Message-Passing Algorithms and Homology

From Thermodynamics to Statistical Learning

Olivier Peltre

2020

Introduction

The problem of describing the statistics of a large number x_i, x_j,... of interacting random variables emerged in physics with Boltzmann's efforts to lay principles of thermodynamics on statistical grounds, and high dimensional statistics are now expected to provide with reasonable and tractable models in artificial intelligence and biology. Aimed at modelling the emergence of collective behaviours in large assemblies of constituents, the prism of statistics hence shows deep analogies between atoms in a crystal, and neurons in a network.

A probability distribution $p(x)$ on the joint variable $x = (x_j)_{j\in\Omega}$ is usually assumed to capture all collective phenomena, although a dimensional curse prohibits the computation of expectation values. Local effects on a small subset $\alpha \subseteq \Omega$ of variables may nonetheless be estimated, as the statistics of the local variable $x_\alpha = (x_i)_{i\in\alpha}$ only involve the marginal distribution $p_\alpha(x_\alpha)$. Spontaneous magnetisation, for instance, is given by the expectation value of a single atomic dipole $x_i = \pm 1$, subject to interactions within an arbitrary large crystalline network. Accessing marginals is also a crucial step of statistical learning: usually appearing in the gradient of a loss function, they are necessary to guide the update of model parameters. The design of efficient algorithms for marginal estimation is hence of great practical importance. *Message-passing algorithms* estimate marginals through a parallelised and asynchronous computing scheme, in which a collection of local units communicate until they eventually reach a consensual state. Understanding their connections with algebraic topology was the first motivation of this thesis.

Gibbs random fields[1] are probabilistic models with a local structure described by a collection X of subsets α, β, γ,... of Ω, over which the global distribution $p(x)$ factorises as a product of local functions. We write $p \in \mathcal{G}(X)$ when there exists a collection of positive factors (f_α) such that:

$$p(x) = \frac{1}{Z} \prod_{\alpha \in X} f_\alpha(x_\alpha) \tag{1}$$

Distributions of this form are more often called *graphical models* in the computer science literature. The hypergraph $X \subseteq \mathcal{P}(\Omega)$ is then represented by the so-called factor graph, depicted in figure 1, formed by joining variable nodes (x_i) with their associated factor nodes (f_α). This factorisation is more conveniently viewed at the level of energies, where the *hamiltonian* H is defined as a sum of local *interaction potentials* (u_α), related to the factors by $u_\alpha = -\ln f_\alpha$:

$$p = \frac{\mathrm{e}^{-H}}{Z} \qquad \text{where} \qquad H(x) = \sum_{\alpha \in X} u_\alpha(x_\alpha) \tag{2}$$

The fundamental Legendre duality between the energy function $H(x)$ and its Gibbs distribution $p(x)$ is related to variational principles on entropy and free energy, of which message-passing algorithms yield approximate solutions.

One of our contributions is to view the Gibbs random field $p \in \mathcal{G}(X)$ as a *homology class* of factors. Introducing mutual dependence of variables, overlapping subsets in X also make the parametrisation

[1] Or *Gibbs distributions*, which are also *Markov random fields* according to the Hammersley-Clifford correspondence. Factorisability however yields a finer characterisation of $\mathcal{G}(X)$ than Markov properties, hence the preferred terminology.

(a) (b)

Figure 1: (a) factor graph and (b) hypergraph representations of $X \subseteq \mathcal{P}(\Omega)$.

of $p(x)$ given by equation (1) ambiguous, two collections of factors (f_α) and (f'_α) defining the same Gibbs random field whenever $\prod_\alpha f_\alpha \simeq \prod_\alpha f'_\alpha$ up to a scaling factor. This ambiguity is resolved by introducing *messages* as collections $(m_{\alpha\beta})$ of local functions $m_{\alpha\beta}(x_\beta)$ for every ordered pair $\alpha \supseteq \beta$ in X, and a *boundary operator* ∂ defining factors from messages through:

$$(\partial m)_\beta = \prod_{\alpha \supseteq \beta} m_{\alpha\beta} \Big/ \prod_{\gamma \subseteq \beta} m_{\beta\gamma} \tag{3}$$

Assuming X is closed under intersection, we show that (f_α) and (f'_α) define the same Gibbs random field if and only if there exists $(m_{\alpha\beta})$ such that $f' \simeq f \cdot \partial m$ up to scaling. In this view, message-passing algorithms explore a homology class of factors by iterating over messages, the homological constraint expressing conservation of the global distribution $p \in \mathcal{G}(X)$.

Beliefs (q_α) are intended to estimate the local marginals (p_α) of the global probability distribution. These local probabilities should in particular satisfy *consistency* conditions which require that q_β is the marginal of q_α for every $\beta \subseteq \alpha$. Of cohomological nature, this constraint shall take the form $dq = 0$ and is expressed by the following set of equations:

$$q_\beta(x_\beta) = \sum_{y \in E_{\alpha \smallsetminus \beta}} q_\alpha(x_\beta, y) \tag{4}$$

Defining local probabilities through the local analog of equation (1):

$$q_\alpha(x_\alpha) = \frac{1}{Z_\alpha} \prod_{\beta \subseteq \alpha} f'_\beta(x_\beta) \tag{5}$$

the specificity of message-passing algorithms is hence to search for consistent beliefs that derive from homologous factors $f' = f \cdot \partial m$. The most general message-passing scheme, *belief propagation*, assumes the following update rule:

$$m_{\alpha\beta}(x_\beta) \leftarrow m_{\alpha\beta}(x_\beta) \cdot \frac{\sum_y q_\alpha(x_\beta, y)}{q_\beta(x_\beta)} \tag{6}$$

Our main contribution is to introduce diffusion equations of the form $\dot{u} = \delta\Phi(u)$ on interaction potentials, which allow to view existing message-passing algorithms as coarse numerical integrators of continuous-time differential equations. The operator δ is the first-degree boundary of a natural homology theory, acting on a collection of energy fluxes $\varphi_{\alpha\beta}(x_\beta)$ by:

$$\delta_\beta \varphi(x_\beta) = \sum_{\alpha \supseteq \beta} \varphi_{\alpha\beta}(x_\beta) - \sum_{\beta \supseteq \gamma} \varphi_{\beta\gamma}(x_\gamma) \tag{7}$$

In addition to revealing their deeply homological character, this approach should dramatically improve the stability[2] of message-passing algorithms. Showing belief propagation equivalent to a time-step-one

[2] Belief propagation has for instance been reported to start converging poorly after several epochs of training restricted Boltzmann machines, a brutal phenomenon that has been compared to phase transitions of the Hopfield model.

explicit Euler scheme of $\dot{u} = \delta\Phi(u)$, a first and highly advisable improvement is to use a smaller time step $\lambda < 1$, which would act as an exponent on the geometric increment of $m_{\alpha\beta}$ in equation (6). As another direction of improvement, we propose a combinatorial correction of messages eliminating their redundancies by extending Möbius inversion formulas to higher degrees. A practical question which shall remain open is whether there exists a notion of optimal transport on Φ bringing interaction potentials to equilibrium.

Our approach reveals that stationary states of message-passing algorithms lie at the intersection of two constraint surfaces, of homological and cohomological nature respectively. The homological constraint is linear at the level of interaction potentials, and expresses conservation of the *total energy*:

$$H(x) = \sum_{\alpha \in X} u_\alpha(x_\alpha) \tag{8}$$

The cohomological constraint $dq = 0$, however, is linear at the level of the effective *Gibbs states*:

$$q_\alpha = \frac{\mathrm{e}^{-U_\alpha}}{Z_\alpha} \quad \text{with} \quad U_\alpha(x_\alpha) = \sum_{\beta \subseteq \alpha} u_\beta(x_\beta) \tag{9}$$

The problem of describing this intersection is hence highly non-linear, and trying to understand how the geometry of the underlying hypergraph X affects the geometry of message-passing equilibria will lead to difficult questions meeting both algebraic topology and singularity theory.

* * *

Out of the six chapters contained in this thesis, chapters 1 to 3 review and develop the algebraic theory we shall rely upon. Energy and information functionals are covered in chapter 4, providing background for the central theorem 4.22 characterising solutions of Kikuchi's cluster variational method [10] i.e. consistent collections of local probabilities which are critical for a generalised Bethe free energy.

Message-passing algorithms are then addressed in chapter 5, as discrete integrators of continuous-time[3] diffusion equations $\dot{u} = \delta\Phi(u)$. The homological picture allows us to give a rigorous proof of the correspondence theorem 5.13 between stationary states of belief propagation and solutions of the cluster variational method, as suggested by Yedidia *et al.* [32]. Our approach more generally characterises all the flux functionals Φ for which such a correspondence holds, while the combinatorics developed in chapter 3 lead us to propose another regularisation of the generalised belief propagation algorithm by a degree-one Möbius inversion[4] on the flux functional Φ.

The geometry of message-passing equilibria is finally studied in chapter 6. We describe a class of *retractable* hypergraphs for which message-passing always converges to the exact marginals of the global probabilistic model to estimate. In general, multiple equilibria may coexist, whose bifurcations are related to singularities of the projection of a smooth manifold of consistent potentials onto their homology classes, and may be tracked in the spectrum of a linearised diffusion operator.

Our first efforts consisted in looking for a formalism in which the elementary operations of message-passing algorithms would fit. The reader is therefore expected to run into some unusual notations and properties, which the following few pages attempt to summarise efficiently. With these in mind, we hope that an informed reader mostly curious of applications might jump directly to chapter 5.

[3] The time step of the integrator may be tuned < 1 to improve stability, analogously to a learning rate.
[4] Möbius inversion eliminates redundancies otherwise counted in the heat flux Φ.

Statistical Systems

A system will be defined by a collection of random variables $x_i, x_j, x_k, \dots$ indexed by labels $i \in \Omega$.

The set of labels Ω in general has an additional geometric structure describing interactions (e.g. graph or hypergraph).

Subsets $\alpha, \beta, \gamma, \dots$ in $\mathcal{P}(\Omega)$ form a partial order for inclusion, usually denoted in descending alphabetical order:

$$\alpha \supseteq \beta \supseteq \gamma$$

Local Spaces

Given a finite set of microstates E_i for every atom/neuron/bit $i \in \Omega$ and a subset of atoms $\alpha \subseteq \Omega$:

- $E_\alpha = \prod_{i\in\alpha} E_i$ set of local microstates: $x_\alpha = (x_i)_{i\in\alpha} \in E_\alpha$
- $A_\alpha = \mathbb{R}^{E_\alpha}$ algebra of local observables: $f_\alpha(x_\alpha) \in A_\alpha$
- $A^*_\alpha \simeq \mathbb{R}^{E_\alpha}$ vector space of local measures: $q_\alpha : f_\alpha \mapsto \langle\, q_\alpha \,|\, f_\alpha \,\rangle \in A^*_\alpha$
- $\Delta_\alpha \subseteq A^*_\alpha$ convex set of positive local probabilities: $p_\alpha > 0$ and $\sum_{x_\alpha} p_\alpha(x_\alpha) = 1$

Local Operators

For every $\alpha \subseteq \Omega$ and $\beta \subseteq \alpha$:

- $\pi^{\beta\alpha} : E_\alpha \longrightarrow E_\beta$ natural projection

$$\pi^{\beta\alpha} : x_\alpha = (x_i)_{i\in\alpha} \mapsto x_\beta = (x_j)_{j\in\beta}$$

- $j_{\alpha\beta} : A_\beta \longrightarrow A_\alpha$ natural extension[5]

$$j_{\alpha\beta}(f_\beta)(x_\alpha) = f_\beta(x_\beta)$$

- $\Sigma^{\beta\alpha} : A^*_\alpha \longrightarrow A^*_\beta$ partial integration = marginal projection

$$\Sigma^{\beta\alpha}(q_\alpha)(x_\beta) = \sum_{x' \in E_{\alpha\smallsetminus\beta}} q_\alpha(x_\beta, x')$$

- $\mathbb{E}^{\beta\alpha}_{p_\alpha} : A_\alpha \longrightarrow A_\beta$ conditional expectation w.r.t. $p_\alpha \in \Delta_\alpha$ given β

$$\mathbb{E}^{\beta\alpha}_{p_\alpha}(f_\alpha)(x_\beta) = \mathbb{E}_{p_\alpha}[f_\alpha \mid x_\beta] = \sum_{x' \in E_{\alpha\smallsetminus\beta}} \frac{p_\alpha(x_\beta, x')\, f_\alpha(x_\beta, x')}{p_\beta(x_\beta)}$$

- $\mathbb{F}^{\beta\alpha} : A_\alpha \xrightarrow{C^\infty} A_\beta$ conditional free energy of α given β = effective energy

$$\mathbb{F}^{\beta\alpha}(f_\alpha)(x_\beta) = -\ln \sum_{x' \in E_{\alpha\smallsetminus\beta}} \mathrm{e}^{-f_\alpha(x_\beta, x')}$$

Local Duality

- natural duality bracket $\langle - | - \rangle : A^*_\alpha \otimes A_\alpha \longrightarrow \mathbb{R}$

$$\langle\, q_\alpha \,|\, f_\alpha \,\rangle = \sum_{x_\alpha \in E_\alpha} q_\alpha(x_\alpha)\, f_\alpha(x_\alpha)$$

- covariance metric $\langle - | - \rangle_{p_\alpha} : A_\alpha \otimes A_\alpha \longrightarrow \mathbb{R}$ induced by a local probability $p_\alpha \in \Delta_\alpha$

$$\Big\langle\, f_\alpha \,\Big|\, g_\alpha \,\Big\rangle_{p_\alpha} = \mathbb{E}_{p_\alpha}\big[f_\alpha\, g_\alpha\big] = \sum_{x_\alpha \in E_\alpha} p_\alpha(x_\alpha)\, f_\alpha(x_\alpha)\, g_\alpha(x_\alpha)$$

Properties

- Adjunction of $\Sigma^{\beta\alpha}$ and $j_{\alpha\beta}$ for the natural duality brackets:

$$\langle\, \Sigma^{\beta\alpha}(q_\alpha) \,|\, f_\beta \,\rangle = \langle\, q_\alpha \,|\, j_{\alpha\beta}(f_\alpha) \,\rangle$$

- Adjunction[6] of $\mathbb{E}^{\beta\alpha}_{p_\alpha}$ and $j_{\alpha\beta}$ for the metric induced by p_α on A_α

$$\Big\langle\, \mathbb{E}^{\beta\alpha}(f_\alpha) \,\Big|\, g_\beta \,\Big\rangle_{\Sigma^{\beta\alpha}(p_\alpha)} = \Big\langle\, f_\alpha \,\Big|\, j_{\alpha\beta}(g_\beta) \,\Big\rangle_{p_\alpha}$$

- Gibbs state conditional expectations[7] from effective energies: $\mathbb{E}^{\beta\alpha}_{p_\alpha} = d\mathbb{F}^{\beta\alpha}_{H_\alpha}$

$$\mathbb{E}_{p_\alpha}[f_\alpha | x_\beta] = \mathbb{F}^{\beta\alpha}(H_\alpha + f_\alpha) - \mathbb{F}^{\beta\alpha}(H_\alpha) + o(f_\alpha) \quad \text{for} \quad p_\alpha = \frac{1}{Z_\alpha}\, \mathrm{e}^{-H_\alpha}$$

[5] $j_{\alpha\beta}$ coincides with the identity map w.r.t. the inclusion $A_\beta \subseteq A_\alpha$, we shall therefore simply write f_β for $j_{\alpha\beta}(f_\beta)$.

[6] $\mathbb{E}^{\beta\alpha}$ is the orthogonal projection of A_α onto $A_\beta \subseteq A_\alpha$ for the covariance metric $\langle - | - \rangle_{p_\alpha}$.

[7] Gibbs state expectation $\mathbb{E}^\alpha_{p_\alpha} = \mathbb{E}^{\varnothing\alpha}_{p_\alpha}$ is the differential of the free energy $\mathbb{F}^\alpha = \mathbb{F}^{\varnothing\alpha}$ while the Fisher information metric is recovered in the second differential of $\mathbb{F}^\alpha$. Shannon entropy S_α is the Legendre transform of $\mathbb{F}^\alpha$ so that $\mathbb{F}^\alpha(H_\alpha) = \min_{p_\alpha \in \Delta_\alpha} \langle\, p_\alpha \,|\, H_\alpha \,\rangle - S_\alpha(p_\alpha)$ for all $H_\alpha \in A_\alpha$ (chapter 4).

Fields

Suppose given a covering $X = \{\alpha, \beta, \gamma, \dots\}$ of Ω by subsets s.t. $\alpha \cap \beta \in X$ for every $\alpha, \beta \in X$

n-Fields[8] are collections $\{f_{\alpha_0 \dots \alpha_n} \in A_{\alpha_n} \mid \alpha_0 \supset \dots \supset \alpha_n \in X\}$ of local observables indexed by n-chains in X

Field Spaces

- $A_0(X) = \prod_\alpha \mathbb{R}^{E_\alpha}$ space of potentials
- $A_1(X) = \prod_{\alpha \supset \beta} \mathbb{R}^{E_\beta}$ space of currents
- $A_n(X) = \prod_{\alpha_0 \supset \dots \supset \alpha_n} \mathbb{R}^{E_{\alpha_n}}$ space of local observable n-fields
- $\Delta_0(X) = \prod_\alpha \Delta_\alpha \subseteq A_0^*(X)$ convex space of positive beliefs
- $\Gamma(X) \subseteq \Delta_0(X)$ convex subset of consistent beliefs: $p \in \Gamma(X)$ iff $p_\beta = \Sigma^{\beta\alpha}(p_\alpha)$ for all $\alpha \supseteq \beta$.

Differential Operators

- $\delta : A_1(X) \longrightarrow A_0(X)$ divergence

$$\delta(\varphi)_\beta(x_\beta) = \sum_{\alpha' \supseteq \beta} \varphi_{\alpha'\beta}(x_\beta) - \sum_{\beta \supseteq \gamma'} \varphi_{\beta\gamma'}(x_{\gamma'})$$

- $d : A_0^*(X) \longrightarrow A_1^*(X)$ differential

$$d(q)_{\alpha\beta}(x_\beta) = q_\beta(x_\beta) - \Sigma^{\beta\alpha}(q_\alpha)(x_\beta)$$

- $\nabla_p : A_0(X) \longrightarrow A_1(X)$ gradient w.r.t. to a consistent belief $p \in \Gamma(X)$

$$\nabla_p(f)_{\alpha\beta}(x_\beta) = f_\beta(x_\beta) - \mathbb{E}^{\beta\alpha}_{p_\alpha}(f_\alpha)(x_\beta)$$

- $\mathcal{D} : A_0(X) \longrightarrow A_1(X)$ effective energy gradient

$$\mathcal{D}(f)_{\alpha\beta}(x_\beta) = f_\beta(x_\beta) - \mathbb{F}^{\beta\alpha}(f_\alpha)(x_\beta)$$

Field Duality

- natural duality bracket $\langle - \mid - \rangle : A_n^*(X) \otimes A_n(X) \longrightarrow \mathbb{R}$

$$\langle q \mid f \rangle = \sum_{\alpha_0 \supset \dots \supset \alpha_n} \langle q_{\alpha_0 \dots \alpha_n} \mid f_{\alpha_0 \dots \alpha_n} \rangle$$

- covariance metric $\langle - \mid - \rangle_p : A_n(X) \otimes A_n(X) \longrightarrow \mathbb{R}$ induced by a consistent $p \in \Gamma(X)$

$$\Big\langle f \,\Big|\, g \Big\rangle_p = \sum_{\alpha_0 \supset \dots \supset \alpha_n} \Big\langle f_{\alpha_0 \dots \alpha_n} \,\Big|\, g_{\alpha_0 \dots \alpha_n} \Big\rangle_{p_{\alpha_n}}$$

Properties

- Adjunction of d and δ for the natural duality bracket

$$\langle dq \mid \varphi \rangle = \langle q \mid \delta\varphi \rangle$$

- Adjunction of ∇_p and δ for the metric induced by $p \in \Gamma(X)$

$$\Big\langle \nabla_p(f) \,\Big|\, \varphi \Big\rangle_p = \Big\langle f \,\Big|\, \delta\varphi \Big\rangle_p$$

- Gibbs State gradient operator $\nabla_p = d\mathcal{D}_H$

$$\nabla_p(f)_{\alpha\beta} = \mathcal{D}(H + f)_{\alpha\beta} - \mathcal{D}(H)_{\alpha\beta} + o(f) \quad \text{for} \quad p_\alpha = \frac{1}{Z_\alpha} \mathrm{e}^{-H_\alpha}$$

[8]The graded vector space $A_\bullet(X)$ of fields is an analog of the space $\Omega^\bullet(\mathbb{R}^3)$ of scalar, vector, ... fields on $\mathbb{R}^3$, or of the space $C_\bullet(K)$ of chains in a simplicial complex K, except here fields have functional coefficients $f_\alpha(x_\alpha)$ instead of scalar coefficients.

Combinatorial Operators

– $\tilde{\zeta}_\Omega : A_0(X) \longrightarrow A_\Omega$ total energy

$$\tilde{\zeta}(h)_\Omega(x_\Omega) = \sum_\alpha h_\alpha(x_\alpha) = H_\Omega(x_\Omega)$$

– $\zeta : A_0(X) \longrightarrow A_0(X)$ zeta transform[9]

$$\zeta(h)_\alpha(x_\alpha) = \sum_{\alpha \supseteq \beta'} h_{\beta'}(x_{\beta'}) = H_\alpha(x_\alpha)$$

– $\mu : A_0(X) \longrightarrow A_0(X)$ Möbius transform[10] : $\mu = \zeta^{-1}$

$$\mu(H)_\alpha(x_\alpha) = \sum_{\alpha \supseteq \beta'} \mu_{\alpha\beta'} \, H_{\beta'}(x_{\beta'}) = h_\alpha(x_\alpha)$$

– $\zeta : A_n(X) \longrightarrow A_n(X)$ extended zeta transform

$$\zeta(\varphi)_{\alpha_0 \dots \alpha_n} = \sum_{\alpha_0 \supseteq \beta_0 \not\supseteq \alpha_1} \cdots \sum_{\alpha_n \supseteq \beta_n} \varphi_{\beta_0 \dots \beta_n} = \Phi_{\alpha_0 \dots \alpha_n}$$

– $\mu : A_n(X) \longrightarrow A_n(X)$ extended Möbius transform: $\mu = \zeta^{-1}$

$$\mu(\Phi)_{\alpha_0 \dots \alpha_n} = \sum_{\alpha_n \supseteq \beta_n} \cdots \sum_{\alpha_0 \supseteq \beta_0 \not\supseteq \beta_1} \mu_{\alpha_0\beta_0} \dots \mu_{\alpha_n\beta_n} \, \Phi_{\beta_0 (\beta_0 \cap \beta_1) \dots (\beta_0 \cap \dots \cap \beta_n)} = \varphi_{\alpha_0 \dots \alpha_n}$$

Properties

– Möbius numbers[11] $c_\alpha \in \mathbb{Z}$ and total energy H_Ω of a potential $h = \mu \cdot H \in A_0(X)$

$$H_\Omega = \sum_\alpha h_\alpha = \sum_\alpha c_\alpha \, H_\alpha = \tilde{\zeta}(h)_\Omega$$

– Gauss formula[12] for a current $\varphi \in A_1(X)$

$$\zeta(\delta\varphi)_\alpha = \sum_{\alpha \supseteq \beta'} \delta(\varphi)_{\beta'} = \sum_{\alpha' \not\subseteq \alpha} \sum_{\alpha \supseteq \beta'} \varphi_{\alpha'\beta'} = \tilde{\zeta}(\varphi)_{\Omega\alpha}$$

– Gauss formula[13] for $\varphi = \mu \cdot \Phi \in A_1(X)$

$$\zeta\Big(\delta(\mu \cdot \Phi)\Big)_\alpha = \sum_{\alpha' \not\subseteq \alpha} c_{\alpha'} \, \Phi_{\alpha'(\alpha \cap \alpha')} = \tilde{\zeta}(\mu \cdot \Phi)_{\Omega\alpha}$$

Main Theorem[14]

– Homology and total energy: for every $u, h \in A_0(X)$

$$\exists \varphi \in A_1(X) \quad \text{s.t.} \quad u = h + \delta\varphi \quad \Leftrightarrow \quad \sum_\alpha h_\alpha = \sum_\alpha u_\alpha$$

[9] « $\zeta(h)_\alpha = \int_{\Lambda^\alpha} h$ » is analogous to a discrete integral of the potential h on the cone $\Lambda^\alpha = \{\beta \subseteq \alpha\}$ below α.

[10] The coefficients $\mu_{\alpha\beta} \in \mathbb{Z}$ are computed inductively by $\mu_{\alpha\alpha} = 1$ and $\mu_{\alpha\gamma} = 1 - \sum_{\alpha \supset \beta' \supset \gamma} \mu_{\alpha\beta'}$ for every $\Omega \supseteq \alpha \supseteq \gamma$.

[11] The coefficients $c_\alpha \in \mathbb{Z}$ are computed inductively $c_\beta = 1 - \sum_{\alpha' \supset \beta} c_{\alpha'} = \sum_{\alpha' \supset \beta} \mu_{\alpha'\beta}$.

[12] « $\zeta(\delta\varphi)_\alpha = \int_{\Lambda^\alpha} \delta\varphi = \int_{d\Lambda^\alpha} \varphi$ » is analogous to a discrete flux integral of the current φ bound into Λ^α.

[13] Möbius inversion on the effective energy gradient $\varphi = \mu \cdot \mathcal{D}(U)$ before updating effective energies by $\dot{U} = \zeta(\delta\varphi)$ is one of our proposed regularisations of message-passing schemes (chapter 5).

[14] This result will rely on the fundamental yet not so widely known *interaction decomposition theorem* 2.8 [9, 16] which consistently decomposes each A_α as a direct sum of interaction subspaces $\oplus_\beta Z_\beta$ for $\beta \subseteq \alpha$.

Energy and Information Functionals

Local Functionals

– $\mathbb{F}^\alpha : A_\alpha \xrightarrow{C^\infty} \mathbb{R}$ local free energy

$$\mathbb{F}^\alpha(H_\alpha) = -\ln \sum_{x_\alpha \in E_\alpha} \mathrm{e}^{-H_\alpha(x_\alpha)}$$

– $S_\alpha : \Delta_\alpha \xrightarrow{C^\infty} \mathbb{R}$ local entropy

$$S_\alpha(p_\alpha) = -\ln \sum_{x_\alpha \in E_\alpha} p_\alpha(x_\alpha)\ \ln p_\alpha(x_\alpha)$$

Legendre Duality

– S_α is the Legendre transform of $\mathbb{F}^\alpha$, and reciprocally

$$\mathbb{F}_\alpha(H_\alpha) = \min_{p_\alpha \in \Delta_\alpha} \Big[\langle\, p_\alpha \,|\, H_\alpha \,\rangle - S_\alpha(p_\alpha) \Big]$$

– $d\mathbb{F}^\alpha : A_\alpha \xrightarrow{C^\infty} \Delta_\alpha$ maps hamiltonians to their Gibbs probability densities

$$d\mathbb{F}^\alpha(H_\alpha) = \langle\, p_\alpha \,|\, - \,\rangle = \mathbb{E}_{p_\alpha}[-] \quad \text{where} \quad p_\alpha = \frac{1}{Z_\alpha}\,\mathrm{e}^{-H_\alpha}$$

– $dS_\alpha : \Delta_\alpha \xrightarrow{C^\infty} A_\alpha/\mathbb{R}$ maps probability densities to their hamiltonians, defined up to additive constants

$$dS_\alpha(p_\alpha) = \langle\, - \,|\, H_\alpha \,\rangle \quad \text{where} \quad H_\alpha \simeq -\ln p_\alpha \mod \mathbb{R}$$

Global Functionals

– $\mathcal{U}_\Omega : \Delta_\Omega \times A_\Omega \xrightarrow{C^\infty} \mathbb{R}$ internal energy

$$\mathcal{U}_\Omega(p_\Omega, H_\Omega) = \langle\, p_\Omega \,|\, H_\Omega \,\rangle = \mathbb{E}_{p_\Omega}[H_\Omega]$$

– $\mathcal{F}_\Omega : \Delta_\Omega \times A_\Omega \xrightarrow{C^\infty} \mathbb{R}$ variational free energy

$$\mathcal{F}_\Omega(p_\Omega, H_\Omega) = \langle\, p_\Omega \,|\, H_\Omega \,\rangle - S_\Omega(p_\Omega)$$

Bethe-Kikuchi Approximation

– $\check{\mathcal{F}} : \Delta_0(X) \times A_0(X) \xrightarrow{C^\infty} \mathbb{R}$ Bethe free energy: constrained to consistent beliefs $p \in \Gamma(X)$

$$\check{\mathcal{F}}(p, H) = \sum_{\alpha \in X} c_\alpha \Big[\langle\, p_\alpha \,|\, H_\alpha \,\rangle - S_\alpha(p_\alpha) \Big]$$

Main Theorems

– Homological invariance of $\check{\mathcal{F}}(p, -)$: for every consistent belief $p \in \Gamma(X)$ and effective hamiltonian $H \in A_0(X)$

$$\check{\mathcal{F}}\Big(p, H + \zeta(\delta\varphi)\Big) = \check{\mathcal{F}}(p, H)$$

– $p \in \Gamma(X)$ is critical for $\check{\mathcal{F}}(-, H)_{|\Gamma(X)}$ iff there exists a current $\varphi \in A_1(X)$ s.t. for all $\alpha \in X$

$$-\ln p_\alpha \simeq H_\alpha + \zeta(\delta\varphi)_\alpha \mod \mathbb{R}$$

Message-Passing as Diffusion

Heat Analogy

Energy density = time-dependent scalar field $u : \mathbb{R} \to \Omega^0(\mathbb{R}^3) = C^\infty(\mathbb{R}^3)$

Heat exchange = time-dependent vector field $\vec{\varphi} : \mathbb{R} \to \Omega^1(\mathbb{R}^3) = C^\infty(\mathbb{R}^3, \mathbb{R}^3)$

- energy conservation: $\dot{u} = \mathrm{div}(\vec{\varphi})$
- heat flux: $\vec{\varphi} = -\lambda\, \vec{\mathrm{grad}}(T)$

Characteristic relation $u = c\,T$ (condensed matter) or non-relationship between temperature and energy.

Message-Passing on Graphs[15]

The algorithm takes the form $\dot{u} = -\Big(\delta \circ \mathcal{D} \circ \zeta\Big)(u)$ on potentials, i.e.

$$\dot{u} = \delta\varphi \qquad \text{where} \qquad \left| \begin{array}{l} \varphi = -\mathcal{D}(U) \\ U = \zeta \cdot u \end{array} \right.$$

- Energy conservation $\dot{u} = \delta\varphi$ dictates the update of effective potentials $u \in A_0(X)$

$$\frac{d}{dt} u_{ij}(x_i, x_j) = -\varphi_{ij\to j}(x_j) - \varphi_{ij\to i}(x_i) \quad \text{and} \quad \frac{d}{dt} u_i(x_i) = \sum_{j' \sim i} \varphi_{ij'\to i}(x_i)$$

- Heat flux $\varphi = -\mathcal{D}(U) \in A_1(X)$ measures the lack of consistency of effective hamiltonians

$$\varphi_{ij\to j}(x_j) = -U_j(x_j) - \ln \sum_{x_i \in E_i} \mathrm{e}^{-U_i(x_i, x_j)}$$

- Effective hamiltonians $U = \zeta(u) \in A_0(X)$ are given by

$$U_{ij}(x_i, x_j) = u_{ij}(x_i, x_j) + u_i(x_i) + u_j(x_j) \quad \text{and} \quad U_i(x_i) = u_i(x_i)$$

- Beliefs $q_{ij} = \frac{1}{Z_{ij}}\, \mathrm{e}^{-U_{ij}}$ and $q_i = \frac{1}{Z_i}\, \mathrm{e}^{-U_i}$ should be normalised at each iteration on graphs with loops[16]

Message-Passing on Hypergraphs[17]

Möbius inversion on the heat flux reads $\dot{u} = -\Big(\delta \circ \mu \circ \mathcal{D} \circ \zeta\Big)(u)$ on potentials, i.e.

$$\dot{u} = \delta\varphi \qquad \text{where} \qquad \left| \begin{array}{l} \varphi = \mu \cdot \Phi \\ \Phi = -\mathcal{D}(U) \\ U = \zeta \cdot u \end{array} \right.$$

- Effective hamiltonians $U = \zeta \cdot u \in A_0(X)$ follow the energy conservation principle: $\dot{U} = \delta^\zeta(\Phi)$ where $\delta^\zeta = \zeta\, \delta\, \zeta^{-1}$

$$\frac{d}{dt} U_\alpha(x_\alpha) = \sum_{\alpha' \not\subseteq \alpha} c_{\alpha'}\; \Phi_{\alpha'(\alpha\cap\alpha')}(x_{\alpha\cap\alpha'})$$

- Extensive heat flux $\Phi = \zeta \cdot \varphi \in A_1(X)$ flows against the effective energy gradient: $\Phi = -\mathcal{D}(U)$

$$\Phi_{\alpha\beta}(x_\beta) = -U_\beta(x_\beta) - \ln \sum_{x' \in E_{\alpha\smallsetminus\beta}} \mathrm{e}^{-U_\alpha(x_\beta, x')}$$

- Beliefs $q_\alpha = \frac{1}{Z_\alpha}$ do not need to be normalised[18]

Correspondence Theorem

Effective hamiltonians $U = H + \zeta \cdot \delta\varphi \in A_0(X)$ are stationary under diffusion

$\Leftrightarrow$ Beliefs $q = \frac{1}{Z}\, \mathrm{e}^{-U} \in \Gamma(X)$ are consistent and critical for the Bethe free energy $\check{\mathcal{F}}(-, H)_{|\Gamma(X)}$

[15] On acyclic graphs (trees) the algorithm converges in finite time, as already stated in Pearl's seminal paper [23]. Substitute $u^{(t+1)} - u^{(t)}$ for $\frac{du}{dt}$ and translate to beliefs to recover the usual belief propagation algorithm.

[16] With loops, the dynamic on potentials is best understood up to additive constants.

[17] On *retractable* hypergraphs $X \subseteq \mathcal{P}(\Omega)$, we show the algorithm to converge in finite time (chapter 6). Note that Möbius inversion of the heat flux only affects additive constants when X is a graph, hence the proposed regularisation only modifies the generalised belief propagation (GBP) algorithm of Yedidia et al. [32]

[18] When $\varnothing \in X$, Möbius inversion of fluxes $\Phi_{\alpha\varnothing} \in \mathbb{R}$ already takes care of regularising normalisation factors.

Index of Notations

Functors:

- $X \subseteq \mathcal{P}(\Omega)$ base hypergraph
- (E, π) microstates, §2.1.1
- (A, j) observables, §2.2.2
- (A^*, Σ) measures, §2.2.3
- (Δ, Σ) probability densities, §2.2.4

Spaces:

- $A_\bullet(X)$ complex of local observables, def. 2.6
- $A^*_\bullet(X)$ complex of local measures, –
- $\Delta_\bullet(X)$ convex subspace of local probabilities, –
- $\Gamma(X)$ convex subspace of consistent local probabilities, def. 2.7
- $\mathcal{C}(X)$ manifold of consistent local hamiltonians, def. 5.5
- $\mathcal{Z}(X)$ manifold consistent local potentials, def. 5.9

Differential Operators:

- δ boundary of $A_\bullet(X)$, §2.2.2
- d differential of $A^*_\bullet(X)$, –
- $\mathcal{D}$ effective energy gradient, §5.2.1
- $\nabla = \mathcal{D}_*$ linearised effective energy gradient –

Combinatorial Operators:

- ζ zeta transform, sections 3.2 and 3.3
- $\mu = \zeta^{-1}$ Möbius transform, §3.2.1 and 3.3.3

Diffusion Operators:

- $\Phi = -\mathcal{D} \circ \zeta$ standard diffusion flux, §5.2.2
- $\mathcal{T} = \delta\Phi$ standard diffusion vector field, –
- $\phi = -\mu \circ \mathcal{D} \circ \zeta$ canonical diffusion flux, §5.3.2
- $\tau = \delta\phi$ canonical diffusion vector field, –

Fields:

- $h, u \in A_0(X)$ interaction potentials, h for reference and u for evolution
- $H, U \in A_0(X)$ local hamiltonians, $H = \zeta \cdot h$ and $U = \zeta \cdot u$
- $\varphi \in A_1(X)$ energy flux
- $q \in \Delta_0(X)$ local beliefs, $q_\alpha = [\mathrm{e}^{-U_\alpha}]$
- $p \in \Gamma(X)$ Gibbs state marginals $p_\alpha = \Sigma^{\alpha\Omega}\big(p_\Omega\big)$

Information Functionals:

- $\mathbb{F}^\alpha$ free energy, §4.1.1
- $\mathbb{F}^{\beta\alpha}$ effective energy, §4.1.2
- S Shannon entropy, §4.2.1
- $\mathcal{F}$ variational free energy, §4.3.2
- $\check{\mathcal{F}}$ Bethe free energy, §4.3.3

Contents

Acknowledgements

I would first like to thank my director, Daniel Bennequin, whose curiosity and breadth of knowledge is only met by his patience, kindness and humility. Working by his side was an invaluable chance and an every day inspiration to look beyond scientific borders. I hope I can one day reach anywhere close to him in the span of these qualities.

This work could never have been possible without the initial year-long turmoil of working groups and discussions I had with Grégoire Sergeant-Perthuis and Juan-Pablo Vigneaux. I feel very lucky that I had them by my side to teach and challenge me. I would like to thank Juan for having shared so much of his knowledge, I hope to count him as a colleague and friend for long, in spite of distances. I wish many luck to Grégoire in finishing his PhD that I am so curious to view complete. He knows how precious his support was to me, in so many different dimensions, and I hope I can make it up to him in the months to come.

I have many precious friends that I could not thank individually. I would only like them to know that I care for them and want to spend more time cultivating the valuable relations too long neglected. They did not need to understand the motives behind this work to be helpful, and they all deserve credit for its completion. I thank each of them warmly. I would however add a particular mention for my saturday library companions, who I was so glad to find and to whom I wish the strongest courage for the difficult work they carry on.

I am obviously very grateful for the wonderful family I always have behind me. Much of the courage I used pursuing this long and demanding work comes from my parents, sisters and brothers in law. I wish to be there for them as much as they have been for me. I would like to dedicate this work to the memory of my grandfather James, who was such an inspiring presence of calm and wisdom in my childhood that I picture him as patient and dedicated a teacher as he was with his grandchildren.

Chapter 1

Homological Algebra

This chapter consists of a short yet self-contained introduction to some of the most remarkable concepts of XX$^{\text{th}}$ mathematics, which happened to take a definite form simultaneously and for the needs of one another: homology and categories. It still seems quite rare that they belong to a common language with the physicist or the computer scientist, and we hope this chapter provides with more than the necessary material[1]. For the more informed reader, this chapter's purpose is to relate our construction to the general theory of simplicial groups.

Categories may be thought of as collections of points and arrows, which describe mathematical objects and their relations, while functors consistently transform categories into other categories. Section 1 reviews these elementary definitions and focuses on providing concrete examples such as the categories of groups, vector spaces, topological spaces, programmable types, *etc.*

The practical use of category theory language is specially remarkable in the characterisation of particular objects by universal properties. Section 2 focuses on the categorical concept of limit, which unifies many constructions such as union and product of sets, sums of vector spaces, inductive and projective limits, *etc.* It should familiarise the reader with commutative diagrams and will help describe homology groups in chapter 2.

Homology provides a general procedure to extract algebraic invariants from topological spaces, while cohomology may be thought of as an abstraction of differential calculus. Section 3 provides with the basic definition of homology groups, which from a purely algebraic point of view, occur in the study of a square-null operator d such that $d^2 = 0$. Motivations are numerous, having probably emerged in formulating physical principles for electromagnetism[2].

1.1 Categories and Functors

1.1.1 Categories

Categories provide with a convenient abstraction of most mathematical constructions and theories. They were introduced by Eilenberg and MacLane [14] to build homological algebra on a rigorous and flexible ground, they have since proven useful in many diverse applications in mathematics, informatics and physics.

[1] Apart from a few proofs, this work should demand little more than a good understanding of the notion of functor, and formulas defining the boundary operator and the differential could talk for themselves.

[2] These principles involve (i) the geometric operator ∂ mapping a subspace to its boundary, which has empty boundary, and (ii) its adjoint differential d acting on fields as gradient, curl, divergence, while d^2 acts as: $\mathrm{div} \circ \mathrm{curl} = \mathrm{curl} \circ \mathrm{grad} = 0$. [[motivate homology - move to 1.3]]

Definition 1.1. *A* category $\mathbf{C}$ *is a class of objects* $A, B, C, \ldots$ *denoted by* $\mathrm{Ob}(\mathbf{C})$ *together with:*

- *a set of arrows* $\mathrm{Hom}(A, B)$ *for every* A, B *in* $\mathrm{Ob}(\mathbf{C})$*,*
- *an identity arrow* $1_A \in \mathrm{Hom}(A, A)$ *for every* A *in* $\mathrm{Ob}(\mathbf{C})$*,*
- *a composite arrow* $gf \in \mathrm{Hom}(A, C)$ *for every* $f \in \mathrm{Hom}(A, B)$ *and* $g \in \mathrm{Hom}(B, C)$

satisfying the following axioms:

(i) Identity: *for every* $f : A \to B$

$$f = f \cdot 1_A = 1_B \cdot f \tag{1.1}$$

(ii) Associativity: *for every* $f : A \to B$*,* $g : B \to C$ *and* $h : C \to D$

$$h(gf) = (hg)f \tag{1.2}$$

The first example of a category is **Set**, the category whose objects are sets and arrows are functions. When each object of $\mathbf{C}$ may be viewed as a set and each arrow $f \in \mathrm{Hom}(A, B)$ induces a function $f \in B^A$ of the underlying sets, the category $\mathbf{C}$ is called concrete, Equivalently, a concrete category $\mathbf{C}$ is a subcategory of **Set**. Although the following definitions make sense in any category, they respectively correspond to bijections, injections, and surjections in **Set** as in most examples of concrete categories.

