Report No.
UCB/SEMM-[2026/02]

Structural Engineering
Mechanics and Materials

Simplified Notes on Coarse-graining, Renormalization Group, and Asymptotic Distributions

By

Ziqi Wang

[September 2026]

Department of Civil and Environmental Engineering
University of California, Berkeley

Simplified Notes on Coarse-graining, Renormalization Group, and Asymptotic Distributions

Ziqi Wang[a,*]

[a]*Department of Civil and Environmental Engineering, University of California, Berkeley, United States*

Abstract

These pedagogical notes introduce the core concepts and formulation of coarse-graining and the renormalization group. Section 1 outlines the foundational idea, while Sections 2 and 3 apply the renormalization group ansatz to derive Gaussian and extreme value distributions. By showing how a single ansatz unifies these classical limit laws, we hope to encourage readers to extend this approach to complex multiscale systems beyond the reach of standard asymptotic distributions.

Keywords: coarse-graining, renormalization group, asymptotic distributions

1. General Intuitions and Formulations

Consider a system characterized by a scalar random field $X : \Omega \times \mathcal{Z} \to \mathbb{R}$, where Ω is the sample space and $\mathcal{Z}$ is an index set. We define a *coarse-graining* scheme to represent the system at varying levels of granularity. Starting with the fine-scale field $X_0 \equiv X$, we proceed to a sequence of coarse-grained fields $X_s : \Omega \times \mathcal{Z}_s \to \mathbb{R}$, where the index set $\mathcal{Z}_s$ maintains an invariant cardinality across all scaling steps $s \in \mathbb{N}$. Let $\mathcal{M}_s$ denote the model governing X_s. We aim to study how the model $\mathcal{M}_s$ evolves as the observation scale grows. This evolution in the space of models is referred to as the *Renormalization Group (RG) flow*.

- **Remark 1:** A key requirement of RG theory is that the models $\mathcal{M}_s$ must remain comparable across different scales s. This often necessitates spatial coordinate rescaling or continuous field formulations. For instance, in a finite, discrete spatial field where coarse-graining is defined as an average over an expanding window, the raw cardinality of $\mathcal{Z}_s$ decreases with s, eventually collapsing to a single point (the global spatial mean). To maintain a consistent model space, we may reformulate X as a continuous spatial field (or a countably infinite field), such that coarse-graining reduces the effective resolution rather than the domain itself. This allows $\mathcal{M}_s$ to be defined on a fixed mathematical setting for all s; then, the scale transformation can be described by a coherent evolution of field theories.

*Corresponding author
Email address: ziqiwang@berkeley.edu (Ziqi Wang)

- **Remark 2:** The model $\mathcal{M}_s$ can take various forms, including probability distributions, characteristic functions, spectral representations, or physics-based/data-driven generative models that reproduce the statistics of X_s.

- **Remark 3:** A central feature of RG flows is the existence of fixed points. These represent universal multiscale behaviors where the system becomes statistically scale-invariant under further coarse-graining. This dynamical flow picture generalizes classical asymptotic limit theorems to complex multiscale systems.

- **Remark 4:** In the language of dynamical systems, the behavior of an RG flow near a fixed point is governed by the linearized stability of the flow operator:

  - *Relevant parameters* correspond to unstable eigenvalues (expanding directions). These grow along the flow and dictate the macroscopic phase of the system.
  - *Irrelevant parameters* correspond to stable eigenvalues (contracting directions). These transients decay to zero as the scale increases.
  - *Marginal parameters* correspond to unit eigenvalues, typically requiring higher-order nonlinear terms to determine their stability.

- **Remark 5:** RG offers a universal geometric language to describe *phases* and *phase transitions*. The space of all possible models for a given phenomenon is partitioned into distinct basins of attraction under the RG flow. A macroscopic *phase* (such as a liquid, a magnet, or a Gaussian limit regime) corresponds to a basin governed by a *stable fixed point* (or sink). Separating the basins are critical surfaces (phase boundaries) that contain *unstable fixed points* (or saddle points). The relevant eigen-directions at these critical fixed points are orthogonal to the phase boundary and push perturbations away from criticality.

We represent an RG flow by

$$\mathcal{M}_{s+1} = \mathcal{R}\left(\mathcal{M}_s\right) , \tag{1}$$

where $\mathcal{R}$ is the RG operator transforming the model from scale s to $s+1$. Some literature explicitly separates coarse-graining from a subsequent coordinate dilation and rescaling step. This rescaling may preserve physical dimensions and prevent parameter blow-up. Here, all such operations are absorbed into $\mathcal{R}$.

A fixed point under the RG flow, denoted $\mathcal{M}^\star$, satisfies the invariance condition

$$\mathcal{M}^\star = \mathcal{R}\left(\mathcal{M}^\star\right) . \tag{2}$$

To examine local stability, we perturb the fixed point by a small increment $\delta\mathcal{M}_s$. Expanding Eq. (1) around $\mathcal{M}^\star$ gives:

$$\mathcal{M}_{s+1} = \mathcal{R}(\mathcal{M}^\star + \delta\mathcal{M}_s) = \mathcal{M}^\star + D\mathcal{R}_{\mathcal{M}^\star}(\delta\mathcal{M}_s) + \mathcal{O}(\|\delta\mathcal{M}_s\|^2) ,$$

where $D\mathcal{R}_{\mathcal{M}^\star}$ is the Fréchet derivative (Appendix A) of $\mathcal{R}$ at $\mathcal{M}^\star$ defined on a suitable Banach space. Dropping second-order terms yields the linearized iteration:

$$\delta\mathcal{M}_{s+1} \approx D\mathcal{R}_{\mathcal{M}^\star}(\delta\mathcal{M}_s)\,. \tag{3}$$

