_{k=0}^{n} \frac{h^k}{k!} \frac{C^n(h) - I}{h^n} \end{aligned}$$

If the C_n-semigroup possesses uniform continuity for $t \in \mathbb{R}_{\geq 0}$, then,

$$C_n(t) = e^{(e^{Ah})t} := \sum_{m=0}^{\infty} \frac{t^m}{m!} \left(\sum_{k=0}^{n} \frac{h^k}{k!} A^k \right)^m$$

where t is the global parameter governing the semigroup's evolution in time and the inner sum $\sum_{k=0}^{n} \frac{h^k}{k!} A^k$ represents the local action up to n-orders over an infinitesimal h.

> **Aside:** In a way the recursive operation of the exponential extends the semigroup across different measures. The first application, e^{Ah}, yields the local semigroup action at a set of measure-0. The second, $e^{(e^{Ah})t}$ yields the action across a trajectory set of measure-1, We could apply once more for a manifold of measure-2,
>
> $$\exp(A, h, 3) = e^{(e^{(e^{Ah})t_1})t_2}$$
>
> I Should probably explore this elsewhere though as a formalization of multiparameter semigroups and investigate connections to concepts in other fields like the exponential map in Riemannian geometry, which takes a tangent vector (local direction) and maps it to a geodesic (global trajectory). This could provide a nice supplemental perspective when investigating locality and curvature across dimensions.

Returning to the inner sum, in the limit

$$\begin{aligned} \lim_{h\to 0^+} C(h) &= \lim_{h\to 0^+} \sum_{k=0}^{n} \frac{h^k}{k!} A^k \\ H &= \lim_{h\to 0^+} \sum_{k=0}^{n} \frac{h^k}{k!} \frac{C^n(h) - I}{h^n} \end{aligned}$$

we call H the *Hille operator*. Upon application, $Hx(t)$, this operator yields a special case of the

Hille series:

$$
\begin{aligned}
Hx(t) &= \lim_{h\to 0^+} \sum_{k=0}^{n} \frac{h^k}{k!}\frac{C^k(h) - I}{h^k} x(t) \\
&= \lim_{h\to 0^+} \sum_{k=0}^{n} \frac{h^k}{k!}\frac{C^k(h)x(t) - x(t)}{h^k} \\
&= \lim_{h\to 0^+} \sum_{k=0}^{n} \frac{h^k}{k!}\frac{x(t+kh) - x(t)}{h^k} \\
x(t+h) &= \lim_{h\to 0^+} \sum_{k=0}^{\infty} \frac{\Delta_h^k x(t)}{k!}
\end{aligned}
$$

Again, the notation can be a little confusing here. In our formulation we identify h as both the span over which we apply the operator and the higher-order infinitesimals. This offers the simplification of the time dependent and infinitesimal terms to yield the local expansion at t. Whereas in the more general Hille series,

$$x(a+t) = \lim_{h\to 0^+} \sum_{k=0}^{\infty} \frac{t^k}{k!}\frac{\Delta_h^k x(a)}{h^k}$$

t is the independent time parameter and h is the infinitesimal step size. The simplification merely arises from the fact that, for a one parameter semigroup, $a = t \in \mathbb{R}_{\geq 0}$

The result is H provides a way to generalize the semigroup dynamics to include higher-order effects via the Hille like expansion, defining the locally continuous action of a one-parameter nonlinear semigroup. Uniform continuity is obtained upon recursively applying the matrix exponential yielding the general class of C_n-semigroups. Where n is not finite, we have the strongly nonlinear C_∞-semigroup.

3.3 The resolvent of C_0^n

The resolvent

$$R(\lambda, H(h)) = (\lambda I - H(h))^{-1}$$

is important because it tells us whether the Hille operator $H(h)$ has an inverse for certain values of λ. This can help define the spectrum and growth behavior of the C_0^n semigroup.

We are interested in the resolvent set that provides a well defined inverse,

$$\rho(H) = \left\{ n \in \mathbb{Z}, \lambda \in \mathbb{C} : \lim_{h\to 0^+} (\lambda I - H(h))^{-1} \right\}$$

We have

$$
\begin{aligned}
R(\lambda, H(t)) &= \lim_{h\to 0^+} (\lambda I - H(t))^{-1} \\
&= \lim_{h\to 0^+} \left(\lambda I - \sum_{n} \frac{t^n}{n!}\frac{A_h^n}{h^n} \right)^{-1}
\end{aligned}
$$

so,

$$0 \neq \lim_{h \to 0^+} |\lambda I - H(t)|$$

We want $\lambda \in \rho(H(t))$ where

$$\begin{aligned}\rho(H(t)) &= \mathbb{C}/\sigma\,(H(t)) \\ &= \mathbb{C}/\sigma \left(\lim_{h \to 0^+} \sum_n \frac{t^n}{n!} \frac{A_h^n}{h^n} \right)\end{aligned}$$

There are a number of options to resolve $\sigma(H(t))$ in non-empirical settings. These include, among others, methods from functional calculus, operator, theory, and perturbation theory.

4 An Empirical Hille Operator

In an empirical setting the coefficients, a_n, of the operator,

$$H(t) = \lim_{h \to 0^+} \sum_n a^n \frac{\Delta_h^n}{h^n}$$

may not ensure convergence. In such cases we have to consider how to treat the distribution of coefficients,

$$P(a_n) \sim f(n)$$

assuming it is some function of the order n. In this case the only required constraint is normalization

$$\sum_n P(a_n) = 1$$

Beyond this many natural systems may or may not exhibit further constraints including bounded higher order moments. How should we, *a priori*, evaluate methods to estimate $P(a_n)$

If we seek the least biased estimate of the dynamics described by $H(t)$ then we should adopt a maximum entropy approach. Following such a solution we may assess the spectra of the ensemble, $\sigma(< H(t)x(t) >)$, where

$$< H(t)x(t) >= \sum_n P(a_n)\, a_n \frac{\Delta_h^n x(t)}{h^n}$$

4.1 Maximum Entropy Estimates of $H(t)$

Given a set of trajectory data, $\{X_t\}_t \geq 0$, we propose the following application of the empirical Hille operator,

$$\begin{aligned} H(t)x(t) &= \sum_n P(a_n)\, a_n \frac{\Delta_h^n x(t)}{h^n} \\ &= \sum_n P(a_n)\, a_n \frac{\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} C_h^k(t)}{h^n} \\ &= \sum_n P(a_n)\, a_n \frac{\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} X_t^{-1} X_{t+kh}}{h^n} \\ &= \sum_n X_t^{-1} P(a_n)\, a_n \frac{\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} X_{t+kh}}{h^n} \end{aligned}$$

The entropy of $P(a_n)$ is given by

$$S(P) = -\sum_n P(a_n) \log P(a_n)$$

where,

$$\sum_n P(a_n) = 1$$

We seek the solution to the Lagrangian,

$$\mathcal{L} = -\sum_n P(a_n) \log P(a_n) + \lambda \left(\sum_n P(a_n) - 1 \right)$$