Definition 1.2. *Let* $\mathbf{C}$ *be a category and* $f : A \to B$ *a morphism.*

- f *is an* isomorphism *if there exists* $g : B \to A$ *such that* $gf = 1_A$ *and* $fg = 1_B$*.*
- f *is a* monomorphism *if for all* X *and* $u, u' : X \to A$*,* $fu = fu'$ *implies* $u = u'$*.*
- f *is an* epimorphism *if for all* Y *and* $v, v' : B \to Y$*,* $vf = v'f$ *implies* $v = v'$*.*

A category may have *terminal* objects, satisfying one of the conditions of the following definition. These very special objects are also called *universal* as a terminal object of a given kind, when it exists, is always defined up to isomorphism. Universal objects are related to the existence of certain *limits*, and describe many fundamental constructions in algebra and geometry[3].

Definition 1.3. *Let* $\mathbf{C}$ *be a category.*

- *an object* I *is* initial *in* $\mathbf{C}$ *if there is a unique arrow* $I \to A$ *for every object* A *in* $\mathbf{C}$*,*
- *an object* F *is* final *in* $\mathbf{C}$ *if there is a unique arrow* $A \to F$ *for every object* A *in* $\mathbf{C}$*,*
- *an object* O *is* null *in* $\mathbf{C}$ *if it is both final and initial.*

Proposition 1.4. *If* T *and* T' *are terminal objects of the same kind, then* T *is isomorphic to* T'*.*

Proof. When T is a terminal object, the axioms imply that 1_T is the unique arrow of $\mathrm{Hom}(T, T)$. If T' is terminal of the same kind, the arrows $T \to T'$ and $T' \to T$ must then compose as 1_T and $1_{T'}$. □

Definition 1.5. *For any category* $\mathbf{C}$*, its* dual *or* opposite *category* $\mathbf{C}^{op}$ *has the same objects as* $\mathbf{C}$ *and, for every arrow* $f : A \to B$ *in* $\mathbf{C}$*, a reversed arrow* $f^{op} : B \to A$ *in* $\mathbf{C}^{op}$*.*

In the following fundamental examples, we give an initial and a terminal object when they exist. In many interesting examples, the set of morphisms between two objects is also an object of the category. This is not true in general and we precise when it is the case[4].

Examples of Categories:

[3] Such constructions are called universal, a few of which being the object of section 1.2. The construction of the tensor algebra $(T(V), \otimes)$ from a vector space V is a classical example that is not exposed here.

[4] The existence of a *hom-object* is a defining property of cartesian categories.

1. A partial order $(X, \geq)$ is a category with a unique arrow $x \to y$ whenever $x \geq y$.

   The identity axiom $x \geq x$ expresses the reflexivity of the order relation, while the transitivity asking that if $x \geq y$ and $y \geq z$, then $x \geq z$, is given by the existence of compositions.

   Initial and final elements correspond to maximal and minimal elements respectively.

2. The category **Set** whose objects are sets and arrows are functions.

   The set of arrows $\mathrm{Hom}(A, B)$ is itself a set, denoted B^A.

   The empty set $\varnothing$ is initial and the point $\bullet = \{\varnothing\}$ is final in **Set**.

3. The category $\mathbf{Alg}_{\mathbb{K}}$ of unital algebras over a field $\mathbb{K}$ whose arrows are algebra morphisms.

   The field $\mathbb{K}$ is both initial and final in $\mathbf{Alg}_{\mathbb{K}}$, it is a null object.

4. The category **Top** whose objects are topological spaces and arrows are continuous functions.

   The point $\bullet$ is final in **Top**.

5. The category **Types** whose objects are variable types and arrows are programs[5].

   The set of arrows $\mathrm{Hom}(\mathtt{a}, \mathtt{b})$ represents the programs with input of type `a` and output of type `b`. It is itself a type, denoted by $(\mathtt{a} \to \mathtt{b})$.

   The empty or *bottom* type $\perp$ is initial, while the unit or *top* type $\top$ is final. An arrow of type $\mathtt{a} \to \perp$ represents a program which does not terminate.

6. For every object X of a category $\mathbf{C}$, the category $\mathbf{C}_X$ above X has arrows $f : A \to X$ as objects. A morphism $\varphi : f \to g$ in $\mathbf{C}_X$ between $f : A \to X$ and $g : B \to X$ is an arrow $\varphi : A \to B$ in $\mathbf{C}$ such that the following diagram commutes:

$$\begin{tikzcd} A \arrow[r, "\varphi"] \arrow[dr, "f"'] & B \arrow[d, "g"] \\ & X \end{tikzcd} \tag{1.3}$$

   There is a similar category $\mathbf{C}^X$ below X defined by reversing arrows, which amounts to reading the above definition in $\mathbf{C}^{op}$.

1.1.2 Functors

Just as morphisms describe relations between objects in a category, functors describe relations between categories by bringing every object to an object and every arrow to an arrow.

Definition 1.6. *A covariant functor T from two categories $\mathbf{C}$ and $\mathbf{C}'$ is defined by:*

– *An object $T(A)$ of $\mathbf{C}'$ for every object A of $\mathbf{C}$*

– *An arrow $T(f) : T(A) \to T(B)$ in $\mathbf{C}'$ for every arrow $f : A \to B$ in $\mathbf{C}$.*

satisfying the following axioms:

(i) $T(1_A) = 1_{T(A)}$,

(ii) $T(fg) = T(f) \cdot T(g)$

Examples of Functors:

[5]This example is motivated by functional programming, although types also aim to provide with a constructivist and rigorous ground for mathematical logic. See for instance the Curry-Howard "proofs as programs" correspondence and Martin-Löf's theory of types.

1. When $(X, \geq)$ and $(Y, \geq)$ are partially ordered sets, a functor from X to Y is an order-preserving map from X to Y.

   This defines the category **Ord** of ordered sets with order-preserving map as morphisms.

2. For every object X of a category $\mathbf{C}$, there are canonical functors $\mathrm{Hom}(-, X)$ and $\mathrm{Hom}(X, -)$ from $\mathbf{C}$ to $\mathbf{Set}$.

   The pull-back of $f : A \to B$ is the map $f^* : \mathrm{Hom}(B, X) \to \mathrm{Hom}(A, X)$ defined by $f^*(u) = u \circ f$ for every $u : B \to X$. $\mathrm{Hom}(-, X)$ is a contravariant functor.

   The push-out of $g : C \to D$ is the map $g_* : \mathrm{Hom}(X, C) \to \mathrm{Hom}(X, D)$ defined by $g_*(v) = g \circ v$ for every $u : X \to C$. $\mathrm{Hom}(X, -)$ is a covariant functor.

3. A contravariant functor from **Top** to **Alg** is defined by associating to each topological space Ω the algebra $C(\Omega)$ of real continuous functions on Ω.

   For every arrow $\varphi : \Omega \to \Omega'$ in **Top**, its pullback $\varphi^* : C(\Omega') \to C(\Omega)$ defined by $\varphi^*(f) = f \circ \varphi$ is an algebra morphism.

   This is a particular case of the previous example, as $C(\Omega) = \mathrm{Hom}(\Omega, \mathbb{R})$ in **Top**.

4. The endofunctor `list` of **Types** associates to each type `a` the type `[a]` of lists whose elements are of type `a`.

   Any function $\texttt{f} : \texttt{a} \to \texttt{b}$ induces a function $\texttt{map(f)} : [\texttt{a}] \to [\texttt{b}]$ returning the list of images under `f` of the input list's elements.

$$\begin{aligned} \texttt{map(f)} &: [\texttt{x},\ \ldots \texttt{xs}] \mapsto [\texttt{f(x)},\ \ldots \texttt{map(f)(xs)}] \\ &: [\,] \mapsto [\,] \end{aligned} \tag{1.4}$$

5. When $\mathbf{C}$ is a concrete category, there is a canonical forgetful functor from $\mathbf{C}$ to $\mathbf{Set}$.

1.1.3 Natural Transformations

Relations between functors are described by natural transformations, also called functor morphisms, as they allow to view functors as objects of a category.

Definition 1.7. *Let T, T' be two functors from $\mathbf{C}$ to $\mathbf{C}'$. A natural transformation Φ from T to T' is a collection of morphisms $\Phi(A) : T(A) \to T(A')$ in $\mathbf{C}'$ for all A in $\mathbf{C}$, such that the diagram:*

$$\begin{array}{ccc} T(A) & \xrightarrow{\Phi(A)} & T'(A) \\ \downarrow{\scriptstyle T(f)} & & \downarrow{\scriptstyle T'(f)} \\ T(B) & \xrightarrow{\Phi(B)} & T'(B) \end{array} \tag{1.5}$$

is commutative for all $f : A \to B$ in $\mathbf{C}$.

Given two categories $\mathbf{C}$ and $\mathbf{C}'$, the functor category $[\mathbf{C}, \mathbf{C}']$ has functors $T : \mathbf{C} \to \mathbf{C}'$ as objects and natural transformations $\Phi : T \to T'$ as morphisms, where:

- The identity $1_T : T \to T$ is defined by $1_T(A) = 1_{T(A)}$ for all object A of $\mathbf{C}$,
- The composition of $\Phi : T \to T'$ and $\Psi : T' \to T''$ is defined by $(\Psi \circ \Phi)(A) = \Psi(A) \circ \Phi(A)$.

Examples: [[Hom$(-, X)$ and Yoneda]] [[adjunction example]]

1.2 Limits and Colimits

Universal properties allow for an abstract definition of limits, unifying some simple constructions such as sums and products of sets with more elaborate ones, such as inductive and projective limits. [[cf Dwyer-Spalinski [5], H. Cartan]]

1.2.1 Definition

A *diagram* of shape $\mathbf{D}$ in a category $\mathbf{C}$ consists of a functor $C : \mathbf{D} \to \mathbf{C}$ where $\mathbf{D}$ is a small[6] category describing the diagram shape. It is a collection $C(f) : C_\alpha \to C_\beta$ of arrows in $\mathbf{C}$ for all $f : \alpha \to \beta$ in $\mathbf{D}$.

A *cone* over C in $\mathbf{C}$ is an object S of $\mathbf{C}$ and a collection of morphisms $\varphi_\alpha : S \to C_\alpha$ for all $\alpha \in \mathbf{D}$ such that the following diagram in $\mathbf{C}$ commutes for every $f : \alpha \to \beta$ in $\mathbf{D}$:

$$\begin{tikzcd}
& S \arrow[dl, "\varphi_\alpha"'] \arrow[dr, "\varphi_\beta"] & \\
C_\alpha \arrow[rr, "C(f)"] & & C_\beta
\end{tikzcd} \tag{1.6}$$

In other words, (S, φ) extends the functor C to the category $\mathbf{D}_0$ preceding $\mathbf{D}$ with an initial element. A morphism between two cones (S, φ) and (S', φ') over C is a morphism $\psi : S \to S'$ in $\mathbf{C}$ such that the following diagram commutes for all $\alpha \in \mathbf{D}$:

$$\begin{tikzcd}
S \arrow[r, "\psi"] \arrow[dr, "\varphi_\alpha"'] & S' \arrow[d, "\varphi'_\alpha"] \\
& C_\alpha
\end{tikzcd} \tag{1.7}$$

A *limit* of a diagram $C : \mathbf{D} \to \mathbf{C}$ is a final element (L, λ) in the category of cones over C. When a limit L exists, it is defined up to isomorphisms in $\mathbf{C}$ by the universal property requiring that for every cone (S, φ) over C there be a unique morphism $\psi : S \to L$ factorising S through L.

$$\begin{tikzcd}
& S \arrow[d, dashed, "\psi"] \arrow[ddl, bend right, "\varphi_\alpha"'] \arrow[ddr, bend left, "\varphi_\beta"] & \\
& L \arrow[dl, "\lambda_\alpha"'] \arrow[dr, "\lambda_\beta"] & \\
C_\alpha \arrow[rr, "C(f)"] & & C_\beta
\end{tikzcd} \tag{1.8}$$

Definition 1.8. *A category* $\mathbf{C}$ *is called* complete *when every small diagram* $C : \mathbf{D} \to \mathbf{C}$ *has a limit. When it exists, we denote by* $\lim_{\mathbf{D}} C$ *the limit of* C *defined up to isomorphism.*

A *colimit* of a diagram $C : \mathbf{D} \to \mathbf{C}$, is reciprocally defined by reversing arrows. It is an initial element in the category of cones under C, made of extensions of C to the category $\mathbf{D}_1$ appending a final element to $\mathbf{D}$. When it exists, the universal property satisfied by a colimit L' of C is represented by the diagram:

$$\begin{tikzcd}
C_\alpha \arrow[rr, "C(f)"] \arrow[dr, "\lambda'_\alpha"] \arrow[ddr, bend right, "\varphi'_\alpha"'] & & C_\beta \arrow[dl, "\lambda'_\beta"'] \arrow[ddl, bend left, "\varphi'_\beta"] \\
& L' \arrow[d, dashed, "\psi'"] & \\
& S' &
\end{tikzcd} \tag{1.9}$$

[6] A category is said *small* when the class of its objects actually forms a set.

Definition 1.9. *A category* **C** *is called* cocomplete *when every small diagram* $C : \mathbf{D} \to \mathbf{C}$ *has a colimit. When it exists, we denote by* $\operatorname{colim}_{\mathbf{D}} C$ *the colimit of* C *defined up to isomorphism.*

We give a few examples of limits below, although the following paragraphs will illustrate much better the universality of limits.

Examples:

1. When **D** is the empty category, limits and colimits of the empty diagram in **C** are initial and final objects of **C** respectively.

2. Any object A of a category **C** defines a diagram, whose shape **D** is the category with only one object and its identity map. The limit and colimit are both represented by $1_A : A \to A$.

3. Let $u \in \mathbb{R}^{\mathbb{N}}$ denote a sequence of real numbers. The induced set map defines a functor between the partial orders $(\mathcal{P}(\mathbb{N}), \subseteq)$ and $(\mathcal{P}(\mathbb{R}), \subseteq)$ associating to a subset $S \subseteq \mathbb{N}$ its direct image $u(S) \subseteq \mathbb{R}$.

   Consider now its restriction $\tilde{u}$ to subsets of the form $S_n = \{n, n+1, \dots\}$ for $n \in \mathbb{N}$. The limit of $\tilde{u}$ is the largest subset $L \subseteq \mathbb{R}$ such that $L \subseteq u(S_n)$ for all n. It consists of all the accumulation points of u.

4. in some abelian category e.g. **Vect** [[Kernel and Image]].

1.2.2 Sums and Products

Any two objects A, A' in a category **C** define a diagram of shape a category **D** with two objects and identities as morphisms. When it exists, a final cone over A and A' defines their *product* $A \times A'$, satisfying the universal property depicted by:

$$\begin{tikzcd} & & A \\ X \arrow[r, dashed] \arrow[urr, bend left] \arrow[drr, bend right] & A \times A' \arrow[ur, "\pi"] \arrow[dr, "\pi'"'] & \\ & & A' \end{tikzcd} \tag{1.10}$$

Their *sum* or *coproduct* $A \sqcup A'$ is reciprocally defined as an initial cone under A and A'.

$$\begin{tikzcd} A \arrow[dr, "j"] \arrow[drr, bend left] & & \\ & A \sqcup A' \arrow[r, dashed] & Y \\ A' \arrow[ur, "j'"'] \arrow[urr, bend right] & & \end{tikzcd} \tag{1.11}$$

Examples:

1. In **Set** and **Top**, the product of A and A' is their cartesian product $A \times A'$, while their sum is the disjoint union $A \sqcup A'$.

2. In **Grp** the product and coproduct of G and G' coincide as $G \times G'$. In **Vect**, the product and the sum of V and V' also coincide as $V \oplus V'$. This is a general property of abelian categories.

3. In the category **Com** of unital commutative algebras, the coproduct of A and A' is their tensor product $A \otimes A'$, with canonical injections $1 \otimes -$ and $- \otimes 1$.

1.2.3 Pushouts and Pullbacks

Consider the diagram shape given by $\mathbf{D} : \alpha \to \beta \leftarrow \alpha'$. The limit of this kind of diagrams, when it exists, defines the pullback or fibered product $A \times_B A'$ of A and A' over B:

$$\begin{tikzcd} & & A \arrow[dr, "v"] & \\ X \arrow[r, dashed] \arrow[urr, bend left] \arrow[drr, bend right] & A \times_B A' \arrow[ur, "\pi"] \arrow[dr, "\pi'"'] & & B \\ & & A' \arrow[ur, "v'"'] & \end{tikzcd} \tag{1.12}$$

The pushout or amalgated sum $A \sqcup_B A'$ of A and A' over B, when it exists, is defined by the dual universal property:

$$\begin{tikzcd} & A \arrow[dr, "j"] \arrow[drr, bend left] & & \\ B \arrow[ur, "u"] \arrow[dr, "u'"'] & & A \sqcup_B A' \arrow[r, dashed] & Y \\ & A' \arrow[ur, "j'"'] \arrow[urr, bend right] & & \end{tikzcd} \tag{1.13}$$

Note that the morphisms are implicit in the notations $A \times_B A'$ and $A \sqcup_B A'$ although the resulting objects depends on them.

Examples:

1. In **Set** as in **Top**, the fibered product of $v : A \to B$ and $v' : A' \to B$ is defined by:

$$A \times_B A' = \{(x, x') \in A \times A' \mid v(x) = v'(x')\} \tag{1.14}$$

   while the pushout of $u : B \to A$ and $u' : B \to A'$ is the quotient:

$$A \sqcup_B A' = A \sqcup A' \,/\, \big(u(y) \sim u'(y)\big)_{y \in B} \tag{1.15}$$

2. In the category of commutative algebras **Com**, the pushout $A \otimes_B A'$ of $u : B \to A$ and $u' : B \to A'$ is the quotient of $A \otimes A'$ by the equivalence relation generated by the action of all $b \in B$:

$$\big(a \cdot u(b) \otimes a'\big) \sim \big(a \otimes u'(b) \cdot a'\big) \tag{1.16}$$

1.2.4 Sheaves and Cosheaves

Definition 1.10. *A* presheaf *over a topological space* $(\Omega, \mathcal{T}_\Omega)$ *is a functor* $F : \mathcal{T}_\Omega \to \mathbf{Set}$*, associating:*

- *to each open subset* $U \subseteq \Omega$ *a set* $F(U)$ *of sections over* U*,*
- *to each ordered pair* $V \subseteq U$ *a restriction map* $\rho_{VU} : F(U) \to F(V)$*.*

A presheaf is thus contravariant[7] *from* $(\mathcal{T}_\Omega, \subseteq)$ *to* $\mathbf{Set}$*, and covariant from* $\mathcal{T}_\Omega^{op} = (\mathcal{T}_\Omega, \supseteq)$ *to* $\mathbf{Set}$*.*

The fiber of F at $x \in \Omega$ is defined as the colimit of F over the partial order $(\mathcal{V}_x, \supseteq)$ of neighborhoods containing x and denoted by F_x:

$$F_x = \operatorname*{colim}_{\mathcal{V}_x} F \tag{1.17}$$

For all U containing x, the image $s(x)$ of a section $s \in F(U)$ under the canonical map $F(U) \to F_x$ is called the germ of s at x.

When $\mathcal{U}_{\Omega'}$ is an open covering of $\Omega' \subseteq \Omega$ closed under intersection, the limit of F over $(\mathcal{U}_{\Omega'}, \supseteq)$ is the set of compatible sections on $\mathcal{U}_{\Omega'}$:

$$\lim_{\mathcal{U}_{\Omega'}} F \simeq \left\{ (s_U) \in \prod_{U \in \mathcal{U}_{\Omega'}} F(U) \;\middle|\; \forall U, V \in \mathcal{U}_{\Omega'} \;\; (s_U)_{|U \cap V} = (s_V)_{|U \cap V} \right\} \tag{1.18}$$

A presheaf F over Ω is a sheaf when for all such covering $\mathcal{U}_{\Omega'}$ of Ω', the sections of F over Ω' are in one-to-one correspondence with the compatible sections on $\mathcal{U}_{\Omega'}$.

Definition 1.11. *A* sheaf F *over* Ω *is a presheaf such that:*

$$F(\Omega') \simeq \lim_{\mathcal{U}_{\Omega'}} F \tag{1.19}$$

for every open covering $\mathcal{U}_{\Omega'}$ *of* $\Omega' \subseteq \Omega$ *closed under intersection.*

Morphisms of sheaves are defined as natural transformations of functors, and the category of sheaves over Ω is naturally defined by inclusion in the functor category $[\mathcal{T}_\Omega^{op}, \mathbf{Set}]$. When F is a sheaf, it is customary to denote it by $F(\Omega)$ with a slight abuse[8] of notations. Note that the sheaf axiom implies that the diagram:

$$\begin{tikzcd} F(U \cup V) \arrow[r] \arrow[d] & F(V) \arrow[d] \\ F(U) \arrow[r] & F(U \cap V) \end{tikzcd} \tag{1.20}$$

is a pullback square in $\mathbf{Set}$, for all open $U, V \subseteq \Omega$. When limits exist, sheaves may be defined in any category, and one is often mostly interested with sheaves of abelian groups, rings, modules, *etc.* fibered products and limits coinciding with those coming from sets.

There is a dual notion of cosheaf, although seemingly less common.

Definition 1.12. *A* pre-cosheaf *over a topological space* $(\Omega, \mathcal{T}_\Omega)$ *in a category with colimits* $\mathbf{C}$ *is a covariant functor* $G : \mathcal{T}_\Omega \to \mathbf{C}$ *associating:*

- *to each open subset* $U \subseteq \Omega$ *an object* $G(U)$ *in* $\mathbf{C}$
- *to each ordered pair* $V \subseteq U$ *a morphism* $j_{UV} : G(V) \to G(U)$ *in* $\mathbf{C}$

Definition 1.13. *A* cosheaf *over* Ω *is a pre-cosheaf such that:*

$$G(\Omega') \simeq \operatorname*{colim}_{\mathcal{U}_{\Omega'}} G \tag{1.21}$$

for every open covering $\mathcal{U}_{\Omega'}$ *of* $\Omega' \subseteq \Omega$ *closed under intersection.*

[7] As a subcategory of $\mathbf{Set}$, this is the right choice of arrows on $\mathcal{T}_\Omega$.
[8] $F(U)$ may not be the image of $F(\Omega)$ under $\rho_{U\Omega}$.

The category of $\mathbf{C}$-valued cosheaves on Ω is similarly defined as a subcategory of $[\mathcal{T}_\Omega, \mathbf{C}]$. The cosheaf axiom implies that $G(\varnothing)$ is initial in $\mathbf{C}$, and that the diagram:

$$\begin{array}{ccc} G(U \cap V) & \longrightarrow & G(V) \\ \downarrow & & \downarrow \\ G(U) & \longrightarrow & G(U \cup V) \end{array} \tag{1.22}$$

is a pushout square in $\mathbf{C}$ for all open $U, V \subseteq \Omega$.

Examples:

1. The space $C(\Omega)$ of real continuous functions over a topological space Ω defines the fundamental example of a sheaf of algebras, with obvious restrictions.

2. Suppose given a numerable set Ω with finite sets E_i for all $i \in \Omega$, and let $E_\alpha = \prod_{i \in \alpha} E_i$ for all $\alpha \subseteq \Omega$ Then E_Ω defines a sheaf of sets over the discrete topological space Ω.

3. When $f \in C(\Omega, \Omega')$ is a continuous map of topological spaces, the map $U' \mapsto f^{-1}(U')$ is a cosheaf of sets over Ω'.

4. Given E_Ω as above, let $A_\alpha = \mathbb{R}^{E_\alpha}$ denote the algebra of real functions on E_α, for all $\alpha \subseteq \Omega$. Then for all $\beta, \beta' \subseteq \Omega$ we have $A_{\beta \cup \beta'} = A_\beta \otimes_{A_{\beta \cap \beta'}} A_{\beta'}$ and A is a cosheaf of algebras over Ω.

1.3 Differential Structures

Simplices generalise the usual figures of point, segment, triangle, tetrahedron, *etc.* They correspond to elementary objects in topology and geometry, as any n-dimensional manifold may be triangulated by simplices of dimension n. They also carry the fundamental combinatorial properties of differential calculus, which will motivate the much more algebraic definition of simplicial objects.

1.3.1 Simplicial Complexes

Given an affine space E and $n+1$ affinely independent points $P_0, \dots, P_n$, the convex polyhedron generated by those points is called the n-simplex of vertices P_i:

$$S = \Big\{ M \in E \;\Big|\; \overrightarrow{OM} = \sum_{i=0}^{n} \lambda_i \, \overrightarrow{OP}_i \quad \text{for } \lambda_j \geq 0 \text{ and } \sum_i \lambda_i = 1 \Big\} \tag{1.23}$$

Note that for any choice of origin $O \in E$, the barycentric coordinates λ_i of M are uniquely determined independently of O. Barycentric coordinates identify points of a simplex with probability measures on its set of vertices.

Definition 1.14. *The* topological n-simplex *over the set of $n+1$ vertices Ω is defined by:*

$$|S_\Omega| = \Big\{ \lambda : \Omega \to \mathbb{R}^+ \;\Big|\; \sum_{i \in \Omega} \lambda_i = 1 \Big\} \tag{1.24}$$

It is a convex subset of $\mathbb{R}^\Omega$, and a topological space for the topology induced by $\mathbb{R}^\Omega$.

Let S denote the simplex of vertices Ω. A q-face S' of S is defined by a set of $q+1$ vertices $\Omega' \subseteq \Omega$, such that $S' \subseteq S$ consists of the barycentric coordinates λ that vanish on $\Omega - \Omega'$. The interior of S for the topology of $\mathbb{R}^\Omega$ consists of all the non-vanishing barycentric coordinates $\lambda > 0$, and coincides with the complement of all the proper faces of S within S.

There is an equivalence $\Omega \mapsto |S_\Omega|$ between the categories of finite sets and topological simplices ordered by inclusion, where $\mathcal{P}(\Omega)$ is identified with the set of faces of $|S_\Omega|$. It is therefore natural to identify a simplex with its set of vertices. Every $f : \Omega \to \Omega'$ induces a continuous map $f_* : |S_\Omega| \to |S'_\Omega|$ defined by:

$$(f_*\lambda)_j = \sum_{i \in \Omega \mid f(i)=j} \lambda_i \tag{1.25}$$

and sending every face of $|S_\Omega|$ to a face of $|S_{\Omega'}|$. Simplices hence define a covariant functor $\mathbf{Set}_\mathrm{f} \to \mathbf{Top}$, which could be extended to the larger category of measurable spaces.

Definition 1.15. *An* abstract simplicial complex (Ω, K) *is a finite set of vertices* Ω *together with a collection of faces* $K \subseteq \mathcal{P}(\Omega)$ *made of finite subsets of* Ω*, such that for all* $\alpha \in K$*, every* $\beta \subseteq \alpha$ *is also in* K*.*

The n-skeleton K_n of a simplicial complex K consists of all its n-faces, *i.e.* faces having exactly $n+1$ vertices. The abstract simplex S_Ω is the trivial simplicial complex $(\Omega, \mathcal{P}(\Omega))$ having all possible faces. A simplicial complex (Ω, K) is essentially a reunion of abstract simplices S_α, for α in K.

The topological space $|K|$ associated to a simplicial complex (Ω, K) is obtained by gluing the simplices of K along their intersecting faces:

$$|K| = \operatorname*{colim}_{\alpha \in K} |S_\alpha| \tag{1.26}$$

The inductive limit, taken over the functor $\alpha \mapsto |S_\alpha|$, is essentially a reunion in the ambient topological simplex $|S_\Omega|$.

Definition 1.16. *A* simplicial morphism $f : (\Omega, K) \to (\Omega', K')$ *is a map of sets* $f : \Omega \to \Omega'$ *such that for all face* α *of* K*, its image* $f(\alpha) \subseteq \Omega'$ *is a face of* K'*.*

Simplical complexes form a category $\mathbf{KS}$. A simplicial map $f : K \to K'$ induces for all $\alpha \in K$ a continuous function $f_*^\alpha : |S_\alpha| \to |S_{f(\alpha)}|$. These maps extend to a map of topological spaces $f_* : |K| \to |K'|$ and topological realisation defines a covariant functor $\mathbf{KS} \to \mathbf{Top}$.

The definition of a simplicial complex K with vertices in Ω could be naturally extended when Ω is numerable and more generally, when Ω is a measurable space.

1.3.2 Simplicial Objects

For every $n \in \mathbb{N}$, denote by $[n] = \{0, \dots, n\}$ the total order with $n+1$ elements, and by $\big([n], S_n\big)$ the abstract n-simplex with ordered vertices.

Definition 1.17. *The* simplicial category $\mathbf{\Delta}$ *is defined by:*

- *objects:* $[n]$ *for any* n *in* $\mathbb{N}$*,*
- *morphisms:* $[m] \to [n]$ *order-preserving map.*

Equivalently, $\mathbf{\Delta}$ *is the subcategory of* **Ord** *with objects* $[n]$ *for* $n \in \mathbb{N}$*.*

An ordered n-simplex σ in a simplicial complex (Ω, K) is a simplicial map $\sigma : \big([n], S_n\big) \to (\Omega, K)$. The ordered n-simplex $\sigma = (\sigma_0, \dots, \sigma_n)$ is said non-degenerate when $\sigma_i \neq \sigma_j$ for $i \neq j$ and the underlying set map is injective.

A simplicial complex (Ω, K) then defines a contravariant functor $\vec{K} : \mathbf{\Delta}^{op} \to \mathbf{Set}$ where:

- $\vec{K}_n = \vec{K}\big([n]\big)$ is the set of ordered n-simplices in K,
- $t^* : \vec{K}_n \to \vec{K}_m$ is defined for all $t : [m] \to [n]$ by the pullback $\sigma \mapsto \sigma \circ t$.

Denoting by $|\sigma| = \mathrm{Im}(\sigma)$ the image of an ordered n-simplex, we have $|t^*\sigma| \subseteq |\sigma|$ for all $t : [m] \to [n]$.

A simplicial map $f : (\Omega, K) \to (\Omega', K')$ induces a natural transformation $f_* : \vec{K} \to \vec{K}'$, defined by the pushforward $\sigma \mapsto f \circ \sigma$. The natural transformation f_* is a morphism in the category of functors $[\mathbf{\Delta}^{op}, \mathbf{Set}]$ and the assignment $(\Omega, K) \mapsto \vec{K}$ defines a covariant functor $\mathbf{KS} \to [\mathbf{\Delta}^{op}, \mathbf{Set}]$. The set of ordered simplices $\vec{K}$ of a simplicial complex (Ω, K) is the fundamental example of a simplicial set, and motivates the following more general definition of simplicial objects in an arbitrary category.

Definition 1.18. *A* simplicial object *in a category* $\mathbf{C}$ *is a functor* $X : \mathbf{\Delta}^{op} \to \mathbf{C}$.

Simplicial objects in $\mathbf{C}$ form a category $[\mathbf{\Delta}^{op}, \mathbf{C}]$ with natural transformations as morphisms. Given a simplicial object X, we denote $X\big([n]\big)$ by X_n and for $t : [m] \to [n]$ we denote $X(t)$ by $t^* : X_n \to X_m$.

A category $\mathbf{C}$ defines a simplicial set $N(\mathbf{C})$ called its nerve, defined by:

$$N_n(\mathbf{C}) = \mathrm{Hom}([n], \mathbf{C}) \tag{1.27}$$

An ordered n-simplex $\sigma \in N_n(\mathbf{C})$ is a covariant functor $\sigma : \big([n], \leq\ \big) \to \mathbf{C}$. Equivalently, it is a commutative diagram of $n+1$ objects $\sigma_0, \dots, \sigma_n$ with arrows $\sigma_{ij} : \sigma_i \to \sigma_j$ for all $i < j$.

Given a simplicial set $X : \mathbf{\Delta}^{op} \to \mathbf{Set}$, the group of chains $\mathbb{Z}[X] : \mathbf{\Delta}^{op} \to \mathbf{Ab}$ is the simplicial abelian group freely generated by X, with:

$$\mathbb{Z}_n[X] = \bigoplus_{\sigma \in X_n} \mathbb{Z} \cdot e_\sigma \tag{1.28}$$

and every map $t : [m] \to [n]$ inducing a group morphism $t^* : \mathbb{Z}_n[X] \to \mathbb{Z}_m[X]$ defined by $t^*(e_\sigma) = e_{t^*\sigma}$. Chain groups thus define a functor $\mathbb{Z}[\,\cdot\,]$ from simplicial sets to simplicial abelian groups.

For all $n \in \mathbb{N}$ and $0 \leq i \leq n$, consider the i-th face map $\partial^i : [n-1] \to [n]$ defined as the injection of $[n-1]$ whose image misses the i-th vertex in $[n]$. Note that face maps generate all injective maps of $\mathbf{\Delta}$ and satisfy the following fundamental commutation relations:

$$\partial^i \circ \partial^j = \partial^{j-1} \circ \partial^i \quad \text{for} \quad i < j \tag{1.29}$$

where the sums of indices $i + j$ and $(j-1) + i$ have reversed parity.

Every simplicial abelian group $G : \mathbf{\Delta}^{op} \to \mathbf{Ab}$ has a canonical boundary operator $\partial : G \to G$ where for all degree n, the map $\partial : G_n \to G_{n-1}$ is defined by:

$$\partial = \sum_{i=0}^{n} (-1)^i \ (\partial^i)^* \tag{1.30}$$

The commutation relations of face maps imply that $\partial^2 = \partial \circ \partial = 0$.

Let $G = \mathbb{Z}\big[\vec{K}\big]$ be the group of chains in a simplicial complex K. If σ is an oriented simplex in K, then $\partial\sigma$ is the oriented boundary of σ. The fundamental equation $\partial^2 = 0$ reflects the geometric fact that the boundary of a boundary is empty.

Definition 1.19. *A* cosimplicial object *in a category* $\mathbf{C}$ *is a functor* $Y : \mathbf{\Delta} \to \mathbf{C}$.

Every cosimplicial abelian group $F : \mathbf{\Delta} \to \mathbf{Ab}$ has a canonical coboundary operator $d : F \to F$, defined by the family of maps $d^n : F^n \to F^{n+1}$ with:

$$d^n = \sum_{i=0}^{n+1} (-1)^i \ (\partial^i)_* \tag{1.31}$$

The operator d is called the differential of F and satisfies $d^2 = d \circ d = 0$.

Given a covering $\mathcal{U}$ of a topological space Ω, its Čech nerve is the simplicial set:

$$\check{\mathcal{U}}_n = \{\sigma : [n] \to \mathcal{U} \mid U_\sigma = \sigma_0 \cap \cdots \cap \sigma_n \neq \varnothing\} \tag{1.32}$$

For every $\sigma \in \check{\mathcal{U}}_n$ and $t : [m] \to [n]$, the associated simplex $t^*\sigma$ in $\check{\mathcal{U}}_m$ satisfies $\mathrm{Im}(t^*\sigma) \subseteq \mathrm{Im}(\sigma)$ in $\mathcal{U}$ so that the intersection $U_{t^*\sigma}$ contains U_σ in Ω. When F is a sheaf of abelian groups over Ω, it defines a cosimplicial abelian group $\check{F}(\mathcal{U}) : \mathbf{\Delta} \to \mathbf{Ab}$ where:

$$\check{F}^n(\mathcal{U}) = \bigoplus_{\sigma \in \check{\mathcal{U}}_n} F(U_\sigma) \tag{1.33}$$

For $t : [m] \to [n]$ the map $t_* : \check{F}^m(\mathcal{U}) \to \check{F}^n(\mathcal{U})$ is defined for every $f \in \check{F}^m(\mathcal{U})$ by:

$$(t_* f)_\sigma = (f_{t^*\sigma})_{|U_\sigma} \tag{1.34}$$

and $\check{F}(\mathcal{U})$ is called the group of Čech cochains of F in $\mathcal{U}$.

1.3.3 Homology

Definition 1.20. *A* differential group (G, ∂) *is an abelian group G together with an endomorphism $\partial : G \to G$ satisfying $\partial^2 = 0$. The morphism ∂ is called the* boundary operator *of G.*

Given a simplicial complex (Ω, K), a fundamental example is given by the group of chains $\big(\mathbb{Z}\big[\vec{K}\big], \partial\big)$. When (G, ∂) is any differential group, its boundary operator defines the two following subgroups:

- a *cycle* is an element of $Z(G) = \mathrm{Ker}(\partial)$,
- a *boundary* is an element of $B(G) = \mathrm{Im}(\partial)$,

The rule $\partial^2 = 0$ implies that every boundary is a cycle and $B(G) \subseteq Z(G)$.

Definition 1.21. *The* homology group *of (G, ∂) is the quotient group $H(G) = \mathrm{Ker}(\partial)/\mathrm{Im}(\partial)$.*

More generally, $x, x' \in G$ are said homologous when there exists $y \in G$ such that $x' = x + \partial y$, we then write $x \sim x'$ and denote by $[x]$ the class of $x \in G$ for this equivalence relation. Homology groups consist of the equivalence classes of cycles.

When (Ω, K) is a simplicial complex, the homology of its group of chains is denoted by $H(K; \mathbb{Z})$.

Definition 1.22. *A* morphism of differential groups $f : (G, \partial) \to (G', \partial')$ *is a map of abelian groups $f : G \to G'$ such that the following diagram is commutative:*

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ \downarrow{\scriptstyle\partial} & & \downarrow{\scriptstyle\partial'} \\ G & \xrightarrow{f} & G' \end{array} \tag{1.35}$$

In particular, f sends $Z(G)$ in $Z(G')$ and $B(G)$ in $B(G')$.

Differential groups form a category which we denote by $\mathbf{Ab}_\partial$. Given a map of simplicial sets $f : X \to X'$, the group morphism $f_* : \mathbb{Z}[X] \to \mathbb{Z}[X']$ commutes with boundary operators and chain groups define a covariant functor from simplicial sets to $\mathbf{Ab}_\partial$.

The following proposition expresses that homology defines a functor $H : \mathbf{Ab}_\partial \to \mathbf{Ab}$, and we get in particular by left composition with the group of chains, a functor $\mathbf{KS} \to \mathbf{Ab}$. Functoriality is a fundamental property of homology, as it was introduced to yield algebraic invariants of topological spaces, the homology groups of two homeomorphic spaces being isomorphic.

Proposition 1.23. *A map of differential groups $f : (G, \partial) \to (G', \partial')$ induces a morphism in homology denoted by $[f] : H(G) \to H(G')$.*

Proof. A morphism f sends $Z(G)$ to $Z(G')$ and $B(G)$ to $B(G')$, hence $f : Z(G) \to Z(G')/B(G')$ factors through $Z(G) \to Z(G)/B(G)$. □

Definition 1.24. *Two maps $f, f' : (G, \partial) \to (G', \partial')$ are* homotopic *when there exists a map $h : G \to G'$ of abelian groups such that $f - f' = \partial' h + h\partial$.*

We write $f \sim f'$ when f and f' are homotopic. When h is a homotopy from f to f', the sum of the two outer paths coincides with the inner arrow of the following diagram:

$$\begin{tikzcd} & G' \arrow[d, "\partial'"] \\ G \arrow[ur, "h"] \arrow[r, "f-f'"] \arrow[d, "\partial"] & G' \\ G \arrow[ur, "h"'] & \end{tikzcd} \tag{1.36}$$

The homotopy relationship vanishes in homology.

Proposition 1.25. *When $f, f' : (G, \partial) \to (G', \partial')$ are homotopic, their induced maps in homology $[f], [f'] : H(G) \to H(G')$ coincide.*

Proof. Let $h : G \to G'$ denote a homotopy between f and f', so that $f - f' = \partial' h + h\partial$.
For all $z \in Z(G)$ we have $f(z) - f'(z) = \partial' h(z) \in B(G)$, so that $[f(z)] = [f'(z)]$ in $H(G')$. □

When (G', ∂) is a subgroup of (G, ∂) with $\partial(G') \subseteq G'$, the boundary operator of G factors to the quotient onto G/G' where it induces a boundary ∂'. The relative homology of the pair (G, G') is defined as $H(G, G') = H(G/G')$ More precisely, let us denote by:

- $Z(G, G') = \partial^{-1}(G')$ the set of relative cycles,
- $B(G, G') = G' + \partial G$ the set of relative boundaries.

Then $H(G, G')$ is the quotient of $Z(G, G')$ by $B(G, G')$. Noting that ∂ sends $Z(G, G')$ to $Z(G')$ and $B(G, G')$ to $B(G')$, there is a canonical morphism in homology $[\partial] : H(G, G') \to H(G')$.

The injection $f : G' \to G$ and the projection $g : G \to G/G'$ induce maps in homology:

$$H(G') \xrightarrow{[f]} H(G) \xrightarrow{[g]} H(G, G') \tag{1.37}$$

Proposition 1.26. *Given a subgroup (G', ∂) of (G, ∂), the homology sequence of the pair (G, G') is exact:*

$$\begin{tikzcd} & H(G) \arrow[dr, "{[g]}"] & \\ H(G') \arrow[ur, "{[f]}"] & & H(G, G') \arrow[ll, "{[\partial]}"] \end{tikzcd} \tag{1.38}$$

Proof. Prove that:

- $\mathrm{Im}([\partial]) = \mathrm{Ker}([f])$,
- $\mathrm{Im}([f]) = \mathrm{Ker}([g])$,
- $\mathrm{Im}([g]) = \mathrm{Ker}([\partial])$.

□

When (G,∂) and (G',∂') are graded differential groups, the endomorphisms from G to G' form a graded differential complex $\big(L(G,G'),\delta\big)$ with:

$$L_n(G,G') = \prod_k L(G_k, G_{k+n}) \tag{1.39}$$

where the boundary of $f \in L_n(G,G')$ is defined by:

$$\delta(f) = \partial' f - (-1)^n f\partial \tag{1.40}$$

In particuar, a 0-cycle is a morphism of differential groups:

$$\delta(f) = 0 \quad \Leftrightarrow \quad \partial' f = f\partial \tag{1.41}$$

While a 0-boundary is homotopic to zero:

$$\delta(h) = \partial' h + h\partial \tag{1.42}$$

Chapter 2

Statistical Systems

The main goal of this chapter is to introduce the differential complex $A_\bullet(X)$ of local observables and compute its homology. *Statistical system* should here be understood in a double sense. In physics, it commonly refers to a pair (E_Ω, H_Ω) where $E_\Omega = \prod_{i \in \Omega} E_i$ is the configuration space of the system and $H_\Omega : E_\Omega \to \mathbb{R}$ is the hamiltonian, which induces the statistics as a function of temperature. Our localisation procedure on a hypergraph $X \subseteq \mathcal{P}(\Omega)$ also leads to an *inductive system* $(A_\alpha)_{\alpha \in X}$ of algebras, where X shall be chosen large enough for the hamiltonian to decompose as a sum of local interactions potentials $(h_\alpha)_{\alpha \in X}$.