Stability is determined by the spectral properties of $D\mathcal{R}_{\mathcal{M}^\star}$. We consider the eigen-decomposition

$$D\mathcal{R}_{\mathcal{M}^\star}(\phi_i) = \lambda_i\,\phi_i\,, \tag{4}$$

where ϕ_i are the eigenfunctions (directions in model space) and λ_i are the associated eigenvalues. The magnitude of λ_i classifies the behavior along eigen-directions:

- **Relevant operators ($|\lambda_i| > 1$):** Perturbations grow exponentially with s. These define unstable directions that drive the system away from the fixed point.
- **Irrelevant operators ($|\lambda_i| < 1$):** Perturbations decay to zero. The microscopic details along these directions vanish at large scales.
- **Marginal operators ($|\lambda_i| = 1$):** The linear approximation is inconclusive; or it is truly marginal (for toy models).

Finally, we note two classical theorems on the existence and uniqueness of fixed points:

- **Banach fixed-point theorem:** If $\mathcal{R}$ is a contraction mapping on a Banach space, there exists a unique, globally stable fixed point $\mathcal{M}^\star$, and the flow converges to it exponentially.
- **Schauder fixed-point theorem:** If $\mathcal{R}$ continuously maps a compact, convex subset K into itself, there exists at least one fixed point. This guarantees the existence of fixed points even when $\mathcal{R}$ is not a contraction mapping.

2. Gaussian Fixed Points and Beyond

Consider a random field

$$X_0 : \Omega \times \mathbb{N} \to \mathbb{R}$$

with $\mathbb{E}[X_0(j)] = 0$ and $\mathbb{V}ar\,[X_0(j)] = 1$ for all $j \in \mathbb{N}$*. The law of $X_0(j)$ may vary with j; the correlation structure is discussed later. The countably infinite index set makes this setting convenient for illustrating the RG viewpoint on the Central Limit Theorem (CLT).

Define the coarse-graining map by

$$X_s(j) = \frac{X_{s-1}(2j-1) + X_{s-1}(2j)}{\sqrt{2}}, \qquad j \in \mathbb{N}\,,$$

*The zero mean and unit variance assumption is not restrictive: other finite-mean and finite-variance cases can be obtained by linear transformations that preserve the distribution.

for scaling step $s \in \mathbb{N}$. That is, we sequentially aggregate neighboring pairs and normalize so that the variance stays at order one.

Let the law of $X_s(j)$ be represented by its characteristic function

$$\varphi_s(j,t) = \mathbb{E}\left[e^{\mathrm{i}tX_s(j)}\right].$$

The coarse-graining relation suggests

$$\varphi_s(j,t) = \mathbb{E}\left[\exp\Big(\mathrm{i}t\frac{X_{s-1}(2j-1)+X_{s-1}(2j)}{\sqrt{2}}\Big)\right].$$

2.1. Independent

Assume first the (simplest) independent case: beyond some scale the neighboring variables are independent. Then

$$\varphi_{s+1}(j,t) = \varphi_s\Big(2j-1,\frac{t}{\sqrt{2}}\Big)\,\varphi_s\Big(2j,\frac{t}{\sqrt{2}}\Big).$$

A fixed point $\varphi^\star$ satisfies

$$\varphi^\star(t) = \varphi^\star\Big(\frac{t}{\sqrt{2}}\Big)^2.$$

Note that we have dropped the dependence on j. Even if the initial laws $\varphi_0(j,t)$ vary with j, the coarse-graining map mixes neighboring indices and progressively washes out inhomogeneities. Consequently, if the RG flow converges, the fixed-point must be independent of j. The continuous solution is the Gaussian characteristic function (Appendix B)

$$\varphi^\star(t) = \exp\Big(-\frac{t^2}{2}\Big).$$

To reveal the relevant directions of the Gaussian fixed point, consider the log-characteristic function (the cumulant generating function)

$$K_s(t) \coloneqq \log\varphi_s(t) = \sum_{m=1}^{\infty}\frac{(\mathrm{i}t)^m}{m!}\,\kappa_m^{(s)},$$

where $\kappa_m^{(s)}$ is the m-th cumulant at scale s. For sufficiently large scales the RG map satisfies

$$K_{s+1}(t) = \mathcal{R}(K_s(t)) = 2\,K_s\Big(\frac{t}{\sqrt{2}}\Big),$$

and matching the coefficients of the power series yields

$$\kappa_m^{(s+1)} = 2^{\,1-m/2}\,\kappa_m^{(s)}.$$

Thus all cumulants with $m > 2$ decay exponentially and are therefore *irrelevant*. To make the RG structure more explicit, consider a perturbation of the cumulant generating function around the Gaussian fixed point,

$$K_s(t) = K^\star(t) + \delta K_s(t), \qquad K^\star(t) = -\frac{1}{2}\,t^2.$$

Applying the RG operator

$$K_{s+1}(t) = \mathcal{R}(K_s(t)) \approx K^\star(t) + 2\delta K_s\Big(\frac{t}{\sqrt{2}}\Big)$$

yields the linearized RG recursion:

$$\delta K_{s+1}(t) \approx D\mathcal{R}_{K^\star}(\delta K_s(t)) = 2\,\delta K_s\Big(\frac{t}{\sqrt{2}}\Big).$$

$D\mathcal{R}_{K^\star}$ is a dilation operator acting on functions of t. To analyze its eigenstructure, adopt the monomial basis

$$\phi_m(t) = t^m, \qquad m = 1, 2, \ldots$$

which is natural because $K_s(t)$ is conventionally expanded in this basis. The eigen-equation reads

$$D\mathcal{R}_{K^\star}(\phi_m(t)) = 2\,\phi_m\Big(\frac{t}{\sqrt{2}}\Big) = 2^{1-m/2}\,t^m = \lambda_m\,\phi_m(t),$$

so each $\phi_m(t) = t^m$ is an eigen-direction with eigenvalue

$$\lambda_m = 2^{1-m/2}.$$

This yields the RG classification:

- *Relevant:* $|\lambda_m| > 1$, occurring only for $m = 1$. This corresponds to the mean.
- *Marginal:* $|\lambda_m| = 1$, occurring only for $m = 2$. This is associated with the variance.
- *Irrelevant:* $|\lambda_m| < 1$, holding for all $m \geq 3$. These correspond to higher-order cumulants, which decay exponentially under coarse-graining.

Perturbing the mean shifts the expectation of the coarse-grained variables, producing a macroscopic effect; hence, $m = 1$ is relevant. However, enforcing a zero-mean field (or re-centering at each scale) eliminates this relevant operator. The variance direction ($m = 2$) is marginal because the $1/\sqrt{2}$ normalization preserves variance under iteration. Higher-order cumulants ($m \geq 3$) are irrelevant, as they decay exponentially under coarse-graining.

2.2. Local correlation

Next, we consider the correlated case, where $X_0(j)$ may exhibit dependence across different indices. We first focus on *local correlations*, assuming the field has a finite correlation length* $r \in \mathbb{N}$. Concretely, $X_0(j)$ and $X_0(j+r+1)$ are independent for any j, as their separation distance exceeds r. Under this assumption, we claim that sequential coarse-graining causes the block variables being increasingly independent, so that the CLT–RG analysis developed for independent variables eventually applies. This claim is intuitively expected, but we elaborate its structure below.

*In statistical physics, the term *correlation length* is commonly used even when the underlying notion concerns general statistical dependence rather than linear correlation. We adopt this convention here.

Consider two contiguous blocks $A = \{1, \ldots, N\}$ and $B = \{N+1, \ldots, 2N\}$, and their normalized averages

$$X_A = \frac{1}{\sqrt{N}} \sum_{j \in A} X_j, \qquad X_B = \frac{1}{\sqrt{N}} \sum_{k \in B} X_k.$$

Their covariance is

$$\mathbb{C}ov\left[X_A, X_B\right] = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=N+1}^{2N} \mathbb{C}ov\left[X_j, X_k\right].$$

If the random field has finite correlation length r, then only pairs (j,k) with $|j-k| \leq r$ contribute to the double sum. Since each of the r sites near the boundary of A can correlate with at most r sites in B, the number of contributing pairs is $\mathcal{O}(r^2)$*. Hence

$$\left|\mathbb{C}ov\left[X_A, X_B\right]\right| = \mathcal{O}\left(\frac{r^2}{N}\right).$$

Because $\mathbb{V}ar\left[X_A\right]$ and $\mathbb{V}ar\left[X_B\right]$ are $\mathcal{O}(r)$, the correlation coefficient satisfies

$$\left|\mathrm{Corr}(X_A, X_B)\right| = \frac{\left|\mathbb{C}ov\left[X_A, X_B\right]\right|}{\sqrt{\mathbb{V}ar\left[X_A\right]\ \mathbb{V}ar\left[X_B\right]}} = \mathcal{O}\left(\frac{r}{N}\right).$$

Thus, the correlation between the block variables decays as r/N; sequential coarse-graining therefore progressively washes out correlation. A similar analysis can be applied to higher-order moments, such as $\mathbb{E}\left[X_A^p X_B^q\right]$ for $p, q \in \mathbb{N}$, from which we can conclude that coarse-graining progressively washes out dependence.

2.3. Non-local correlation

We now consider a random field $X_0(j)$ with *long-range correlations*, for example

$$\mathbb{C}ov\left[X_0(j), X_0(k)\right] \sim |j-k|^{-\alpha}, \qquad j \neq k,$$

for some $\alpha > 0$. Define contiguous block variables as before. Then the covariance between the blocks satisfies

$$\mathbb{C}ov\left[X_A, X_B\right] = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \mathbb{C}ov\left[X_j, X_{N+k}\right] = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{N} |N+k-j|^{-\alpha}.$$

Introduce the offset $m = |N+k-j|$. For each $m = 1, \ldots, 2N-1$, the number of pairs (j,k) giving the same m is

$$\#\mathrm{pairs}(m) = N - |N-m|.$$

Hence we can rewrite the sum as

$$\mathbb{C}ov\left[X_A, X_B\right] = \frac{1}{N} \sum_{m=1}^{2N-1} \left(N - |N-m|\right) m^{-\alpha}.$$

*The precise count is $r(r+1)/2$.

Similarly, for block A (block B will be the same), we have

$$\mathbb{V}ar\left[X_A\right] = \frac{1}{N}\sum_{j,k=1}^{N}\mathbb{C}ov\left[X_j\,,X_k\right] = \frac{1}{N}\sum_{j=1}^{N}\mathbb{V}ar\left[X_j\right] + \frac{1}{N}\sum_{j\neq k}\mathbb{C}ov\left[X_j\,,X_k\right]\,.$$

Using $\mathbb{V}ar\left[X_j\right] = 1$ and $\mathbb{C}ov\left[X_j\,,X_k\right] = |j-k|^{-\alpha}$ for $j \neq k$, we get

$$\mathbb{V}ar\left[X_A\right] = 1 + \frac{2}{N}\sum_{m=1}^{N-1}(N-m)m^{-\alpha},$$

where for each $m = 1,\ldots,N-1$, there are $2(N-m)$ off-diagonal pairs with $|j-k| = m$.