With an emphasis on functoriality, section 1 presents natural constructions leading to statistics, probability densities being defined from the algebra of observables[1]. This approach, common in the field of operator algebras, somehow differs from the usual probabilistic definitions, but has the considerable advantage of unifying classical statistics with quantum states.

The construction of the complex $A_\bullet(X)$ essentially lifts the functor $A : X^{op} \to \mathbf{Alg}$ of local observables to a functor on the nerve $\check{A} : N(X)^{op} \to \mathbf{Alg}$, associating to any ordered chain $\alpha \supseteq \ldots \supseteq \gamma$ a copy of A_γ. The simplicial structure of the nerve will provide $A_\bullet(X)$ with a boundary operator δ, whose action $A_1(X) \to A_0(X)$ will describe the dynamic of message-passing algorithms. Section 2 describes such constructions in their generality[2].

Specialising to the setting where $A_\alpha = \mathbb{R}^{E_\alpha}$ is the algebra of functions on the cartesian product $E_\alpha = \prod_{i \in \alpha} E_i$, we compute the homology of $A_\bullet(X)$ in section 3. Its acyclicity shall come as a consequence of the supposedly well-known *interaction decomposition* theorem. This fundamental result is recalled and proved by means of harmonic analysis, following an original proof by Matŭs.

2.1 Global Statistics

The main purpose of this short and informal section is to introduce the fundamental structures involved with statistics, along with their notations:

Microstates	Observables	Measures	States
E	$A = C(E)$	A^*	$\Delta \subseteq A^*$
Top	**Alg**	**Vect**	**Conv**

[1]Instead of introducing measureable spaces and probability measures before observables as measurable functions.

[2]Thanks to a thorough investigation by D. Bennequin, we became recently aware that the construction of a complex by the same lifting to the nerve was already considered by Grothendieck and Verdier in SGA-4-V [29], see also [18].

In this picture, most columns are related by functors. In particular, the space Δ of statistical states can be functorially defined from a set of microstates E or from a C*-algebra of observables A.

A fundamental component of statistical physics is the Gibbs state map $\rho : A \to \Delta$ defined by:

$$\rho(H) = \left[\, \mathrm{e}^{-H} \,\right] \tag{2.1}$$

where H computes the energy of the system and the bracket denotes normalisation. Both theoretical and computational problems with the normalisation factor $Z(H) = \int_E \mathrm{e}^{-H}$ arise when E gets large. It involves a computation of exponential complexity in the dimension of E, while the study of phase transitions requires to let the number of atoms go to infinity.

One may thus be lead to give up global observations, and decide that only small enough regions of the global system may be simultaneously observed. This approach underlies the present work and will rely heavily on functoriality. What follows could then be thought of as a description of local models for statistics, which one may join consistently to cover larger systems. One should hope that such a localisation procedure still efficiently describes collective phenomena.

We also hope that the following general discussion may give perspective on possible extensions of the present work to the continuous and quantum settings. See [17, 26] for good quantum references.

2.1.1 Microscopic States

In classical probability theory, one starts with a measurable set E describing all possible outcomes of an experiment. Consider for instance a physical system of N atoms, labelled by $i, j, \dots$, each of which having degrees of freedom in E_i. A configuration of the full system is given by an element of the cartesian product:

$$E = \prod_{i=1}^{N} E_i \tag{2.2}$$

A configuration is also called a microscopic state of the system. This is a classical point of view, only valid for statistical systems at high enough temperatures[3].

In what follows, we shall keep the notation E for configuration spaces. In most applications covered by this thesis, it is enough to view E as an object of the category $\mathbf{Set}_\mathrm{f}$ of finite sets. However, some of our constructions may gain generality by considering topological spaces in **Top**.

Starting with a set of microscopic states is a classical point of view, although somehow artificial and arbitrary. It is only valid at high enough temperatures as quantum mechanics provides an argument of nature against the fiction of microscopic states. Classical probabilities and quantum states will both be naturally described by the states of an algebra of observables.

2.1.2 Observables

In quantum mechanics, one starts with a C*-algebra[4] A of observables, describing all possible linear combinations of measurements that may be performed on a system. Classical statistics also fit very nicely in this framework, by restricting oneself to commutative algebras of observables.

Given a topological space E describing classical microscopic states, we let:

$$A = C(E) \tag{2.3}$$

[3] A quantum description becomes necessary at low temperatures.

[4] A C*-algebra is an algebra over $\mathbb{C}$ with (i) a continuous and complete norm $|\cdot|$ and (ii) an antilinear involution $*$ such that $|a^*a| = |a|^2$.

denote the commutative algebra of continuous and bounded real functions over E, equipped with the infinite norm $||u||_\infty = \sup_{x\in E} |u(x)|$. A classical observable is just a function of the microscopic states. This assignment defines a contravariant functor $C : \mathbf{Top}^{op} \to \mathbf{Alg}$, as any continuous map $\varphi : E \to E'$ has a pull-back $\varphi^* : A' \to A$ defined by:

$$\big(\varphi^* u\big)(x') = u\big(\varphi(x)\big) \tag{2.4}$$

In most of this work, the algebra of observables $A = C(E)$ will be commutative and given by such a procedure. We however emphasize that once given the algebra, one may very well forget about the underlying set.

One may argue that physical observables are often unbounded, as is the case of the position and momentum coordinates q and p. We prefer to view these as infinitesimal generators of a group of invertible observables, as would be the case in quantum mechanics, although such a discussion is not in the scope of this thesis. We will be mostly interested in the finite setting where:

$$A = \mathbb{R}^E \tag{2.5}$$

is a finite dimensional vector space, isomorphic to the multiplicative Lie group $G = (\mathbb{R}_+^*)^E$ of strictly positive observables, and could be viewed as the abelian Lie algebra of G. Restricting to finite configuration spaces will leave aside most technical difficulties, greater generality is only mentioned here for the sake of perspective.

At the quantum level, the prototype of a C*-algebra is given by a Von Neumann algebra:

$$A \subseteq B(\mathcal{H}) \tag{2.6}$$

of bounded operators over a complex Hilbert space $\mathcal{H}$, with complex adjunction $a^* = \bar{a}^t$ as involution, although most constructions can be carried on the algebra most naturally, without any reference to a particular Hilbert space.

Every C*-algebra A has a positive cone A^+ defined by:

$$a \geq 0 \quad \Leftrightarrow \quad \exists b \in A \quad \text{with} \quad b^* b = a \tag{2.7}$$

Any positive element $a \geq 0$ is self-adjoint and satisfies $a^* = a$; its spectrum is contained in $\mathbb{R}^+$. The above would also describe positive functions of $C(E)$. Positivity will be a fundamental concept when defining the states of the algebra.

2.1.3 Linear Forms and Measures

Given an algebra with a continuous norm A, its topological dual A^* is the vector space of continuous linear forms on A. The topological dual defines a contravariant functor $\mathbf{Alg}^{op} \to \mathbf{Vect}$ as any linear map $T : A \to A'$ has an adjoint map $T^* : A'^* \to A^*$ defined for all $\lambda \in A'^*$ and $a \in A$ by:

$$\langle\, T^*\lambda \,|\, a \,\rangle = \langle\, \lambda \,|\, Ta \,\rangle \tag{2.8}$$

The duality comes from the underlying vector space and is common enough not to be discussed. We only briefly review some classical constructions and notations.

When $A = C(E)$ is the real algebra of continuous and bounded functions over E, its dual A^* is the space of Borel measures of finite mass on E, equipped with the L^1-norm:

$$\langle\, \lambda \,|\, f \,\rangle = \int_{x\in E} f(x) \cdot \lambda(dx) \tag{2.9}$$

for all $\lambda \in A^*$ and $f \in A$, with $\big|\langle\, \lambda \,|\, f \,\rangle\big| \leq ||f||_\infty \cdot ||\lambda||_1$.

A continuous map $\varphi : E \to E'$ induces a map of algebras $\varphi^* : A' \to A$ by pull-back. Its adjoint map is the push-forward of measures $\varphi_* : A^* \to A'^*$, defined for every $\lambda \in A^*$ and every measurable subset $S' \subseteq E'$ by:

$$(\varphi_* \lambda)(S') = \int_{x \in \varphi^{-1}(S')} \lambda(dx) \tag{2.10}$$

When E is finite, the push-forward of a measure $\lambda \in A^*$ is given by its weight on each point $x' \in E'$:

$$(\varphi_* \lambda)(x') = \sum_{\varphi(x) = x'} \lambda(x) \tag{2.11}$$

In applications, the set maps we will consider are projections of the form $\varphi : E_1 \times E_2 \to E_1$. The pushforward of φ is then called the marginal projection on E_1, or partial integration along E_2.

When $A = B(\mathcal{H})$ is the algebra of bounded operators on a Hilbert space, equipped with a continuous trace operator $\mathrm{Tr} : A \to \mathbb{C}$, one may define the hermitian scalar product of a and b in A by $\mathrm{Tr}(a^* b)$. A linear form $\lambda \in A^*$, that is also continuous for the hermitian norm induced, may be represented by an element of $B(\mathcal{H})$ such that:

$$\langle \lambda \,|\, a \rangle = \mathrm{Tr}(\lambda^* a) \tag{2.12}$$

This point of view is the most commonly used in quantum statistics.

2.1.4 Statistical States

A state of a unital involutive algebra A is a linear form $\omega \in A^*$ satisfying the two following axioms:

(i) $\omega(a^* a) \geq 0$ for all $a \in A$ (positivity)

(ii) $\omega(1) = 1$ (normalisation)

The states of A form a convex subset of linear forms $\Delta \subseteq A^*$.

When $A = C(E)$ is commutative and A^* is the space of finite mass measures on E, the above axioms define positive measures of mass 1, *i.e.* probability densities on E and we have $\Delta = \mathrm{Prob}(E)$. According to Gelfand's theorem, any commutative C*-algebra is isomorphic to a complex algebra $C(E, \mathbb{C})$ of continuous bounded functions over a compact space E, called the spectrum of A.

When A is a generic C*-algebra, the Gelfand-Naimark-Segal construction associates to each state $\omega \in \Delta$ a Hilbert space representation $\mathcal{H}_\omega$ and a unit cyclic vector $\psi_\omega \in \mathcal{H}_\omega$ such that for all $a \in A$:

$$\omega(a) = \langle \psi_\omega \,|\, a \cdot \psi_\omega \rangle \tag{2.13}$$

In quantum mechanics, this expression traditionally defines the mean value of a self-adjoint observable a when the system is in the state $\psi_\omega \in \mathcal{H}_\omega$. For every self-adjoint a, the spectral projections of ψ_ω define a probability distribution on the spectrum of a. This may be viewed as a consequence of Gelfand's theorem, as the commutative C*-algebra generated by a and a^* is isomorphic to $C(\mathrm{Sp}\, a)$.

When $A \subseteq B(\mathcal{H})$ is already represented on a Hilbert space and is equipped with a trace operator, operators of $B(\mathcal{H})$ are mapped to linear forms. Any positive operator $\rho \in B(\mathcal{H})_+$ such that $\mathrm{Tr}(\rho) = 1$ then defines a state of A by letting for all $a \in A$:

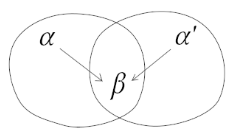
$$\rho(a) = \mathrm{Tr}(\rho a) \tag{2.14}$$

This picture can lead to confusion with the previous one, as $\mathcal{H}$ is not the GNS representation of ρ. The operator ρ may however be viewed as a vector of the Hilbert space $\mathcal{H} \otimes \mathcal{H}^*$, of which the GNS representation $\mathcal{H}_\rho$ is a subspace. In statistical quantum mechanics, ρ is called the *density matrix*.

2.2 Systems

This section introduces the differential and module structures on which relies the present work. The theory will be treated abstractly to keep as much generality as possible, and deals with what one may call systems of algebraic structures, *i.e.* a particular type of functors.

One should still read the theory with the contents of the previous section in mind, to which it aims to be applied. The main idea is to localise the previous structures from a global set of variables $\Omega = \{i, j, k, \dots\}$ to a covering of Ω by smaller regions $X = \{\alpha, \beta, \gamma, \dots\} \subseteq \mathcal{P}(\Omega)$. This leads to the definition of local configuration spaces, local algebras of observables, *etc.* related by morphisms every time a region is contained in another. Giving $(X, \supseteq)$ a category structure by agreeing that a unique arrow $\alpha \to \beta$ exists whenever α contains β, we will get functors[5] from X to **Set**, **Alg**, *etc.*



$$\begin{tikzcd}
A_\alpha & & A_{\alpha'} \\
& A_\beta \arrow[ul, "j_{\alpha\beta}"] \arrow[ur, "j_{\alpha'\beta}"'] &
\end{tikzcd}$$

The main result of this section is that we may define a chain complex $\big(A_\bullet(X), \delta\big)$ of observables. Its boundary operator δ will play a crucial role in describing the Lagrange multipliers of the cluster variation method and in defining transport equations that generalise belief propagation. This construction was already considered by Grothendieck and Verdier under the name of *canonical projective resolution for presheaves* [29, 18]. Their motivations having been more abstract, aiming at the unification of all known homology and cohomology theories, we believe the present work has the benefit of providing a simple and concrete application of this complex. The interaction decomposition theorem 2.8 will also allow us to clarify the structure of homology groups in our setting.

2.2.1 Systems of Sets

Given a numerable set of atoms Ω and a configuration space E_i for all $i \in \Omega$, let:

$$E_\alpha = \prod_{i \in \alpha} E_i \tag{2.15}$$

denote the configuration space of $\alpha \subseteq \Omega$. There is a projection $\pi^{\beta\alpha} : E_\alpha \to E_\beta$ for every $\alpha \supseteq \beta$, forgetting the state of atoms outside β. This will consist of our fundamental example of a projective system of sets, for any collection of regions $X \subseteq \mathcal{P}(\Omega)$.

$$\begin{tikzcd}
E_\alpha \arrow[dr, "\pi^{\beta\alpha}"'] & & E_{\alpha'} \arrow[dl, "\pi^{\beta\alpha'}"] \\
& E_\beta &
\end{tikzcd} \tag{2.16}$$

When X covers Ω, the system efficiently keeps all the available information, as the global configuration space can be recovered as the projective limit $E_\Omega = \lim_{\alpha \in X} E_\alpha$.

Definition 2.1. *A* system of sets E *over a partial order* X *is a covariant functor* $E : X \to$ **Set**. *Denoting by* $\varphi_* : E_\alpha \to E_\beta$ *the map induced by* $\varphi : \alpha \to \beta$ *in* X, *a system is said:*

[5] Functors with a partial order as source category are often called systems in the literature, as they were considered long before the categorical language became common use.

– injective *when φ_* is an injection for all φ.*

– projective *when φ_* is a surjection for all φ,*

The functor category $[X, \mathbf{Set}]$ of systems over X has natural transformations $\eta : E \to E'$ as morphisms.

A functor $t : X \to X'$ induces a pull-back functor $t^* : [X', \mathbf{Set}] \to [X, \mathbf{Set}]$ defined by $t^* E' = E' \circ t$. This allows to compare the categories of systems over different partial orders X and X' and to define a global category of systems of sets over an arbitrary partial order.

Definition 2.2. *We denote by $\{\mathbf{Set}\}$ the category of systems of sets, with:*

– *objects (X, E) where X is a partial order and E is a system of sets over X,*

– *morphisms $(t, \eta) : (X, E) \to (X', E')$ where $t : X \to X'$ is a functor and $\eta : E \to t^* E'$ is a natural transformation.*

We introduce the notation $\{\mathbf{Set}\}$ to avoid confusion with the larger category of functors $[-, \mathbf{Set}]$ as the source category X is restricted to partial orders. As a subcategory of the latter, it should be thought of in the same way, and the partial order hypothesis will have no influence until the next section.

Examples.

1. A single set E is a system over the point category $\{\bullet\}$.

2. The restriction of a system over X to a subcategory $Y \subseteq X$ is naturally mapped into the original system. This provides with a trivial example of morphism in $\{\mathbf{Set}\}$.

3. Given an equivalence relation $\sim$ in X, any system E over X induces a system $\bar{E}$ over the quotient space $\bar{X} = (X/\sim)$ defined by:
$$\bar{E}_{[\alpha]} = \bigsqcup_{\alpha' \sim \alpha} E_{\alpha'} \tag{2.17}$$
Denoting by $p : X \to \bar{X}$ the quotient map, the natural transformation from E to $p^* \bar{E}$ is canonically defined by inclusion of E_α in the disjoint union $\bar{E}_{[\alpha]}$.

2.2.2 Systems of Abelian Groups

The category $\{\mathbf{Ab}\}$ of abelian group systems is defined by restricting to functors $G : X \to \mathbf{Ab}$. To such a system G, we shall associate chain and cochain complexes denoted by $\big(G_\bullet(X), \delta\big)$ and $\big(G^\bullet(X), d\big)$ respectively. Their construction extends G to a functor on the nerve[6] of X and in the following we denote by $\bar{\alpha} \in N_p(X)$ a p-chain $\alpha_0 \to \ldots \to \alpha_p$ in X.

Consider the contravariant functor $\hat{G} : N(X)^{op} \to \mathbf{Ab}$ defined by $\hat{G}_{\bar{\alpha}} = G_{\alpha_0}$ for all $\bar{\alpha} \in N_p(X)$. For every subchain $\bar{\beta}$ of $\bar{\alpha}$, the map $\hat{G}_{\bar{\alpha}} \to \hat{G}_{\bar{\beta}}$ is induced by $G_{\alpha_0} \to G_{\beta_0}$ as $\alpha_0 \to \beta_0$. The simplicial set structure of $N(X)$ thus makes $\hat{G}$ a simplicial abelian group, and $\hat{G}$ defines a chain complex $G_\bullet(X)$ equipped with a boundary operator $\delta : G_{n+1}(X) \to G_n(X)$, where:
$$G_n(X) = \prod_{\bar{\alpha} \in N_n(X)} \hat{G}_{\bar{\alpha}} \tag{2.18}$$

Reciprocally, a covariant functor $\check{G} : N(X) \to \mathbf{Ab}$ is defined by letting $\check{G}_{\bar{\alpha}} = G_{\alpha_p}$ for $\bar{\alpha} \in N_p(X)$. When $\bar{\beta}$ is a subchain of degree $k \leq p$ of $\bar{\alpha}$, we have a map $\check{G}_{\bar{\beta}} \to \check{G}_{\bar{\alpha}}$ as $\beta_k \to \alpha_p$. Hence $\check{G}$ is a cosimplicial abelian group and defines a cochain complex $G^\bullet(X)$ with a differential $d : G^n(X) \to G^{n+1}(X)$, where:
$$G^n(X) = \prod_{\bar{\alpha} \in N_n(X)} \check{G}_{\bar{\alpha}} \tag{2.19}$$

[6] The nerve of a category is defined in paragraph 1.3.2.

Dual constructions of course arise when $G : X^{op} \to \mathbf{Ab}$ is a cosystem over X. In this case we still let $\hat{G}_{\bar{\alpha}} = G_{\alpha_0}$ and $\check{G}_{\bar{\alpha}} = G_{\alpha_p}$ to define functors $\hat{G} : N(X) \to \mathbf{Ab}$ and $\check{G} : N(X)^{op} \to \mathbf{Ab}$. Applications will involve both covariant and contravariant functors of abelian groups but their extension to the nerve will mostly be done through $\check{G}$. The following table might be useful:

$*$	$\big(G_\bullet(X), \delta\big)$	$\big(G^\bullet(X), d\big)$
$G : X \to \mathbf{Ab}$	$\hat{G}$	$\check{G}$
$G : X^{op} \to \mathbf{Ab}$	$\check{G}$	$\hat{G}$

Fundamental Examples.

1. When $E : X \to \mathbf{Set}$ is a system of sets over X, it defines a system of algebras $A : X^{op} \to \mathbf{Alg}$ by letting $A_\alpha = \mathbb{R}^{E_\alpha}$ for all $\alpha \in X$. In the chain complex $(A_\bullet(X), \delta)$, a 1-chain φ is defined by a collection of local observables $\varphi_{\alpha\beta} \in \mathbb{R}^{E_\beta}$ while its boundary $\delta\varphi$ is given by:
$$(\delta\varphi)_\beta = \sum_{\alpha' \to \beta} \varphi_{\alpha'\beta} - \sum_{\beta \to \gamma'} [\varphi_{\beta\gamma'}]_\beta \tag{2.20}$$
where $[\varphi_{\beta\gamma}]_\beta \in \mathbb{R}^{E_\beta}$ denotes the pullback of $\varphi_{\beta\gamma} \in \mathbb{R}^{E_\gamma}$ under the map $E_\beta \to E_\gamma$.

2. When $E : X \to \mathbf{Set}$ and $A = \mathbb{R}^E$, duality defines a system of vector spaces $A^* : X \to \mathbf{Vect}$. In the cochain[7] complex $(A^*_\bullet(X), d)$, a 0-cochain q is defined by a collection of linear forms $q_\alpha \in L(\mathbb{R}^{E_\alpha}, \mathbb{R})$ while its differential dq is given by:
$$(dq)_{\alpha\beta} = q_\beta - [q_\alpha]_\beta \tag{2.21}$$
where $[q_\alpha]_\beta \in L(\mathbb{R}^{E_\beta}, \mathbb{R})$ denotes the pushforward of $q_\alpha \in L(\mathbb{R}^{E_\alpha}, \mathbb{R})$ by the map $E_\alpha \to E_\beta$.

The difference of incoming and departing fluxes defining $\delta\varphi$ in (2.20) recalls the classical *divergence* operator of differential geometry. Acting on a smooth vector field $\vec{\varphi} : \mathbb{R}^3 \to \mathbb{R}^3$, the divergence is intrinsically related to the fundamental *Gauss formula*:
$$\int_V \mathrm{div}(\vec{\varphi})\, dv = \int_{\partial V} \vec{\varphi} \cdot \vec{n}\, ds \tag{2.22}$$
Integrating $\mathrm{div}(\vec{\varphi})$ on a volume $V \subseteq \mathbb{R}^3$ yields the outbound[8] flux of $\vec{\varphi}$ through the boundary of V.

Proposition 2.3 played a crucial role in understanding the topological structure underlying message-passing algorithms. To serve as counterparts for the integration[9] supports of (2.22), let us denote by:

– $\Lambda^\alpha = \{\beta \in X \mid \alpha \to \beta\}$ the cone below α in X,

– $d\Lambda^\alpha = \{\alpha'\beta' \in N_1(X) \mid \alpha' \notin \Lambda^\alpha,\ \beta' \in \Lambda^\alpha\}$ the coboundary of Λ^α.

Proposition 2.3 (Gauss Formula)**.** *Given a cosystem $G : X^{op} \to \mathbf{Ab}$ of abelian groups, let $(G_\bullet(X), \delta)$ denote the associated chain complex. For every $\varphi \in G_1(X)$, we have:*
$$\sum_{\beta' \in \Lambda^\alpha} (\delta\varphi)_{\beta'} = \sum_{\alpha'\beta' \in d\Lambda^\alpha} \varphi_{\alpha'\beta'} \tag{2.23}$$

Proof. In the sum of $\delta\varphi$ over Λ^α, each term $\varphi_{\beta'\gamma'}$ is counted twice with opposite signs if $\beta' \in \Lambda^\alpha$. □

[7] Although it is a cochain complex, we write degrees as indices for $A^*_\bullet(X)$ as it is the dual vector space of $A_\bullet(X)$.

[8] With the sign convention from physics. In Hodge theory where $d^* = -\mathrm{div}$, an *inbound* flux is measured instead.

[9] The sum over Λ^α is the *zeta transform* of chapter 3. Proposition 2.3 hence provides with a beautiful illustration of the analogy between Möbius inversion formulas and the fundamental theorem of calculus already noted by Rota in [24].

2.2.3 Systems of Rings and Modules

The category $\{\mathbf{Ring}\}$ of ring systems is similarly defined by restricting to functors $R : X \to \mathbf{Ring}$. To avoid confusion in applications to come, it will be more convenient to consider cosystems over X, *i.e.* functors $R : X^{op} \to \mathbf{Ring}$. This subsection, mostly inspired by Kodaira [12], explores the different products and module structures one may generalise from the usual theory with scalar coefficients. Proofs are left to the appendix.

Given a cosystem of rings R over X, we give a natural ring structure $\big(R^\bullet(X), +, \wedge\big)$ to its associated cochain complex, by defining the exterior[10] product of a p-field φ with a k-field ψ as the $(p+k)$-field:

$$(\varphi \wedge \psi)_{\alpha\dots\beta\dots\gamma} = \varphi_{\alpha\dots\beta} \cdot \big[\psi_{\beta\dots\gamma}\big]_\alpha \tag{2.24}$$

where $[\psi_{\beta\dots\gamma}]_\alpha$ denotes the image of $\psi_{\beta\dots\gamma} \in R_\beta$ in R_α. The exterior product is compatible with the differential and we have the graded Leibniz rule:

$$d(\varphi \wedge \psi) = d\varphi \wedge \psi + (-1)^{|\varphi|}\, \varphi \wedge d\psi \tag{2.25}$$

denoting by $|\varphi|$ the degree of φ, and $\big(R^\bullet(X), +, \wedge, d\big)$ is a differential ring.

There is a natural notion of module over a ring system which makes any ring system a module over itself. The dual notion of comodule will also be of interest to us, they both give a differential module structure to one of the associated complexes.

Definition 2.4. *We call module over $R : X^{op} \to \mathbf{Ring}$ any cosystem of abelian groups $M : X^{op} \to \mathbf{Ab}$ such that:*

- *M_α is an R_α-module for all $\alpha \in X$,*
- *$[r_\beta]_\alpha \cdot [m_\beta]_\alpha = [r_\beta \cdot m_\beta]_\alpha$ for all $\alpha \to \beta$ in X.*

where $[m_\beta]_\alpha$ denotes the image of m_β in M_α.

When M is a module over R, the cochain complex $M^\bullet(X)$ inherits a module structure over $R^\bullet(X)$ for the action of $\varphi \in R^p(X)$ on $m \in M^k(X)$ extending the exterior product:

$$(\varphi \times m)_{\alpha\dots\beta\dots\gamma} = \varphi_{\alpha\dots\beta} \cdot \big[m_{\beta\dots\gamma}\big]_\alpha \tag{2.26}$$

and the differential ∇ on $M^\bullet(X)$ then satisfies the graded module Leibniz rule:

$$\nabla(\varphi \times m) = d\varphi \times m + (-1)^{|\varphi|}\, \varphi \times \nabla m \tag{2.27}$$

so that $\big(M^\bullet(X), \nabla\big)$ is a differential module over $\big(R^\bullet(X), d\big)$.

Definition 2.5. *We call comodule over $R : X^{op} \to \mathbf{Ring}$ any system of abelian groups $M : X \to \mathbf{Ab}$ such that:*

- *M_α is an R_α-module for all $\alpha \in X$,*
- *$r_\beta \cdot [m_\alpha]_\beta = [r_\beta \cdot m_\alpha]_\beta$ for all $\alpha \to \beta$ in X.*

where $[m_\alpha]_\beta$ denotes the image of m_α in M_β and $\cdot$ the action of any R_β such that $\alpha \to \beta$ on M_α.

When M is a comodule over R, the action of $\varphi \in R^k(X)$ on $m \in M_n(X)$ is the chain of $M_{n-k}(X)$ given by the interior product:

$$(\varphi \lrcorner\, m)_{\beta\dots\gamma} = \sum_{\alpha'\dots\beta} \varphi_{\alpha'\dots\beta} \cdot \big[m_{\alpha'\dots\beta\dots\gamma}\big]_\beta \tag{2.28}$$

[10] With scalar coefficients, $\wedge$ is called the Alexander product in [12]. It may also be called a cup-product [[ref cap-cup-products]]

and by $(\varphi \lrcorner m) = 0$ when $n < k$. Note that $\lrcorner$ actually defines a right action as $(\varphi \wedge \psi) \lrcorner m = \psi \lrcorner (\varphi \lrcorner m)$. The boundary of $M_\bullet(X)$ satisfies the graded module Leibniz rule:

$$\delta(\varphi \lrcorner \psi) = d\varphi \lrcorner \psi + (-1)^{|\varphi|} \varphi \lrcorner \delta\psi \tag{2.29}$$

and $\big(M_\bullet(X), \delta\big)$ is a differential module over $\big(R^\bullet(X), d\big)$.

Examples.

1. Consider the constant ring system $\mathbb{Z}$ over X, and for every $\alpha \in X$, the 1-cochain $\lambda^\alpha \in \mathbb{Z}^1(X)$ defined by:
$$\lambda^\alpha_{\alpha'\beta'} = \begin{cases} 1 & \text{if} \quad \alpha' = \alpha \\ 0 & \text{if} \quad \alpha' \neq \alpha \end{cases} \tag{2.30}$$
An easy computation shows that $d\lambda^\beta \in \mathbb{Z}^2(X)$ may be written for all $\beta \in X$ as:
$$d\lambda^\beta = \sum_{\alpha \to \beta} \lambda^\alpha \wedge \lambda^\beta \tag{2.31}$$

2. Any cosystem of algebras $A : X^{op} \to \mathbf{Alg}$ is a comodule over $\mathbb{Z}$. For every $\beta \in X$ we then denote by i_β the degree -1 endomorphism of $A_\bullet(X)$ defined by $i_\beta(\varphi) = \lambda^\beta \lrcorner \varphi$. For every $\bar\gamma \in N(X)$, if $\beta \not\to \gamma_0$ we have $i_\beta(\varphi)_{\bar\gamma} = 0$ and otherwise:
$$i_\beta(\varphi)_{\bar\gamma} = \varphi_{\beta\bar\gamma} \tag{2.32}$$
The previous example and the graded module Leibniz rule give the following formula:
$$\delta(i_\beta \varphi) = i_\beta\Big(\sum_{\alpha \to \beta} i_\alpha \varphi \Big) - i_\beta(\delta\varphi) \tag{2.33}$$
which will be very useful in the proof of proposition 3.13.

2.3 Local Statistics

Localising the statistical structures, this section restricts to a special case of systems, that essentially derive from a system of sets given by cartesian products of atomic configuration spaces. They are closely related to what has been called *free sheaves* in [3], where a more general setting is considered. These very common systems however cover all the applications we shall be interested in.

A fundamental consequence of this hypothesis is the so-called *interaction decomposition theorem*. It asserts that each algebra of observables splits as a direct sum $A_\alpha = \bigoplus_\beta Z_\beta$, where the interaction vector spaces Z_β may be chosen consistently over X and play the role of independent generators[11]. This well-known yet subtle result may be given a significant number of proofs and calls for greater generality[12], but the simple setting we consider allows for a beautiful proof via harmonic analysis.

The main result of this chapter will then be the acyclicity of the chain complex of observables $A_\bullet(X)$. We moreover show that the homology class of $h \in A_0(X)$ is completely determined by its global sum $H_\Omega = \sum_\alpha h_\alpha$ in A_Ω. In a physical terminology, one would say that two potentials h and h' are homologous if and only if they define the same global hamiltonian H_Ω.

[11] This fact motivates the « free » terminology.
[12] See for instance [9, 27, 16, 3].

2.3.1 Statistical System

From now on, we consider a finite set of atoms Ω with finite sets of microstates E_i for all $i \in \Omega$. The configuration space E_α of any region $\alpha \subseteq \Omega$ is defined as:

$$E_\alpha = \prod_{i \in \alpha} E_i \tag{2.34}$$

For every $\alpha \supseteq \beta$, we denote the canonical projection by $\pi^{\beta\alpha} : E_\alpha \to E_\beta$. Alternatively, one could say that $E : \big(\mathcal{P}(\Omega), \supseteq\big) \to \mathbf{Set}$ defines a sheaf of finite sets over the finite topological space Ω.

Given these sets of microstates, local algebras of observables are defined by $A_\alpha = \mathbb{R}^{E_\alpha}$ and for every $\alpha \supseteq \beta$, the canonical injection $j_{\alpha\beta} : A_\beta \to A_\alpha$ is the pull-back of $\pi^{\beta\alpha}$ sending a function on E_β to its cylindrical extension on E_α. The multiplicative structure of E is carried to A by:

$$A_\alpha = \bigotimes_{i \in \alpha} A_i \tag{2.35}$$

as A_α is linearly generated by the Dirac masses (δ_{x_α}) on E_α, which may be written as pure tensors of the form $\delta_{x_i} \otimes \cdots \otimes \delta_{x_j}$, and the extension map $j_{\alpha\beta}$ may then be viewed as the tensor multiplication $1_{\alpha \smallsetminus \beta} \otimes -$ with the unit of $A_{\alpha \smallsetminus \beta}$. Better suited generators (χ_{k_α}) will be introduced in the next section.

In the following, we suppose chosen a covering $X \subseteq \mathcal{P}(\Omega)$ closed under intersection, *i.e.* such that:

$$\alpha \in X \quad \text{and} \quad \beta \in X \quad \Rightarrow \quad \alpha \cap \beta \in X \tag{2.36}$$

We then restrict E to a system $E : X \to \mathbf{Set}$ and A to a cosystem $A : X^{op} \to \mathbf{Alg}$ over X. The closure hypothesis is fundamental for the interaction decomposition theorem to hold, it will also be useful for the generalised combinatorics we propose in chapter 3.

Definition 2.6. *We denote by:*

- $\big(A_\bullet(X), \delta\big)$ *the chain complex of* local observables*,*
- $\big(A^*_\bullet(X), d\big)$ *the cochain complex of* local measures*,*
- $\Delta_\bullet(X) \subseteq A^*_\bullet(X)$ *the convex subspace of* local probabilities*.*

This localisation procedure will lead us to represent the global hamiltonian $H_\Omega \in A_\Omega$ by a homology class of interaction potentials $(h_\alpha) \in A_0(X)$ satisfying:

$$H_\Omega = \sum_{\alpha \in X} h_\alpha \tag{2.37}$$

The global Gibbs state $p_\Omega \in \Delta_\Omega$ would also be ideally represented by its local marginals $(p_\alpha) \in \Delta_0(X)$ or an approximation of the latter. Marginals of p_Ω are said consistent as p_β is the marginal of p_α for every $\beta \subseteq \alpha$, and the following more general notion of cohomology class in $\Delta_0(X)$ will substitute for the global probabilities of Δ_Ω.

Definition 2.7. *The convex subspace* $\Gamma(X)$ *of* consistent local probabilities *is defined by:*

$$\Gamma(X) = \Delta_0(X) \cap \mathrm{Ker}(d) \tag{2.38}$$

We also denote by $\mathring{\Gamma}(X)$ *the space of consistent positive local probabilities.*

The image of Δ_Ω in $\Delta_0(X)$ is in general a strict convex polytope of $\Gamma(X)$, as although any consistent $q \in \Gamma(X)$ may always be extended to a global measure $q_\Omega \in A^*_\Omega$, the non-negativity of q_Ω is not insured. This was already noticed by Vorob'ev who first characterised the simplicial complexes X having the property that any consistent family of probabilities in $\Gamma(X)$ may be extended[13] to Δ_Ω. They essentially coincide with the *retractable* hypergraphs on which we show message-passing algorithms to be exact in chapter 6.

[13] See also [1] for a sheaf theoretic approach of this problem, and relations to the notion of contextuality. The complex we consider is related to a barycentric subdivision to the complex of [1], as the categorical nerve subdivides the Čech nerve of a covering, and the isomorphism of homology groups is proved in [3].

2.3.2 Interaction Decomposition

For every $\alpha \supseteq \beta$ in X, the algebra A_β is naturally embedded in A_α. Therefore A_α contains observables that may be split as a sum of observables on strict subregions of α. We call them *boundary* observables and write:

$$B_\alpha = \sum_{\alpha \supset \beta} A_\beta \tag{2.39}$$

We say that Z_α is an *interaction subspace* of A_α if it is a supplement of B_α. One may then write:

$$A_\alpha = Z_\alpha \oplus \Big(\sum_{\alpha \supset \beta} A_\beta \Big) \tag{2.40}$$

and continue this procedure inductively, which is the content of the interaction decomposition theorem.

Theorem 2.8. *Given an interaction subspace $Z_\alpha \subseteq A_\alpha$ for every $\alpha \in X$, one has for all α:*

$$A_\alpha = \bigoplus_{\alpha \supseteq \beta} Z_\beta \tag{2.41}$$

It will be useful to consider the following rewording of the theorem. In this point of view, interaction subspaces define a cosystem of vector spaces $Z : X^{op} \to \mathbf{Vect}$ with trivial maps, embedded in A, and we denote by $Z_0(X)$ their direct sum viewed as a subspace of $A_0(X)$.

Corollary 2.9. *Given a choice of interaction subspaces, we have a projection $P : A_0(X) \to Z_0(X)$ given by:*

$$P(u)_\beta = \sum_{\alpha \supseteq \beta} P^{\beta\alpha}(u_\alpha) \tag{2.42}$$

where $P^{\beta\alpha} : A_\alpha \to Z_\beta$ is a projection of A_α onto Z_β, vanishing on Z_γ for every $\gamma \neq \beta$.

The theorem asserts that a direct sum decomposition holds for *any* choice of interaction subspaces. As supplements of boundary observables, they are not defined intrinsically, although unless explicit mention of the opposite, Z_α will from now on be assumed orthogonal to B_α for the canonical scalar product of $\mathbb{R}^{E_\alpha}$. This choice will play a particular role in describing the high temperature limit, see for instance 5.11 and 6.29.

Definition 2.10. *The* canonical interaction subspaces $Z_\alpha \subseteq A_\alpha$ *are defined as orthogonal supplements of boundary observables for the canonical scalar product of $A_\alpha \simeq \mathbb{R}^{E_\alpha}$:*

$$Z_\alpha = \Big(\sum_{\beta \subset \alpha} \mathrm{Im}\,(j_{\alpha\beta}) \Big)^{\perp} \quad \Leftrightarrow \quad Z_\alpha = \bigcap_{\beta \subset \alpha} \mathrm{Ker}\,(\Sigma^{\beta\alpha}) \tag{2.43}$$

Instead of proving 2.8 in its generality[14], inspired by Matůs, we show how harmonic analysis gives an enlightening perspective on the canonical interaction decomposition. This seems to be the most natural construction of the projections $P^{\beta\alpha}$ and is the one we implemented in javascript.

Chose an ordering of E_j for all $j \in \Omega$, so that each E_α may be identified with the finite torus:

$$E_\alpha \simeq \prod_{j \in \alpha} \frac{\mathbb{Z}}{N_j \mathbb{Z}} \tag{2.44}$$

and consider for every $\alpha \in X$ the complexified algebra of observables $\tilde{A}_\alpha = \mathbb{C}^{E_\alpha}$. It is a fundamental result of abelian group theory that the spectrum $\hat{E}_\alpha = \mathrm{Hom}(E_\alpha, \mathbb{C}^*)$ of E_α defines an orthormal

[14] [[appendix: proof by action of permutations]]

basis of $\tilde{A}_\alpha$ for its canonical hermitian product, where the so-called characters of $\hat{E}_\alpha$ are plane waves generating a discrete Fourier transform on $\tilde{A}_\alpha$.

By duality, there is a natural injection $\hat{E}_\beta \to \hat{E}_\alpha$ by pull-back of the projection $\pi^{\beta\alpha} : E_\alpha \to E_\beta$. Consider the subset $\hat{F}_\alpha \subseteq \hat{E}_\alpha$ of characters defined by:

$$\hat{F}_\alpha = \bigcup_{\beta \subset \alpha} \hat{E}_\beta \tag{2.45}$$

It is easily seen that $\hat{F}_\alpha$ is an orthonormal basis of the subspace $\tilde{B}_\alpha$ of complex boundary observables. Its complement $\hat{G}_\alpha$ then provides with an orthonormal basis of the interaction subspace $\tilde{Z}_\alpha = \tilde{B}_\alpha^\perp$. The interaction decomposition theorem here amounts to the observation that $\hat{E}_\alpha$ is recovered as the disjoint union:

$$\hat{E}_\alpha = \bigsqcup_{\beta \subseteq \alpha} \hat{G}_\beta \tag{2.46}$$

A boundary observable in $\tilde{A}_\alpha$ is spanned by plane waves with wave vectors tangent to some E_β with $\beta \subset \alpha$, and the spectral support of an interaction observable is in the complement of such wave vectors.

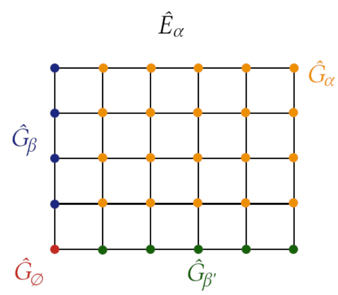


Figure 2.1: Spectral decomposition $\hat{E}_\alpha = \bigsqcup_{\beta \subseteq \alpha} \hat{G}_\beta$.