We consider the large-N behavior. Define:

$$S_1(N;\alpha) = \sum_{m=1}^{N-1} m^{-\alpha}, \qquad S_2(N;\alpha) = \sum_{m=1}^{N-1} m^{1-\alpha}.$$

Then

$$\mathbb{V}ar\left[X_A\right] = 1 + 2S_1(N;\alpha) - \frac{2}{N}S_2(N;\alpha).$$

For large N, we have

$$S_1(N;\alpha) \sim \begin{cases} \dfrac{N^{1-\alpha}}{1-\alpha}\,, & 0<\alpha<1\,, \\ \log N\,, & \alpha = 1\,, \\ \zeta(\alpha), & \alpha > 1\,. \end{cases} \qquad S_2(N;\alpha) \sim \begin{cases} \dfrac{N^{2-\alpha}}{2-\alpha}\,, & 0<\alpha<2\,, \\ \log N\,, & \alpha = 2\,, \\ \zeta(\alpha-1)\,, & \alpha > 2\,, \end{cases}$$

Consequently

$$\mathbb{V}ar\left[X_A\right] \sim \begin{cases} 1 + \dfrac{2}{(1-\alpha)(2-\alpha)}\,N^{1-\alpha}, & 0<\alpha<1\,, \\ -1 + 2\log N, & \alpha = 1\,, \\ 1 + 2\zeta(\alpha), & \alpha > 1\,. \end{cases}$$

The S_2 term is negligible when $\alpha > 1$. It is seen that the variance is order-one when $\alpha > 1$; it grows logarithmically at $\alpha = 1$, and grows algebraically for $0 < \alpha < 1$.

To analyze the large-N behavior of the covariance term, notice that

$$N - |N-m| = \begin{cases} m, & 1 \le m \le N, \\ 2N-m, & N < m \le 2N-1. \end{cases}$$

We split the covariance sum at $m = N$ and obtain

$$\mathbb{C}ov\left[X_A\,,X_B\right] = \frac{1}{N}\sum_{m=1}^{N} m^{1-\alpha} + \frac{1}{N}\sum_{m=N+1}^{2N-1}(2N-m)m^{-\alpha}\,.$$

For $1 \leq m \leq N$, let $m \approx pN$ with $p \in (0,1)$. For $N < m \leq 2N-1$, let $m \approx N + pN$, again with $p \in (0,1)$. The integral approximation of the sum is

$$\mathbb{C}ov\left[X_A, X_B\right] \sim N^{1-\alpha}\left[\int_0^1 p^{1-\alpha}\, dp + \int_0^1 (1-p)\,(1+p)^{-\alpha}\, dp\right], \qquad N \to \infty.$$

The behavior separates into the regimes:

$$\mathbb{C}ov\left[X_A, X_B\right] \sim \begin{cases} c(\alpha)N^{1-\alpha} \to \infty & 0 < \alpha < 1, \\ \dfrac{\log N}{N} \to 0, & \alpha = 1, \\ c(\alpha)N^{1-\alpha} \to 0, & \alpha > 1. \end{cases}$$

The $\alpha = 1$ case is analyzed using the exact sum instead of the integral approximation. This result provides the leading–order scaling of the correlation coefficient:

$$|\mathrm{Corr}[X_A, X_B]| = \begin{cases} \mathcal{O}(1) & 0 < \alpha < 1, \\ \mathcal{O}\left(\dfrac{1}{N}\right), & \alpha = 1, \\ \mathcal{O}\left(N^{1-\alpha}\right), & \alpha > 1. \end{cases}$$

We conclude that:

- $\alpha \geq 1$: block correlations decay algebraically (power-law) to zero; coarse-graining slowly suppresses dependence, and the Gaussian fixed point is stable.
- $0 < \alpha < 1$: block correlations approach a nonzero constant, so long-range dependence persists under coarse-graining. Non-Gaussian fixed points can emerge.

2.4. Lévy fixed point

In deriving the Gaussian fixed point, we relied on expanding the characteristic function (and its logarithm) to at least second-order. A critical limitation of this derivation is that the second-order expansion does not exist if the variance is infinite. This scenario is common in distributions with power-law tails. To address this issue, we generalize the coarse-graining map into:

$$X_s(j) = \frac{X_{s-1}(2j-1) + X_{s-1}(2j)}{b}, \qquad j \in \mathbb{N}$$

where $b > 0$ is selected to match the scale of the numerator. The fixed point equation for the characteristic function becomes

$$\varphi^\star(t) = \varphi^\star\left(\frac{t}{b}\right)^2.$$

To achieve scale invariance, it is found (Appendix C) that b is:

$$b = 2^{1/\alpha},$$

where $\alpha \in (0, 2]$ is a characteristic exponent of the fixed-point distribution. Assuming a symmetric distribution, the fixed-point is the Lévy stable form:

$$\varphi^\star(t) = \exp(-c|t|^\alpha),$$

with $c > 0$ a parameter. The Gaussian fixed point is recovered for $\alpha = 2$.

The linearized RG recursion around the Lévy fixed point is given by:

$$\delta K_{s+1}(t) = D\mathcal{R}_{K^\star}(\delta K_s(t)) = 2\,\delta K_s\Big(\frac{t}{2^{1/\alpha}}\Big)\,.$$

Here, $|t|^\alpha$ is an eigenfunction with an eigenvalue of 1, i.e., a marginal operator that is preserved under coarse-graining.

3. Extreme Value Fixed Points

The Gaussian fixed point relied on summing neighboring states; we now consider coarse-graining via block maximization. Due to the identity $\min\{x_i\} = -\max\{-x_i\}$, we focus on the maximum.

Consider a random field $X_0 : \Omega \times \mathbb{N} \to \mathbb{R}$ with independent and identically distributed components. We define the coarse-graining map for block maxima as

$$X_s(j) = \frac{\max\{X_{s-1}(2j-1),\, X_{s-1}(2j)\} - c_s}{b_s}\,, \qquad j \in \mathbb{N}\,,$$

where $b_s > 0$ and $c_s \in \mathbb{R}$ are scale and location parameters chosen to prevent the distribution from degenerating or shifting to infinity under repeated maximization.

Let $F_s(x)$ denote the cumulative distribution function (CDF) at scale s. Coarse-graining yields

$$F_{s+1}(x) = F_s(b_s x + c_s)^2\,.$$

At a fixed point $s \to \infty$, $b = b_\infty$, and $c = c_\infty$, the invariant CDF $F^\star(x)$ satisfies

$$F^\star(x) = F^\star(bx + c)^2\,.$$

This equation shares the same structure of the Gaussian case (Section 2.1). Let $L_s(x) \coloneqq \log F_s(x)$, we have:

$$L_{s+1}(x) = \mathcal{R}(L_s(x)) = 2\,L_s(b_s x + c_s)\,,$$
$$L^\star(x) = 2\,L^\star(bx + c)\,.$$

Depending on b, the fixed-point equation admits three classes of solutions, corresponding to the Generalized Extreme Value theory (Appendix D):

- **Gumbel fixed point ($b = 1, c = \log 2$):** The condition $L^\star(x) = 2\,L^\star(x + \log 2)$ admits the exponential solution

$$L^\star(x) = -e^{-x} \implies F^\star(x) = \exp\left(-e^{-x}\right), \qquad x \in \mathbb{R}\,.$$

  This fixed point governs distributions with exponential tails (e.g., Gaussian, exponential).

- **Fréchet fixed point ($b = 2^{1/\alpha}, c = 0$ for $\alpha > 0$):** The condition $L^\star(x) = 2\,L^\star(2^{1/\alpha}x)$ yields the power-law solution

$$L^\star(x) = -x^{-\alpha} \implies F^\star(x) = \exp\left(-x^{-\alpha}\right), \qquad x > 0\,.$$

  This fixed point governs heavy-tailed distributions with power-law decay.

- **Weibull fixed point ($b = 2^{-1/\alpha}, c = 0$ for $\alpha > 0$):** The condition $L^\star(x) = 2\,L^\star(2^{-1/\alpha}x)$ yields the bounded-support solution

$$L^\star(x) = -(-x)^{\alpha} \implies F^\star(x) = \exp\left(-(-x)^{\alpha}\right), \qquad x < 0\,.$$

  This fixed point governs distributions with a finite upper bound.

Now we perform stability analysis of the linear operator

$$\delta L_{s+1}(x) \approx D\mathcal{R}_{L^\star}\left(\delta L_s(x)\right) = \delta L_s(bx + c)\,.$$

The eigenfunctions $\phi(x)$ and eigenvalues λ satisfy

$$D\mathcal{R}_{L^\star}\left(\phi(x)\right) = 2\phi(bx + c) = \lambda\phi(x)\,.$$

To solve this eigenvalue problem, we transform to the canonical coordinate y, which shifts the origin to the fixed point of the affine transformation $T(x) = bx + c$. For $b \neq 1$, $y = x - x_0$ where $x_0 = \frac{c}{1-b}$; for $b = 1$, $y = e^{-x}$. In coordinate y, the affine transformation simplifies to pure dilation $T(y) = by = 2^\gamma y$, where $\gamma \in \mathbb{R}$ is a shape parameter.

Adopting monomial eigenfunctions $\phi_p(y) = y^p$ for exponent $p \in \mathbb{R}$, Eq. (3) yields the eigenvalue spectrum:

$$D\mathcal{R}_{L^\star}\left(\phi_p(y)\right) = 2\left(2^\gamma y\right)^p = 2^{1+\gamma p}y^p \implies \lambda_p = 2^{1+\gamma p}\,. \tag{5}$$

Mapping the canonical exponent p back to physical coordinates x yields the spectrum for each universality class:

- **Gumbel ($\gamma \to 0$, $b = 1$):** Setting $y = e^{-x}$ converts the translation $x \to x + \log 2$ into dilation $y \to y/2$. Canonical monomials $y^p = (e^{-x})^p = e^{-px}$ define physical exponential perturbations $\phi_k(x) = e^{-kx}$ (setting $p = k \geq 0$), yielding the discrete spectrum:

$$\lambda_k = 2^{1-k}\,.$$

- **Fréchet ($\gamma = 1/\alpha > 0$, $b = 2^{1/\alpha}$):** For $y = x > 0$, physical power-law tail perturbations are expressed as $\phi_\beta(x) = x^{-\beta}$, where $\beta > 0$ dictates the perturbation's decay rate. Matching $y^p = x^{-\beta}$ sets $p = -\beta$. Substituting $\gamma = 1/\alpha$ into Eq. (5) gives:

$$\lambda_\beta = 2^{1+(1/\alpha)(-\beta)} = 2^{1-\beta/\alpha}\,.$$

- **Weibull ($\gamma = -1/\alpha < 0$, $b = 2^{-1/\alpha}$):** For $y = -x > 0$ ($x < 0$), boundary perturbations scaling near the origin take the form $\phi_\beta(x) = (-x)^\beta$, setting $p = \beta$. Substituting $\gamma = -1/\alpha$ into Eq. (5) gives:
$$\lambda_\beta = 2^{1+(-1/\alpha)\beta} = 2^{1-\beta/\alpha}\,.$$

The magnitude of $\lambda_p = 2^{1+\gamma p}$ partitions the space of log-CDF perturbations into three regimes under coarse-graining:

1. **Relevant operator ($\lambda_p > 1 \iff 1 + \gamma p > 0$):** These perturbations grow exponentially as λ_p^s under iteration. In heavy-tailed distributions ($\gamma > 0$), any perturbation with a decay rate slower than the fixed point ($\beta < \alpha$) alters the asymptotic tail exponent, driving the flow away from the reference fixed point.
2. **Marginal operator ($\lambda_p = 1 \iff 1 + \gamma p = 0$):** The mode $p^\star = -1/\gamma$ yields $\lambda = 1$; this corresponds to scale re-parameterizations ($x \to bx$). Infinitesimal location shifts ($x \to x + \epsilon$) generate marginal or irrelevant modes depending on the class.
3. **Irrelevant operator ($\lambda_p < 1 \iff 1 + \gamma p < 0$):** Any perturbations yielding $1 + \gamma p < 0$.