More concretely, define for every $k_\alpha \in E_\alpha$ the character $\chi_{k_\alpha} \in \hat{E}_\alpha$ by:

$$\chi_{k_\alpha}(x_\alpha) = \mathrm{e}^{i\langle\, k_\alpha \,|\, x_\alpha \,\rangle} \quad \text{where} \quad \langle\, k_\alpha \,|\, x_\alpha \,\rangle = \sum_{j \in \alpha} \frac{k_j x_j}{2\pi N_j} \tag{2.47}$$

The character $\chi_{k_\beta} \in \hat{E}_\beta$ is extended to $\hat{E}_\alpha$ by letting $k_j = 0$ for $j \notin \beta$. Identifying $\hat{E}_\alpha$ with E_α, the discrete Fourier transform defines a unitary transformation of $\tilde{A}_\alpha$. For all $u_\alpha \in \tilde{A}_\alpha$, write:

$$u_\alpha = \sum_{k_\alpha \in E_\alpha} \hat{u}_\alpha(k_\alpha) \cdot \chi_{k_\alpha} \tag{2.48}$$

The orthogonal projection $P^{\beta\alpha} : \tilde{A}_\alpha \to \tilde{Z}_\beta$ is then simply given by selecting only the modes of $G_\beta \subseteq E_\beta$, complement of $\bigcup_{\beta \subset \gamma} E_\gamma$:

$$P^{\beta\alpha}(u_\alpha) = \sum_{k_\beta \in G_\beta} \hat{u}_\alpha(k_\beta) \cdot \chi_{k_\beta} \tag{2.49}$$

The real Fourier transform on A_α would obviously give a similar explicit construction of real interaction subspaces. This construction appealed to us for both its great conceptual simplicity and ease of implementation. It however strongly relies on the multiplicative structure of the underlying system of sets.

The following proposition gives a more abstract characterisation of interaction subspaces, as other approaches[15] would be necessary for general inductive systems of vector spaces. Note that it is not true in general that the inductive limit splits as a direct sum.

Proposition 2.11. *The inductive limit of A is isomorphic to the direct sum of interaction subspaces:*

$$Z_0(X) \simeq \operatorname*{colim}_{\alpha \in X} A_\alpha \tag{2.50}$$

Proof. Let (V, f) denote a cone[16] below A defined by a collection of consistent maps $f_\beta : A_\beta \to V$. Identify (V, f) with a map $f : A_0(X) \to V$ and suppose given $\tilde{f} : Z_0(X) \to V$ such that $\tilde{f} \circ P = f$. Then for every $v_\alpha \in Z_\alpha \subseteq A_0(X)$, we have $\tilde{f}(v_\alpha) = f(v_\alpha)$. This defines a unique map $\tilde{f} : Z_0(X) \to V$ factorising f through P, and $Z_0(X)$ satisfies the universal property of the colimit. □

The inclusions of each A_α into the global algebra A_Ω induce a map $\zeta_\Omega : A_0(X) \to A_\Omega$ on the direct sum, given by:

$$\zeta_\Omega : h \longmapsto \sum_{\alpha \in X} h_\alpha \tag{2.51}$$

We denote by $B_\Omega \subseteq A_\Omega$ its image. It follows from the previous proposition that we have a universal map from $Z_0(X)$ to B_Ω, factorising ζ_Ω through P:

$$\begin{tikzcd} A_0(X) \arrow[r, "P"] \arrow[dr, "\zeta_\Omega"'] & Z_0(X) \arrow[d, dashed] \\ & B_\Omega \subseteq A_\Omega \end{tikzcd} \tag{2.52}$$

Chosing a global interaction subspace to write $A_\Omega = Z_\Omega \oplus B_\Omega$, the interaction decomposition theorem asserts that $A_\Omega = Z_\Omega \oplus Z_0(X)$ so that $Z_0(X)$ and B_Ω are isomorphic. This allows for two different representations of the inductive limit, a local one, and a global one.

Proposition 2.12. *The two following maps are equivalent representations of* colim A*:*

(i) $P : A_0(X) \longrightarrow Z_0(X)$

(ii) $\zeta_\Omega : A_0(X) \longrightarrow B_\Omega$

2.3.3 Homology and Cohomology

The global hamiltonian H_Ω of a physical system is typically given as a sum of local interactions:

$$H_\Omega = \sum_{\alpha \in X} h_\alpha \tag{2.53}$$

and we will see that $H_\Omega = \zeta_\Omega(h)$ represents a unique homology class in $A_0(X)$. The interaction decomposition theorem is essential to compute the homology of $A_\bullet(X)$. We shall first characterise the first homology of $A_0(X)$ and the first cohomology of $A_0^*(X)$, both isomorphic to $Z_0(X)$, before proving the acyclicity of the whole complex $A_\bullet(X)$.

Theorem 2.13. *The interaction projection $P : A_0(X) \to Z_0(X)$ induces an isomorphism in the first homology group of observable fields:*

$$\frac{A_0(X)}{\delta A_1(X)} \simeq Z_0(X) \tag{2.54}$$

[15] See [3] where the decomposition is carried in a much more general setting.
[16] See section 1.2 for the categorical definition of limits.

Proof. A Gauss formula on the cone over $\beta \in X$ first ensures that for all $\varphi \in A_1(X)$, we have:

$$P(\delta\varphi)_\beta = \sum_{\alpha' \to \beta} P^{\beta\alpha'}(\delta_{\alpha'}\varphi) = \sum_{\alpha' \to \beta} \sum_{\beta' \not\to \beta} P^{\beta\alpha'}(\varphi_{\alpha'\beta'}) = 0 \tag{2.55}$$

as $P^{\beta\alpha'}(A_{\beta'})$ is non-zero if and only if β' contains β. Hence P vanishes on boundaries and we denote the induced quotient map by $[P]$. Reciprocally, given $u \in A_0(X)$ we define $\varphi \in A_1(X)$ by $\varphi_{\alpha\beta} = P^{\beta\alpha}(u_\alpha)$ and consider its boundary:

$$\delta_\beta\varphi = \sum_{\alpha' \to \beta} \varphi_{\alpha'\beta} - \sum_{\beta \to \gamma'} \varphi_{\beta\gamma'} = P(u)_\beta - u_\beta \tag{2.56}$$

When $P(u) = 0$, the above gives $u = -\delta\varphi$ so that $\mathrm{Ker}(P) = \delta A_1(X)$. Hence $[P]$ is injective and induces an isomorphism between $H_0(X) = A_0(X)/\delta A_1(X)$ and its image $Z_0(X)$. □

It is now a simple consequence of the previous subsection that total energy, viewed as a global observable of A_Ω, is a maximal homological invariant of $A_0(X)$:

Corollary 2.14. *Two observable fields $h, h' \in A_0(X)$ are homologous if and only if:*

$$\sum_{\alpha \in X} h_\alpha = \sum_{\alpha \in X} h'_\alpha \tag{2.57}$$

and $\zeta_\Omega : A_0(X) \to B_\Omega$ induces an isomorphism in homology.

Theorem 2.13 implies that the first cohomology of $A_0^*(X)$ is also isomorphic to $Z_0(X)$ by duality. It is however very instructive to construct the isomorphism explicity, as this representation of consistent measures already involves a fundamental automorphism ζ of $A_0(X)$, the *zeta transform*, main object of the next chapter. We shall later rewrite theorem 2.15 as:

$$\mathrm{Ker}(d) \;\simeq\; \zeta' \cdot Z_0(X) \tag{2.58}$$

Erratum: The action of ζ' is defined by (2.59) below. Although it requires to rescale the injection of A_β into A_α by volumic factors, it still fits under the definition of zeta transforms given by (3.16). Without the volumic terms, one has $\mathrm{Ker}(\nabla) = \zeta \cdot Z_0(X)$ as per (5.11), where ∇ is also an adjoint of δ, but where partial integrations have been replaced by conditional expectations for the uniform measure.

Theorem 2.15. *A collection of measures $(q_\alpha) \in A_0^*(X)$ is consistent if and only if there exists a collection of interaction potentials $(u_\alpha) \in Z_0(X)$ such that for all $\alpha \in X$:*

$$q_\alpha = \sum_{\alpha \supseteq \beta} \frac{|E_\beta|}{|E_\alpha|} \langle\, u_\beta \,|\, - \,\rangle_\alpha \tag{2.59}$$

where $\langle - | - \rangle_\alpha$ denotes the canonical scalar product of $\mathbb{R}^{E_\alpha}$.

Proof. As $\Sigma^{\beta\alpha}$ is the orthogonal projection onto $\mathbb{R}^{E_\beta} \subseteq \mathbb{R}^{E_\alpha}$ for the canonical scalar product of $\mathbb{R}^{E_\alpha}$, we have $\Sigma^{\beta\alpha}(Z_\gamma) = Z_\gamma$ for every $\gamma \subseteq \beta$ and $\Sigma^{\beta\alpha}(Z_\gamma) = 0$ otherwise. Hence for every $q \in A_0^*(X)$ of the form (2.59) one has:

$$\Sigma^{\beta\alpha}(q_\alpha) = \sum_{\alpha \supseteq \gamma} \Sigma^{\beta\alpha}\left(\frac{|E_\gamma|}{|E_\alpha|}\, u_\gamma\right) = \sum_{\beta \supseteq \gamma} \frac{|E_\gamma|}{|E_\beta|}\, u_\gamma = q_\beta \tag{2.60}$$

Reciprocally, given a consistent $q \in A_0^*(X)$ one may recover $u \in Z_0(X)$ by Möbius inversion, see 3.1. It however suffices here to conclude by dimension using theorem 2.13. □

Let us now prove the acyclicity of $A_\bullet(X)$ by constructing an explicit retraction of $A_\bullet(X)$ to $Z_0(X)$. For every $\alpha \in X$, we denote by $\mathbf{z}_\alpha$ and $\mathbf{b}_\alpha$ the coherent projectors onto Z_α and B_α, so that:

$$\mathrm{id}_{A_\alpha} = \mathbf{z}_\alpha \oplus \mathbf{b}_\alpha \tag{2.61}$$

They induce projectors $\mathbf{z}$ and $\mathbf{b}$ onto the subspaces $Z_\bullet(X)$ and $B_\bullet(X)$ of $A_\bullet(X)$.

Denoting by d the adjoint of δ for the canonical metric of $A_\bullet(X)$, consider the homotopy $\eta = \mathbf{z}\circ d\circ \mathbf{b}$. Explicitly, $\eta : A_p(X) \to A_{p+1}(X)$ acts on $\varphi \in A_p(X)$ by:

$$\eta(\varphi)_{\bar\alpha\beta} = (-1)^{p+1}\, \mathbf{z}_\beta(\varphi_{\bar\alpha}) \tag{2.62}$$

Proposition 2.16. *The map $\eta = \mathbf{z}\, d\, \mathbf{b}$ defines a homotopy between the identity of $A_\bullet(X)$ and the homogeneous extension of the interaction projection $P : A_\bullet(X) \to Z_0(X)$:*

$$\eta\delta + \delta\eta = 1 - P \tag{2.63}$$

Proof. First, notice that $B_\bullet(X)$ is stable under δ so that $\mathbf{z}\delta\mathbf{b} = 0$ and δ splits in the triangular form:

$$\delta = \mathbf{b}\delta\mathbf{b} + \mathbf{b}\delta\mathbf{z} + \mathbf{z}\delta\mathbf{z} \tag{2.64}$$

As $\eta = \mathbf{z}d\mathbf{b}$, we then have:

$$\delta\eta + \eta\delta = (\mathbf{b}\delta\mathbf{z})\eta + \eta(\mathbf{b}\delta\mathbf{z}) + \mathbf{z}(\delta\eta + \eta\delta)\mathbf{b} \tag{2.65}$$

As maps between $Z_\bullet(X)$ and $B_\bullet(X)$, the interaction theorem implies that $\mathbf{b}\delta\mathbf{z}$ inverts η on non-zero degrees. More precisely, for every $p \geq 1$ and $\varphi \in A_p(X)$ we have:

$$(\mathbf{b}\delta\mathbf{z})(\varphi)_{\bar\beta} = (-1)^p \sum_{\bar\beta \supset \gamma'} \mathbf{z}_\gamma(\varphi_{\bar\beta\gamma'}) \tag{2.66}$$

Denoting by $\mathbf{z}_0$ the homogeneous projection onto $Z_0(X)$ induced by $\mathbf{z}$, we may write:

$$(\mathbf{b}\delta\mathbf{z})\eta = \mathbf{b} \quad \text{and} \quad \eta(\mathbf{b}\delta\mathbf{z}) = \mathbf{z} - \mathbf{z}_0 \tag{2.67}$$

so that $(\mathbf{b}\delta\mathbf{z})\eta + \eta(\mathbf{b}\delta\mathbf{z}) = 1 - \mathbf{z}_0$ and it only remains to show that $\mathbf{z}(\delta\eta + \eta\delta)\mathbf{b} = \mathbf{z}_0 - P$.

On the zero degree, we have for all $u \in A_0(X)$ and $\beta \in X$:

$$(\mathbf{z}\delta\eta)(u)_\beta = -\sum_{\alpha' \supset \beta} \mathbf{z}_\beta(u_{\alpha'}) \tag{2.68}$$

which is precisely $\mathbf{z}_0(u)_\beta - P(u)_\beta$.

On higher degrees, we have for every $\varphi \in A_p(X)$ and $\bar\beta \in N_p(X)$:

$$(\mathbf{z}\delta\eta)(\varphi)_{\bar\beta} = (-1)^{p+1} \sum_{k=0}^{p} (-1)^k \sum_{\partial_k \bar\alpha' = \bar\beta} \mathbf{z}_{\alpha'_{p+1}}(\varphi_{\alpha'_0 \dots \alpha'_p}) \tag{2.69}$$

The last summand $k = p+1$ of $\delta\eta(\varphi)_{\bar\beta}$ is valued in $B_{\bar\beta}$ and truncated by $\mathbf{z}_{\beta_p}$ which enforces $\alpha'_{p+1} = \beta_p$. On the other hand, we have:

$$(\eta\delta\mathbf{b})(\varphi)_{\bar\beta} = \mathbf{z}_{\beta_p}\left((-1)^p \sum_{k=0}^{p} (-1)^k \sum_{\delta_k \bar\beta' = \beta_0 \dots \beta_{p-1}} \mathbf{b}_{\beta'_p}(\varphi_{\bar\beta'}) \right) \tag{2.70}$$

Whenever $\beta_p \not\subset \beta'_p$, we have $\mathbf{z}_{\beta_p} \circ \mathbf{b}_{\beta'_p} = 0$. We may hence assume that $\beta_p \subset \beta'_p$ and let $\bar\alpha' = \bar\beta'\beta_p$ so that for every $k \leq p$ we have:

$$\delta_k \bar\alpha' = (\delta_k \bar\beta')\beta_p = \bar\beta \tag{2.71}$$

Comparing with the above, we see that $\mathbf{z}\delta\eta + \eta\delta\mathbf{b}$ vanishes on non-zero degrees, while it is equal to $\mathbf{z}\delta\eta = \mathbf{z}_0 - P$ otherwise, which finishes to show that:

$$\delta\eta + \eta\delta = 1 - \mathbf{z}_0 + \mathbf{z}_0 - P = 1 - P \tag{2.72}$$

□

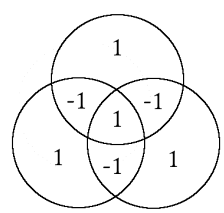
Theorem 2.17. *The homology groups of $A_\bullet(X)$ are given by:*

$$H_0(X) \simeq Z_0(X) \quad \text{and} \quad H_n(X) = 0 \quad \text{for} \quad n \geq 1 \tag{2.73}$$

Chapter 3

Combinatorics

In this chapter, we review the classical theory of Dirichlet convolution and Möbius inversion. They give algebraic foundations to many combinatorial problems, starting with the so-called inclusion-exclusion principles, generalising the usual formula $|E_i \cup E_j| = |E_i| + |E_j| - |E_i \cap E_j|$. The following structures however really emerged from considerations of number theory.



These methods give insight on the conceptual distinction between intensive and extensive quantities, stemming from the definition internal hamiltonian of a region $\alpha \subseteq \Omega$ as a sum of local interactions:

$$H_\alpha = \sum_{\alpha \supseteq \beta} h_\beta \tag{3.1}$$

The extensivity of H may be thought of as a consequence of the quick decay of h on large regions[1]. Möbius inversion will show the relation between H and h to be bijective, a fundamental correspondence which will be written as:

$$H = \zeta \cdot h \quad \Leftrightarrow \quad h = \mu \cdot H \tag{3.2}$$

In section 1, we start by reviewing the classical theory of incidence algebra. In section 2 we investigate some interesting properties and exhibit a locally cohomological character of the zeta transform $h \mapsto \zeta \cdot h$. In the last and exploratory section, we propose an extension of ζ and μ as two reciprocal endomorphisms on the whole complex of observable fields $A_\bullet(X)$. This extension of ζ to higher degrees is done consistently with the local cocyle properties satisfied on the zero degree, while the degree one Möbius transform will allow for various enhancements of the belief propagation algorithm.

[1]This remark shall be made more precise in the next chapter, where the extensivity of entropy *a priori* justifies the Bethe approximation scheme.

3.1 Dirichlet Convolution

The convolution product on the set of arithmetic functions $f : \mathbb{N}^* \to \mathbb{C}$ was first introduced by Dirichlet in 1837 to prove his theorem on arithmetic progressions. Relying on the partial order structure of $\mathbb{N}^*$ induced by divisibility, it is given by:

$$(f * g)(n) = \sum_{d|n} f(d)\, g(n/d) \tag{3.3}$$

A meromorphic function $\hat{f}$ may be associated to every arithmetic function f of subpolynomial growth by analytic extension of its Dirichlet generating series, defined for all $s \in \mathbb{C}$ such that $\mathrm{Re}(s)$ is large enough by:

$$\hat{f}(s) = \sum_{n \in \mathbb{N}^*} \frac{f(n)}{n^s} \tag{3.4}$$

The assignment $f \mapsto \hat{f}$ interlaces Dirichlet convolution with the usual product on $\mathbb{C}$ and we have for every arithmetic functions f and g:

$$\widehat{f * g} = \hat{f} \cdot \hat{g} \tag{3.5}$$

The Riemann zeta function $\hat{\zeta}$ extends the generating series of the arithmetic function ζ, defined by $\zeta(n) = 1$ for all $n \in \mathbb{N}^*$. The classical Möbius inversion formula states that ζ has an inverse for $*$ given by $\mu(n) = (-1)^k$ if n is the product of k distinct primes, and zero otherwise. This in turn yields the coefficients of the generating series of $\hat{\mu} = 1/\hat{\zeta}$.

Dirichlet convolution and Möbius inversion formulas were then studied in more general contexts, notably by Fréchet, and a fundamental and seminal reference on the subject is the general treatment of Möbius functions given by Rota in [24]. Dirichlet convolution may not only be defined on a locally finite partial order but on any locally finite category. We however restrict to the former and refer the interested reader to [13] for greater generality.

3.1.1 Incidence Ring

Given a locally finite[2] partially ordered set X and a unital ring R, let us denote by $R_1(X)$ the R-module of R-valued functions on the set $N_1 X$ of 1-chains in X:

$$R_1(X) = \big\{ \varphi : X \times X \to R \mid \alpha \not\to \beta \Rightarrow \varphi_{\alpha\beta} = 0 \big\} = R^{N_1 X} \tag{3.6}$$

It is equipped with the Dirichlet convolution product $*$ defined for every $\varphi, \psi \in R_1(X)$ by:

$$(\varphi * \psi)_{\alpha\gamma} = \sum_{\alpha \to \beta' \to \gamma} \varphi_{\alpha\beta'} \cdot \psi_{\beta'\gamma} \tag{3.7}$$

This product is associative. It has a unit $\mathbf{1}$, which is the Kronecker symbol[3] defined by $\mathbf{1}_{\alpha\alpha} = 1$ and $\mathbf{1}_{\alpha\beta} = 0$ if $\alpha \neq \beta$.

The unital ring $\big(R_1(X), +, *\big)$ is called the incidence ring of X, or incidence algebra when R is a field. Combinatorics are however mostly interested with relative integers and it is remarkable that Möbius inversion can be carried out on $\mathbb{Z}$.

[2] X is locally finite if the set of non-degenerate chains from any element to an other is finite.
[3] We prevent from using the usual Kronecker symbol notation δ to avoid confusion with codifferentials.

3.1.2 Incidence Modules

Let M denote a module on R, and denote by $M_1(X)$ the space of M-valued functions on N_1X. Then $M_1(X)$ is a module on the incidence ring $R_1(X)$ for the left action defined for $m \in M_1(X)$ by:

$$(\varphi \cdot m)_{\alpha\gamma} = \sum_{\alpha \to \beta' \to \gamma} \varphi_{\alpha\beta'} \cdot m_{\beta'\gamma} \tag{3.8}$$

when M is a left R-module, and symetrically, for the right action:

$$(m \cdot \psi)_{\alpha\gamma} = \sum_{\alpha \to \beta' \to \gamma} m_{\alpha\beta'} \cdot \psi_{\beta'\gamma} \tag{3.9}$$

when M is a right R-module.

Let us now denote by $M_0(X)$ the space of M-valued functions on X. It is also equipped with a left action of $R_1(X)$ defined for every $n \in M_0(X)$ by:

$$(\varphi \cdot n)_{\alpha} = \sum_{\alpha \to \beta'} \varphi_{\alpha\beta'} \cdot n_{\beta'} \tag{3.10}$$

when M is a left module, and otherwise with the right action:

$$(n \cdot \psi)_{\beta} = \sum_{\alpha' \to \beta} n_{\alpha'} \cdot \psi_{\alpha'\beta} \tag{3.11}$$

Following Rota, one may relate these two actions and append to X initial and terminal elements 0 and 1 when necessary, defining a possibly larger poset $\bar{X}$. To $m \in M_1(X)$ associate $n \in M_0(X)$ by $n_\alpha = m_{\alpha 1}$. This interlaces the two left actions of $R_1(\bar{X})$ on when M is a left module. Similarly, letting $n_\beta = m_{0\beta}$ interlaces the two right actions.

Notice that when $M = R = \mathbb{R}$, the vector spaces $R_1(X)$ and $\mathbb{R}_0(X)$ have a canonical scalar product for which the left and right actions define adjoint endomorphisms.

3.1.3 Systems

Given a cosystem $R : X^{op} \to \mathbf{Ring}$, we may similarly equip the space of upper 1-fields $\mathbb{R}^1(X)$ with a convolution product. There is a ring morphism $R_\alpha \leftarrow R_\beta$ for every $\alpha \to \beta$ in X, and for $r_\beta \in R_\beta$ we denote by $[r_\beta]_\alpha$ its image in R_α. Then for $\varphi, \psi \in R_1(X)$, let:

$$(\varphi * \psi)_{\alpha\gamma} = \sum_{\alpha \to \beta' \to \gamma} \varphi_{\alpha\beta'} \cdot [\psi_{\beta'\gamma}]_\alpha \tag{3.12}$$

By duality, we may similarly define the convolution product $*$ on the space of lower fields $R_1(X)$ when $R : X \to \mathbf{Ring}$ is a system of rings.

When $M : X^{op} \to \mathbf{Ring}$ is a module over the cosystem R, we may extend the left action of $R^1(X)$ on the space of upper 1-fields $M^1(X)$. We have a morphism $M_\alpha \leftarrow M_\beta$ for every $\alpha \to \beta$ in X, and for every $m_\beta \in M_\beta$, we still denote by $[m_\beta]_\alpha$ its image in M_α. The action is then given by:

$$(\varphi \cdot \psi)_{\alpha\gamma} = \sum_{\alpha \to \beta' \to \gamma} \varphi_{\alpha\beta'} \cdot [m_{\beta'\gamma}]_\alpha \tag{3.13}$$

And a left action of $R^1(X)$ can similarly be defined on $M_0(X)$ by:

$$(\varphi \cdot m)_{\alpha} = \sum_{\alpha \to \beta'} \varphi_{\alpha\beta'} \cdot [m_{\beta'}]_\alpha \tag{3.14}$$

The action of $\mathbb{Z}_1(X)$ on the module $A_0(X)$ of observable fields, falls into this case. It is the main example we shall be interested in, before considering functors M assigning to each M_α the space of functionals on $A_0(\Lambda^\alpha)$ or on the set Σ_α of states of A_α.

If R is a fixed ring, then $R^1(X) = R_1(X)$ has a natural action on the space of lower fields $M_1(X)$:

$$(\varphi \cdot m)_{\alpha\gamma} = \sum_{\alpha \to \beta'} \varphi_{\alpha\beta'} \cdot m_{\beta'\gamma} \tag{3.15}$$

Although functoriality of M does not appear, this case is worth mentioning as it covers for instance the action of $\mathbb{Z}_1(X)$ on $A_1(X)$.

Dual constructions take place when M is a covariant functor. In particular, this is the case when $M = A^*$ is the module of linear forms on observables, and the right action of $\mathbb{Z}_1(X)$ on $A_0^*(X)$ is the adjoint of its left action on $A_0(X)$.

3.2 Properties of the Zeta Transform

The zeta function $\zeta \in \mathbb{Z}_1(X)$ is defined by $\zeta_{\alpha\beta} = 1$ for all $\alpha \to \beta$ in X. Given a cosystem of abelian groups $M : X^{op} \to \mathbf{Ab}$, the left action of ζ on $M_0(X)$ is given by:

$$(\zeta \cdot f)_\alpha = \sum_{\alpha \to \beta} [f_\beta]_\alpha \tag{3.16}$$

This section exposes elementary properties of the endomorphism thus defined on $M_0(X)$. The first one, bijectivity, is given by the classical Möbius inversion formulas. The next one, locality, is also significant however obvious it is. As an easy consequence of the previous chapter, we then introduce a local cocycle property satisfied by the zeta transform.

3.2.1 Möbius Inversion

The fundamental theorem of Möbius inversion states that ζ has an inverse $\mu \in \mathbb{Z}_1(X)$ for Dirichlet convolution, it naturally implies the bijectivity of the zeta transform on any module.

Theorem 3.1 (Möbius - Rota)**.** *The inverse of $\zeta \in \mathbb{Z}_1(X)$ is the Möbius function $\mu \in \mathbb{Z}_1(X)$:*

$$\mu = \sum_{k \geq 0} (-1)^k (\zeta - \mathbf{1})^k \tag{3.17}$$

Proof. The n-th power of ζ for $*$ counts the number of n-chains between any two elements $\alpha, \beta \in X$:

$$\zeta^n_{\alpha\beta} = \sum_{\alpha = \alpha_0 \to \dots \to \alpha_n = \beta} 1 \tag{3.18}$$

and $(\zeta - \mathbf{1})^n_{\alpha\beta}$ similarly counts the number of non-degenerate n-chains from α to β. Because X is locally finite, there exists N such that $(\zeta - \mathbf{1})^N_{\alpha\beta} = 0$ and $\zeta = \mathbf{1} + (\zeta - \mathbf{1})$ is invertible. □

Note that more practical identities arise from the relations $\mu * \zeta = \zeta * \mu = \mathbf{1}$, allowing to compute the values of μ inductively. Starting from $\mu_{\alpha\alpha} = 1$, one may for instance iterate over γ under α:

$$\mu_{\alpha\gamma} = 1 - \sum_{\alpha \geq \beta' > \gamma} \mu_{\alpha\beta'} \tag{3.19}$$

The Möbius function μ is closely related to diverse inclusion-exclusion principles. Consider for example the problem of finding a collection of integers $(c_\alpha) \in \mathbb{Z}_0(X)$ such that for all $\beta \in X$:

$$\sum_{\alpha' \to \beta} c_{\alpha'} = 1 \tag{3.20}$$

One may compute these coefficients inductively starting from maximal cells. Note that this expression involves the right action of ζ, and as by Möbius inversion $c \cdot \zeta = 1$ is equivalent to $c = 1 \cdot \mu$ we have:

$$c_\beta = \sum_{\alpha' \to \beta} \mu_{\alpha'\beta} \tag{3.21}$$

We call $c = 1 \cdot \mu$ and $\bar{c} = \mu \cdot 1$ the right and left Möbius numbers.

Given $X \subseteq \tilde{X}$, let us say that X is *full* if for all $\alpha \in X$, every $\beta \in \tilde{X}$ such that $\beta \subseteq \alpha$ is also in X.

Proposition 3.2. *If X is full in $\tilde{X}$, then the restriction of the Möbius function of $\tilde{X}$ to X is the Möbius function of X.*

Proposition 3.3. *If $\tilde{X} = \{\Omega\} \sqcup X$ completes X with an initial element Ω and $\tilde{\mu} \in \mathbb{Z}_1(\tilde{X})$ denotes the Möbius function on $\tilde{X}$, then for every $\beta \in X$ we have $c_\beta = -\tilde{\mu}_{\Omega\beta}$.*

Proof. The Möbius inversion formula in $\tilde{X}$ between Ω and $\beta \in X$ gives:

$$\sum_{\Omega \to \alpha \to \beta} \tilde{\mu}_{\alpha\beta} = \tilde{\mu}_{\Omega\beta} + \sum_{\alpha \to \beta} \mu_{\alpha\beta} = \tilde{\mu}_{\Omega\beta} + c_\beta = 0 \tag{3.22}$$

□

Examples.

1. Let $X = \mathbb{N}$ denote the standard total order of integers. The Möbius function is given by $\mu_{\alpha\alpha} = 1$ and $\mu_{\alpha\beta} = -1$ if $\beta = \alpha + 1$, zero otherwise. We recover the classical finite differences formula:

$$U_\alpha = \sum_{\beta=0}^{\alpha} u_\beta \quad \Leftrightarrow \quad u_\alpha = U_\alpha - U_{\alpha-1} \tag{3.23}$$

   which is the discrete version of the fundamental theorem of calculus.

2. Let $X = \mathcal{P}(\Omega)$ for a finite set Ω and denote by $|\alpha|$ the cardinal of $\alpha \subseteq \Omega$. The Möbius function is then given by:

$$\mu_{\alpha\beta} = (-1)^{|\alpha|-|\beta|} \tag{3.24}$$

   Suppose given now for each $i \in \Omega$ a measurable subset $\mathcal{A}_i$ of some probability space, relate to:

$$\mathbb{P}(\cup_{i\in\alpha}\mathcal{A}_i) = \sum_{\beta\subseteq\alpha} (-1)^{|\beta|}\mathbb{P}(\cap_{j\in\beta}\mathcal{A}_j) \tag{3.25}$$

3. Total interaction $P^{\beta\alpha}$ and marginal projections $\Sigma^{\beta\alpha}$.

$$\Sigma^{\beta\alpha} = \sum_{\beta\to\gamma'} P^{\gamma'\alpha} \quad \Leftrightarrow \quad P^{\beta\alpha} = \sum_{\beta\to\gamma'} \mu_{\beta\gamma'} \cdot \Sigma^{\gamma'\alpha} \tag{3.26}$$

   This is a property of the canonical metric and the marginal projections, beware that in general:

$$\mathbb{E}^{\beta\alpha} \neq \sum_{\beta\to\gamma'} P^{\gamma'\alpha} \tag{3.27}$$

   as $\bigoplus_{\gamma\notin\Lambda^\beta} Z_\gamma$ might not be the orthogonal to A_β in A_α.

3.2.2 Locality

For every $\alpha \in X$, we denote by Λ^α the cone under α consisting of those $\beta \in X$ such that $\alpha \to \beta$. As subset of X, there is a natural restriction map:

$$\mathrm{r}_\alpha : M_0(X) \longrightarrow M_0(\Lambda^\alpha) \tag{3.28}$$

For every $f \in M_0(X)$ the value of $(\zeta \cdot f)_\alpha$ only depends on the values f_β for $\alpha \to \beta$. In other words, ζ commutes with the restriction to Λ^α and we have $\mathrm{r}_\alpha \circ \zeta = \zeta \circ \mathrm{r}_\alpha$.

$$\begin{tikzcd} M_0(X) \arrow[r, "\mathrm{r}_\alpha"] \arrow[d, "\zeta"'] & M_0(\Lambda^\alpha) \arrow[d, "\zeta"] \\ M_0(X) \arrow[r, "\mathrm{r}_\alpha"'] & M_0(\Lambda^\alpha) \end{tikzcd} \tag{3.29}$$

This locality property comes from the form of the action of $\mathbb{Z}_1(X)$ on $M_0(X)$ and is absolutely not specific to ζ. Such local endomorphisms preserve the sheaf structure[4] of $M_0(X)$.

Denoting by $i_\alpha : M_0(X) \to M_\alpha$ the evaluation on α, a consequence of the locality of ζ is that we may factorise $i_\alpha \circ \zeta$ through r_α:

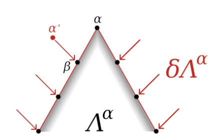
$$\begin{tikzcd} M_0(X) \arrow[r, "\mathrm{r}_\alpha"] \arrow[dr, "i_\alpha \zeta"'] & M_0(\Lambda^\alpha) \arrow[d, dashed, "\zeta_\alpha"] \\ & M_\alpha \end{tikzcd} \tag{3.30}$$

We denote by ζ_α the factorised map, and will now show that ζ_α induces a map in homology.

3.2.3 Local Cocycle Property

In this paragraph, we relate the Gauss formula 2.3 to properties of the zeta transform, fortifying its analogy with integral calculus underlined by Rota. Recall that the coboundary of Λ^α was defined as the set of 1-chains $d\Lambda^\alpha = \left\{ \alpha'\beta' \in N_1(X) \,\middle|\, \beta' \in \Lambda^\alpha \text{ and } \alpha' \notin \Lambda^\alpha \right\}$ so that for every $\varphi \in M_1(X)$:

$$\sum_{\beta' \in \Lambda^\alpha} \delta_{\beta'} \varphi = \sum_{\alpha'\beta' \in d\Lambda^\alpha} \varphi_{\alpha'\beta'}$$



In chapter 5, we shall rely on this Gauss formula to relate belief propagation to a transport equation. Its right hand side may also be related to the zeta transform, as expressed by the following equivalent proposition, while extending the zeta transform to $M_1(X)$ will allow for the even more natural formula (3.55) of the next section.

Proposition 3.4. *For every $\varphi \in M_1(X)$, we have:*

$$\zeta(\delta\varphi)_\alpha = \sum_{\alpha' \notin \Lambda^\alpha} \zeta(i_{\alpha'}\varphi)_\alpha \tag{3.31}$$

In particular, $i_\alpha \zeta$ vanishes on $\delta M_1(\Lambda^\alpha)$.

[4] For the restriction maps $M_0(Y) \to M_0(Z)$ induced by every $Z \subseteq Y \subseteq X$.

It follows from proposition 3.4 property that ζ_α factorises through the canonical projection of $M_0(\Lambda^\alpha)$ onto its quotient by $\delta M_1(\Lambda^\alpha)$:

$$\begin{tikzcd} M_0(\Lambda^\alpha) \arrow[r] \arrow[dr, "\zeta_\alpha"'] & \dfrac{M_0(\Lambda^\alpha)}{\delta M_1(\Lambda^\alpha)} \arrow[d, dashed] \\ & M_\alpha \end{tikzcd} \tag{3.32}$$

and in particular, ζ_α induces a map in homology $[\zeta_\alpha] : H_0(\Lambda^\alpha) \longrightarrow M_\alpha$. In the case of the complex of local observables $A_\bullet(X)$ defined in 2.6, it follows from theorem 2.14 that $[\zeta_\alpha]$ actually induces an isomorphism between the homology classes of $A_0(\Lambda^\alpha)$ and its image $B_\alpha \subseteq A_\alpha$. This fact is a consequence of the interaction decomposition theorem, and in general, $[\zeta_\alpha]$ may fail to be injective.

3.3 Higher Degree Combinatorics

In this section, we extend the zeta transform to a full endomorphism of the chain complex $M_\bullet(X)$ satisfying higher degree analogs of the previous section properties. Our initial motivation was to search for a higher degree diffusion equation that would generalise belief propagation. This program appeared only feasible for the linearised algorithm, although it soon became clear to us that the degree 1 zeta transform was already hidden in the degree 0 algorithm with Dirichlet boundary conditions, and that modifying the degree 0 algorithm to perform a degree 1 Möbius inversion on its messages seemed deeply natural to eliminate their redundancies.

In proving the invertibility of our extension of ζ, we will need to assume that $X \subseteq \mathcal{P}(\Omega)$ is closed under intersection[5] and forms a semi-lattice. For every $\beta \in X$ and $\alpha_0 \dots \alpha_n \in N_n(X)$, we denote by $\beta \cap (\alpha_0 \dots \alpha_n)$ the possibly degenerate chain $(\beta \cap \alpha_0) \dots (\beta \cap \alpha_n)$, while it will be implicit in our notations that fields vanish on non-ordered sequences.

3.3.1 Extended Zeta Transform

For every $\alpha \in X$ and $\beta \subseteq \alpha$, we denote by Λ^α_β the complement of Λ^β in Λ^α. When $\varphi \in M_1(X)$ is of degree one, we shall define $\zeta(\varphi)_{\alpha\beta}$ as the flux passing from Λ^α to Λ^β:

$$\zeta(\varphi)_{\alpha\beta} = \sum_{\beta' \in \Lambda^\alpha_\beta} \sum_{\gamma' \in \Lambda^\beta} \varphi_{\beta'\gamma'} \tag{3.33}$$

More generally, a chain $\alpha_0 \dots \alpha_n$ in $N_n(X)$ yields a sequence of cones $\Lambda^{\alpha_0} \supseteq \dots \supseteq \Lambda^{\alpha_n}$ and we simply define ζ by summing in the interspaces.

Definition 3.5. *We extend the zeta transform to $\zeta : M_\bullet(X) \to M_\bullet(X)$ by letting for every $\varphi \in M_n(X)$:*

$$\zeta(\varphi)_{\alpha_0 \dots \alpha_n} = \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \sum_{\beta_1 \in \Lambda^{\alpha_1}_{\alpha_2}} \cdots \sum_{\beta_n \in \Lambda^{\alpha_n}} \varphi_{\beta_0 \dots \beta_n} \tag{3.34}$$

From the above formula, it clearly appears that the definition of ζ involves an inductive extension to higher degrees and the following proposition will greatly ease computations.

Proposition 3.6. *For every $n \geq 1$, the action of ζ on $M_n(X)$ is related to that on $M_{n-1}(X)$ by:*

$$\zeta(\varphi)_{\alpha_0 \dots \alpha_n} = \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \zeta(i_{\beta_0}\varphi)_{\alpha_1 \dots \alpha_n} \tag{3.35}$$

[5] The same hypothesis was necessary for the interaction decomposition to hold, section 2.3.2.

3.3.2 Locality and Colocality

For any $\alpha \in X$, the zeta transform still commutes with the restriction to Λ^α:

$$\begin{tikzcd} M_\bullet(X) \arrow[r, "\mathrm{r}_\alpha"] \arrow[d, "\zeta"'] & M_\bullet(\Lambda^\alpha) \arrow[d, "\zeta"] \\ M_\bullet(X) \arrow[r, "\mathrm{r}_\alpha"'] & M_\bullet(\Lambda^\alpha) \end{tikzcd} \tag{3.36}$$

The locality of ζ is better precised by the following proposition, that makes use of the semi-lattice structure of X and will soon come to use in extending Möbius inversion.

Proposition 3.7. *For every $\beta \in X$ and $\alpha_0 \dots \alpha_n \in N_n(X)$, we have:*

$$\zeta(\mathrm{r}_\beta \varphi)_{\alpha_0 \dots \alpha_n} = \zeta(\varphi)_{\beta \cap (\alpha_0 \dots \alpha_n)} \tag{3.37}$$

Proof. As $\mathrm{r}_\beta\varphi$ is supported inside Λ^β, the sums over $\Lambda^{\alpha_i}_{\alpha_{i+1}}$ may be restricted to $\Lambda^\beta \cap \Lambda^{\alpha_i}_{\alpha_{i+1}} = \Lambda^{\beta\cap\alpha_i}_{\beta\cap\alpha_{i+1}}$. □

Locality also implies that for every $n \geq 1$, composing the partial evaluation $i_{\alpha_0} : M_n(X) \to M_{n-1}(\Lambda^{\alpha_0})$ with ζ factorises through r_{α_0}:

$$\begin{tikzcd} M_n(X) \arrow[r, "\mathrm{r}_{\alpha_0}"] \arrow[dr, "i_{\alpha_0}\zeta"'] & M_n(\Lambda^{\alpha_0}) \arrow[d, dashed] \\ & M_{n-1}(\Lambda^{\alpha_0}) \end{tikzcd} \tag{3.38}$$

A more subtle observation on the support of ζ is given by the following dual property, expressing the independence of $i_{\alpha_1} i_{\alpha_0}\zeta$ with respect to fields supported inside Λ^{α_1}.

Proposition 3.8. *For every $\alpha_0 \dots \alpha_n \in N_n(X)$ and $\varphi \in M_n(\Lambda^{\alpha_1})$, we have $\zeta(\varphi)_{\alpha_0\dots\alpha_n} = 0$.*

The sum defining $\zeta(\varphi)_{\alpha_0\dots\alpha_n}$ indeed only involves evaluations in the first argument of φ inside $\Lambda^{\alpha_0}_{\alpha_1} = \Lambda^{\alpha_0} \smallsetminus \Lambda^{\alpha_1}$ and does not depend on the restriction of φ to Λ^{α_1}. The kernel of $i_{\alpha_1} i_{\alpha_0}\zeta$ hence contains $M_n(\Lambda^{\alpha_1})$ and we have the following factorisation:

$$\begin{tikzcd} M_n(X) \arrow[r] \arrow[dr, "i_{\alpha_1} i_{\alpha_0}\zeta"'] & \dfrac{M_n(\Lambda^{\alpha_0})}{M_n(\Lambda^{\alpha_1})} \arrow[d, dashed] \\ & M_{n-2}(\Lambda^{\alpha_1}) \end{tikzcd} \tag{3.39}$$

3.3.3 Extended Möbius Transform

The definition of μ will make use of the semi-lattice structure of X. For every unordered sequence $\beta_0, \dots, \beta_n \in X$, let us denote by $[\beta_0 \dots \beta_n]_\cap$ the n-chain $\beta_0(\beta_0 \cap \beta_1) \dots (\beta_0 \cap \dots \cap \beta_n)$.

Definition 3.9. *We extend the Möbius transform to $\mu : M_\bullet(X) \to M_\bullet(X)$ by letting for all $\Phi \in M_n(X)$:*

$$\mu(\Phi)_{\alpha_0\dots\alpha_n} = \sum_{\beta_n \in \Lambda^{\alpha_n}} \mu_{\alpha_n\beta_n} \cdots \sum_{\beta_1 \in \Lambda^{\alpha_1}_{\beta_2}} \mu_{\alpha_1\beta_1} \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\beta_1}} \mu_{\alpha_0\beta_0} \cdot \Phi_{[\beta_0\dots\beta_n]_\cap} \tag{3.40}$$

In contrast with ζ, the supports of the sums defining μ do depend on the summation variables β_i. We may still give an inductive construction of μ, allowing to conveniently prove reciprocity with ζ. For every $\alpha_0 \in X$ and $n \geq 1$, let us introduce the map $\nu_{\alpha_0} : M_n(X) \to M_{n-1}(X)$ defined by:

$$\nu_{\alpha_0}(\Phi)_{\alpha_1 \dots \alpha_n} = \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \mu_{\alpha_0 \beta_0} \cdot \Phi_{\beta_0 \cap (\alpha_0 \dots \alpha_n)} \tag{3.41}$$

and extend this map to $\nu_{\alpha_0} : M_0(X) \to M_{\alpha_0}$ by letting:

$$\nu_{\alpha_0}(\Phi) = \sum_{\beta_0 \in \Lambda^{\alpha_0}} \mu_{\alpha_0 \beta_0} \cdot \Phi_{\beta_0} \tag{3.42}$$

The action of $\mu \in M_n(X)$ then deduces from that on $M_{n-1}(X)$ by the relation $i_{\alpha_0} \circ \mu = \mu \circ \nu_{\alpha_0}$ as expressed by the following proposition.

Proposition 3.10. *For every $\alpha_0 \dots \alpha_n \in N_n(X)$ and $\Phi \in M_n(X)$, we have:*

$$\mu(\Phi)_{\alpha_0 \dots \alpha_n} = \nu_{\alpha_n} \dots \nu_{\alpha_0}(\Phi) \tag{3.43}$$

Proof. In the last sum over $\beta_0 \in \Lambda^{\alpha_0}_{\beta_1}$ defining $\mu(\Phi)_{\alpha_0 \dots \alpha_n}$, we may recognise $\nu_{\alpha_0}(\Phi)_{[\beta_1 \dots \beta_n]_\cap}$ □

With this characterisation of μ, we may now generalise the Möbius inversion formula to $M_\bullet(X)$. Having two reciprocal endomorphisms ζ and μ acting on all degrees will allow to conjugate non-homogeneous operators and consider[6] for instance $\delta^\zeta = \zeta \circ \delta \circ \mu$ or $\nabla^\mu = \mu \circ \nabla \circ \zeta$.

Theorem 3.11. *The Möbius transform is the inverse of the zeta transform.*

The proof[7] of the inversion theorem will come as an easy consequence of the following lemma:

Lemma 3.12. *For every α_0 in X, we have:*

$$\nu_{\alpha_0} \circ \zeta = \zeta \circ i_{\alpha_0} \tag{3.44}$$

The above extends to $M_0(X) \to M_{\alpha_0}$ by agreeing that $\zeta \circ i_{\alpha_0} = i_{\alpha_0}$,

Proof. Injecting the recurrence relation defining ζ, we may write $\nu_{\alpha_0}\zeta(\varphi)_{\alpha_0 \dots \alpha_n}$ as:

$$\sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \mu_{\alpha_0 \beta_0} \cdot \zeta(\varphi)_{\beta_0 \cap (\alpha_0 \dots \alpha_n)} = \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \mu_{\alpha_0 \beta_0} \sum_{\gamma_0 \in \Lambda^{\beta_0} \cap \Lambda^{\alpha_0}_{\alpha_1}} \zeta(i_{\gamma_0}\varphi)_{\beta_0 \cap (\alpha_1 \dots \alpha_n)} \tag{3.45}$$

As $i_{\gamma_0}\varphi$ is supported in Λ^{γ_0}, we have $\zeta(i_{\gamma_0}\varphi)_{\beta_0 \cap (\alpha_1 \dots \alpha_n)} = \zeta(i_{\gamma_0}\varphi)_{\gamma_0 \cap (\alpha_1 \dots \alpha_n)}$ in virtue of prop.3.7. Both β_0 and γ_0 running over $\Lambda^{\alpha_0}_{\alpha_1}$ with the only additional condition that $\beta_0 \to \gamma_0$, we have:

$$\nu_{\alpha_0}\zeta(\varphi)_{\alpha_0 \dots \alpha_n} = \sum_{\gamma_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \left(\sum_{\alpha_0 \to \beta_0 \to \gamma_0} \mu_{\alpha_0 \beta_0} \right) \zeta(i_{\gamma_0}\varphi)_{\gamma_0 \cap (\alpha_1 \dots \alpha_n)} \tag{3.46}$$

Recognising the classical Möbius inversion formula $(\mu * \zeta)_{\alpha_0 \gamma_0} = \mathbf{1}_{\alpha_0 \gamma_0}$, we get:

$$\nu_{\alpha_0}\zeta(\varphi)_{\alpha_0 \dots \alpha_n} = \zeta(i_{\alpha_0}\varphi)_{\alpha_1 \dots \alpha_n} \tag{3.47}$$

When φ is of degree zero the identity reduces to the $\nu_{\alpha_0}\zeta(\varphi) = \mu\zeta(\varphi)_{\alpha_0} = \varphi_{\alpha_0}$. □

Proof of theorem 3.11. For every $\alpha_0 \dots \alpha_n$ in $N_n(X)$, we have:

$$\begin{aligned} i_{\alpha_n} \dots i_{\alpha_0} \circ \mu \circ \zeta &= \nu_{\alpha_n} \dots \nu_{\alpha_0} \circ \zeta \\ &= \zeta \circ i_{\alpha_n} \dots i_{\alpha_0} \\ &= i_{\alpha_n} \dots i_{\alpha_0} \end{aligned} \tag{3.48}$$

□

[6] See chapter 5.
[7] [[appendix: add explicit computation on $A_1(X)$.]]

3.3.4 Local Cocycle Property

The following formula is the higher degree analog of proposition 3.4.

Proposition 3.13. *For every $\alpha_0 \dots \alpha_n \in N_n(X)$ and $\psi \in M_{n+1}(X)$ we have:*

$$\zeta(\delta\psi)_{\alpha_0\dots\alpha_n} = \sum_{\alpha_0' \notin \Lambda^{\alpha_0}} \zeta(i_{\alpha_0'}\psi)_{\alpha_0\dots\alpha_n} \tag{3.49}$$

In particular, if $\psi \in M_n(\Lambda^{\alpha_0})$, then $i_{\alpha_0}\zeta(\delta\psi) = 0$.

Proof. We will prove the relation by induction on the degree n. Let us first recall the module Leibniz rule for the interior product $\lrcorner$, in section 2.2 example 2:

$$\delta \circ i_\beta = i_\beta \circ \Big(\sum_{\alpha \in X} i_\alpha \Big) - i_\beta \circ \delta \tag{3.50}$$

For every $\psi \in M_{n+1}(X)$ and by construction of ζ, we may thus rewrite $\zeta(\delta\psi)_{\alpha_0\dots\alpha_n}$ as:

$$\begin{aligned}\sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \zeta(i_{\beta_0}\delta\psi)_{\alpha_1\dots\alpha_n} = & \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \sum_{\alpha_0' \in X} \zeta(i_{\beta_0} i_{\alpha_0'}\psi)_{\alpha_1\dots\alpha_n} \\ & - \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \zeta(\delta i_{\beta_0}\psi)_{\alpha_1\dots\alpha_n}\end{aligned} \tag{3.51}$$

We may now use the proposition on degree $n-1$ to rewrite the summand of the second term. Because $i_{\beta_0}\psi$ is supported inside $\Lambda^{\beta_0} \subseteq \Lambda^{\alpha_0}$, observing that $(X \setminus \Lambda^{\alpha_1}) \cap \Lambda^{\beta_0} = \Lambda^{\alpha_0}_{\alpha_1} \cap \Lambda^{\beta_0}$ we get:

$$\zeta(\delta i_{\beta_0}\psi)_{\alpha_1\dots\alpha_n} = \sum_{\alpha_1' \in \Lambda^{\alpha_0}_{\alpha_1}} \zeta(i_{\alpha_1'} i_{\beta_0}\psi)_{\alpha_1\dots\alpha_n} = \zeta(i_{\beta_0}\psi)_{\alpha_0\dots\alpha_n} \tag{3.52}$$

Swapping the first two sums in the previous expression of $\zeta(\delta\psi)_{\alpha_0\dots\alpha_n}$ we are left with the difference:

$$\zeta(\delta\psi)_{\alpha_0\dots\alpha_n} = \sum_{\alpha_0' \in X} \zeta(i_{\alpha_0'}\psi)_{\alpha_0\dots\alpha_n} - \sum_{\beta_0 \in \Lambda^{\alpha_0}_{\alpha_1}} \zeta(i_{\beta_0}\psi)_{\alpha_0\dots\alpha_n} \tag{3.53}$$

In virtue of prop.3.8, we have $\zeta(i_{\beta_0}\psi)_{\alpha_0\dots\alpha_n} = 0$ for every $\beta_0 \in \Lambda^{\alpha_1}$. We get the desired formula on degree n by rewriting the second sum over $\beta_0 \in \Lambda^{\alpha_0}$. □

For every $\alpha_0 \in X$, the local cocyle property similarly implies that we have the factorisation:

$$\begin{tikzcd} M_n(X) \arrow[r] \arrow[dr, "i_{\alpha_0}\zeta"'] & \dfrac{M_n(\Lambda^{\alpha_0})}{\delta M_{n+1}(\Lambda^{\alpha_0})} \arrow[d, dashed] \\ & M_{n-1}(\Lambda^{\alpha_0}) \end{tikzcd} \tag{3.54}$$

This will come as an essential feature of ζ when defining higher degree transport equations generalising belief propagation. Letting a field $\varphi \in M_n(X)$ evolve up to a boundary term $\delta\psi$, the partial evaluations $i_{\alpha_0}\Phi$ of its zeta transform $\Phi = \zeta(\varphi)$ shall only depend on the generalised messages ψ coming from the outside of Λ^{α_0}.

Note that when X contains a maximal element Ω, proposition 3.13 may be rewritten as:

$$\zeta(\delta\psi)_{\bar\alpha} = \zeta(\psi)_{\Omega\bar\alpha} \tag{3.55}$$

In general, formula (3.55) could serve as a natural definition for the notation $\zeta(\psi)_{\Omega\bar\alpha}$. One could also define $\tilde X = \{\Omega\} \sqcup X$ by prepending X with an initial element and extending the module system M by $M_\Omega = \operatorname{colim}_\alpha M_\alpha$. This point of view will be very useful in understanding the canonical diffusion flux in chapter 5 and proving proposition 5.32.

Theorem 3.14. *We have the commutation relation:*

$$\tilde{\zeta} \circ \delta = i_\Omega \circ \tilde{\zeta} \tag{3.56}$$

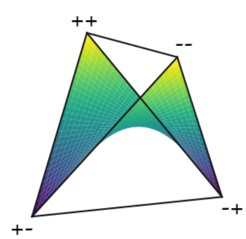
so that $\tilde{\zeta} : M_\bullet(X) \to M_\bullet(\tilde{X})$ *defines a morphism of chain complexes between* $\big(M_\bullet(X), \delta\big)$ *and* $\big(M_\bullet(\tilde{X}), i_\Omega\big)$ *on positive degrees, while* $M_\bullet(\tilde{X})$ *is naturally extended by* $M_{-1}(\tilde{X}) = M_\Omega$.

Note that $M_0(X)$ is then naturally isomorphic to the 0-cycles of $M_0(\tilde{X})$. The extended complex $M_\bullet(\tilde{X})$ is acyclic as $i_\Omega \varphi = 0$ implies $\varphi = i_\Omega(e_\Omega \varphi)$, although this require $M_\bullet(X)$ to be acyclic.

Chapter 4

Energy and Information

Many laws of nature remarkably take the form of variational principles. Since the lagrangian formulation of mechanics they have become a fundamental constituent of most physical theories, from thermodynamics to modern quantum field theory. As likelihood optimisation problems, they are also now a central occupation in data science and artificial intelligence, leading one to wonder if such variational principles should not help in understanding the self-organisation of biological systems. Set aside the theoretical beauty of variational principles remains the challenge of designing efficient computations of their solutions, while local and parallel optimisation algorithms in return provide with good abstract models[1] for neuronal interactions.



Entropy generates such variational principles and could be seen as the main object of this chapter. Legendre duality has long been a classical technique in thermodynamics, transforming variational principles into differently constrained ones. Considering Legendre duality as a smooth correspondence between observables and statistical states will motivate the formal study of free energy, viewed as a functional of the hamiltonian, given in section 1. The functorial properties satisfied by effective energy, a conditional form of free energy, will be essential to understand the structure of belief propagation in chapter 5. In section 2 we relate mutual informations to a combinatorial localisation of entropy. Giving a fine description on the structure of interactions and correlations, mutual informations are fundamental in understanding the content of Bethe's approximation of entropy. The main result of this chapter concludes section 3: we give a homological description of the solutions of Kikuchi's cluster variation method, which consists in estimating marginals of the Gibbs state by solving a variational principle on the local approximation of entropy.

[1] See for instance [6] for models of active inference and neuronal message-passing relying on belief propagation.

4.1 Free and Effective Energies

In this section, we first review some fundamental properties of free energy, seen as a functional on observables, although it is more customarily defined as a function of the inverse temperature $\theta = \frac{1}{kT}$ once given the hamiltonian H governing the system. Free energy is then defined from the partition function $Z(\theta) = \sum \mathrm{e}^{-\theta H}$ by letting $F(\theta) = -\frac{1}{\theta} \ln Z(\theta)$. It is well known that both functionals encode many physical properties of the system: the partition function may be seen as the Laplace transform of the spectral density of H and its derivatives yield the moments of H in the Gibbs state $[\mathrm{e}^{-\theta H}]$.

Temperature simply acts as an energy scaling and we here argue that given a region $\alpha \in X$, free energy is best seen as a smooth function $\mathbb{F}^\alpha \in C^\infty(A_\alpha)$ of the hamiltonian itself. More generally, for every subregion $\beta \subseteq \alpha$, partial integration along the fibers of $E_\alpha \to E_\beta$ in the sum defining $\mathbb{F}^\alpha$ will yield a smooth map $\mathbb{F}^{\beta\alpha} \in C^\infty(A_\alpha, A_\beta)$ which we call effective energy. We finally show that our definition of free energy simply generates Gibbs propability densities through its differential $\mathbb{F}^\alpha_* : A_\alpha \to A^*_\alpha$, while the differential of effective energies will yield conditional expectations in the Gibbs state.

4.1.1 Free Energy

Definition 4.1. *For every $\alpha \in X$, we call free energy the functional $\mathbb{F}^\alpha \in C^\infty(A_\alpha)$ defined by:*

$$\mathbb{F}^\alpha(H_\alpha) = -\ln\Big(\sum_{E_\alpha} \mathrm{e}^{-H_\alpha}\Big) \tag{4.1}$$

This definition agrees with the free energy of an isolated system governed by the hamiltonian H_α at inverse temperature $\theta = 1$. One may reintroduce the temperature dependency by setting:

$$\mathbb{F}^\alpha_\theta(H_\alpha) = \theta^{-1}\mathbb{F}^\alpha(\theta H_\alpha) \tag{4.2}$$

A fundamental property of free energy is additivity along non-interacting subsystems, which may be thought of as a weaker form of extensivity.

Proposition 4.2 (Additivity)**.** *If α is the disjoint union of $\beta_1, \dots, \beta_n$ and H_α splits as $\sum_i H_{\beta_i}$, then:*

$$\mathbb{F}^\alpha\Big(\sum_i H_{\beta_i}\Big) = \sum_i \mathbb{F}^{\beta_i}(H_{\beta_i}) \tag{4.3}$$

Proof. When $\alpha = \sqcup_i \beta_i$ we have: $-\ln \sum_{E_\alpha} \prod_{i=1}^n \mathrm{e}^{-H_{\beta_i}} = -\ln \prod_{i=1}^n \sum_{E_{\beta_i}} \mathrm{e}^{-H_{\beta_i}}$ □

The additivity of $\mathbb{F}^\alpha$ along constants may be seen as a particular case of the previous proposition, identifying $\mathbb{R}$ with $A_\varnothing$. For every $H_\alpha \in A_\alpha$ and $\lambda \in \mathbb{R}$, we have:

$$\mathbb{F}^\alpha(H_\alpha + \lambda) = \mathbb{F}^\alpha(H_\alpha) + \lambda \tag{4.4}$$

One should however note that $\mathbb{F}^\alpha(0) \neq 0$ and free energy contains an entropic contribution $-\ln|E_\alpha|$. It is hence important for the previous proposition that the β_i's actually cover α, and one should beware that $\mathbb{F}^\alpha(H_{\beta_i})$ and $\mathbb{F}^{\beta_i}(H_\alpha)$ differ by the entropic term $\mathbb{F}^\alpha(0) - \mathbb{F}^{\beta_i}(0)$. One should thus define a reduced free energy functional $\tilde{\mathbb{F}}^\alpha$ by subtracting the entropic term to get $\tilde{\mathbb{F}}^\alpha \circ j_{\alpha\beta} = \tilde{\mathbb{F}}^\beta$.

The free energy of H_α is by definition the additive constant we subtract from H_α to renormalise the Gibbs density e^{-H_α}:

$$\big[\,\mathrm{e}^{-H_\alpha}\,\big] = \mathrm{e}^{-H_\alpha + \mathbb{F}^\alpha(H_\alpha)} \tag{4.5}$$

The smooth hypersurface $\{\mathbb{F}^\alpha = 0\}$ given by the image of $H_\alpha \mapsto H_\alpha - \mathbb{F}^\alpha(H_\alpha)$ is hence diffeomorphic to the space $\mathring{\Delta}_\alpha$ of non-vanishing probability densities. The Gibbs state is however even more naturally recovered through the differential of $\mathbb{F}^\alpha$.

Proposition 4.3 (Gibbs Expectations)**.** *The differential $d\mathbb{F}^\alpha_\theta : A_\alpha \to A^*_\alpha$ of free energy is given by:*

$$d\mathbb{F}^\alpha_\theta(H_\alpha) = [\mathrm{e}^{-\theta H_\alpha}] \tag{4.6}$$

We shall denote by $\mathbb{E}^\alpha_\theta \in \Omega^1(A_\alpha)$ the differential of $\mathbb{F}^\alpha_\theta$ viewed as a 1-form over A_α.

Proof. Using $de^x = \mathrm{e}^x\, dx$ and $d\ln(y) = \frac{dy}{y}$ we have for every perturbation f_α of $H_\alpha \in A_\alpha$:

$$\mathbb{F}^\alpha(H_\alpha + f_\alpha) = \mathbb{F}^\alpha(H_\alpha) + \frac{\sum f_\alpha\, \mathrm{e}^{-H_\alpha}}{\sum \mathrm{e}^{-H_\alpha}} + o\,(f_\alpha) \tag{4.7}$$

The linear term is precisely the expectation $\mathbb{E}^\alpha[f_\alpha]$ relative to the Gibbs state $[\mathrm{e}^{-H_\alpha}]$. The formula at generic temperatures easily follows from the chain rule applied to $H_\alpha \mapsto \theta H_\alpha$. □

The previous proposition generalises the usual properties expected from thermodynamic potentials, which yield thermodynamic variables through their derivatives. Fixing the hamiltonian H_α and considering free energy as a sole function of inverse temperature, we would recover internal energy as:

$$\mathscr{U}_\alpha(\theta) = \mathbb{E}^\alpha_\theta[H_\alpha] = \frac{d(\theta \mathbb{F}^\alpha_\theta)}{d\theta} \tag{4.8}$$

Proposition 4.4 (Integral Form)**.** *The free energy of $H_\alpha \in A_\alpha$ is given by the integral formula:*

$$\mathbb{F}^\alpha(H_\alpha) = \mathbb{F}^\alpha(0) + \int_{\theta=0}^{1} \mathbb{E}^\alpha_\theta[H_\alpha]\, d\theta \tag{4.9}$$

where $\mathbb{E}^\alpha_\theta$ denotes expectation in the Gibbs state $[\mathrm{e}^{-\theta H_\alpha}]$.

Proof. This is the fundamental theorem of calculus applied along the path $\theta \mapsto \theta H_\alpha$. □

Another important property of free energy is concavity, as it will allow us to view Shannon entropy as the Legendre transform of free energy in section 3.

Proposition 4.5 (Concavity)**.** *For every $U_\alpha, V_\alpha \in A_\alpha$ and $t \in [0,1]$, we have:*

$$\mathbb{F}^\alpha\big(t\, U_\alpha + (1-t)\, V_\alpha\big) \;\geq\; t\, \mathbb{F}^\alpha(U_\alpha) + (1-t)\, \mathbb{F}^\alpha(V_\alpha) \tag{4.10}$$

Proof. Both $x \mapsto \mathrm{e}^{-x}$ and $y \mapsto -\ln(y)$ are convex while the latter is order reversing. □

A consequence of the concavity of free energy is the negative signature of its second differential, explicitly given by the following proposition:

Proposition 4.6 (Fisher Metric)**.** *The second differential of free energy $D^2\mathbb{F}^\alpha : A_\alpha \to S^2 A^*_\alpha$ yields the opposite covariance of observables with respect to the Gibbs state:*

$$-\,D^2\mathbb{F}^\alpha(f_\alpha, g_\alpha) = \mathbb{E}^\alpha[f_\alpha \cdot g_\alpha] - \mathbb{E}^\alpha[f_\alpha] \cdot \mathbb{E}^\alpha[g_\alpha] \tag{4.11}$$

This is also the Fisher information metric for the exponential parametrisation of $\mathring{\Delta}_\alpha$ by A_α.

Proof. $D^2\mathbb{F}^\alpha(f_\alpha, g_\alpha)$ is the variation of $\mathbb{E}^\alpha[f_\alpha]$ along a perturbation g_α of the hamiltonian and:

$$-\frac{\partial}{\partial g_\alpha}\left(\frac{\sum f_\alpha\, \mathrm{e}^{-H_\alpha - g_\alpha}}{\sum \mathrm{e}^{-H_\alpha - g_\alpha}}\right) = \frac{\sum f_\alpha\, g_\alpha\, \mathrm{e}^{-H_\alpha}}{\sum \mathrm{e}^{-H_\alpha}} - \frac{\big(\sum f_\alpha\, \mathrm{e}^{-H_\alpha}\big)\big(\sum g_\alpha\, \mathrm{e}^{-H_\alpha}\big)}{\big(\sum \mathrm{e}^{-H_\alpha}\big)^2} \tag{4.12}$$

□

Note that Gibbs states induce a non-degenerate metric $\mathbb{E}^\alpha[f_\alpha \cdot g_\alpha]$, its image under the projection $f_\alpha \mapsto f_\alpha - \mathbb{E}^\alpha[f_\alpha]$ being the covariance bilinear form. The exponential map $(\mathrm{e}^-)_\alpha$ is a diffeomorphism from the level hypersurface $\{\mathbb{F}^\alpha = 0\} \subseteq A_\alpha$ to the manifold $\mathring{\Delta}_\alpha$ of non-degenerate probability densities, while the previous projection is precisely the orthogonal projection onto the tangent space of $\{\mathbb{F}^\alpha = 0\}$.

4.1.2 Effective Energy

Partial integration defines a linear map $\Sigma^{\beta\alpha} : A_\alpha \to A_\beta$, associating to f_α the observable on β:

$$\Sigma^{\beta\alpha} f_\alpha \;:\; x_\beta \longmapsto \sum_{y \in E_{\alpha \smallsetminus \beta}} f_\alpha(x_\beta, y) \tag{4.13}$$

Partial integration is however not an algebra morphism and should rather be seen as an interlacing of the natural map $A_\alpha^* \to A_\beta^*$ with identifications of each algebra with its dual vector space. Effective energy, as defined below, may then be thought of analogously:

$$\begin{tikzcd} A_\alpha \arrow[r, "\Sigma^{\beta\alpha}"] & A_\beta \arrow[d, "(-\ln)_\beta"] \\ A_\alpha \arrow[u, "(\mathrm{e}^-)_\alpha"] \arrow[r, "\mathbb{F}^{\beta\alpha}"] & A_\beta \end{tikzcd} \tag{4.14}$$

Definition 4.7. *For all $\alpha \to \beta$ in X, we call effective energy the map $\mathbb{F}^{\beta\alpha} \in C^\infty(A_\alpha, A_\beta)$ defined by:*

$$\mathbb{F}^{\beta\alpha}(H_\alpha) = -\ln\Big(\Sigma^{\beta\alpha}\big(\,\mathrm{e}^{-H_\alpha}\,\big)\Big) \tag{4.15}$$

Note that the free energy $\mathbb{F}^\alpha$ is recovered as the effective energy of vacuum $\mathbb{F}^{\varnothing\alpha}$. More generally, $\mathbb{F}^{\beta\alpha}$ may be thought of as the conditional free energy of α given the microstate of β. This conditioning is functorial, as expressed by this first remarkable property.

Proposition 4.8 (Functoriality). *For every $\alpha \to \beta \to \gamma$ in X, we have in $C^\infty(A_\alpha, A_\gamma)$:*

$$\mathbb{F}^{\gamma\beta} \circ \mathbb{F}^{\beta\alpha} = \mathbb{F}^{\gamma\alpha} \tag{4.16}$$

Effective energies hence define a system $\mathbb{F} : X \to \mathbf{Mfd}$ of differentiable manifolds over X.

Proof. We have $\mathbb{F}^{\gamma\beta} \circ \mathbb{F}^{\beta\alpha}(H_\alpha) = -\ln\big(\Sigma^{\gamma\beta} \circ \Sigma^{\beta\alpha}(\mathrm{e}^{-H_\alpha})\big)$ while $\Sigma^{\gamma\beta} \circ \Sigma^{\beta\alpha} = \Sigma^{\gamma\alpha}$. ☐

The following weaker commutative diagram expresses that effective energy describes Gibbs states marginalisation at the level of hamiltonians. Consistency of the Gibbs states will hence transpose to a notion of effective consistency on hamiltonians, requiring that $H_\beta = \mathbb{F}^{\beta\alpha}(H_\alpha)$ for every $\alpha \to \beta$ in X.

$$\begin{tikzcd} A_\alpha \arrow[r, "\mathbb{F}^{\beta\alpha}"] \arrow[d, "{[\mathrm{e}^-]_\alpha}"'] & A_\beta \arrow[d, "{[\mathrm{e}^-]_\beta}"] \\ \mathring{\Delta}_\alpha \arrow[r, "\Sigma^{\beta\alpha}"] & \mathring{\Delta}_\beta \end{tikzcd} \tag{4.17}$$

Proposition 4.9 (Marginalisation). *For every $\alpha \to \beta$ in X we have:*

$$\Sigma^{\beta\alpha}\big[\,\mathrm{e}^{-H_\alpha}\,\big] = \big[\,\mathrm{e}^{-\mathbb{F}^{\beta\alpha}(H_\alpha)}\,\big] \tag{4.18}$$

The following proposition expresses that information on the microstates outside the support of a hamiltonian does not affect effective energy. More precisely, given a hamiltonian $H_\alpha \in A_\alpha$ and $\beta \subseteq \alpha$, extending H_α to $\alpha' = \alpha \sqcup \gamma$ and letting $\beta' = \beta \sqcup \gamma$, we have the commutative diagram:

$$\begin{tikzcd} A_{\alpha'} \arrow[r, "\mathbb{F}^{\beta'\alpha'}"] & A_{\beta'} \\ A_\alpha \arrow[u, "j_{\alpha'\alpha}"] \arrow[r, "\mathbb{F}^{\beta\alpha}"] & A_\beta \arrow[u, "j_{\beta'\beta}"'] \end{tikzcd} \tag{4.19}$$

Proposition 4.10. *If $\beta \subseteq \alpha$ and $H_\alpha \in A_\alpha$, then for every $\beta' \subseteq \alpha'$ such that $\alpha \smallsetminus \beta = \alpha' \smallsetminus \beta'$:*

$$\mathbb{F}^{\beta'\alpha'}(H_\alpha) = \mathbb{F}^{\beta\alpha}(H_\alpha) \tag{4.20}$$

Proof. $\Sigma^{\beta'\alpha'}$ and $\Sigma^{\beta\alpha}$ involve a sum over the same variables and e^{-H_α} only depends on those in α. □

Just like conditional independence generalises the notion of independence, additivity of free energy along uninteracting subsystems generalises to the following conditional form:

Proposition 4.11 (Pairwise Conditional Additivity)**.** *If $\alpha = \beta \cup \beta'$ and $\beta \cap \beta' = \gamma$, we have:*

$$\mathbb{F}^{\gamma\alpha}(H_\beta + H_{\beta'}) = \mathbb{F}^{\gamma\beta}(H_\beta) + \mathbb{F}^{\gamma\beta'}(H_{\beta'}) \tag{4.21}$$

Proof. We have $-\ln\big(\Sigma^{\gamma\alpha}(\mathrm{e}^{-H_\beta}\,\mathrm{e}^{-H_{\beta'}})\big) = -\ln\big(\Sigma^{\gamma\beta}(\mathrm{e}^{-H_\beta})\,\Sigma^{\gamma\beta'}(\mathrm{e}^{-H'_\beta})\big)$ as $\Sigma^{\gamma\alpha}$ involves a sum over the variables in the disjoint union $(\beta \smallsetminus \gamma) \sqcup (\beta' \smallsetminus \gamma)$, while e^{-H_β} is independent of the variables in $\beta' \smallsetminus \gamma$ and reciprocally. □

In particular, $\mathbb{F}^{\beta\alpha}$ is additive along A_β so that for every $H_\alpha \in A_\alpha$ and $H_\beta \in A_\beta$ we have:

$$\mathbb{F}^{\beta\alpha}(H_\alpha + H_\beta) = \mathbb{F}^{\beta\alpha}(H_\alpha) + H_\beta \tag{4.22}$$

As before, one should beware that $\mathbb{F}^{\beta\alpha}(0) \neq 0$ as effective energy contains the entropic term $-\ln|E_{\alpha\smallsetminus\beta}|$. The reduced effective energy $\tilde{\mathbb{F}}^{\beta\alpha}$ defined by subtracting this term to $\mathbb{F}^{\beta\alpha}$ is a smooth section of $j_{\alpha\beta}$. Considering larger coverings, we may now rewrite conditional additivity as follows:

Proposition 4.12 (Conditional Additivity)**.** *If α is the union of $\beta_1, \dots, \beta_n$ and H_α splits as $\sum_i H_{\beta_i}$, then denoting by γ_i the intersection of β_i with $\bigcup_{j\neq i} \beta_j$ and by γ the reunion $\bigcup_i \gamma_i$ we have:*

$$\mathbb{F}^{\gamma\alpha}\Big(\sum_i H_{\beta_i}\Big) = \sum_i \mathbb{F}^{\gamma_i\beta_i}(H_{\beta_i}) \tag{4.23}$$

Proof. Reasoning by induction on n, let $\tilde\beta_1 = \beta_1 \cup \gamma$ and $\tilde\beta_2 = \bigcup_{j=2}^n \beta_j \cup \gamma$ with hamiltonians $\tilde H_{\tilde\beta_1} = H_{\beta_1}$ and $\tilde H_{\tilde\beta_2} = \sum_{i>1}^n H_{\beta_i}$. Pairwise conditional additivity gives $\mathbb{F}^{\gamma\alpha}(\sum_i H_{\beta_i}) = \mathbb{F}^{\gamma\tilde\beta_1}(\tilde H_{\tilde\beta_1}) + \mathbb{F}^{\gamma\tilde\beta_2}(\tilde H_{\tilde\beta_2})$ as $\tilde\beta_1 \cap \tilde\beta_2 = \gamma$, while $\mathbb{F}^{\gamma\tilde\beta_1}(H_{\beta_1})$ and $\mathbb{F}^{\gamma_1\beta_1}(H_{\beta_1})$ coincide as conditional free energies of $\tilde\beta_1 \smallsetminus \gamma = \beta_1 \smallsetminus \gamma_1$. The induction hypothesis applied to $\mathbb{F}^{\gamma\tilde\beta_2}(\tilde H_{\tilde\beta_2})$ then terminates the proof. □

Effective energy is still a concave functional in the sense below, and its point-wise Legendre transform above $x_\beta \in E_\beta$ will generate the conditional Shannon entropy $S_\alpha(p_\alpha|x_\beta)$ defined in section 2.

Proposition 4.13 (Concavity)**.** *For every $U_\alpha, V_\alpha \in A_\alpha$ and $t \in [0,1]$, we have:*

$$\mathbb{F}^{\beta\alpha}\big(t\,U_\alpha + (1-t)\,V_\alpha\big) \;\geq\; t\,\mathbb{F}^{\beta\alpha}(U_\alpha) + (1-t)\,\mathbb{F}^{\beta\alpha}(V_\alpha) \tag{4.24}$$

the inequality being understood in the partial order of functions on E_β.

While the Gibbs state was recovered through the differential of free energy, the differential of effective energy carries the effects of Gibbs state conditioning.

Proposition 4.14 (Conditional Expectations)**.** *The differential $d\mathbb{F}^{\beta\alpha} : A_\alpha \to A_\alpha^* \otimes A_\beta$ of effective energy is the conditional expectation given the microstate of β with respect to the Gibbs state on α:*

$$d\mathbb{F}^{\beta\alpha}(H_\alpha) \;:\; \left\{ \begin{array}{ccc} A_\alpha & \longrightarrow & A_\beta \\ f_\alpha & \longmapsto & \mathbb{E}^\alpha[\, f_\alpha \mid \beta\,] \end{array} \right. \tag{4.25}$$

We denote by $\mathbb{E}^{\beta\alpha} \in \Omega^1(A_\alpha, A_\beta)$ the differential of $\mathbb{F}^{\beta\alpha}$ viewed as an A_β-valued 1-form over A_α.

Proof. Using again $de^x = \mathrm{e}^x\, dx$ and $d\ln(y) = \frac{dy}{y}$ we now have for every perturbation f_α of H_α:

$$\mathbb{F}^{\beta\alpha}(H_\alpha + f_\alpha) = \mathbb{F}^{\beta\alpha}(H_\alpha) + \frac{\Sigma^{\beta\alpha}(f_\alpha\, \mathrm{e}^{-H_\alpha})}{\Sigma^{\beta\alpha}(\mathrm{e}^{-H_\alpha})} + o\,(f_\alpha) \tag{4.26}$$

Rescaling the fraction by the normalisation factor $\Sigma^{\varnothing\alpha}(\mathrm{e}^{-H_\alpha})$, denoting by p_α the Gibbs state on α and by p_β its marginal on β, the linear term rewrites as:

$$\Sigma^{\beta\alpha}\Big(f_\alpha \cdot \frac{p_\alpha}{p_\beta}\Big) = \mathbb{E}_{p_\alpha}[f_\alpha \mid \beta] \tag{4.27}$$

As according to the Bayes rule, p_α/p_β is the conditional probability on α given the microstate of β. □

Conditional expectation has a simple geometrical characterisation which is worth recalling. For the riemannian metric on A_α induced by the Gibbs state:

$$(f_\alpha\,,\, g_\alpha) = \mathbb{E}^\alpha[f_\alpha \cdot g_\alpha] \tag{4.28}$$

$\mathbb{E}^{\beta\alpha}$ is just the orthogonal projection of A_α onto the subspace A_β of observables depending only on the state of β. One might hope for a consistent orthogonal splitting of A_α as[2] $\bigoplus Z_\beta$, defining each interaction subspace Z_β by orthogonality with B_β. However, as soon as β and β' interact with each other, correlations will imply that Z_β is no longer orthogonal to $Z_{\beta'}$ for the metric on A_α.

Proposition 4.15 (Integral Form)**.** *The effecive energy of $H_\alpha \in A_\alpha$ on β is given by the integral:*

$$\mathbb{F}^{\beta\alpha}(H_\alpha) = \mathbb{F}^{\beta\alpha}(0) + \int_{\theta=0}^{1} \mathbb{E}_\theta^{\beta\alpha}[H_\alpha]\, d\theta \tag{4.29}$$

where $\mathbb{E}_\theta^{\beta\alpha}$ denotes conditional expectation with respect to the Gibbs state $[\mathrm{e}^{-\theta H_\alpha}]$.

Proof. This is again the fundamental theorem of calculus applied along the path $\theta \mapsto \theta H_\alpha$. □

4.2 Information Quantities

This section introduces the Shannon entropy, also called Shannon information. We then relate mutual information quantities with a combinatorial decomposition of entropy into summands, at the core of Bethe's approximation of entropy.

4.2.1 Shannon Entropy

Given a finite set E_α, we denote by Δ_α the space of probability measures on E_α. Shannon entropy is the concave functional $S_\alpha : \Delta_\alpha \to \mathbb{R}$ defined by:

$$S_\alpha(p_\alpha) = -\sum_{E_\alpha} p_\alpha \ln(p_\alpha) \tag{4.30}$$

It reaches its global maximum $S_\alpha([1_\alpha]) = \ln|E_\alpha|$ on the uniform measure, which may be called the Boltzmann entropy of E_α.

[2] See the interaction decomposition theorem, section 2.3.2.

Given a surjection $\pi^{\beta\alpha} : E_\alpha \to E_\beta$, and $p_\alpha \in \Delta_\alpha$, let us denote by $p_\beta = \Delta^{\beta\alpha}(p_\alpha)$ its marginal distribution on E_β and by $p_{\alpha|x_\beta}$ its conditional probability distribution given any $x_\beta \in E_\beta$. The conditional entropy of p_α given E_β is defined as:

$$S_\alpha(p_\alpha \mid \pi^{\beta\alpha}) = \sum_{x_\beta \in E_\beta} p_\beta(x_\beta) \cdot S_\alpha(p_{\alpha|x_\beta}) \tag{4.31}$$

When $E_\alpha = E_\beta \times E_\gamma$ and $p_\alpha = p_\beta \otimes p_\gamma$ then an easy computation shows that $S_\alpha(p_\alpha \mid \pi^{\beta\alpha}) = S_\gamma(p_\gamma)$.

The fundamental property of entropy is the chain rule:

$$S_\alpha(p_\alpha) = S_\beta(p_\beta) + S_\alpha(p_\alpha \mid \pi^{\beta\alpha}) \tag{4.32}$$

Shannon showed that entropy is essentially haracterised by this functional equation, up to a multiplicative constant and under an additional monotonicity condition with respect to the cardinal of E_α. When $p_\alpha = \otimes_{\beta'} p_{\beta'}$ is a tensor product of independent probabilities on $E_\alpha = \prod_{\beta'} E_{\beta'}$, we have in particular:

$$S_\alpha(p_\alpha) = \sum_{\beta'} S_{\beta'}(p_{\beta'}) \tag{4.33}$$

expressing that entropy is additive along independent systems. [See JPV, Bennequin-Baudot, Leinster]

4.2.2 Möbius Inversion

For every $\alpha \in X$, let us denote by $\mathscr{F}_\alpha = C^\infty(\Delta_\alpha)$ the vector space of smooth functionals on Δ_α. For every $\alpha \to \beta$, the pullback of the marginal projection $\Delta_\alpha \to \Delta_\beta$ defines an injection $\mathscr{F}_\alpha \leftarrow \mathscr{F}_\beta$, so that $\mathscr{F}$ is a contravariant functor of vector spaces on X. In particular, $\tilde{\mathbb{Z}}_1(X)$ acts on $\mathscr{F}_0(X)$ and for every collection of functionals $(\mathscr{L}_\alpha)$ we have the equivalence:

$$\mathscr{L}_\alpha = \sum_{\alpha \to \beta'} [\ell_{\beta'}]_\alpha \quad \Leftrightarrow \quad \ell_\alpha = \sum_{\alpha \to \beta'} \mu_{\alpha\beta'} \cdot [\mathscr{L}_{\beta'}]_\alpha \tag{4.34}$$

We call $\ell \in \mathscr{F}_0(X)$ the combinatorial localisation of $\mathscr{L}$.

A particular case of functionals is given by the expectation values of observables. Given $H \in A_0(X)$, let $\mathcal{U} \in \mathscr{F}_0(X)$ be defined for every $\alpha \in X$ by:

$$\mathcal{U}_\alpha(p_\alpha) = \langle\, p_\alpha \,|\, H_\alpha \,\rangle = \mathbb{E}_{p_\alpha}[H_\alpha] \tag{4.35}$$

The natural map from $A_0(X)$ to $\mathscr{F}_0(X)$ is a morphism of $\tilde{\mathbb{Z}}_1(X)$-modules. When $H = \zeta \cdot h$, the localisation of $\mathcal{U}$ as $\zeta \cdot u$ defines another field of functionals $u \in \mathscr{F}_0(X)$ where:

$$u_\alpha(p_\alpha) = \langle\, p_\alpha \,|\, h_\alpha \,\rangle = \mathbb{E}_{p_\alpha}[h_\alpha] \tag{4.36}$$

The internal energy of a system $\mathcal{U}_\Omega$ is the expectation value of the global hamiltonian H_Ω, typically given as a sum $\sum_\alpha h_\alpha$ of local interactions with $h_\Omega = 0$. This remark will prove the Bethe approximation scheme to be exact on internal energy, as $\mathcal{U}_\Omega = \sum_\alpha u_\alpha$ with $u_\Omega = 0$.

4.2.3 Mutual Informations

Suppose given n random variables $\mathcal{X}_1, \dots \mathcal{X}_n$. In this paragraph, we introduce the various amounts of information following Hu Kuo Ting in [8]. First, denote by $S(\mathcal{X}_i \cup \mathcal{X}_j)$ the entropy of their joint law $(\mathcal{X}_i, \mathcal{X}_j)$. If the variables were pairwise independent, the chain rule of entropy would give for all distinct i, j:

$$S(\mathcal{X}_i \cup \mathcal{X}_j) = S(\mathcal{X}_i) + S(\mathcal{X}_j) \tag{4.37}$$

But more generally, denoting by $S(\mathcal{X}_i|\mathcal{X}_j)$ the conditional entropy of $(\mathcal{X}_i, \mathcal{X}_j)$ given $\mathcal{X}_j$, we have:

$$S(\mathcal{X}_i \cup \mathcal{X}_j) = S(\mathcal{X}_i|\mathcal{X}_j) + S(\mathcal{X}_j) \tag{4.38}$$

We may define a quantity $S(\mathcal{X}_i \cap \mathcal{X}_j)$ as the difference $S(\mathcal{X}_i) - S(\mathcal{X}_i|\mathcal{X}_j)$. These relations naturally remind of those satisfied by an additive function on sets, where independent variables correspond to disjoint subsets, and conditioning describes set difference. Note however that in the following theorem, the order of logicals operation matters, see [2] for examples.