Our RG scheme for extreme value distributions assumes independent and identically distributed (i.i.d.) random variables. We make the following remarks regarding its non-i.i.d. generalization.

- **Short-range correlations:** For short-range dependent variables, extreme events occur in clusters. Once the coarse-graining scale exceeds the correlation length (cluster size), statistical independence between effective blocks is restored, and the i.i.d. RG results apply.
- **Non-identical distributions:** When random variables are from different distributions, the RG flow is governed by the dominant tail. Provided no individual subset destabilizes the scale sequence, the system flows toward the standard fixed point corresponding to the heaviest tail.
- **Strong dependence:** When long-range correlations or hierarchical structures are present (e.g., log-correlated fields, spin glasses, or branching processes), large fluctuations remain strongly coupled across scales. This breaks the discrete scaling equation $L_{s+1}(x) = 2L_s(bx + c)$, giving rise to non-standard limit distributions.

Appendix A. Fréchet derivative

Let the RG map $\mathcal{R}$ act on a Banach space $\mathcal{X}$ of models. The linearization of $\mathcal{R}$ around a fixed point $\mathcal{M}^\star$ is described by the Fréchet derivative $D\mathcal{R}_{\mathcal{M}^\star} : \mathcal{X} \to \mathcal{X}$. For any small perturbation $\delta\mathcal{M} \in \mathcal{X}$, the derivative is defined by

$$\mathcal{R}(\mathcal{M}^\star + \delta\mathcal{M}) = \mathcal{R}(\mathcal{M}^\star) + D\mathcal{R}_{\mathcal{M}^\star}(\delta\mathcal{M}) + o(\|\delta\mathcal{M}\|).$$

In this sense $D\mathcal{R}(\mathcal{M}^\star)$ provides the *best linear approximation* of $\mathcal{R}$ near $\mathcal{M}^\star$. $D\mathcal{R}_{\mathcal{M}^\star}$ captures how infinitesimal deviations evolve under coarse-graining and provides the basis for identifying relevant, irrelevant, and marginal directions. For different choices of the model, the Fréchet derivative can take different forms, such as Jacobian matrices for discretized models, or integral operators for continuous models.

Appendix B. Solving for Gaussian

Consider the scalar fixed-point equation for characteristic functions

$$\varphi^\star(t) = \varphi^\star\Big(\frac{t}{\sqrt{2}}\Big)^2, \qquad \varphi^\star(0) = 1\,.$$

First note that $\varphi^\star(t)$ cannot vanish for any $t \neq 0$. If $\varphi^\star(t_0) = 0$ for some t_0, then $\varphi^\star(t_0/2^{n/2}) = 0$ for all $n \geq 0$ by iterating the equation, which (by continuity) would force $\varphi^\star(0) = 0$, contradicting $\varphi^\star(0) = 1$. Hence $\varphi^\star(t) \neq 0$ for all t, so the cumulant generating function

$$K^\star(t) \coloneqq \log \varphi^\star(t)$$

is well-defined. The fixed-point identity of $K^\star$ is

$$K^\star(t) = 2\,K^\star\Big(\frac{t}{\sqrt{2}}\Big)\,.$$

For $t \neq 0$ define the ratio

$$h(t) \coloneqq \frac{K^\star(t)}{t^2}\,.$$

Then from the fixed-point relation

$$h(t) = \frac{K^\star(t)}{t^2} = \frac{2\,K^\star(t/\sqrt{2})}{t^2} = \frac{K^\star(t/\sqrt{2})}{(t/\sqrt{2})^2} = h\Big(\frac{t}{\sqrt{2}}\Big)\,.$$

Hence h is invariant under the scaling $t \mapsto t/2^{1/2}$. By iterating,

$$h(t) = h\Big(\frac{t}{2^{s/2}}\Big) \quad \text{for all } s \geq 0\,.$$

Therefore, $h(t)$ must be a constant function with value

$$h(t) = \lim_{\tau \to 0} h(\tau)\,.$$

Expanding $\varphi^\star$ near zero gives

$$\varphi^\star(t) = 1 - \frac{t^2}{2} + o(t^2)\,,$$

so

$$K^\star(t) = \log \varphi^\star(t) = -\frac{t^2}{2} + o(t^2)\,,$$

and therefore

$$\lim_{\tau\to 0} h(\tau) = \lim_{\tau\to 0} \frac{K^\star(\tau)}{\tau^2} = -\frac{1}{2}\,.$$

Thus $K^\star(t) = h(t)\, t^2 = -\frac{t^2}{2}$, and

$$\varphi^\star(t) = \exp\Big(-\frac{t^2}{2}\Big)\,.$$

A key assumption in the above derivation is that the characteristic function is at least twice differentiable at the origin. If this does not hold, suggesting the variance is infinite, we may have a Lévy fixed point.