Theorem 4.16 (Hu Kuo Ting)**.** *Given random variables $\mathcal{X}_1, \dots, \mathcal{X}_n$ and their probability distributions, there exists sets $A_1, \dots, A_n$ and an additive real function φ on the algebra generated by those sets, such that for all operation Q generated by $\cup, \cap$ and $-$ which 1) forms collections of unions 2) takes successive intersections of these unions and 3) subtracts one of them, one has:*

$$S(Q(\mathcal{X}_1, \dots, \mathcal{X}_n)) = \varphi(Q(A_1, \dots, A_n)) \tag{4.39}$$

All quantities obtained this way are generically called amounts of information by Hu Kuo Ting. Of particular importance are the mutual informations, appearing in the right hand side of:

$$S(\mathcal{X}_{i_1} \cup \dots \cup \mathcal{X}_{i_k}) = \sum_{1 \le p \le k} S(\mathcal{X}_{i_p}) - \sum_{1 \le p < q \le k} S(\mathcal{X}_{i_p} \cap \mathcal{X}_{i_q}) + \dots + (-1)^{k+1}\, S(\mathcal{X}_{i_1} \cap \dots \cap \mathcal{X}_{i_k}) \tag{4.40}$$

We will denote the mutual information of $\mathcal{X}_{i_1}, \dots \mathcal{X}_{i_k}$ by:

$$I(\mathcal{X}_{i_1}, \dots, \mathcal{X}_{i_k}) = S(\mathcal{X}_{i_1} \cap \dots \cap \mathcal{X}_{i_k}) \tag{4.41}$$

The following theorems will best express the significance of mutual informations. Although very similar in appearance, the second theorem is a much more recent result than the first, proved in [2].

Theorem 4.17 (Hu)**.** *$\mathcal{X}_1, \dots \mathcal{X}_n$ form a Markov chain if and only if for all $1 \le i_1 < \dots < i_k \le n$:*

$$I(\mathcal{X}_{i_1}, \dots \mathcal{X}_{i_k}) = I(\mathcal{X}_{i_1}, \mathcal{X}_{i_k}) \tag{4.42}$$

Theorem 4.18 (Bennequin)**.** *$\mathcal{X}_1, \dots, \mathcal{X}_n$ are independent if and only if for all $1 \le i_1, \dots, i_k \le n$:*

$$I(\mathcal{X}_{i_1}, \dots, \mathcal{X}_{i_k}) = 0 \tag{4.43}$$

Let us give an explicit definition of mutual informations as functionals of the probability distributions, by relating Hu Kuo Ting's construction to a Möbius inversion on entropy functionals. Denoting by $(\mathcal{X}_i)_{i \in \Omega}$ a finite set of random variables, and by p_α the joint law of $(\mathcal{X}_i)_{i \in \alpha}$ for every $\alpha \subseteq \Omega$, each joint entropy may be expressed as:

$$S_\alpha(p_\alpha) = \sum_{\alpha \supseteq \beta'} s_{\beta'}(p_{\beta'}) \tag{4.44}$$

Comparing with Hu Kuo Ting's formula, the mutual information $I_\beta(p_\beta)$ is then given by $(-1)^{|\beta|+1}\, s_\beta$. The entropy summands given by Möbius inversion on $\mathcal{P}(\Omega)$ satisfy:

$$s_\alpha(p_\alpha) = \sum_{\alpha \supseteq \beta'} (-1)^{|\alpha| - |\beta'|}\, S_{\beta'}(p_{\beta'}) \tag{4.45}$$

while mutual informations are given by:

$$I_\alpha(p_\alpha) = \sum_{\alpha \supseteq \beta'} (-1)^{|\beta'|+1}\, S_{\beta'}(p_{\beta'}) \tag{4.46}$$

Fixing $\alpha \in X$, this expression may also be seen as a Möbius inversion on the opposite partial order spanned by each $\mathcal{X}_i$ for $i \in \alpha$, viewed as a maximal element.

4.2.4 Bethe Entropy

As first recognised by Morita in [20], Bethe's approximation of entropy is essentially a truncation of the previous Möbius inversion procedure. Consider a general covering $X \subseteq \mathcal{P}(\Omega)$ not containing Ω, and let $\bar{X} = \{\Omega\} \cup X$. Through Möbius inversion on $\bar{X}$, entropy can still be exactly localised by $s \in \mathscr{F}_0(\bar{X})$ with:

$$s_\alpha = \sum_{\alpha \to \beta'} \mu_{\alpha\beta'} \cdot [S_{\beta'}]_\alpha \tag{4.47}$$

In particular, the global entropy S_Ω is recovered as:

$$S_\Omega = s_\Omega + \sum_{\alpha \in X} [s_\alpha]_\Omega \tag{4.48}$$

The Bethe approximation of entropy $\check{S}_\Omega = S_\Omega - s_\Omega$, according to the previous paragraph, may be seen to only neglect all mutual informations of the form $I_\omega(p_\omega)$ for $\omega \in \mathcal{P}(\Omega)$ not contained in any $\alpha \in X$. Intuitively, s_Ω is expected[3] to be small when large enough regions are taken in X to cover Ω, by extensivity of entropy.

The approximate entropy $\check{S}_\Omega(p_\Omega)$ only depends on the marginal distributions $(p_\alpha)_{\alpha \in X}$ of p_Ω, so that $\check{S}_\Omega$ factors through the canonical map $\Delta_\Omega \to \Gamma(X)$. We call Bethe entropy the functional $\check{S}$ defined on $\Delta_0(X)$ by:

$$\check{S} = \sum_{\alpha \in X} s_\alpha = \sum_{\beta \in X} c_\beta \cdot S_\beta \tag{4.49}$$

Each term s_α acting on Δ_α, so that the restriction of $\check{S}$ to the image of Δ_Ω in $\Delta_0(X)$ coincides with $\check{S}_\Omega$. Only the restriction of Bethe entropy to $\Gamma(X)$ shall be relevant, it is however important to remark that $\Gamma(X)$ is in general larger than the image of Δ_Ω.

4.3 Variational Principles

There has been a longstanding practice in thermodynamics to characterise the equilibrium state of a system subject to various kinds of constraints through variational principles expressed in terms of macroscopic variables such as pressure, volume, temperature, *etc.* The statistical counterpart of these variational principles leads to the same variety of functionals expressed in terms of the hamiltonian H_Ω and the statistical state p_Ω of the system, while macroscopic variables such as pressure and temperature arise as Lagrange multipliers associated to volume and energy constraints.

In this section, we suppose given the hamiltonian $H_\Omega = \sum_\alpha h_\alpha$ as a sum of interactions over X. We briefly recall two classical variational principles characterising the Gibbs equilibrium state, the first on entropy, the second performing the Legendre transform to free energy. The second one we shall approximate by Kikuchi's cluster variation method, before giving the announced homological characterisation of its solutions. This will justify the use of message-passing algorithms in the next chapter, iterating over Lagrange multipliers associated to the consistency constraint.

4.3.1 Maximal Entropy

The simplest variational principle characterising the Gibbs distribution states that once the internal or mean energy $\mathcal{U}_\Omega$ of the system is fixed, equilibrium is reached at the maximally entropic probability density. Internal energy is the smooth functional $\mathcal{U}_\Omega : \Delta_\Omega \to \mathbb{R}$ giving the expectation value of the hamiltonian:

$$\mathcal{U}_\Omega(q_\Omega) = \langle\, q_\Omega \mid H_\Omega \,\rangle = \mathbb{E}_{p_\Omega}[H_\Omega] \tag{4.50}$$

[3] Schlijper proved this procedure to converge to the true entropy per lattice point for the Ising 2D model in [25]

Let us denote by $\lambda_\infty = \inf(H_\Omega)$ the minimal energy and by $\lambda_0 = \mathcal{U}_\Omega([1])$ the mean energy for the uniform distribution on E_Ω.

Theorem 4.19 (Maximal Entropy Principle)**.** *Given $\lambda \in]\lambda_\infty, \lambda_0[$ the constrained variational problem:*

$$S_\Omega(p_\Omega) = \max_{\{\mathcal{U}_\Omega = \lambda\}} S_\Omega \tag{4.51}$$

has a unique solution given by $p_\Omega = \left[\, \mathrm{e}^{-\theta H_\Omega} \,\right]$ for some inverse temperature $\theta \in \mathbb{R}_+^*$.

Proof. The cotangent space of $\mathring{\Delta}_\Omega \subseteq A_\Omega^*$ is isomorphic to the quotient of A_Ω by additive constants, generating the normalisation constraint $\langle\, p_\Omega \,|\, 1 \,\rangle = 1$. The inverse temperature then arises as a Lagrange multiplier for the energy constraint $\mathcal{U}_\Omega = \lambda$, as for every positive $p_\Omega \in \mathring{\Delta}_\Omega$:

$$\frac{\partial S_\Omega}{\partial p_\Omega} = \theta \cdot \frac{\partial \mathcal{U}_\Omega}{\partial p_\Omega} \mod \mathbb{R} \quad \Leftrightarrow \quad -\ln(p_\Omega) = \theta H_\Omega \mod \mathbb{R} \tag{4.52}$$

Reciprocally, consider the smooth path $\theta \mapsto [\mathrm{e}^{-\theta H_\Omega}]$ described by the Gibbs states in $\mathring{\Delta}_\Omega$ for $\theta \in]0, +\infty[$. When the temperature goes to zero and $\theta \to \infty$ the Gibbs distribution goes to the uniform distribution on the minimisers of H_Ω and $\mathcal{U}_\Omega \to \lambda_\infty$. When the temperature goes to infinity and $\theta \to 0$, the Gibbs distribution goes to the uniform measure and $\mathcal{U}_\Omega \to \lambda_0$. ☐

Gibbs states as a function of a generalised inverse temperature $\theta \in \mathbb{R}$ are one-parameter subgroups of $\mathring{\Delta}_\Omega$ for its multiplicative structure, while $\theta \in \mathbb{R}_+$ restricts to semi-groups, as pictured by the figure below.

At the macroscopic level, the form of the maximal entropy principle suggests that the equilibrium entropy $S(\mathcal{U})$ be naturally defined as a function of internal energy $\mathcal{U}$, while the temperature T measures a kind of inverse entropic susceptibility:

$$\frac{1}{kT} = \theta = \frac{\partial S}{\partial \mathcal{U}} \tag{4.53}$$

The Legendre transform essentially consists in a change of variables, defining an equivalent functional parametrised by the derivative of the original one. We would here recover the so-called free entropy[4]:

$$\Psi(\theta) = S(\mathcal{U}) - \theta\mathcal{U} \tag{4.54}$$

The equilibrium free energy recovered as $F(\theta) = -\Psi(\theta)/\theta$ being more commonly used. It is remarkable that the statistical free energy $\mathbb{F}^\Omega(H_\Omega)$ may be defined from the Shannon entropy $S_\Omega(p_\Omega)$ in a perfectly similar manner.

4.3.2 Thermal Equilibrium

Free energy variational principles describe a system interacting with a thermostat, exchanging arbitrary amounts of energy without modifying the temperature of the latter. Such variational principles are very natural as they for instance describe interaction with the atmosphere. Let us call *variational free energy* the smooth bifunctional $\mathcal{F}_\Omega : \Delta_\Omega \times H_\Omega \to \mathbb{R}$ given by:

$$\mathcal{F}_\Omega(p_\Omega, H_\Omega) = \langle\, p_\Omega \,|\, H_\Omega \,\rangle - S_\Omega(p_\Omega) \tag{4.55}$$

This functional generates the Legendre transform of S_Ω by minimisation of $\mathcal{F}_\Omega(\,-\,, H_\Omega)$ and does yield the equilibrium free energy $\mathbb{F}^\Omega$ at inverse temperature[5] $\theta = 1$.

[4] Free entropy is also called the Massieu potential, as introduced in his 1869 note [15].

[5] One may reintroduce temperature dependency by letting $\mathcal{F}^\theta_\Omega(p_\Omega, H_\Omega) = \theta^{-1}\mathcal{F}_\Omega(p_\Omega, \theta H_\Omega)$, recovering the classical formula $\mathcal{F}^\theta_\Omega = \mathcal{U}_\Omega - \theta^{-1} S_\Omega$. At fixed temperatures one may however always assume $\theta = 1$ up to a choice of units.

Theorem 4.20 (Minimal Free Energy Principle)**.** *For every $H_\Omega \in A_\Omega$, the variational problem:*

$$\mathcal{F}_\Omega(p_\Omega, H_\Omega) = \min_{\Delta_\Omega} \mathcal{F}_\Omega(\,-\,, H_\Omega) \tag{4.56}$$

has a unique solution given by the Gibbs state $[\mathrm{e}^{-H_\Omega}]$ *reaching the equilibrium free energy* $\mathbb{F}^\Omega(H_\Omega)$.

Proof. Describing the cotangent space of $\mathring{\Delta}_\Omega$ by the quotient $A_\Omega/\mathbb{R}$, a critical $p_\Omega \in \mathring{\Delta}_\Omega$ satisfies:

$$\frac{\partial \mathcal{F}_\Omega}{\partial p_\Omega} = H_\Omega + \ln(p_\Omega) = 0 \mod \mathbb{R} \tag{4.57}$$

Reciprocally, concavity of entropy implies that $p_\Omega = [\mathrm{e}^{-H_\Omega}]$ indeed minimises $\mathcal{F}_\Omega(\,-\,, H_\Omega)$. □

The theorem really states that free energy is the Legendre transform of Shannon entropy. Although not bijective, the Legendre duality between hamiltonians and statistical states is fundamental:

$$p_\Omega = [\mathrm{e}^{-H_\Omega}] \quad \Leftrightarrow \quad H_\Omega = -\ln(p_\Omega) \mod \mathbb{R} \tag{4.58}$$

and defines a surjective abelian group morphism $A_\Omega \to \mathring{\Delta}_\Omega$. The computability of H_Ω however does not at all imply that of p_Ω, as normalising the Gibbs density would require to compute the equilibrium free energy $\mathbb{F}^\Omega(H_\Omega)$, involving an integral over E_Ω of exponential complexity in the cardinal of Ω.

4.3.3 Cluster Variation Method

Introduced by Kikuchi in [10], the cluster variation method seeks to approximate the marginals $(p_\alpha) \in \Gamma(X)$ of the global Gibbs state p_Ω by a consistent collection of local probabilities $(q_\alpha) \in \Gamma(X)$, obtained through a variational principle on a local approximation of free energy.

Definition 4.21. Bethe free energy *is the smooth bifunctional* $\check{\mathcal{F}} : \Delta_0(X) \times A_0(X) \to \mathbb{R}$ *defined by:*

$$\check{\mathcal{F}}(p, H) = \sum_{\beta \in X} c_\beta \Big(\langle\, p_\beta \mid H_\beta \,\rangle - S_\beta(p_\beta) \Big) \tag{4.59}$$

Because of the Möbius numbers c_β appearing in its definition, the Bethe free energy $\check{\mathcal{F}}$ is no longer convex in general, and $\check{\mathcal{F}}(\,-\,, H)$ may have a great multiplicity[6] of critical points inside the space $\mathring{\Gamma}(X)$ of consistent positive densities. We provide with a rigorous characterisation of the critical points of $\check{\mathcal{F}}(\,-\,, H)$ constrained to $\mathring{\Gamma}(X)$, by showing that they bear a homological relationship with the reference hamiltonian field H.

Theorem 4.22. *A positive and consistent statistical field* $p \in \mathring{\Gamma}(X)$ *is critical for the constrained Bethe free energy* $\check{\mathcal{F}}(\,-\,, H)_{|\Gamma(X)}$ *if and only if there exists* $\varphi \in A_1(X)$ *such that:*

$$-\ln(p) \simeq H + \zeta \cdot \delta\varphi \mod \mathbb{R}_0(X) \tag{4.60}$$

The proof of theorem 4.22 shall conclude this chapter, the flux term $\delta\varphi$ essentially appearing as Lagrange multipliers associated to the consistency constraint $dp = 0$. The correspondence theorem 5.13 between stationary states of belief propagation and critical points of $\check{\mathcal{F}}$ shall come as an easy consequence of 4.22, once message-passing algorithms have been related with transport equations. A crucial combinatorial argument in proving 4.22 is contained in the following proposition:

[6] For numerical studies see [31, 21, 11], A first mathematical proof of multiplicity is given by Bennequin in [4].

Proposition 4.23. *For every $H \in A_0(X)$ and $q \in A_0^*(X)$ such that $dq = 0$ we have:*

$$\langle\, q \,|\, \mu \cdot H \,\rangle = \langle\, q \,|\, c\, H \,\rangle \tag{4.61}$$

In particular $c\,H \in \mathrm{Im}(\delta)$ if and only if $\mu \cdot H \in \mathrm{Im}(\delta)$, or equivalently if $H \in \zeta \cdot \mathrm{Im}(\delta)$.

Proof. Using $q_\beta = \Sigma^{\beta\alpha}(q_\alpha)$ and $c_\beta = \sum_{\alpha \to \beta} \mu_{\alpha\beta}$ we have:

$$\langle\, q \,|\, \mu \cdot H \,\rangle = \sum_{\alpha \to \beta} \mu_{\alpha\beta} \,\langle\, \Sigma^{\beta\alpha}(q_\alpha) \,|\, H_\beta \,\rangle = \sum_\beta c_\beta \,\langle\, q_\beta \,|\, H_\beta \,\rangle = \langle\, q \,|\, c\, H \,\rangle \tag{4.62}$$

In particular $\mu \cdot H \perp \mathrm{Ker}(d) \Leftrightarrow c\,H \perp \mathrm{Ker}(d)$, while $\mathrm{Im}(\delta) = \mathrm{Ker}(d)^\perp$ as d is the adjoint of δ. □

Homological invariance of $\check{\mathcal{F}}$ with respect to interaction potentials also follows from 4.23. This interesting property sheds light on the form of the Lagrange multipliers appearing in theorem 4.22

Proposition 4.24. *For every consistent $p \in \Gamma(X)$ and every $U = H + \zeta \cdot \delta\varphi$ in $A_0(X)$ we have:*

$$\check{\mathcal{F}}(p, U) = \check{\mathcal{F}}(p, H) \tag{4.63}$$

Proof. By orthogonality of $\mathrm{Ker}(d)$ with $\mathrm{Im}(\delta)$ and using the lemma, the internal energy terms satisfy:

$$\langle\, p \,|\, c\,U \,\rangle = \langle\, p \,|\, \mu \cdot U \,\rangle = \langle\, p \,|\, \mu \cdot H \,\rangle + \langle\, p \,|\, \delta\varphi \,\rangle = \langle\, p \,|\, c\,H \,\rangle \tag{4.64}$$

□

Recall that homologous interaction potentials $u = h + \delta\varphi$ define the same global hamiltonian as h:

$$H_\Omega = \sum_{\alpha \in X} h_\alpha = \sum_{\beta \in X} c_\beta H_\beta \tag{4.65}$$

and by 2.14, one furthermore has $\sum_\alpha u_\alpha = H_\Omega$ if and only if u is homologous to h. Hence a subtle difference between the homological invariance of 4.24 and the following proposition lies in the assumption that $p \in \Gamma(X)$ is given by the marginals of a global probability distribution[7] $p_\Omega \in \Delta_\Omega$. In that case, it is well-known that the Bethe approximation yields an exact measure of internal energy.

Proposition 4.25. *Given $p_\Omega \in \Delta_\Omega$ of image p in $\Gamma(X)$, for any local hamiltonians $H_\alpha \in A_0(X)$ of global hamiltonian $H_\Omega \in A_\Omega$ we have:*

$$\check{\mathcal{F}}(p, H) = \langle\, p_\Omega \,|\, H_\Omega \,\rangle - \check{S}(p) \tag{4.66}$$

Proof. When $p_\beta = \Sigma^{\beta\Omega}(p_\Omega)$ for every $\beta \in X$, by definition of $\check{\mathcal{F}}$ and H_Ω we have:

$$\check{\mathcal{F}}(p, H) = \sum_{\beta \in X} \langle\, p_\Omega \,|\, c_\beta H_\beta \,\rangle - \check{S}(p) = \langle\, p_\Omega \,|\, H_\Omega \,\rangle - \check{S}(p) \tag{4.67}$$

□

The Bethe approximation also counts degrees of freedom properly, as measured by the maximum entropy reaches in the high temperature limit. The following proposition further justifies its soundness. [[proved in Kikuchi? differently in [32]]]

[7]The image of $\Delta_\Omega(X)$ in general forms a strict convex polytope of $\Gamma(X)$. Its boundaries are the image of the positivity constraints on the global density, see [30] and [1].

Proposition 4.26. *When X is a $\cap$-closed covering*[8] *of Ω, the high temperature limit $\check{\mathcal{F}}([1], H)$ of the Bethe free energy coincides with that of the true free energy $\mathcal{F}([1_\Omega], H_\Omega) = \langle H_\Omega \rangle - \ln |E_\Omega|$.*

Proof. According to 4.25, we only need to show that $\check{S}([1]) = \ln |E_\Omega| = \sum_i \ln |E_i|$. The assumption on X implies that for every $i \in \Omega$, there exists a minimal $\beta_i \in X$ containing i so that:

$$\sum_{\alpha \in X} c_\alpha \ln |E_\alpha| = \sum_{\alpha \in X} c_\alpha \sum_{i \in \alpha} \ln |E_i| = \sum_{i \in \Omega} \ln |E_i| \sum_{\alpha \supseteq \beta_i} c_\alpha = \sum_{i \in \Omega} \ln |E_i| \tag{4.68}$$

□

Before proceding to the proof of theorem 4.22, we finally introduce the slightly finer characterisation of critical points of $\check{\mathcal{F}}$ given by 4.28. It will be especially useful in proving the correspondence theorem 5.13 in the next chapter.

Definition 4.27. *Denote by δ' the truncation of the boundary δ to $X \smallsetminus \{\varnothing\}$:*

$$\delta'_\alpha \varphi = \delta_\alpha \varphi \quad \text{if} \quad \alpha \neq \varnothing \quad \text{and} \quad \delta'_\varnothing \varphi = 0 \tag{4.69}$$

Theorem 4.28. *Assuming $H_\varnothing = 0$, a consistent statistical state $p \in \mathring{\Gamma}(X)$ is critical for the constrained Bethe free energy $\check{\mathcal{F}}(-, H)_{|\Gamma(X)}$ if and only if there exists $\varphi \in A_1(X)$ such that:*

$$-\ln(p) = H + \zeta \cdot \delta' \varphi \tag{4.70}$$

Proof of theorem 4.22. To account for the normalisation constraints $\langle p_\alpha \,|\, 1 \rangle = 1$, we may describe the cotangent space at p of $\mathring{\Delta}_0(X) \subseteq A_0^*(X)$ as the quotient $A_0(X)/\mathbb{R}_0(X)$ and write:

$$\frac{\partial \check{\mathcal{F}}}{\partial p} \simeq \sum_{\beta \in X} c_\beta \big(H_\beta + \ln(p_\beta) \big) \mod \mathbb{R}_0(X) \tag{4.71}$$

The flux term comes as a collection of Lagrange multipliers for the consistency constraints, a consistent $p \in \mathring{\Gamma}(X)$ being critical if and only if the differential of $\check{\mathcal{F}}(-, H)$ vanishes on $\mathrm{Ker}(d) = \mathrm{Im}(\delta)^\perp$ or:

$$c\big(H + \ln(p)\big) \in \mathrm{Im}(\delta) + \mathbb{R}_0(X) \tag{4.72}$$

Proposition 4.23 is crucial[9] to state that the above is equivalent to:

$$H + \ln(p) \in \zeta \cdot \mathrm{Im}(\delta) + \mathbb{R}_0(X) \tag{4.73}$$

□

Proof of theorem 4.28. First note that $\zeta \cdot \delta' \varphi \simeq \zeta \cdot \delta \varphi \mod \mathbb{R}_0(X)$, so that we only need to reduce the Lagrange multipliers of 4.22 to the form given by 4.28. Assume there exists $\psi \in A_1(X)$ and $\lambda \in \mathbb{R}_0(X)$ such that:

$$-\ln(p) = H + \zeta \cdot \delta \psi + \lambda \tag{4.74}$$

Define $\varphi \in A_1(X)$ by letting $\varphi_{\alpha\beta} = \psi_{\alpha\beta}$ for non-empty β, and otherwise letting:

$$\varphi_{\alpha\varnothing} = \psi_{\alpha\varnothing} - \sum_{\alpha \supseteq \beta} \mu_{\alpha\beta} \, \lambda_\beta \tag{4.75}$$

Then $\zeta(\delta\varphi)_\alpha = \zeta(\delta\psi)_\alpha + \lambda_\alpha$ for all non-empty α, while $-\ln(p_\varnothing) = 0$. □

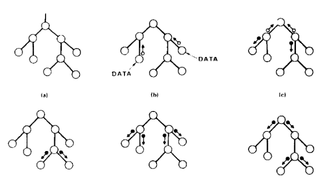
[8] The $\cap$-closure of X is in fact equivalent to the «region-graph» condition of [32]. Note that X was assumed closed under $\cap$ since section 2.3, as this property was necessary for the interaction decomposition to hold.

[9] As noticed by D. Bennequin, the original proof given in [32] is problematic when there exists β such that $c_\beta = 0$.

Chapter 5

Message-passing and Diffusion

In this chapter, we consider dynamics over the space of statistical fields that enlarge a class of bayesian inference algorithms known as *message-passing algorithms.* The generalised belief propagation algorithm introduced by Yedidia, Freeman and Weiss in [33] is the most interesting of these processes and reviewed in section 1. For a general study of belief propagation on networks and their relations to statistical physics, see [22].



The main contribution of this thesis is to reveal the homological character of message-passing algorithms. Among the deep consequences of this new view stands a collection of conservation laws. Section 2 shows how, at the level of energies, their discrete dynamics approximate a continuous-time transport equation of the form $\dot{u} = \delta\Phi(u)$. Recovering belief propagation through a naïve Euler scheme of time step 1, the present approach yields new algorithms at smaller time scales. Shorter time steps seem highly advisable to avoid unstable behaviour.

Finally, building on the higher degree combinatorics of chapter 3, we propose a combinatorial enhancement of the algorithm and its continuous version, through a degree-one Möbius inversion on the energy flux. We show in section 3 how this canonical diffusion eliminates redundancies and natively allows to enforce Dirichlet boundary conditions, a fundamental component of learning.

5.1 Belief Propagation

Belief propagation was first introduced as a decoding algorithm by Gallager in his 1960 PhD thesis on low-density parity-check codes [7]. In his setting, a collection of code bits and parity-check bits, transmitted through a noisy communication channel, exchange messages to update the belief on their true value until parity-check consistency is achieved. The algorithm was later rediscovered by Pearl in 1982 [23] to perform exact inference on bayesian trees, then generalised by Yedidia *et al.* in 2001 [33] on more general coverings $X \subseteq \mathcal{P}(\Omega)$.

This introductory section first defines the algorithm in its general form, before reviewing some of its remarkable properties, a starting point of this work having been the theorem of [32] establishing a correspondence between fixed points of the algorithm and critical points of a Bethe free energy.

5.1.1 Algorithm

Given $X \subseteq \mathcal{P}(\Omega)$, generalised belief propagation assumes priors are described by a collection (f_α) of strictly positive observables indexed by $\alpha \in X$. The dynamic takes place on a set of messages $(m_{\alpha\beta})$ according to the update rule:

$$m_{\alpha\beta} \leftarrow m_{\alpha\beta} \cdot \Sigma^{\beta\alpha} \left(\frac{\prod\limits_{\beta' \in \Lambda^\alpha \smallsetminus \Lambda^\beta} f_{\beta'} \times \prod\limits_{\alpha'\beta' \in d\Lambda^\alpha \smallsetminus d\Lambda^\beta} m_{\alpha'\beta'}}{\prod\limits_{\beta'\gamma' \in d\Lambda^\beta \smallsetminus d\Lambda^\alpha} m_{\beta'\gamma'}} \right) \tag{5.1}$$

where $\Lambda^\alpha \subseteq X$ still stands for the cone of subregions of α and $d\Lambda^\alpha$ for its coboundary[1]. Denoting by $G_\alpha \subseteq A_\alpha$ the multiplicative group of strictly positive observables for every $\alpha \in X$, the algorithm thus iterates over $m \in G_1(X)$ given $f \in G_0(X)$. This crude yet classical formula justifies the alternative name of *sum-product algorithm*.

Priors and messages serve to define a collection of beliefs (q_α) according to the formula:

$$q_\alpha = \left[\prod_{\beta' \in \Lambda^\alpha} f_{\beta'} \times \prod_{\alpha'\beta' \in d\Lambda^\alpha} m_{\alpha'\beta'} \right] \tag{5.2}$$

the normalisation bracket projecting G_α onto the subspace $\mathring{\Delta}_\alpha$ of non-vanishing probability densities. The dynamic of the beliefs is bound to that of the messages, now following the much nicer update rule:

$$m_{\alpha\beta} \leftarrow m_{\alpha\beta} \cdot \frac{\Sigma^{\beta\alpha}(q_\alpha)}{q_\beta} \tag{5.3}$$

Stationary states of the algorithm hence correspond to consistent[2] beliefs (q_α) searched within a particular subspace defined by the priors (f_α). We shall relate this subspace to a homology class of $A_0(X)$ in section 2. Another result not yet stated to our knowledge, is that stationarity of beliefs implies stationarity of messages. This will justify to focus on the dynamic over $q \in \mathring{\Delta}_0(X)$ following the update rule:

$$q_\alpha \leftarrow \left[q_\alpha \times \prod_{\alpha'\beta' \in d\Lambda^\alpha} \frac{\Sigma^{\beta'\alpha'}(q_{\alpha'})}{q_{\beta'}} \right] \tag{5.4}$$

We will denote by $\mathrm{BP} : \mathring{\Delta}_0(X) \to \mathring{\Delta}_0(X)$ the associated smooth map, the algorithm reading $q \leftarrow \mathrm{BP}(q)$. It will only be a matter of preference whether to keep track of the messages $m \in G_1(X)$ over time, the relevance of the final messages being questionable as the assignment $(f, m) \mapsto q$ is not injective.

5.1.2 Properties

Belief propagation was initially considered on graphs, which we describe by coverings $X \subseteq \mathcal{P}(\Omega)$ containing the vertex $\{i\}$ for all $i \in \Omega$, and otherwise consisting only of edges of the form $\alpha = \{i, j\}$. In general, the underlying graph of Gallager's algorithm contains loops, but Pearl's algorithm, describing hierchical data structures, was however restricted to acyclic graphs. *Retractable hypergraphs* $X \subseteq \mathcal{P}(\Omega)$ shall be defined in chapter 6 to generalise the acyclic property of trees and the following theorem will then extend Pearl's result. On such retractable systems, the fundamental property justifying interest in belief propagation is the parallelised computation it provides for exact bayesian inference.

[1] $d\Lambda^\alpha$ is the set of ordered pairs $\alpha' \supseteq \beta'$ such that $\alpha' \notin \Lambda^\alpha$ and $\beta' \in \Lambda^\alpha$ (see 3.2.3). Our formula stays as close as possible as that of [32] but we already write the products over set differences of cone coboundaries.

[2] The beliefs (q_α) are consistent when q_β is the marginal of q_α whenever β is contained in α.

Theorem 5.1. *Assume $X \subseteq \mathcal{P}(\Omega)$ is retractable*[3]*. Given positive priors (f_α) and initial messages $(m_{\alpha\beta})$, the beliefs (q_α) iterated through* BP *converge to the exact marginals of the Markov field:*

$$p_\Omega = \left[\prod_{\alpha \in X} f_\alpha \right] \tag{5.5}$$

Moreover, convergence is reached in finite time less or equal to the diameter of X.

When X is a graph with loops, in general when X is an unretractable hypergraph, belief propagation still performs approximate bayesian inference surprisingly well. The following theorem of [32] beautifully bridges bayesian learning with statistical physics and greatly motivated the present work. It expresses that stationary states of belief propagation actually compute the critical points of a Bethe free energy approximation, which according to Kikuchi's cluster variation method, should give a good estimate of the marginals of the Markov field $p_\Omega = [\mathrm{e}^{-H_\Omega}]$.

Theorem 5.2. *For any $X \subseteq \mathcal{P}(\Omega)$, the fixed points of* BP *with priors (f_α) are in one-to-one correspondence with the critical points of $\check{\mathcal{F}}_H$ with respect to the local hamiltonians:*

$$H_\alpha = \sum_{\beta' \in \Lambda^\alpha} - \ln f_{\beta'} \tag{5.6}$$

It was recognised by D. Bennequin that the proof given by Yedidia *et al.* is problematic when there exists $\alpha \in X$ such that the Möbius number c_α vanishes, and we shall correct their proof in section 2.

In the general setting, the uniqueness of equilibria is not maintained. Although convergence of the algorithm remains an open question, the existence of at least one fixed point is insured.

Theorem 5.3. *For any $X \subseteq \mathcal{P}(\Omega)$ and set of priors, there exists a fixed point of* BP*.*

Proof. The topology of $\Delta_0(X)$ is that of a product of spheres and the proof given in [4] relies on Brouwer's fixed point theorem, after having shown that BP : $\mathring{\Delta}_0(X) \to \mathring{\Delta}_0(X)$ keeps away from the boundary of $\Delta_0(X)$. □

5.2 Statistical Diffusions

We now introduce an ordinary differential equation of the form $\dot{u} = \delta\Phi(u)$ on interaction potentials. This transport equation is very reminiscent of the heat equation and similar diffusions[4]. Trying to understand the geometric nature of message-passing formulas and unveiling their connections with algebraic topology was the first motivation of this thesis: we relate belief propagation to a coarse integrator of this diffusion flow.

This homological picture completes the correspondence between the stationary states of belief propagation and the critical points of Bethe free energy $\check{\mathcal{F}}(\,-\,, H)$, as described by theorem 4.22. Both lie at the intersection of a homology class of interaction potentials[5] with the space of consistent beliefs, through the non-linear correspondence summarised in table 5.1:

$$[\mathrm{e}^{-U}] \in \mathring{\Gamma}(X) \quad \text{and} \quad U \in H + \zeta \cdot \mathrm{Im}(\delta) \tag{5.7}$$

[3] See definition 6.3.

[4] Apart from the non-linearity of the algorithm, the main difference lies in the central role of the zeta transform defining local hamiltonians from interaction potentials. See section 6.3.

[5] Technically $U \in H + \zeta \cdot \mathrm{Im}(\delta)$ should be understood up to additive constants. We shall discuss how to best deal with normalisation constraints later on.

Behind this correspondence is an interesting form of duality, exchanging constraints with degrees of freedom. The free energy variational problem looks for constrained $q \in \mathring{\Gamma}(X)$ such that variations are generated by Lagrange multipliers, while the algorithm iterates over $\varphi \in A_1(X)$ until beliefs eventually reach a consistent equilibrium state.

interaction potentials	$u = h + \delta\varphi$
local hamiltonians	$U = \zeta \cdot u$
beliefs	$q = \left[\, \mathrm{e}^{-U} \,\right]$

Table 5.1: Interaction potentials, local hamiltonians, and local Gibbs states – or beliefs.

Decomposing the dynamic into elementary operations, the vector field we introduce is of the form $\mathcal{T} = \delta \circ (-\mathcal{D}) \circ \zeta$ on the vector space $A_0(X)$ of interaction potentials:

$$\begin{tikzcd} A_0(X) \arrow[r, "\zeta"] & A_0(X) \arrow[dl, "-\mathcal{D}"] \\ A_1(X) \arrow[u, "\delta"] & \end{tikzcd} \tag{5.8}$$

The definition of the non-linear map $\mathcal{D}$ is the object of the first subsection. Preparing for the study of other flux functionals, we then characterise a family of transport equations of the form $\dot{u} = \delta\Phi(u)$ sharing the common property of yielding critical points of Bethe free energy at equilibrium. Specialising to the case where $\Phi = -\mathcal{D} \circ \zeta$, we show how explicit Euler schemes on $A_0(X)$ generalise the discrete dynamic of BP on $\mathring{\Delta}_0(X)$ to arbitrary time scales.

5.2.1 Effective Energy Gradient and Consistency

The effective energy gradient $\mathcal{D} \in C^\infty\big(A_0(X), A_1(X)\big)$ will be essential in defining the currents carrying energy from one region to another. Cancellation of these currents will then define equilibrium as a collection of effective hamiltonians $U \in A_0(X)$ whose effective energies are consistent with one another.

Definition 5.4. *We call effective energy gradient the smooth map $\mathcal{D} : A_0(X) \to A_1(X)$ defined by:*

$$\mathcal{D}(H)_{\alpha\beta} = H_\beta - \mathbb{F}^{\beta\alpha}(H_\alpha) \tag{5.9}$$

where $\mathbb{F}^{\beta\alpha}(H_\alpha) = -\ln \Sigma^{\beta\alpha}(\mathrm{e}^{-H_\alpha})$ is the effective energy as defined in 4.1.2.

The zero locus of $\mathcal{D}$ is naturally diffeomorphic to the space of consistent positive measures by 4.9. There will be two different ways to account for the normalisation constraints.

Definition 5.5. *We say that a collection of local hamiltonians $U \in A_0(X)$ is:*

- consistent *if $\mathcal{D}(U) = 0$,*
- projectively consistent *if $\mathcal{D}(U) \in \mathbb{R}_1(X)$.*

We denote by $\mathcal{C}(X) \subseteq \mathcal{C}'(X)$ the spaces of consistent and projectively consistent local hamiltonians.

Note for instance that $0 \in A_0(X)$ is only projectively consistent as:

$$\mathcal{D}(0)_{\alpha\beta} = \ln|E_\beta| - \ln|E_\alpha| \tag{5.10}$$

There is however a unique consistent $U = \ln|E| \in \mathbb{R}_0(X)$ such that $U_\varnothing = 0$, accounting for the entropic contributions in the high temperature limit.

Proposition 5.6. *Consistent probability densities may be parametrised by local hamiltonians via:*

(i) $\mathcal{C}'(X)$ *is the inverse image of* $\mathring{\Gamma}(X)$ *under* $U \mapsto [\mathrm{e}^{-U}]$ *and:*

$$\mathring{\Gamma}(X) \simeq \mathcal{C}'(X)/\mathbb{R}_0(X) \tag{5.11}$$

(ii) $\{U \in \mathcal{C}(X) \mid U_\varnothing = \lambda\}$ *is diffeomorphic to* $\mathring{\Gamma}(X)$ *under* $U \mapsto \mathrm{e}^{-(U-\lambda)}$ *for all* $\lambda \in \mathbb{R}$ *and:*

$$\mathring{\Gamma}(X) \simeq \mathcal{C}(X)/\mathbb{R} \tag{5.12}$$

Proof. (i) Letting $p = [\mathrm{e}^{-U}]$, for all $\alpha \supseteq \beta$ one has $p_\beta = \Sigma^{\beta\alpha}(p_\alpha)$ if and only if $U_\beta = \mathbb{F}^{\beta\alpha}(U_\alpha) \mod \mathbb{R}$. As $[\mathrm{e}^{-U}] = [\mathrm{e}^{-U'}]$ if and only if $U' \in U + \mathbb{R}_0(X)$, the quotient $\mathcal{C}'(X)/\mathbb{R}_0(X)$ is diffeomorphic to $\mathring{\Gamma}(X)$. (ii) $\mathcal{D}(U) = 0$ implies consistency and $\Sigma^{\varnothing\alpha}(\mathrm{e}^{-U_\alpha}) = \mathrm{e}^{-U_\varnothing} = \mathrm{e}^{-\lambda}$ holds a single normalisation factor. Reciprocally when $p \in \mathring{\Gamma}(X)$, letting $U = -\ln(p) + \lambda$ we have $\mathcal{D}(U) = 0$ with $U_\varnothing = -\ln(1) + \lambda = \lambda$. □

The manifold of consistent local hamiltonians $\mathcal{C}(X)$ has a natural riemannian structure. Given any consistent positive $p \in \mathring{\Gamma}(X)$, consider the inner product on $A_0(X)$ defined by:

$$\langle\, f \mid g \,\rangle_p = \sum_{\alpha \in X} \mathbb{E}_{p_\alpha}[f_\alpha \cdot g_\alpha] \tag{5.13}$$

As $\mathcal{C}'(X)$ is mapped onto $\mathring{\Gamma}(X)$, a metric of the form (5.13) is naturally associated to any $U \in \mathcal{C}'(X)$. The tangent space $T_U\mathcal{C}(X)$ is moreover described by the cohomology of a differential ∇, adjoint of δ for the metric induced by $[\mathrm{e}^{-U}]$, a characterisation which will be particularly useful in section 6.3.

Definition 5.7. *We denote*[6] *by* $\nabla = \mathcal{D}_*$ *the linearised effective energy gradient:*

$$\nabla(H)_{\alpha\beta} = H_\beta - \mathbb{E}^{\beta\alpha}[H_\alpha] \tag{5.14}$$

viewed as the smooth map $\nabla : \mathcal{C}'(X) \to \mathrm{Hom}\big(A_0(X), A_1(X)\big)$*, letting* $\mathbb{E}^{\beta\alpha} = d\mathbb{F}^{\beta\alpha}$ *as per 4.14.*

With this notation, we have $T\mathcal{C}(X) = \mathrm{Ker}(\nabla)$ as a subbundle of $TA_0(X)$ restricted above $\mathcal{C}(X)$. As the following suggests, ∇ naturally extends to a differential acting on all degrees of $A_\bullet(X)$.

Proposition 5.8. *Given* $p = [\mathrm{e}^{-H}] \in \Gamma(X)$*, we have* $\nabla = \delta^*$ *for the metric induced by* p*.*

Proof. This is the consequence of the fact that the conditional expectation $\mathbb{E}^{\beta\alpha} : A_\alpha \to A_\beta$ is adjoint to the canonical extension $j_{\alpha\beta} : A_\beta \to A_\alpha$ for the metrics induced by p_α and p_β:

$$\sum_{E_\beta} g_\beta \cdot p_\beta \cdot \mathbb{E}^{\beta\alpha}[f_\alpha] = \sum_{E_\alpha} g_\beta \cdot p_\alpha \cdot f_\alpha \tag{5.15}$$

Standard computations then show that $\langle\, \nabla f \mid \varphi \,\rangle_p = \langle\, f \mid \delta\varphi \,\rangle_p$ for all $f \in A_0(X)$ and $\varphi \in A_1(X)$. □

Acting on local hamiltonians, the effective energy gradient $\mathcal{D}$ will be generally precomposed by ζ. Writing the hamiltonian $H = \zeta \cdot h$ as a sum of local interactions, one may think of $\mathcal{D}(H)_{\alpha\beta}$ as the effective contribution of $\Lambda^\alpha \smallsetminus \Lambda^\beta$ to the energy of Λ^β:

$$\mathcal{D}(\zeta \cdot h)_{\alpha\beta} = \mathbb{F}^{\beta\alpha}\Bigg(\sum_{\beta' \in \Lambda^\alpha \smallsetminus \Lambda^\beta} h_{\beta'}\Bigg) \tag{5.16}$$

To complete the picture on consistency, let us finally parametrise $\mathring{\Gamma}(X)$ by interaction potentials.

Definition 5.9. *We denote by:*

[6]To avoid burdening notations, we will often leave the dependency of ∇ in $p \in \mathring{\Gamma}(X)$ or $U \in \mathcal{C}'(X)$ implicit.

– $\mathcal{Z}(X) = \mu \cdot \mathcal{C}(X)$ *the manifold of* consistent interaction potentials*,*

– $\mathcal{Z}'(X) = \mu \cdot \mathcal{C}'(X)$ *the manifold of* projectively consistent interaction potentials*.*

Proposition 5.10. *Consistent probability densities may be parametrised by interaction potentials via:*

(i) $\mathcal{Z}'(X)$ *is the inverse image of* $\mathring{\Gamma}(X)$ *under* $u \mapsto [\mathrm{e}^{-\zeta \cdot u}]$ *and:*

$$\mathring{\Gamma}(X) \simeq \mathcal{Z}'(X)/\mathbb{R}_0(X) \tag{5.17}$$

(ii) $\{u \in \mathcal{Z}(X) \mid u_\varnothing = \lambda\}$ *is diffeomorphic to* $\mathring{\Gamma}(X)$ *under* $u \mapsto \mathrm{e}^{\lambda - \zeta \cdot u}$ *for all* $\lambda \in \mathbb{R}$ *and:*

$$\mathring{\Gamma}(X) \simeq \mathcal{Z}(X)/\mathbb{R} \tag{5.18}$$

Proof. This follows from 5.6 as $\zeta \cdot \mathbb{R}_0(X) = \mathbb{R}_0(X)$, while for $u = \mu \cdot U$ one has $u_\varnothing = U_\varnothing$. □

The origin $\bar{0} \in \mathcal{Z}(X)$, associated to the uniform state $p = [1]$, describes the high temperature limit where all variables are independent. The following characterisation of $T_{\bar{0}}\mathcal{Z}(X)$ should be seen as a consequence of this independency[7].

Proposition 5.11. *Letting* $\bar{0} = \mu \cdot \ln |E| \in \mathcal{Z}(X)$*, one has:*

$$T_{\bar{0}}\mathcal{Z}(X) = Z_0(X) \tag{5.19}$$

where $Z_0(X) \subseteq A_0(X)$ *is the image of the canonical interaction decomposition defined in 2.3.2.*

Proof. This is a consequence of [[Ker$(d) = \zeta \cdot Z_0(X)$ in 2.3.3]] □

5.2.2 Diffusions and Correspondence Theorems

Given a smooth *flux functional* $\Phi : A_0(X) \to A_1(X)$, we consider the transport equation:

$$\dot{u} = \delta\Phi(u) \tag{5.20}$$

Let us mention two main differences of the present approach with usual message-passing algorithms:

1. (5.20) is an ordinary differential equation, and $\dot{u} = \frac{du}{dt}$ represents the derivative of $u : \mathbb{R} \to A_0(X)$ with respect to a continuous time variable[8].
2. (5.20) is a dynamic on the vector space of interaction potentials. A dynamic is induced on the multiplicative group of positive beliefs by letting $q = [\mathrm{e}^{-\zeta \cdot u}]$ but we do not take the usual point of view of a dynamic over messages.

Difference 1 consists in a significant improvement for the stability of such algorithms, belief propagation being recovered through the unreasonably coarse finite difference approximation $\dot{u}(t) \simeq u(t+1) - u(t)$. Difference 2 makes the homological character of message-passing more apparent. Note that a dynamic on $\varphi \in A_1(X)$ may be recovered as $\dot{\varphi} = \Phi(h + \delta\varphi)$ for some initial interaction potentials $h \in A_0(X)$.

In this paragraph, we show the announced correspondence between stationary points of belief propagation and critical points of Bethe free energy. Although this correspondence is made more

[7] The conditional expectation $\mathbb{E}^{\beta\alpha}$ is always adjoint to the inclusion $j_{\alpha\beta}$ for the metric induced by the probability density, as orthogonal projection of A_α onto A_β. The effect of interactions between β and β' is however that A_β is no longer orthogonal to $A_{\beta'}$ in A_α. Although one could define interaction subspaces as $Z_\alpha = \cap_{\beta \subset \alpha} \mathrm{Ker}(\mathbb{E}^{\beta\alpha})$ they would hence not satisfy $\mathbb{E}^{\beta\alpha}(Z_{\beta'}) = \{0\}$.

[8] The possibility to differentiate w.r.t. continuous time was probably occluded by the common preference for the multiplicative point of view of «beliefs». Ironically, there is absolutely no originality in the additive point of view of «energies» as Gallagher's electronic apparatus was already logarithmic – additions are simpler than multiplications for humans and machines all the same.

coherent by viewing belief propagation as a dynamic over beliefs rather than on messages, this approach required us to prove that stationarity of beliefs implied stationarity of messages. This will come as a consequence of the *faithfulness* of the flux functional $\Phi = -\mathcal{D} \circ \zeta$, given by proposition 5.17.

More generally, the following definitions help characterise the class of flux functionals admissible for (5.20) to seek critical points of a Bethe free energy. Morally, consistency of Φ will imply that critical points are stationary under $\dot{u} = \delta\Phi(u)$, while the stronger faithfulness property is necessary for the converse to hold. The existence and design of an *optimal* flux functional satisfying these properties is a natural question then to be raised.

Definition 5.12. *A smooth flux functional $\Phi : A_0(X) \to A_1(X)$ will be said:*

- consistent *if $u \in \mathcal{Z}(X) \Rightarrow \Phi(u) = 0$*
- faithful *if $u \in \mathcal{Z}(X) \Leftrightarrow \delta\Phi(u) = 0$*
- projectively faithful *if $u \in \mathcal{Z}'(X) \Leftrightarrow \delta\Phi(u) \in \mathbb{R}_0(X)$*
- locally faithful *if the restriction of Φ to a neighbourhood of $\mathcal{Z}(X)$ is faithful.*

Suppose now given reference interaction potentials $h \in A_0(X)$ such that $h_\varnothing = 0$, and let $H = \zeta \cdot h$. The correspondence is essentially a rephrasing of theorem 4.28, characterising critical points of $\check{\mathcal{F}}(\,-\,, H)$, all difficulties being kept hidden behind the faithfulness assumption. To account for normalisation[9], we still denote by δ' the truncation of δ to $X \smallsetminus \{\varnothing\}$ as defined in 4.27.

Theorem 5.13. *Assume Φ is faithful. Then for all $u \in A_0(X)$, the following are equivalent:*

(i) *There exists φ such that $u = h + \delta'\varphi$ is stationary for $\dot{u} = \delta'\Phi(u)$.*

(ii) *The beliefs $\mathrm{e}^{-\zeta \cdot u}$ are critical for $\check{\mathcal{F}}(\,-\,, H)$ constrained to $\mathring{\Gamma}(X)$.*

Lemma 5.14. *For all $\varphi \in A_1(X)$ if $\delta'\varphi = 0$ then $\delta\varphi = 0$.*

Proof. From the global Gauss formula $\sum_\beta \delta_\beta\varphi = 0$, if $\delta_\alpha\varphi = 0$ for all $\alpha \neq \varnothing$ then $\delta_\varnothing\varphi = 0$, see 2.3. □

Proof of theorem 5.13. Recall that $q \in \mathring{\Gamma}(X)$ is critical for $\check{\mathcal{F}}(\,-\,, H)$ constrained to $\Gamma(X)$ if and only if there exists Lagrange multipliers $\varphi \in A_1(X)$ such that $-\ln(q) = H + \zeta \cdot \delta'\varphi$ by theorem 4.28.

- Assume q is critical and let $\zeta \cdot u = -\ln(q)$. Then there exists φ such that $u = h + \delta'\varphi$ while $\mathcal{D}(\zeta \cdot u) = 0$ by consistency. We have $\delta'\Phi(u) = \delta\Phi(u) = 0$ by faithfulness of Φ and u is stationary.
- Reciprocally, assume $u = h + \delta'\varphi$ is stationary. According to the lemma, $\delta'\Phi(u) = 0$ implies $\delta\Phi(u) = 0$ so that $\mathcal{D}(\zeta \cdot u) = 0$ by faithfulness of Φ. Letting $U = \zeta \cdot u$, it follows that $\mathcal{D}(U) = 0$ with $U_\varnothing = 0$ hence $q = \mathrm{e}^{-U} \in \mathring{\Gamma}(X)$ is consistent and critical for the constrained free energy.

□

Theorem 5.15. *Assume Φ is projectively faithful. For all $u \in A_0(X)$, the following are equivalent:*

(i) *There exists φ such that $u = h + \delta'\varphi$ is projectively stationary for $\dot{u} = \delta'\Phi(u)$.*

(ii) *The beliefs $[\mathrm{e}^{-\zeta \cdot u}]$ are critical for $\check{\mathcal{F}}(\,-\,, H)$ constrained to $\mathring{\Gamma}(X)$.*

Proof.

- If $q = [\mathrm{e}^{-\zeta \cdot u}] \in \mathring{\Gamma}(X)$ is critical for $\check{\mathcal{F}}(-, H)_{|\Gamma(X)}$, there exists φ such that $u = h + \delta'\varphi$ by 4.28. Consistency of q implies $u \in \mathcal{Z}'(X)$ by 5.10 and $\delta\Phi(u) \in \mathbb{R}_0(X)$ by projective faithfulness of Φ.
- Reciprocally, assume $u = h + \delta'\varphi$ is projectively stationary. As $\delta\Phi(u) \in \mathbb{R}_0(X)$ implies $u \in \mathcal{Z}'(X)$ by projective faithfulness, $q = [\mathrm{e}^{-\zeta \cdot u}]$ is consistent and critical for the constrained free energy.

[9]Although belief propagation has to normalise each step, scaling factors being otherwise unstable on graphs with loops, first experiments suggest that this normalisation procedure might be superfluous for finer integrators of (5.20).

$\square$

Let us now show that the flux functional inducing belief propagation is faithful. The algorithm shall be recovered as a coarse integrator of the differential equation:

$$\dot{u} = \delta\varphi \qquad \text{where} \qquad \left| \begin{array}{l} \varphi = -\mathcal{D}(U) \\ U = \zeta \cdot u \end{array} \right. \tag{5.21}$$

Transport takes place at the level of effective interaction potentials u, an energy flux φ balancing the effective hamiltonians $U = \zeta \cdot u$ until they reach effective consistency. The evolution of u being restricted to a single homology class, let us emphasise that total energy is conserved along any integral curve of (5.21). Hence for any reference interaction potentials h homologous to $u(0)$, one has:

$$\sum_\alpha u_\alpha(t) = \sum_\alpha h_\alpha \tag{5.22}$$

This is a direct consequence of 2.9 and obviously holds along integral curves of (5.20) as well.

Definition 5.16. *We call:*

- standard diffusion flux *the functional $\Phi : A_0(X) \to A_1(X)$ defined by $\Phi = -\mathcal{D} \circ \zeta$,*
- standard diffusion *the vector field $\mathcal{T}$ on $A_0(X)$ defined by $\mathcal{T} = \delta\Phi$.*

Proposition 5.17. *The standard diffusion flux is faithful.*

Proof. The proposition claims that $\delta\mathcal{D}(H) = 0$ implies $\mathcal{D}(H) = 0$ for any local hamiltonians H. Denoting by d the adjoint of δ for the canonical metric of $A_0(X)$, we have for every $v, H \in A_0(X)$:

$$\langle\, v \,|\, \delta\mathcal{D}(H) \,\rangle = \langle\, d(v) \,|\, \mathcal{D}(H) \,\rangle \tag{5.23}$$

Assuming that $\delta\mathcal{D}(H) = 0$ and letting $v = \mathrm{e}^{-H}$ the above integration by parts formula gives:

$$\langle\, d(\mathrm{e}^{-H}) \,|\, \mathcal{D}(H) \,\rangle = \sum_{\alpha\beta \in N_1(X)} \Big\langle\, \mathrm{e}^{-H_\beta} - \Sigma^{\beta\alpha}(\mathrm{e}^{-H_\alpha}) \,\Big|\, H_\beta + \ln \Sigma^{\beta\alpha}(\mathrm{e}^{-H_\alpha}) \Big\rangle_{E_\beta} = 0 \tag{5.24}$$

It then follows by monotonicity of $y \mapsto -\ln(y)$ that the differences $v_\beta - \Sigma^{\beta\alpha}(v_\alpha)$ and $H_\beta - \mathbb{F}^{\beta\alpha}(H_\alpha)$ have opposite signs. As $v_\beta = \Sigma^{\beta\alpha}(v_\alpha)$ is equivalent to $H_\beta = \mathbb{F}^{\beta\alpha}(H_\alpha)$, we do have $\mathcal{D}(H) = 0$. $\square$

In the next paragraph, we shall enforce normalisation constraints through a Möbius inversion on the flux of Φ outbound to $\varnothing$. We introduce the flux functional Φ' defined by $\Phi'(u)_{\alpha\beta} = \Phi(u)_{\alpha\beta}$ when $\beta \neq \varnothing$ and otherwise by:

$$\Phi'(u)_{\alpha\varnothing} = \sum_{\alpha \to \beta'} \mu_{\alpha\beta} \cdot \Phi(u)_{\beta\varnothing} \tag{5.25}$$

The faithfulness of Φ' does not come as an easy consequence of the faithfulness of Φ.

Proposition 5.18. *For every $H \in A_0(X)$ we have the equivalence:*

$$\delta\mathcal{D}(H) \in \mathbb{R}_0(X) \quad \Leftrightarrow \quad \exists \lambda \in \mathbb{R}_0(X) \quad \text{s.t.} \quad \mathcal{D}(H - \lambda) = 0 \tag{5.26}$$

Proof. Letting $\Delta = \delta d$ denote the laplacian on $\mathbb{R}_0(X)$, we have $\mathrm{Im}(\Delta) = \mathrm{Im}(\delta) + \mathrm{Im}(d)$ by Hodge decomposition. By additivity of effective energy along constants, one has $\mathcal{D}(H - \lambda) = \mathcal{D}(H) - d\lambda$ for all $\lambda \in \mathbb{R}_0(X)$ so that:

$$\delta\mathcal{D}(H - \lambda) = \delta\mathcal{D}(H) - \Delta(\lambda) \tag{5.27}$$

If $\delta\mathcal{D}(H) = \lambda' \in \mathbb{R}_0(X)$, then $\lambda' \in \mathrm{Im}(\delta) \subseteq \mathrm{Im}(\Delta)$ and there exists $\lambda \in \mathbb{R}_0(X)$ such that $\Delta(\lambda) = \lambda'$. It follows that $\delta\mathcal{D}(H - \lambda) = 0$. The faithfulness property 5.17 of $\mathcal{D} \circ \zeta$ then implies $\mathcal{D}(H - \lambda) = 0$. $\square$

Proposition 5.19. *The flux Φ' defined by (5.25) is projectively faithful.*

Proof. From proposition 5.18, projective stationarity $\delta\Phi(u) \simeq \delta\Phi'(u) \simeq 0 \mod \mathbb{R}_0(X)$ implies the projective consistency of u as $\mathcal{D}(\zeta \cdot u) = \mathcal{D}(\lambda) \in \mathbb{R}_1(X)$ for some $\lambda \in \mathbb{R}_0(X)$. $\square$

5.2.3 Euler Schemes and Belief Propagation

The ordinary differential equation (5.21) defines a continuous-time transport $\dot{u} = \delta\Phi(u)$ of energy at the level of interaction potentials. Its flow may be approximated by common methods of numerical integration, but belief propagation is related to the time-step-1 explicit Euler scheme $u \leftarrow u + \delta\Phi(u)$, which approximates the flow of the vector field $\mathcal{T} = \delta\Phi$ by:

$$e^{n\mathcal{T}} \simeq (1 + \mathcal{T})^n \tag{5.28}$$

We warn against the use of such a coarse scheme[10] as 1 may actually be met by the integrator's Lipschitz bound. A straightforward refinement is to reduce the time step to $0 < \lambda < 1$, yielding new and better-behaved belief propagation algorithms, while higher order integrators could also prove useful.

Defining effective hamiltonians as $U(t) = H + \zeta \cdot \delta\varphi(t)$, it is the differential equation $\dot{\varphi} = -\mathcal{D}(U)$ that is more precisely related to the multiplicative algorithm[11] of equations (5.2) and (5.3). The crucial ingredient in recovering (5.2) is the Gauss formula of proposition 2.3, which gives on every cone Λ^α:

$$U_\alpha(t) = H_\alpha + \sum_{\alpha'\beta' \in d\Lambda^\alpha} \varphi_{\alpha'\beta'}(t) \tag{5.29}$$

Equation (5.29) is, up to additive constants, the logarithm of (5.2) defining beliefs from messages. Approximating the evolution $\dot{\varphi} = -\mathcal{D}(U)$ by the finite difference iteration $\varphi \leftarrow \varphi - \mathcal{D}(U)$ then yields: the logarithm of the message update rule (5.3):

$$\varphi_{\alpha\beta} \leftarrow \varphi_{\alpha\beta} - \ln\left(\frac{\Sigma^{\beta\alpha}\, e^{-U_\alpha}}{e^{-U_\beta}}\right) \tag{5.30}$$

Faithfulness of $\Phi = -\mathcal{D} \circ \zeta$ justifies our choice to drop messages out of memory and focus on the transport equation $\dot{u} = \delta\Phi(u)$ instead. Accounting for normalisation could be done by projecting the evolution of u onto $A_0(X)/\mathbb{R}_0(X)$. As a different point of view, we show that splitting the flux Φ' of (5.25) as $\Phi' = \Phi'_{int} + \Phi'_{out}$, where Φ'_{out} gathers all the flux terms $(\Phi'_{\alpha\varnothing})$ directed to $\varnothing$, allows to naturally enforce normalisation at each step. This only requires to replace δ by its truncation δ' to $X \setminus \{\varnothing\}$, as defined in 4.27, and prepares for the more general boundary conditions considered in the next section.

Proposition 5.20. *Under the correspondence $q = e^{-\zeta \cdot u}$ belief propagation is a splitting scheme for the transport equation $\dot{u} = \delta'\Phi'(u)$ associated to the decomposition $\Phi' = \Phi'_{int} + \Phi'_{out}$, each term being integrated through an explicit Euler scheme of time step 1.*

Proof. Consider the evolution $\dot{U} = \mathcal{X}_{int}(U) + \mathcal{X}_{out}(U)$ induced on the effective hamiltonians $U = \zeta \cdot u$, where $\mathcal{X}_{int}(U) = \zeta \cdot \delta\Phi'_{int}(u)$ and $\mathcal{X}_{out}(U) = \zeta \cdot \delta'\Phi'_{out}(u)$. Applying Gauss formulas on $\Lambda^\alpha \setminus \{\varnothing\}$, one may view $\mathcal{X}_{int}$ and $\mathcal{X}_{out}$ bound into and out of $\Lambda^\alpha \setminus \{\varnothing\}$ respectively.

$$\begin{aligned}
\mathcal{X}_{int}(U)_\alpha &= \sum_{\alpha'\beta' \in d\Lambda^\alpha} \Phi'_{int}(u)_{\alpha'\beta'} \qquad \text{with} \qquad \Phi'_{int}(u)_{\alpha'\beta'} = \mathbb{F}^{\beta'\alpha'}(U_{\alpha'} - U_{\beta'}) \\
\mathcal{X}_{out}(U)_\alpha &= -\sum_{\beta' \in \Lambda^\alpha_\varnothing} \Phi'_{out}(u)_{\beta'\varnothing} \qquad \text{with} \qquad \Phi'_{out}(u)_{\beta'\varnothing} = \sum_{\beta' \to \gamma'} \mu_{\beta'\gamma'}\, \mathbb{F}^{\gamma'}(U_{\gamma'})
\end{aligned} \tag{5.31}$$

Möbius inversion formulas yield $\mathcal{X}_{out}(U)_\alpha = -\mathbb{F}^\alpha(U_\alpha)$. Consider then the splitting scheme:

$$U(n+1) \simeq \big(1 + \mathcal{X}_{out}\big) \circ \big(1 + \mathcal{X}_{int}\big)\big(U(n)\big) \tag{5.32}$$

The first step $U \leftarrow U + \mathcal{X}_{int}(U)$ is, up to constants the logarithm of the belief update rule (5.4). The second step $U \leftarrow U + \mathcal{X}_{out}(U)$ corresponds to normalising[12] the belief $q_\alpha = [e^{-U_\alpha}] = e^{-U_\alpha + \mathbb{F}^\alpha(U_\alpha)}$. $\square$

[10] Consider for instance the real ODE $\dot{y} = -ay$ and the behaviour of $e^{-an\tau} \simeq (1 - a\tau)^n$ for different values of $a\tau$.

[11] Letting $U = -\ln(q)$ and $\varphi = -\ln(m)$. Reference hamiltonians are related to priors by $H = \zeta \cdot h$ with $h = -\ln(f)$.

[12] The total outbound flux $\delta_\varnothing \Phi'_{out} = \sum_\alpha c_\alpha\, \mathbb{F}^\alpha(U_\alpha)$ is a Bethe approximation of the total free energy $\mathbb{F}^\Omega(U_\Omega)$.

Approximating the flow of $\mathcal{T}$ by $\mathrm{e}^{n\mathcal{T}} \simeq (1+\mathcal{T})^n$ may lead to serious convergence and stability issues in regimes where the norm of $\mathcal{T}$ and its Lipschitz bound are not strictly smaller than 1. In simple examples with cycles, $\mathcal{T}_*$ may for instance presents periodic eigenvalues of the form $\mathrm{e}^{i2\pi/n}$.

A straightforward and recommendable improvement of BP is to reduce the time scale of the explicit Euler scheme and approximate the flow by $\mathrm{e}^{n\cdot\lambda\mathcal{T}} \simeq (1+\lambda\mathcal{T})^n$. Iterating the non-linear operator $(1+\lambda\mathcal{T})$ corresponds to updating messages according to:

$$m_{\alpha\beta} \leftarrow m_{\alpha\beta} \cdot \left(\frac{\Sigma^{\beta\alpha}(q_\alpha)}{q_\beta} \right)^\lambda \tag{5.33}$$

It is to be expected that reasonably small values of λ around 0.5 may already change the algorithm's behaviour dramatically. In some regimes, belief propagation has been reported to converge poorly after undergoing a kind of «phase transition». Changing the time scale may probably overcome this limitation.

Definition 5.21. *For every $\lambda > 0$, we call* belief propagation of time scale λ *the algorithm iterating over a collection $(q_\alpha) \in \mathring{\Delta}_0(X)$ of beliefs according to the update rule:*

$$q_\alpha \leftarrow \left[q_\alpha \;\times \prod_{\alpha'\beta' \in d\Lambda^\alpha} \left(\frac{\Sigma^{\beta'\alpha'}(q_{\alpha'})}{q_{\beta'}} \right)^\lambda \right] \tag{5.34}$$

We denote by $\mathrm{BP}_\lambda : \mathring{\Delta}_0(X) \to \mathring{\Delta}_0(X)$ the smooth map inducing the above dynamic.

The following identity is a straightforward consequence of the homological character of BP and conservation of the total energy. It was already known in particular cases, but not stated as a general fact to our knowledge.

Proposition 5.22 (Conservation)**.** *Let $q \in \mathring{\Delta}_0(X)^{\mathbb{N}}$ denote a sequence of beliefs iterated from BP_λ for some $\lambda > 0$. Then the quantity:*

$$q_\Omega(t) = \prod_{\alpha \in X} q_\alpha(t)^{c_\alpha} \tag{5.35}$$

remains constant in $G_\Omega = (\mathbb{R}_+^*)^{E_\Omega}$ up to a scaling factor.

5.3 Canonical Diffusion

As expressed by theorem 5.13, stationary states of any transport equation derived from a faithful flux functional solve the problem of finding consistent pseudo-marginals critical for a Bethe free energy. Therefore it remains a practical and theoretical open question whether $\Phi = -\mathcal{D} \circ \zeta$ is optimal among faithful fluxes, and if not whether a better-behaved flux may be designed. Completing diagram 5.8 with a Möbius inversion on degree one, we introduce a homological vector field $\tau = \delta\phi$ associated to the flux $\phi = \mu \circ (-\mathcal{D}) \circ \zeta$:

$$\begin{array}{ccc} A_0(X) & \xrightarrow{\;\zeta\;} & A_0(X) \\ {\scriptstyle\delta}\uparrow & & \downarrow{\scriptstyle -\mathcal{D}} \\ A_1(X) & \xleftarrow[\;\mu\;]{} & A_1(X) \end{array} \tag{5.36}$$

The symmetry of diagram 5.36 is of great appeal. It involves conjugation of operators on $A_\bullet(X)$ by the extended transforms ζ and μ, which shall have interesting cohomological consequences on the linearised dynamic. We claim that ϕ behaves better than the standard flux Φ for three main reasons:

(i) the flux bound into Λ^α is a Bethe approximation of the total effective energy of $X \smallsetminus \Lambda^\alpha$.

(ii) the flux from Λ^α to a subcone Λ^β is the effective energy of $\Lambda^\alpha \smallsetminus \Lambda^\beta$.

(iii) the algorithm $u \leftarrow u + \delta\varphi(u)$ restricted to a cone $\Lambda^\alpha \subseteq X$ converges in one step.

The standard flux Φ brings redundancies and (i) expresses that effective contributions of neighbouring regions are properly counted. In addition (ii) will allow for a natural enforcement of Dirichlet boundary conditions, the outbound flux taking care of reaching consistency with the boundary. Such boundary conditions, fixing the state of an exterior subset of variables, are a fundamental constituent of learning.

We prepare this section by introducing the differential calculus we shall use in presence of boundary, which relies on a boundary operator $\mathring{\delta}$ truncating δ to interior variables. Investigating some fundamental properties of the canonical flux, we then show that $\phi = \mu \circ (-\mathcal{D}) \circ \zeta$ satisfies a local faithfulness condition and prove proposition 5.32 supporting claims (i) and (ii) above, before going through some of the algorithms that generalise belief propagation by approximate integration of $\dot{u} = \delta\phi(u)$.

5.3.1 Calculus with Boundary

In contrast with differential geometry, there is no intrinsic notion of boundary on $X \subseteq \mathcal{P}(\Omega)$ and deciding which variables belong to the boundary and which do not is either arbitrary or dictated by experience[13]. Assuming $\partial\Omega \subseteq \Omega$ describes a set of variables whose state is given by the exterior, the following compatibility condition on X will define the boundary ∂X at the level of regions.

Definition 5.23 (Boundary)**.** *Given a subset of variables $\partial\Omega \subseteq \Omega$, let us denote by:*

- $\partial\alpha = \alpha \cap \partial\Omega$ *the boundary of a region* $\alpha \subseteq \Omega$,
- $\partial X = \{\partial\alpha \mid \alpha \in X\}$ *the boundary of* $X \subseteq \mathcal{P}(\Omega)$.

We say that a covering $X \subseteq \mathcal{P}(\Omega)$ is adapted to the boundary $\partial\Omega$ *whenever $\partial X \subseteq X$.*

As a consequence, the boundary $\partial\Lambda^\alpha$ of the cone $\Lambda^\alpha \subseteq X$ is the cone $\Lambda^{\partial\alpha} \subseteq \partial X$ for every $\alpha \in X$. In other words, every α has only one maximal subregion[14] $\partial\alpha$ belonging to ∂X, with $\partial\alpha = \varnothing$ when α does not contain any exterior variable. We finally define α to be *interior* whenever it is not contained in ∂X, equivalently, when α contains at least one variable of $\Omega \smallsetminus \partial\Omega$:

Definition 5.24. *If $X \subseteq \mathcal{P}(\Omega)$ is adapted to $\partial\Omega$, we denote by $\mathring{X} = X \smallsetminus \partial X$ the* interior *of X.*

To the splitting $X = \mathring{X} \sqcup \partial X$, we associate the direct sum decomposition $A_0(X) = A_0(\mathring{X}) \oplus A_0(\partial X)$ and write $u = u_{|\mathring{X}} + u_{|\partial X}$ for every $u \in A_0(X)$. When enforcing Dirichlet boundary conditions, we shall restrict the evolution of the interaction potentials u to the interior of X, their trace $u_{|\partial X}$ on the boundary describing input data or exterior stimuli. The evolution of $u_{|\mathring{X}}$ shall bear a homological character through the following definition. Although a general harmonic theory with boundary may be developed on the whole complex $A_\bullet(X)$, we solely focus on degrees zero and one for now.

Definition 5.25. *We call* interior divergence *the map $\mathring{\delta} : A_1(X) \to A_0(\mathring{X})$ truncating δ to $\mathring{X}$, defined by $\mathring{\delta}(\varphi) = \delta(\varphi)_{|\mathring{X}}$ for every $\varphi \in A_1(X)$.*

The following proposition is the analog of the integration by parts formula in differential geometry. Given a submanifold $V \subseteq \mathbb{R}^3$ with boundary ∂V, one has for every scalar field u and vector field $\vec{\varphi}$:

$$\int_V \vec{\mathrm{grad}}(u) \cdot \vec{\varphi}\, dv = -\int_V u \,\mathrm{div}(\vec{\varphi})\, dv + \int_{\partial V} u\, (\vec{\varphi} \cdot \vec{n})\, ds \tag{5.37}$$

denoting by $\vec{n}$ the outbound unit normal vector on ∂V. The formal adjunction of $\vec{\mathrm{grad}}$ with $-\mathrm{div}$ is tweaked by a boundary term representing the integral of the outbound flux of φ against u.

[13] In practice, the set of fixed boundary variables of a neural network depends on the training or testing context.

[14] This contrasts with the topological setting, as one may for instance take the boundary of the 2-simplex to be the triangle formed by its three edges. As there exists locally consistent pseudo-marginals on the triangle that do not have a consistent global extension, such a notion of boundary would be problematic when trying to enforce globally inconsistent Dirichlet boundary conditions.

Proposition 5.26 (Integration by parts)**.** *Let ∇ denote the adjoint of δ for a given metric. Then for every $u \in A_0(X)$ and every $\varphi \in A_1(X)$ we have:*

$$\langle \nabla u \,|\, \varphi \rangle = \langle u \,|\, \mathring{\delta}\varphi \rangle + b(u, \varphi) \tag{5.38}$$

where $b(u, \varphi) = \langle u \mid \mathbf{1}_{\partial X} \,|\, \delta\varphi \rangle$ denotes the scalar product of the restrictions of u and $\delta\varphi$ to ∂X.

Proof. In absence of boundary we have the classical integration by parts formula $\langle \nabla u \,|\, \varphi \rangle = \langle u \,|\, \delta\varphi \rangle$ expressing the adjunction of ∇ with δ. The above simply consists of introducing a splitting of $\langle u \,|\, \delta\varphi \rangle$ as $\langle u \mid \mathbf{1}_{\mathring{X}} + \mathbf{1}_{\partial X} \,|\, \delta\varphi \rangle$ in presence of a boundary. Note that the boundary term:

$$b(u, \varphi) \;=\; \sum_{\beta \in \partial X} \Big\langle u_\beta \;\Big|\; \sum_{\alpha \to \beta} \varphi_{\alpha\beta} - \sum_{\beta \to \gamma} \varphi_{\beta\gamma} \Big\rangle \tag{5.39}$$

does present a formal analogy with 5.37 despite the unusual «thickness» of the boundary ∂X. The analogy becomes clearer if one assumes $\nabla u = 0$ on ∂X as the above reduces to:

$$b(u, \varphi) \;=\; \sum_{\beta \in \partial X} \Big\langle u_\beta \;\Big|\; \sum_{\alpha \in \mathring{X}} \varphi_{\alpha\beta} \Big\rangle \tag{5.40}$$

by duality of ∇ with δ on $A_\bullet(\partial X)$, and represents the integral of the outbound flux of φ against u. $\square$

As proposition 5.26 suggests, the differential calculus of $\mathring{\delta}$ shall only differ from that of δ by the appearance of boundary terms representing energy fluxes leaving $\mathring{X}$ through ∂X. It will be useful to treate those seperately and we decompose $A_1(X)$ as the direct sum $A_1^{int}(X) \oplus A_1^{out}(X)$ according to:

Definition 5.27. *For every $\varphi \in A_1(X)$, we introduce the splitting $\varphi = \varphi^{int} + \varphi^{out}$ defined by:*

- $\varphi^{int}_{\alpha\beta} = \varphi_{\alpha\beta}$ *for every* $\beta \in \mathring{X}$,
- $\varphi^{out}_{\alpha\beta} = \varphi_{\alpha\beta}$ *for every* $\beta \in \partial X$.

We respectively call φ^{int} and φ^{out} the interior *and* outbound *components of φ.*

Note that $\delta : A_1(X) \to A_0(X)$ then takes a block-triangular form as illustrated by the diagram:

$$\begin{tikzcd}
A_1^{int}(X) \arrow[d] & \oplus & A_1^{out}(X) \arrow[dll] \arrow[d, dashed] \\
A_0(\mathring{X}) & \oplus & A_0(\partial X)
\end{tikzcd} \tag{5.41}$$

Full-line arrows represent the components of $\mathring{\delta}$ while the dotted arrow represents the truncation $\delta - \mathring{\delta}$. In particular, we have $\delta(\varphi^{int}) = \mathring{\delta}(\varphi^{int})$ while and the boundary term $b(u, \varphi) = b(u, \varphi^{out})$ of prop. 5.26 depends only on φ^{out}.

Proposition 5.28 (Gauss formula on $\Lambda^\alpha \smallsetminus \partial\Lambda^\alpha$)**.** *For every $\phi \in A_1(X)$ we have:*

$$\zeta(\mathring{\delta}\phi)_\alpha = \tilde{\zeta}(\phi^{int})_{\Omega \to \alpha} - \zeta(\phi^{out})_{\alpha \to \partial\alpha} \tag{5.42}$$

where $\tilde{\zeta}(\phi^{int})_{\Omega \to \alpha} = \displaystyle\sum_{\alpha'\beta' \in d\Lambda^\alpha} \phi^{int}_{\alpha'\beta'}$ denotes the total flux bound from $\mathring{X}$ to the interior of Λ^α.

Proof. By definition of $\mathring{\delta}$ and of the action of ζ on $A_0(X)$ we have:

$$\zeta(\mathring{\delta}\phi)_\alpha \;=\; \sum_{\beta' \in \Lambda^\alpha} \mathring{\delta}_{\beta'}\phi \;=\; \sum_{\beta' \in \Lambda^\alpha \smallsetminus \partial\Lambda^\alpha} \delta_{\beta'}\phi \;=\; \sum_{\beta' \in \Lambda^\alpha_{\partial\alpha}} \Big(\sum_{\alpha' \to \beta'} \phi_{\alpha'\beta'} - \sum_{\beta' \to \gamma'} \phi_{\beta'\gamma'} \Big) \tag{5.43}$$

Terms of the form $\phi_{\alpha'\beta'}$ with $\alpha' \in \Lambda^{\alpha}_{\partial\alpha}$ cancel out the $\phi_{\beta'\gamma'}$ with $\gamma' \in \Lambda^{\alpha}_{\partial\alpha}$ so that we may write:

$$\zeta(\mathring{\delta}\phi)_\alpha = \sum_{\alpha' \in \Lambda^{\Omega}_{\alpha}} \sum_{\beta' \in \Lambda^{\alpha}} \phi^{int}_{\alpha'\beta'} - \sum_{\beta' \in \Lambda^{\alpha}_{\partial\alpha}} \sum_{\gamma' \in \Lambda^{\partial\alpha}} \phi^{out}_{\beta'\gamma'} \tag{5.44}$$

Recognising zeta transforms on degree one, the outbound flux reads $\zeta(\phi^{out})_{\alpha\to\partial\alpha}$ and the inbound flux reads $\tilde{\zeta}(\phi^{int})_{\Omega\to\alpha}$ where ϕ^{int} is extended by zero to $\tilde{X} = \{\Omega\} \cup X$ and $\tilde{\zeta}$ acts on $A_1(\tilde{X})$. □

The zeta transforms in degree one appearing in proposition 5.28 retrospectively provide with a strong motivation for the higher degree combinatorics of chapter 3. They will also justify performing a Möbius inversion on the standard flux functional $\Phi = -\mathcal{D} \circ \zeta$.

5.3.2 Canonical Flux

We now introduce the homological vector field $\tau = \delta\phi$ which we claim to define a canonical diffusion on interaction potentials. Because the evolution of effective hamiltonians $U = \zeta \cdot u$ integrates the energy flux on cones, Gauss formulas applied $\phi = \mu \cdot \Phi$ will shed light on the necessity of performing a Möbius inversion in degree one on the standard diffusion flux[15].

Definition 5.29. *We call:*

- canonical diffusion flux *the smooth functional $\phi = \mu \circ (-\mathcal{D}) \circ \zeta$ defined from $A_0(X)$ to $A_1(X)$,*
- canonical diffusion *the smooth vector field $\tau = \delta\phi$ defined on $A_0(X)$.*

We shall write $\phi = -\mathcal{D}^\mu$ to emphasise on the conjugation of $-\mathcal{D}$ by the Möbius transform. More generally, the following notations will be useful in switching from one point of view to the other.

Definition 5.30. *For every smooth map $T : A_\bullet(X) \to A_\bullet(X)$ we denote by:*

- *$T^\zeta = \zeta \circ T \circ \mu$ the ζ-conjugate of T,*
- *$T^\mu = \mu \circ T \circ \zeta$ the μ-conjugate of T.*

The extensions of ζ and μ to all degrees really allow for two equivalent point of views on $A_\bullet(X)$, of which the associated conjugations seem to be an essential feature. Hence the two equivalent differential equations for the canonical diffusion τ on interaction potentials and its conjugate vector field τ^ζ inducing the evolution of effective hamiltonians:

$$\begin{cases} \dot{u} = \delta\phi \\ \phi = -\mathcal{D}^\mu(u) \end{cases} \quad \Leftrightarrow \quad \begin{cases} \dot{U} = \delta^\zeta(\Phi) \\ \Phi = -\mathcal{D}(U) \end{cases} \tag{5.45}$$

The following theorem best illustrates a first effect of Möbius inversion on the effective energy flux. This correction is necessary for the total flux bound into Λ^α to correctly approximate the global effective energy of $\Omega \setminus \Lambda^\alpha$. Theorem 5.31 also implies one-step convergence to local hamiltonians of $U \leftarrow U - \delta^\zeta\mathcal{D}(U)$ whenever the underlying hypergraph contains a maximal cell[16]. algorithm $u \leftarrow u - \delta\mathcal{D}^\mu(u)$ restricted to $A_0(\Lambda^\alpha)$ below any cell $\alpha \in X$.

Theorem 5.31. *The evolution of effective hamiltonians under $\dot{U} = -\delta^\zeta\mathcal{D}(U)$ reads:*

$$\dot{U}_\alpha = \check{\mathbb{F}}^\Omega(U \mid \alpha) - U_\alpha \tag{5.46}$$

where $\check{\mathbb{F}}^\Omega(U \mid \alpha) = \sum_\omega c_\omega \, \mathbb{F}^\omega(U_\omega \mid \omega \cap \alpha)$ denotes the Bethe approximation of $\mathbb{F}^\Omega(U_\Omega \mid \alpha)$ when X does not contain Ω, and is equal to the latter otherwise.

[15] We believe this combinatorial correction to be a significant improvement of the GBP algorithm of [32], however for the standard BP algorithm on graphs, one has $\phi = \mu \cdot \Phi = \Phi \mod \mathbb{R}_1(X)$ and the correction is useless.

[16] X may for instance contain Ω, but the more practical consequence holds for restrictions of the algorithm to cones $\Lambda^\alpha \subseteq X$, for instance when updating units asynchronously and independently of one another, see section 6.1.

Substituting the expression of the effective energy gradient $\mathcal{D}(U)$ for Φ, the theorem will come as a direct consequence of:

Proposition 5.32. *If $X \subseteq \mathcal{P}(\Omega)$ does not contain Ω, then for every $\Phi \in A_1(X)$ we have:*

$$\delta^\zeta(\Phi)_\alpha = \check{\Phi}_{\Omega\to\alpha} \tag{5.47}$$

where $\check{\Phi}_{\Omega\to\alpha} = \sum_{\omega \notin \Lambda^\alpha} c_\omega \, \Phi_{\omega\to\alpha\cap\omega}$ denotes the Bethe approximation of an expected flux $\Phi_{\Omega\to\alpha}$.