Appendix C. Solving for Lévy

Consider the fixed-point equation for the cumulant generating function:

$$K^\star(t) = 2\, K^\star\Big(\frac{t}{b}\Big)\,.$$

We restrict our attention to *symmetric* distributions, implying $K^\star(t)$ is real-valued. Since $K^\star(t)$ must be a homogeneous function to satisfy scale invariance, we assume a power-law ansatz for the leading-order behavior near the origin:

$$K^\star(t) \sim -c|t|^\alpha \quad \text{as } t \to 0\,,$$

where α is the characteristic exponent and c is a positive scale parameter. For $t \neq 0$, define the scale-invariant ratio

$$h(t) \coloneqq \frac{K^\star(t)}{|t|^\alpha}\,.$$

Substituting the fixed-point relation for $K^\star(t)$ into this definition yields:

$$\begin{aligned} h(t) &= \frac{2\, K^\star(t/b)}{|t|^\alpha} \\ &= \frac{2\, K^\star(t/b)}{b^\alpha\, |t/b|^\alpha} \\ &= \frac{2}{b^\alpha}\, h(t/b)\,. \end{aligned}$$

For the ratio $h(t)$ to be scale-invariant (i.e., independent of t), the prefactor must be unity. This imposes the scaling condition:

$$\frac{2}{b^\alpha} = 1 \implies b = 2^{1/\alpha}\,.$$

It is worth noting that the ansatz $K^\star(t) \sim -c|t|^\alpha$ strictly describes symmetric distributions. The general solution to the fixed-point equation allows for an imaginary component, representing asymmetry (skewness):

$$K^\star(t) = -c|t|^\alpha \left[1 - \mathrm{i}\beta \tan\Big(\frac{\pi\alpha}{2}\Big)\mathrm{sgn}(t)\right] ,$$

where $\beta \in [-1, 1]$ is a skewness parameter. However, the real scaling factor $b = 2^{1/\alpha}$ is determined solely by the homogeneity of the modulus $|t|^\alpha$. Therefore, the derivation of the exponent α remains valid regardless of skewness.

The above algebraic derivation is formally valid for any $\alpha > 0$. However, the resulting characteristic function $\varphi^\star(t) = \exp(K^\star(t))$ must correspond to a valid probability distribution. This requirement restricts the exponent to the interval $\alpha \in (0, 2]$. To understand the upper bound $\alpha \leq 2$, compare the ansatz with the standard Taylor expansion for a distribution with a finite second moment:

$$\varphi^\star(t) = 1 + \mathrm{i}\,\mathbb{E}[X]t - \frac{\mathbb{E}[X^2]}{2}t^2 + o(t^2) .$$

The power-law ansatz implies the expansion:

$$\varphi^\star(t) = 1 - c|t|^\alpha + o(|t|^\alpha) .$$

If we assume $\alpha > 2$, the term $|t|^\alpha$ vanishes faster than t^2 as $t \to 0$. This creates a contradiction: the leading-order decay of any distribution with finite variance is controlled by t^2. Therefore, $\alpha > 2$ is impossible for a non-degenerate distribution. This leaves two regimes:

- $\alpha = 2$: The leading term is quadratic, corresponding to the Gaussian fixed point where $\mathbb{E}[X^2] < \infty$.
- $\alpha < 2$: The term $|t|^\alpha$ dominates t^2 near the origin. This singularity implies that the second derivative at the origin diverges, i.e., $\mathbb{E}[X^2] = \infty$.

By similar reasoning, if $\alpha < 1$, the first derivative diverges, implying an infinite mean.

Appendix D. Deriving the extreme value distribution

We seek non-trivial solutions $L^\star(x) \coloneqq \log F^\star(x)$ satisfying the invariance equation

$$L^\star(x) = 2\, L^\star(bx + c) , \tag{A.1}$$

for constants $b > 0$ and $c \in \mathbb{R}$. Because $F^\star(x)$ is a cumulative distribution function, $L^\star$ must satisfy the following boundary and monotonicity properties:

1. $L^\star(x) \leq 0$ for all $x \in \mathbb{R}$.
2. $L^\star(x)$ is non-decreasing in x.
3. $\lim_{x\to-\infty} L^\star(x) = -\infty$ and $\lim_{x\to+\infty} L^\star(x) = 0$.

The structure of the linear transformation $T(x) = bx + c$ determines the universality class. Equation (A.1) reduces to either pure translation solutions (exponential laws) or pure dilation solutions (power laws, analogous to the Gaussian case).

Appendix D.1. Case 1: Pure translation ($b = 1$)
When $b = 1$, Eq. (A.1) becomes

$$L^\star(x) = 2\, L^\star(x + c)\,.$$

Adopting the exponential ansatz $L(x) \propto -e^{-x}$ fixes $c = \log 2$ and yields

$$L^\star(x) = -e^{-x} \implies F^\star(x) = \exp\left(-e^{-x}\right), \qquad x \in \mathbb{R}\,.$$

This is the Gumbel fixed point. Incidentally, because the fixed-point relation is enforced only at discrete scaling steps, any solution can be modulated by a periodic function. We eliminate the inter-step oscillations by requiring monotonicity and smoothness.

Appendix D.2. Case 2: Affine transformation ($b \neq 1$)
When $b \neq 1$, the linear map $T(x) = bx + c$ possesses a unique fixed point $x_0 = \frac{c}{1-b}$. Shifting the origin via $y = x - x_0$, the transformation simplifies to $T(y) = by$, and Eq. (A.1) reduces to

$$\tilde{L}(y) = 2\, \tilde{L}(by)\,, \tag{A.2}$$

where $\tilde{L}(y) \coloneqq L^\star(y + x_0)$. Depending on whether $b > 1$ or $0 < b < 1$, the domain where $\tilde{L}(y)$ is non-zero splits into two cases:

Subcase 2a: $b > 1$ (Fréchet class). Since $b > 1$, repeated application of $T(y) = by$ drives $y > 0$ further away from 0. To satisfy the boundary conditions, $\tilde{L}(y)$ must be $-\infty$ for $y \leq 0$ and finite for $y > 0$. Setting $b = 2^{1/\alpha}$ for $\alpha > 0$, Eq. (A.2) becomes

$$\tilde{L}(y) = 2\, \tilde{L}\left(2^{1/\alpha} y\right) \implies \tilde{L}\left(2^{1/\alpha} y\right) = \frac{1}{2}\, \tilde{L}(y)\,.$$

The unique monotonic power-law solution on $y > 0$ is

$$\tilde{L}(y) = -A\, y^{-\alpha}\,, \qquad A > 0\,.$$

Setting $x_0 = 0$ ($c = 0$) and $A = 1$ yields

$$L^\star(x) = -x^{-\alpha} \implies F^\star(x) = \exp\left(-x^{-\alpha}\right), \qquad x > 0\,,$$

which is the Fréchet fixed point.

Subcase 2b: $0 < b < 1$ (Weibull class). When $0 < b < 1$, repeated multiplication contracts $y < 0$ toward the origin $y = 0$. In this regime, $\tilde{L}(y) = 0$ for $y \geq 0$ and finite for $y < 0$. Setting $b = 2^{-1/\alpha}$ for $\alpha > 0$, Eq. (A.2) becomes

$$\tilde{L}(y) = 2\, \tilde{L}\left(2^{-1/\alpha} y\right)\,.$$

The unique monotonic power-law solution on $y < 0$ is

$$\tilde{L}(y) = -A(-y)^{\alpha}\,, \qquad A > 0\,.$$

Setting $x_0 = 0$ ($c = 0$) and $A = 1$ yields

$$L^\star(x) = -(-x)^{\alpha} \implies F^\star(x) = \exp\left(-(-x)^{\alpha}\right), \qquad x < 0\,,$$

which is the Weibull fixed point.

Appendix D.3. Tail behavior

To connect the fixed-point parameter b to the tail decay of the parent distribution $F_0(x)$, consider a block of $N = 2^s$ independent random variables. The log-CDF of the unscaled maximum is $NL_0(x)$, where $L_0(x) \coloneqq \log F_0(x) \approx -(1 - F_0(x))$ as x approaches the upper bound of the support.

To establish a non-degenerate limit as $N \to \infty$, we introduce an N-dependent scale parameter $a_N > 0$ and location shift $d_N \in \mathbb{R}$ such that

$$L_N(x) \coloneqq NL_0(a_N x + d_N) = \mathcal{O}(1)\,. \tag{A.3}$$

Under scale doubling $N \to 2N$, the RG scale factor b in $L_{s+1}(x) = 2L_s(bx + c)$ corresponds to the ratio of successive scale parameters:

$$b = \lim_{N\to\infty} \frac{a_{2N}}{a_N}\,. \tag{A.4}$$

The asymptotic decay rate of $L_0(x)$ near the upper boundary determines a_N, which in turn fixes b:

1. Exponential/light tails ($b = 1$, Gumbel class). For parent distributions with exponentially decaying upper tails, $L_0(x) \sim -Ce^{-\lambda x}$ as $x \to \infty$ for $\lambda > 0$. Substituting into Eq. (A.3):

$$NL_0(a_N x + d_N) \sim -NCe^{-\lambda(a_N x + d_N)} = -\left(Ne^{-\lambda d_N}\right) Ce^{-\lambda a_N x}\,.$$

Choosing $d_N = \frac{\log(NC)}{\lambda}$ absorbs N entirely into the shift d_N. Consequently, no multiplicative rescaling is needed ($a_N = 1$ for all N), yielding:

$$b = \frac{a_{2N}}{a_N} = 1\,.$$

2. Power-law/heavy tails ($b > 1$, Fréchet class). For parent distributions with power-law tails, $L_0(x) \sim -Cx^{-\alpha}$ as $x \to \infty$ for $\alpha > 0$. Setting $d_N = 0$, Eq. (A.3) gives:

$$NL_0(a_N x) \sim -NC(a_N x)^{-\alpha} = -\left(Na_N^{-\alpha}\right) Cx^{-\alpha}\,.$$

Stabilizing the prefactor requires $Na_N^{-\alpha} = 1 \implies a_N = N^{1/\alpha}$. Doubling $N \to 2N$ yields:

$$b = \frac{a_{2N}}{a_N} = \frac{(2N)^{1/\alpha}}{N^{1/\alpha}} = 2^{1/\alpha} > 1\,.$$

3. Bounded support ($b < 1$, Weibull class). For parent distributions bounded above at $x_{\max}$, let $\epsilon = x_{\max} - x$. If $L_0(x_{\max} - \epsilon) \sim -C\epsilon^{\alpha}$ as $\epsilon \to 0^+$ for $\alpha > 0$, centering at $x_{\max}$ and setting $d_N = x_{\max}$, we scale the remaining gap $y = x - x_{\max} < 0$:

$$NL_0(x_{\max} + a_N y) \sim -NC(-a_N y)^{\alpha} = -\left(Na_N^{\alpha}\right) C(-y)^{\alpha}\,.$$

Stabilizing the prefactor requires $Na_N^{\alpha} = 1 \implies a_N = N^{-1/\alpha}$. Doubling $N \to 2N$ yields:

$$b = \frac{a_{2N}}{a_N} = \frac{(2N)^{-1/\alpha}}{N^{-1/\alpha}} = 2^{-1/\alpha} < 1\,.$$