Proof. Letting $\phi = \mu \cdot \Phi$ we have $\delta^\zeta(\Phi) = \zeta(\delta\phi)$. Using notations of prop. 5.28, in absence of boundary the Gauss formula on Λ^α reads, following (3.55):

$$\delta^\zeta(\Phi)_\alpha = \tilde{\zeta}(\mu \cdot \Phi)_{\Omega\to\alpha} \tag{5.48}$$

Consider the Möbius transform $\tilde{\phi} = \tilde{\mu} \cdot \Phi$ of the natural extension of Φ to $\tilde{X} = \{\Omega\} \sqcup X$ by zero. Locality of $\tilde{\mu}$ implies that $\tilde{\phi}$ and ϕ coincide on X and only differ on terms of the form $\Omega \to \beta$ where ϕ vanishes. By Möbius inversion on $\tilde{X}$ we have $\tilde{\zeta}(\tilde{\phi})_{\Omega\to\alpha} = \Phi_{\Omega\to\alpha} = 0$ so that:

$$\tilde{\zeta}(\phi)_{\Omega\to\alpha} = -\tilde{\zeta}\big(\tilde{\phi} - \phi\big)_{\Omega\to\alpha} = -\zeta\big(i_\Omega(\tilde{\phi})\big)_\alpha \tag{5.49}$$

From the inductive construction of μ given by proposition 3.10 we have $i_\Omega(\tilde{\phi}) = \mu\big(\tilde{\nu}_\Omega(\Phi)\big)$ which is equivalent to $\zeta\big(i_\Omega(\tilde{\phi})\big) = \tilde{\nu}_\Omega(\Phi)$ by Möbius inversion on $A_0(X)$. Substituting the identity $c_\beta = -\tilde{\mu}_{\Omega\beta}$ given by proposition 3.3 into equation 3.41 defining $\tilde{\nu}_\Omega$, we finally get:

$$-\,\tilde{\nu}_\Omega(\Phi)_\alpha = -\sum_{\beta \in \Lambda^\Omega_\alpha} \tilde{\mu}_{\Omega\alpha} \, \Phi_{\beta\cap(\Omega\to\alpha)} = \sum_{\beta \notin \Lambda^\alpha} c_\beta \, \Phi_{\beta\to\beta\cap\alpha} = \check{\Phi}_{\Omega\to\alpha} \tag{5.50}$$

which gives the desired expression for $\tilde{\zeta}(\phi)_{\Omega\to\alpha} = \tilde{\zeta}(\mu \cdot \Phi)_{\Omega\to\alpha}$.

Note that if $\Omega \in X$, Möbius inversion in $A_1(X)$ would have simply given $\delta^\zeta(\Phi)_\alpha = \Phi_{\Omega\to\alpha}$. □

Another reason for performing Möbius inversion on the energy flux comes with the enforcement of Dirichlet boundary conditions on $A_0(\partial X)$. The outbound flux from Λ^α to its boundary $\partial\Lambda^\alpha = \Lambda^{\partial\alpha}$ now coincides with the effective energy of $\Lambda^\alpha \smallsetminus \partial\Lambda^\alpha$, as expressed by the following theorem. Its effect will be to ensure consistency of $U \in A_0(X)$ with the prescribed values on ∂X.

Theorem 5.33. *The evolution of effective hamiltonians under $\dot{U} = -\mathring{\delta}^\zeta \mathcal{D}(U)$ reads:*

$$\dot{U}_\alpha = \sum_{\omega\cap\alpha \notin \partial X} c_\omega \, \mathbb{F}^\Omega\big(U_\omega - U_{\omega\cap\alpha} \,\big|\, \omega \cap \alpha\big) \;-\; \mathbb{F}^\alpha\big(U_\alpha - U_{\partial\alpha} \big|\, \partial\,\alpha\big) \tag{5.51}$$

The following counterpart of the Gauss formula on $\Lambda^\alpha \smallsetminus \partial\Lambda^\alpha$ will again prove the theorem. It comes as an easy consequence of prop. 5.32 and the lemma below.

Proposition 5.34. *If X does not have a maximal element, then for every $\Phi \in A_1(X)$ we have:*

$$\mathring{\delta}^\zeta(\Phi)_\alpha = \check{\Phi}^{int}_{\Omega\to\alpha} - \Phi^{out}_{\alpha\to\partial\alpha} \tag{5.52}$$

where $\check{\Phi}^{int}_{\Omega\to\alpha} = \sum_{\omega \notin \Lambda^\alpha} c_\omega \, \Phi^{int}_{\omega\to\alpha\cap\omega}$ denotes the Bethe approximation of an expected flux $\Phi^{int}_{\Omega\to\alpha}$.

Lemma 5.35. *For every $\phi = \mu \cdot \Phi \in A_1(X)$, we have $\phi^{out} = \mu \cdot \Phi^{out}$ and $\phi^{int} = \mu \cdot \Phi^{int}$.*

Proof of proposition 5.34. Letting $\phi = \mu \cdot \Phi$, in accordance with lemma 5.35 the Gauss formula with boundary 5.28 gives the following expression for $\zeta(\mathring{\delta}\phi)_\alpha = \mathring{\delta}^\zeta(\Phi)_\alpha$:

$$\tilde{\zeta}(\phi^{int})_{\Omega\to\alpha} - \zeta(\phi^{out})_{\alpha\to\partial\alpha} = \tilde{\zeta}(\mu\cdot\Phi^{int})_{\Omega\to\alpha} - \zeta(\mu\cdot\Phi^{out})_{\alpha\to\partial\alpha} \tag{5.53}$$

The outbound flux reads $\Phi^{out}_{\alpha\to\partial\alpha}$ by Möbius inversion on $A_1(X)$, while it follows from proposition 5.32 applied to $\delta^\zeta(\Phi^{int})_\alpha$ that the inbound flux is the Bethe approximation $\check{\Phi}^{int}_{\Omega\to\alpha}$. □

Proof of lemma 5.35. For all $\alpha \in X$ and $\beta \in \partial X$, we have by definition of the action of μ on $A_1(X)$:

$$\phi^{out}_{\alpha\beta} = \sum_{\beta'\in\Lambda^\beta} \mu_{\beta\beta'} \sum_{\alpha'\in\Lambda^\alpha_{\beta'}} \mu_{\alpha\alpha'}\cdot\Phi_{\alpha'\to\alpha'\cap\beta'} = \big(\mu\cdot\Phi^{out}\big)_{\alpha\beta} \tag{5.54}$$

as $\beta \in \partial X$ implies $\alpha'\cap\beta' \in \partial X$ for all $\beta' \subseteq \beta$ and ϕ^{out} only depends on Φ^{out}. We may then conclude from $\phi^{out} = \mu\cdot\Phi^{out}$ that $\phi^{int} = \phi - \phi^{out}$ coincides with $\mu\cdot\Phi^{int} = \mu\cdot(\Phi - \Phi^{out})$ by linearity of μ. □

Note that the properness of ϕ follows from that of Φ by invertibility of μ, however we were only able to prove a local faithfulness property and the global faithfulness of ϕ will remain an open question.

Definition 5.36. *A flux functional $\phi : A_0(X) \to A_1(X)$ will be said* locally faithful *if there exists an open neighbourhood $\mathcal{V}$ of $\{\mathcal{D}\circ\zeta = 0\} \subseteq A_0(X)$ such that for all $u \in \mathcal{V}$:*

$$\delta\phi(u) = 0 \quad\Leftrightarrow\quad \mathcal{D}(\zeta\cdot u) = 0 \tag{5.55}$$

Proposition 5.37. *The flux functional $\phi = \mu\circ(-\mathcal{D})\circ\zeta$ is locally faithful.*

Proof. Let $u \in A_0(X)$ such that $\mathcal{D}(\zeta\cdot u) = 0$ denote a field of consistent interaction potentials. Writing $\Phi = -\mathcal{D}\circ\zeta$ as before, for every $V = \zeta\cdot v \in A_0(X)$ we have according to propositions 5.32 and 4.14:

$$\zeta\big(\delta\phi(u+v)\big)_\alpha = \check{\Phi}(u+v)_{\Omega\to\alpha} = \sum_{\omega\notin\Lambda^\alpha} c_\omega\, \mathbb{E}^\omega\big[\, V_\omega - V_{\omega\cap\alpha} \,\big|\, \omega\cap\alpha \,\big] + o\,(v) \tag{5.56}$$

where conditional expectations are taken for the consistent statistical field $p = [\mathrm{e}^{-U}]$ with $U = \zeta\cdot u$. Now note that although $p \in \mathring{\Gamma}(X)$ may not derive from a global probability density $p_\Omega \in \Delta_\Omega$, there does exist a global density $q_\Omega \in A^*_\Omega$ such that $p_\alpha = \Sigma^{\alpha\Omega}(q_\Omega)$ for all $\alpha \in X$ by acyclicity of $A^*_\bullet(X)$. We may thus define global «conditional expectation» maps by letting for all $\omega \to \beta$ in X:

$$\mathbb{E}^\Omega[\, V_\omega \mid \beta\,] = \frac{\Sigma^{\beta\Omega}(q_\Omega\cdot V_\omega)}{p_\beta} \tag{5.57}$$

such that $\mathbb{E}^\Omega[\, V_\omega \mid \beta\,]$ coincides with $\mathbb{E}^\omega[\, V_\omega \mid \beta\,]$ as a consequence of $\Sigma^{\omega\Omega}(q_\Omega) = p_\omega$. Observing that $V_{\omega\cap\alpha} = \zeta(v_{|\Lambda^\alpha})_\omega$ and letting $V_\Omega = \tilde{\zeta}(v)_\Omega$, the linearised right hand side of 5.56 now reads:

$$\mathbb{E}^\Omega\Big[\sum_{\omega\in X} c_\omega\,\big(V_\omega - V_{\omega\cap\alpha}\big)\,\Big|\,\alpha\,\Big] \;=\; \mathbb{E}^\Omega\big[\, V_\Omega - V_\alpha \,\big|\, \alpha\,\big] \tag{5.58}$$

Up to second order terms in v, we hence have $\delta\phi(u+v) \simeq 0$ if and only if $V_\alpha \simeq \mathbb{E}^\Omega[V_\Omega \,|\, \alpha]$ for all α, which implies $V_\beta \simeq \mathbb{E}^\alpha[V_\alpha \,|\, \beta]$ for all $\alpha \to \beta$. Hence $\delta\phi$ satisfies the linearised faithfulness condition:

$$\delta\phi(u+v) \simeq 0 \quad\Leftrightarrow\quad \mathcal{D}(U+V) \simeq 0 \tag{5.59}$$

and the tangent spaces of $\{\delta\phi = 0\}$ and $\{\mathcal{D}\circ\zeta = 0\}$ at u coincide. If one could show $\{\delta\phi = 0\}$ to be connected, the global faithfulness of ϕ would follow. □

Chapter 6

Geometry of Equilibria

In this chapter, we study the influence of the geometry of $X \subseteq \mathcal{P}(\Omega)$ on the stationary points of message-passing algorithms, as described by the intersection of the manifold $\mathcal{Z}(X)$ of consistent interaction potentials with homology classes of the form $[h] = h + \delta A_1(X)$.

We shall first extend a well-known uniqueness theorem on trees to a wider class of hypergraphs which we call *retractable*. This constructive procedure moreover the finite-time convergence of a message-passing scheme on such hypergraphs.

We then focus on the relation of consistent pseudo-marginals, obtained by message-passing, with the true marginals of the global distribution they seek to approximate. We introduce a canonical map $T : \mathcal{Z}(X) \to \mathcal{Z}(X)$ which stabilises $h \in \mathcal{Z}(X)$ if and only if h coincides with the true effective potentials induced by the global hamiltonian. Extending again what is known on trees, we then show that T is trivial on retractable hypergraphs, while providing with the simplest example of a graph such that T induces a non-trivial dynamical system.

As numerical studies on graphs have already shown, the number of stationary points grows quickly with the number of loops. We study the appearance of multiple equilibria in $\mathcal{Z}(X) \cap [h]$ through the singularities of the quotient map $\mathcal{Z}(X) \to \mathrm{H}(X)$, which we relate to the spectral properties of a pseudo-laplacian operator L describing the linearised diffusion flow, and finally provide with what we believe to be the first explicit examples of bifurcations on graphs.

6.1 Uniqueness and Retractability

In this section, we prove the uniqueness of a stationary state on a class of *retractable* hypergraphs which generalise trees. These hypergraphs also appear in the pseudo-marginal extension problem, to which Vorob'ev gave a criterion of solvability in [30], and are thus related to what he called *regular* complexes.

Precising first the sheaf structure of the manifold $\mathcal{Z}(X)$ of consistent potentials, we then exhibit an *extensibility* property of $\mathcal{Z}(X)$, on which the construction of the unique equilibrium by successive extension crucially relies.

Note that this section carries out proofs on a generic sheaf denoted by $\mathscr{F}(X)$, mainly purposed to represent $\mathcal{Z}(X)$ or one of its tangent fibers. Generic notations were intended so as to as to leave some space for an extension of the results to higher-degree analogs of $\mathcal{Z}(X)$, e.g. cocycles of $\mathrm{Ker}(\nabla^\mu)$.

6.1.1 Hypergraph Geometry and Sheaves

Definition 6.1. *A hypergraph X will be called a* simple extension *of $X' \subseteq X$ when:*

$$X = \Lambda^\alpha \sqcup_{\Lambda^\beta} X' \tag{6.1}$$

or equivalenlty, when $X = \Lambda^\alpha \cup X'$ and $X \cap X' = \Lambda^\beta$ for some $\alpha \in X$ and $\beta \in X'$.

Definition 6.2. *A hypergraph X is called a* normal extension *of X' if there exists a sequence:*

$$X = X_n \supseteq \ldots \supseteq X_0 = X' \tag{6.2}$$

of simple extensions from each X_i to X_{i+1}.

Definition 6.3. *A hypergraph X is called* retractable *if it is a normal extension of $\{\varnothing\}$.*

We chose instead to remain closer to the vocabulary of Vorob'ev, who called a *normal series* of X any sequence $X = X_n \supseteq \ldots \supseteq X_0$ of simple extensions.

Note that a graph X is a normal extension of X' if and only if it retracts onto X' in the usual sense. In particular, a graph X is retractable if and only if it is a tree, *i.e.* an acyclic graph.

Definition 6.4. *The* Alexandrov topology *of a hypergraph $X \subseteq \mathcal{P}(\Omega)$ is the topology generated by the basis of open neighbourhoods $(\Lambda^\alpha)_{\alpha \in X}$.*

Proposition 6.5. *$Y \subseteq X$ is open for the Alexandrov topology if and only if $\alpha \in Y$ implies $\Lambda^\alpha \subseteq Y$.*

Proposition 6.6. *Given a sub-hypergraph $Y \subseteq X$, the restriction map $\mathrm{r}_{YX} : A(X) \to A(Y)$ commutes with the zeta transforms of $A(X)$ and $A(Y)$ if and only if Y is open in X.*

$$\begin{tikzcd} A(X) \arrow[r, "\zeta_X"] \arrow[d, "\mathrm{r}_{YX}"'] & A(X) \arrow[d, "\mathrm{r}_{YX}"] \\ A(Y) \arrow[r, "\zeta_Y"] & A(Y) \end{tikzcd} \tag{6.3}$$

Proposition 6.7. *The set $\mathcal{Z}(X)$ of consistent interaction potentials forms a subsheaf of $A_0(X)$ for the Alexandrov topology.*

Proof. For every $u \in \mathcal{Z}(X)$ and $Y \subseteq X$, one has $u_{|Y} \in \mathcal{Z}(Y)$ as from $\mathcal{D}_Y(\zeta_Y \cdot u_{|Y}) = \mathcal{D}_X(\zeta_X \cdot u)_{|Y} = 0$ by proposition 6.6. The sheaf colimit property is satisfied by construction as $\mathcal{Z}(X) = \mathrm{colim}_\alpha\, \mathcal{Z}(\Lambda^\alpha)$. □

6.1.2 Cocyclic and Extensible Sheaves

Definition 6.8. *A subsheaf $\mathscr{F}(X)$ of $A(X)$ will be said* cocyclic *if:*

$$\forall u \in A(X) \quad \exists! v = u + \delta\varphi \in \mathscr{F}(X) \tag{6.4}$$

Equivalently, $\mathscr{F}(X)$ is cocyclic if it is a section of the fibration $A(X) \to A(X)/\delta A(X)$.

Letting $\nabla = \delta^*$ for a given metric, the space $\mathrm{Ker}(\nabla) \subseteq A(X)$ of cocycles defines, by orthogonality of $\mathrm{Ker}(\nabla)$ and $\mathrm{Im}(\delta)$, the fundamental example of a globally cocyclic subsheaf of vector spaces. This property is in turn equivalent to the uniqueness of equilibrium of the heat equation $\dot{u} = \delta(\nabla u)$ inside the homology class of any $h \in A_0(X)$.

Definition 6.9. *The subsheaf $\mathscr{F}(X)$ is said* locally cocyclic *if $\mathscr{F}(\Lambda^\alpha)$ is cocyclic for all $\alpha \in X$.*

Definition 6.10. *Given a locally cocyclic subsheaf $\mathscr{F}(X) \subseteq A(X)$, we denote by $T_\alpha : A(X) \to A(X)$ the map extending $T_\alpha : A(\Lambda^\alpha) \to \mathscr{F}(\Lambda^\alpha)$ by zero.*

Proposition 6.11. *The subsheaf $\mathcal{Z}(X)$ of consistent interaction potentials is locally cocyclic.*

Proof. Given $\alpha \in X$ and $u \in A(\Lambda^\alpha)$, let $U = \zeta(u)$ and let $V_\beta = \mathbb{F}^{\beta\alpha}(U_\alpha)$ for every $\beta \subseteq \alpha$. By consistency of V, the associated interaction potentials v are in $\mathcal{Z}(\Lambda^\alpha)$. One also has $v - u \in \delta A_1(\Lambda^\alpha)$ as defining for instance $\varphi \in A_1(\Lambda^\alpha)$ by:

$$\varphi_{\alpha\beta} = \sum_{\beta\to\gamma} \mu_{\beta\gamma}\, \mathbb{F}^{\gamma\alpha}(U_\alpha - U_\gamma) \tag{6.5}$$

and $\varphi_{\beta\gamma} = 0$ for all $\beta \neq \alpha$ yields $U + \zeta \cdot \delta\varphi = V$. [[T_α via one-step iteration of 5.3.2]]

Furthermore, assume $v = u + \delta A_1(\Lambda^\alpha)$ with $u, v \in \mathcal{Z}(\Lambda^\alpha)$. Denoting by $U, V \in A(\Lambda^\alpha)$ the associated local hamiltonians, the Gauss formula 2.3 implies $U_\alpha = V_\alpha$ while consistency gives $V_\beta = U_\beta = \mathbb{F}^{\beta\alpha}(U_\alpha)$ for every $\beta \in \Lambda^\alpha$. Hence u and v coincide on Λ^α by Möbius inversion. □

Belief propagation derives from the flux functional $\Phi(u) = \mathcal{D}(\zeta \cdot u)$ and we shall see that the sheaf $\mathcal{Z}(X) = \{\mathcal{D} \circ \zeta = 0\}$ of manifolds in $A_0(X)$ is in general only locally cocyclic. To prove the uniqueness of its equilibria when the hypergraph X is retractable, we have to show that the locally cocyclic sheaf $\mathcal{Z}(X)$ is then *globally* cocyclic. Our proof is constructive and relies on another property of $\mathcal{Z}(X)$, related to the extension of sections $u \in \mathcal{Z}(X')$ along normal extensions $X \supseteq X'$, under appropriate homological constraints.

Definition 6.12. *Given a subsheaf $\mathscr{F}(X)$ of $A(X)$ and $X' \subseteq X$, we introduce the subsets of $A(X)$:*

- $\mathscr{F}(X, X') = \{u \in A(X) \mid u_{|\Lambda^\alpha} \in \mathscr{F}(\Lambda^\alpha) \text{ for all } \alpha \in X \smallsetminus X'\}$
- $\mathscr{F}(X|X') = \{u \in \mathscr{F}(X, X') + A(X') \mid u_{|X'} \in \mathscr{F}(X')\}$

Note that according to this definition, $\mathscr{F}(X, X) = A(X)$ while $\mathscr{F}(X|X) = \mathscr{F}(X)$.

Definition 6.13. *A locally cocylic sheaf $\mathscr{F}(X)$ of $A(X)$ is said* extensible *if for every $\alpha \to \beta$ in X and every $u \in \mathscr{F}(\Lambda^\alpha|\Lambda^\beta)$, its unique homologous representative $v \in \mathscr{F}(\Lambda^\alpha)$ extends $u_{|\Lambda^\beta}$.*

$\mathscr{F}(X)$ is extensible if and only if the following commutative diagram is commutative for all $\alpha \supseteq \beta$:

$$\begin{tikzcd} \mathscr{F}(\Lambda^\alpha|\Lambda^\beta) \arrow[r, "T_\alpha"] \arrow[dr] & \mathscr{F}(\Lambda^\alpha) \arrow[d] \\ & \mathscr{F}(\Lambda^\beta) \end{tikzcd} \tag{6.6}$$

The subsheaf $\mathscr{F}(X|X')$ is by definition the inverse image of $\mathscr{F}(X') \subseteq A(X')$ under the restriction map $\mathscr{F}(X, X') + A(X') \to A(X')$. In the following commutative diagram, where all horizontal arrows are inclusions and all descending arrows are restrictions to X', the central square is hence a pull-back square.

$$\begin{tikzcd} \mathscr{F}(X) \arrow[r] \arrow[dr] & \mathscr{F}(X|X') \arrow[r] \arrow[d] & \mathscr{F}(X, X') + A(X') \arrow[r] \arrow[d] & A(X) \arrow[dl] \\ & \mathscr{F}(X') \arrow[r] & A(X') & \end{tikzcd} \tag{6.7}$$

As we shall see, whenever X is a normal extension of X' and $\mathscr{F}(X)$ is extensible, there exists reciprocal projections $f : A(X) \to \mathscr{F}(X, X')$ and $g : \mathscr{F}(X|X') \to \mathscr{F}(X)$ preserving the homological constraints.

Proposition 6.14. *The subsheaf $\mathcal{Z}(X)$ of consistent interaction potentials is extensible.*

Proof. Given $\alpha \supseteq \beta$ in X and $u \in \mathscr{Z}(\Lambda^\alpha|\Lambda^\beta)$, let us write $u = h + b$ with $h \in \mathscr{Z}(\Lambda^\alpha)$ and $b \in A(\Lambda^\beta)$ such that $u_{|\Lambda^\beta} \in \mathscr{Z}(\Lambda^\beta)$. Letting $H = \zeta(h)$ and $B = \zeta(b)$ and following the proof of 6.11, the unique homologous $v \in \mathscr{Z}(\Lambda^\alpha)$ is defined by Möbius inversion of the local hamiltonians (V_γ) defined for every $\gamma \in \Lambda^\alpha$ by:

$$V_\gamma = \mathbb{F}^{\gamma\alpha}(H_\alpha + B_\beta) \tag{6.8}$$

For all $\beta \subseteq \alpha$, we have by consistency of H on Λ^α and by consistency of $U = H + B$ on Λ^β:

$$V_\gamma = \mathbb{F}^{\gamma\beta}\Big(\mathbb{F}^{\beta\alpha}(H_\alpha) + B_\beta\Big) = \mathbb{F}^{\gamma\beta}(H_\beta + B_\beta) = H_\gamma + B_\gamma = U_\gamma \tag{6.9}$$

It follows that $v_{|\Lambda^\beta} = u_{|\Lambda^\beta}$ by 6.6, implying commutativity of restrictions with Möbius inversions. □

6.1.3 Forward and Backward Passes

Theorem 6.15. *Assume X is a normal extension of X' and $\mathscr{F}(X) \subseteq A(X)$ is an extensible subsheaf. Then every $u \in \mathscr{F}(X|X')$ admits a unique homologous representative $v \in \mathscr{F}(X)$, which extends $u_{|X'}$.*

Theorem 6.16. *When X is retractable, any extensible subsheaf $\mathscr{F}(X) \subseteq A(X)$ is globally cocyclic.*

Proposition 6.17. *Given a normal sequence $X = X_n \supseteq \dots \supseteq X_0 = X'$ with $X_j = \Lambda^{\alpha_j} \sqcup_{\Lambda^{\beta_j}} X_{j-1}$ for all $n \geq j \geq 1$ and an extensible subsheaf $\mathscr{F}(X) \subseteq A(X)$, the* double pass*:*

$$f = T_{\alpha_n} \circ \dots \circ T_{\alpha_1} \circ \dots \circ T_{\alpha_n} \tag{6.10}$$

maps $A(X)$ to the subspace $\mathscr{F}(X, X')$ formed by those u such that $u_{|\Lambda^\alpha} \in \mathscr{F}(\Lambda^\alpha)$ for all $\alpha \in X \setminus X'$.

Proof. Reasoning by induction on the length of the normal sequence, first consider $X_1 = \Lambda^{\alpha_1} \sqcup_{\Lambda^{\beta_1}} X'$. Given $u \in A(X_1)$, let $v = T_{\alpha_1}(u)$ so that $v_{|\Lambda^{\alpha_1}} \in \mathscr{F}(\Lambda^{\alpha_1})$ by definition of T_{α_1}, and $v \in \mathscr{F}(X_1, X')$. Assume now that $f_{j-1} = T_{\alpha_{j-1}} \dots T_{\alpha_1} \dots T_{\alpha_{j-1}}$ does induce a map from $A(X_{j-1})$ to $\mathscr{F}(X_{j-1}, X')$ for some $1 < j \leq n$, and let $u \in A(X_j)$.

- *forward pass:* letting $\tilde{u} = T_{\alpha_j}(u)$, we have $\tilde{u}_{|\Lambda^{\alpha_j}} \in \mathscr{F}(\Lambda^{\alpha_j})$.
- *induction:* letting $\tilde{v} = f_{j-1}(\tilde{u})$, we have $\tilde{v}_{|X_{j-1}} \in \mathscr{F}(X_{j-1}, X')$. In particular $\tilde{v}_{|\Lambda^{\beta_j}} \in \mathscr{F}(\Lambda^{\beta_j})$ while $\tilde{u}_{|\Lambda^{\alpha_j}} \in \mathscr{F}(\Lambda^{\alpha_j})$ implies $\tilde{v}_{|\Lambda^{\alpha_j}} \in \mathscr{F}(\Lambda^{\alpha_j}|\Lambda^{\beta_j})$.
- *backward pass:* letting $v = T_{\alpha_j}(\tilde{v})$ we have $v_{|\Lambda^{\alpha_j}} \in \mathscr{F}(\Lambda^{\alpha_j})$ extending $\tilde{v}_{|\Lambda^{\beta_j}}$ by extensibility.

Hence $T_{\alpha_j} \dots T_{\alpha_1} \dots T_{\alpha_j}(u)$ lies in $\mathscr{F}(X_j, X')$ for all $1 \leq j \leq n$ and f maps $A(X)$ to $\mathscr{F}(X, X')$. □

Proposition 6.18. *Given a normal sequence $X = X_n \supseteq \dots \supseteq X_0 = X'$ with $X_j = \Lambda^{\alpha_j} \sqcup_{\Lambda^{\beta_j}} X_{j-1}$ for all $n \geq j \geq 1$ and an extensible subsheaf $\mathscr{F}(X) \subseteq A(X)$, the* backward pass*:*

$$g = T_{\alpha_n} \circ \dots \circ T_{\alpha_1} \tag{6.11}$$

maps $\mathscr{F}(X|X')$ to $\mathscr{F}(X)$.

Proof. Assume first that $u \in \mathscr{F}(X_1|X')$, which is equivalent to $u_{|\Lambda^{\alpha_1}} \in \mathscr{F}(\Lambda^{\alpha_1}|\Lambda^{\beta_1})$ and $u_{|X'} \in \mathscr{F}(X')$. The extensibility property ensures that $v_{|\Lambda^{\alpha_1}}$ extends $u_{|\Lambda^{\beta_1}}$ so that v extends $u_{|X'}$ and $v \in \mathscr{F}(X_1)$. Assuming now that $g_{j-1} = T_{\alpha_{j-1}} \dots T_{\alpha_1}$ maps $\mathscr{F}(X_{j-1}|X')$ to $\mathscr{F}(X_{j-1})$ for some $1 < j \leq n$, let $u \in \mathscr{F}(X_j|X')$.

- *induction:* letting $\tilde{v} = g_{j-1}(u)$ we have $\tilde{v}_{|X_{j-1}} \in \mathscr{F}(X_{j-1})$. In particular $\tilde{v}_{|\Lambda^{\beta_j}} \in \mathscr{F}(\Lambda^{\beta_j})$, so that $\tilde{v}_{|\Lambda^{\alpha_j}} \in u_{|\Lambda^{\alpha_j}} + A(\Lambda^{\beta_j})$ and $u_{|\Lambda^{\alpha_j}} \in \mathscr{F}(\Lambda^{\alpha_j}) + A(\Lambda^{\beta_j})$ implies $\tilde{v}_{|\Lambda^{\alpha_j}} \in \mathscr{F}(\Lambda^{\alpha_j}|\Lambda^{\beta_j})$.
- *backward pass:* letting $v = T_{\alpha_j}(\tilde{v})$ we have $v_{|\Lambda^{\alpha_j}} \in \mathscr{F}(\Lambda^{\alpha_j})$ extending $\tilde{v}_{|\Lambda^{\beta_j}}$ by extensibility.

Hence $T_{\alpha_j} \dots T_{\alpha_1}(u)$ lies in $\mathscr{F}(X_j)$ for all $1 \leq j \leq n$ and g maps $\mathscr{F}(X|X')$ to $\mathscr{F}(X)$. □

Proposition 6.19. *Given a normal extension $X \supseteq X'$ and an extensible subsheaf $\mathscr{F}(X) \subseteq A(X)$, the double pass and backward pass maps:*

$$f : A(X) \to \mathscr{F}(X, X') \qquad \text{and} \qquad g : \mathscr{F}(X|X') \to \mathscr{F}(X) \tag{6.12}$$

are uniquely defined, under the constraint that the image of u lies in $u + \delta A(X, X')$.

Proof. Reasoning by induction on the length of a normal sequence, first assume that $X = \Lambda^\alpha \sqcup_{\Lambda^\beta} X'$. If $u, v \in \mathscr{F}(X, X')$ with $v = u + \delta\varphi$ and $\varphi \in A(\Lambda^\alpha)$, then $v_{|\Lambda^\alpha} = u_{|\Lambda^\alpha}$ by cocyclicity of $\mathscr{F}(\Lambda^\alpha)$ and $\delta\varphi = 0$. Assume now that $X = \Lambda^\alpha \sqcup_{\Lambda^\beta} X_1$ with $X_1 \supseteq X'$, and that for all $u, v \in \mathscr{F}(X_1, X')$, if v lies in $u + \delta A(X_1, X')$ then $v = u$. Given $u, v \in \mathscr{F}(X, X')$ with $v = u + \delta\varphi$ and $\varphi \in A(X, X')$, let us write $\delta\varphi = \delta\varphi^+ + \delta(\varphi_{|\Lambda^\beta}) + \delta\varphi^-$ with $\varphi^+ \in A(\Lambda^\alpha)$ and $\varphi^- \in A(X_1, X')$.

- *no forward pass:* letting $\tilde{v} = T_\beta(u + \delta\varphi^-)$, we have $\tilde{v}_{|\Lambda^\alpha} \in \mathscr{F}(\Lambda^\alpha) + \delta A(\Lambda^\beta)$ and $u_{|\Lambda^\beta} \in \mathscr{F}(\Lambda^\beta)$ so that $\tilde{v}_{|\Lambda^\alpha} \in \mathscr{F}(\Lambda^\alpha|\Lambda^\beta)$. Extensibility implies that $v = T_\alpha(u + \delta\varphi^-) = T_\alpha(\tilde{v})$ extends $\tilde{v}_{|\Lambda^\beta}$.
- *induction:* as $v_{|X_1} = \tilde{v}_{|X_1}$ lies in $u_{|X_1} + \delta A(X_1, X')$ we have $v_{|X_1} = u_{|X_1}$ by induction hypothesis.
- *no backward pass:* in particular $\delta(\varphi)_{|X_1} = 0$ so that $v_{|\Lambda^\alpha} = u_{|\Lambda^\alpha} + \delta\varphi^+$ lies in $u_{|\Lambda^\alpha} + \delta A(\Lambda^\alpha)$, and $v_{|\Lambda^\alpha} = u_{|\Lambda^\alpha}$ by local cocyclicity.

Hence $v = u$ for every $u, v \in \mathscr{F}(X, X')$ such that $v \in u + \delta A(X, X')$. It follows that the double pass map $f : A(X) \to \mathscr{F}(X, X')$ associated to any given normal sequence is independent of the sequence. Uniqueness of the backward pass map $g : \mathscr{F}(X|X') \to \mathscr{F}(X)$ is also obtained by applying the same uniqueness argument to any $u, v \in \mathscr{F}(X) \subseteq \mathscr{F}(X, X')$ such that $v \in u + \delta A(X, X')$. □

6.2 Relations with Global Sections

In this section, we show that the unique equilibrium on retractable hypergraphs coincides with the true Gibbs state marginals one seeks to approximates.

In general, we introduce a *global pass* map $T : A_0(X) \to \mathcal{Z}(X)$ defined by the effective energies of the global hamiltonian and their associated interaction potentials and investigate some of its homological properties. As $T(h)$ may not be homologous to h in general, we signal that the true Gibbs state marginals may not be accessible from message-passing schemes. After showing that T coincides with the double pass maps of the previous section on retractable hypergraphs, we give an explicit example where T induces a non-trivial iteration over the manifold of equilibria.

6.2.1 Global Pass

Definition 6.20. *We call* global completion *of $X \subseteq \mathcal{P}(\Omega)$ the hypergraph $\tilde{X} = X \cup \{\Omega\}$.*

Definition 6.21. *Given a global hamiltonian $H_\Omega = \sum_{\alpha \in X} h_\alpha$ in A_Ω, let us define:*

- *the* true effective hamiltonians *by $H^*_\alpha = \mathbb{F}^{\alpha\Omega}(H_\Omega)$ for all $\alpha \in \tilde{X}$,*
- *the* true effective potentials *by $h^* = \mu \cdot H^*$,*
- *the* true Gibbs state marginals *by $p = [\mathrm{e}^{-H^*}]$,*

each of which being consistent, as $\mathcal{D}(H^*) = 0$ implies $h^* \in \mathcal{Z}(\tilde{X})$ and $p \in \Gamma(\tilde{X})$.

Definition 6.22. *We call* global pass *$T : A_0(X) \to A_0(X)$ the smooth map associating to potentials $h \in A_0(X)$ the true effective potentials $h^* \in \mathcal{Z}(X)$ deriving from the total energy $H_\Omega = \sum_\alpha h_\alpha$.*

The global pass resorts to the global hamiltonian and should be thought of as an extrinsic action on $A_0(X)$, uncomputable in practice. The following diagram best illustrates the procedure:

$$\begin{tikzcd} & A_\Omega \arrow[dr, "\mu \circ \mathbb{F}^{-\Omega}"] & \\ A_0(X) \arrow[ur, "\zeta_\Omega"] \arrow[rr, dashed, "T"] & & A_0(X) \end{tikzcd} \tag{6.13}$$

It is good to view T as a map from $A_0(X)$ to itself as we shall see its iterations may not be trivial. Its image is however contained in $\mathcal{Z}(X)$ by definition, while it induces a map in homology:

$$\bar{T} : \mathrm{H}_0(X; A) \longrightarrow \mathcal{Z}(X) \tag{6.14}$$

Whether this map is a diffeomorphism or not is a natural and important question, closely related to the uniqueness or multiplicity of message-passing equilibria. We shall see its answer to depend on the geometry of the underlying hypergraph.

Proposition 6.23. *Assume $\tilde{h} \in \mathcal{Z}(\tilde{X})$ extends $h \in \mathcal{Z}(X)$. Then $\tilde{h}_\Omega = 0$ implies $T(h) = h$.*

Proof. Letting $H_\Omega = \sum_\alpha h_\alpha$ and $\tilde{H} = \tilde{\zeta} \cdot \tilde{h}$, then $\tilde{h}_\Omega = 0$ implies $\tilde{H}_\Omega = H_\Omega$. If furthermore $\tilde{h} \in \mathcal{Z}(\tilde{X})$, then consistency implies that $\tilde{H}$ coincides with the true effective hamiltonians H^* and $T(h) = h$. □

Proposition 6.24. *Given $h \in \mathcal{Z}(X)$ let $\tilde{h}^* = \tilde{T}(h) \in \mathcal{Z}(\tilde{X})$. Then $T(h) = h$ implies $\tilde{h}^*_\Omega = 0$.*

Proof. An immediate consequence of the following proposition. □

Remarks. Due to the apparent reciprocity between propositions 6.23 and 6.24, one may be tempted to close an implication loop and it seems important to mention that:

- $(T(h^*) = h^*) \not\Rightarrow (\tilde{h}^*_\Omega = 0)$?
  Although $T(h^*) = T(h)$ gives $\mathbb{F}^{\alpha\Omega}(H_\Omega) = \mathbb{F}^{\alpha\Omega}(H^*_\Omega)$ for all $\alpha \in X$, this may not imply $H^*_\Omega = H_\Omega$. We do have a global interaction decomposition $A_\Omega = Z_\Omega \oplus B_\Omega$, however $H_\Omega \in B_\Omega$ does not imply $\mathrm{e}^{-H_\Omega} \in B_\Omega$ in general. Hence $\Sigma^{\alpha\Omega}(\mathrm{e}^{-H_\Omega}) = \Sigma^{\alpha\Omega}(\mathrm{e}^{-H^*_\Omega})$ for all $\alpha \in X$ does not imply $\tilde{h}^*_\Omega = 0$.
- $(\tilde{h}^*_\Omega = 0) \not\Rightarrow (T(h) = h)$?
  Even when $h \in \mathcal{Z}(X)$ and its extension by zero $\tilde{h}$ lies in $\mathcal{Z}(\tilde{X})$, assuming h^* homologous to h in $A_0(X)$ does not imply $h^* = h$, as there may exist multiple $u = h + \delta\varphi \in \mathcal{Z}(X)$.

Proposition 6.25. *Given $h \in \mathcal{Z}(X)$, let $\tilde{h}^* = \tilde{T}(h)$ and let $h^* = \tilde{h}^*_{|X}$.*

$$\tilde{h}^*_\Omega = \sum_{\alpha \in X} h_\alpha - \sum_{\alpha \in X} h^*_\alpha \tag{6.15}$$

Proof. Letting $H_\Omega = \sum_\alpha h_\alpha$ and $\tilde{H}^* = \tilde{\zeta} \cdot \tilde{h}^*$, we have $\tilde{H}^*_\Omega = H_\Omega$ by definition of the true effective hamiltonians. Isolating the contribution of Ω to the total energy then gives $H_\Omega = \tilde{h}^*_\Omega + \sum_\alpha h^*_\alpha$. □

6.2.2 Retractable Hypergraphs

Theorem 6.26. *Assume $X \subseteq \mathcal{P}(\Omega)$ is retractable. For every collection of interaction potentials $h \in A_0(X)$ the unique $u = h + \delta\varphi \in \mathcal{Z}(X)$ coincides with the true effective potentials $h^* = T(h)$.*

Equivalently, the theorem asserts that for any $h \in A_0(X)$, extending the unique homologous potentials $u = h + \delta\varphi \in \mathcal{Z}(X)$ given by theorem 6.16 to $\tilde{X} = X \sqcup \{\Omega\}$ does yield a section $\tilde{u} \in \mathcal{Z}(\tilde{X})$. It must then coincide with the unique $\tilde{h}^* = \tilde{h} + \delta\tilde{\varphi} \in \mathcal{Z}(\tilde{X})$ defined from $\tilde{h} \in A_0(\tilde{X})$ extending h, and of which the global pass $h^* \in \mathcal{Z}(X)$ is the restriction:

$$\begin{array}{ccc} A_0(\tilde{X}) & \xrightarrow{\tilde{T}} & \mathcal{Z}(\tilde{X}) \\ \uparrow & & \uparrow \\ A_0(X) & \xrightarrow{f} & \mathcal{Z}(X) \end{array} \tag{6.16}$$

The existence of the right ascending arrow hence implies commutativity in the above diagram, and the proof of the theorem will rely on the following proposition.

Proposition 6.27. *Given $\Omega = \alpha \sqcup_\beta \alpha'$, assume $X = \Lambda^\alpha \sqcup_{\Lambda^\beta} \Lambda^{\alpha'}$ and let $\tilde{X} = X \sqcup \{\Omega\}$. Then:*

$$\mathcal{Z}(\tilde{X}) \cap A_0(X) \simeq \mathcal{Z}(X) \tag{6.17}$$

Equivalently, if $\tilde{U} = \tilde{\zeta} \cdot u \in A_0(\tilde{X})$ with $u \in A_0(X)$, then $\mathcal{D}(\tilde{U}) = 0$ if and only if $\mathcal{D}(\tilde{U})_{|X} = 0$.

Proof. Given any $\tilde{U} \in A_0(\tilde{X})$, first observe that $\mathcal{D}(\tilde{U}) = 0$ is equivalent to $U_\gamma = \mathbb{F}^{\gamma\Omega}(U_\Omega)$ for all $\gamma \in X$. The assumption $X = \Lambda^\alpha \cup \Lambda^{\alpha'}$ implies that $\mathbb{F}^{\gamma\Omega}$ may be factorised either as $\mathbb{F}^{\gamma\alpha} \circ \mathbb{F}^{\alpha\Omega}$ or as $\mathbb{F}^{\gamma\alpha'} \circ \mathbb{F}^{\alpha'\Omega}$ depending on whether $\gamma \in \Lambda^\alpha$ or $\gamma \in \Lambda^{\alpha'}$, so that $\mathcal{D}(\tilde{U})$ is actually equivalent to:

$$\mathcal{D}(\tilde{U})_{\Omega\to\alpha} = 0 \quad \text{and} \quad \mathcal{D}(\tilde{U})_{\Omega\to\alpha'} = 0 \quad \text{and} \quad \mathcal{D}(\tilde{U})_{|X} = 0 \tag{6.18}$$

Assume now that $\tilde{U} = \tilde{\zeta} \cdot u$ with $u \in A_0(X)$. Observing that $\Omega \smallsetminus \alpha = \alpha' \smallsetminus \beta$, applying proposition 4.10 to the effective energy of $\tilde{U}_\Omega - \tilde{U}_\alpha = \tilde{U}_{\alpha'} - \tilde{U}_\beta$ which lies in $A_{\alpha'}$ gives:

$$\mathcal{D}(\tilde{U})_{\Omega\to\alpha} = \mathbb{F}^{\alpha\Omega}\Bigg(\sum_{\gamma \in X \smallsetminus \Lambda^\alpha} u_\gamma\Bigg) = \mathbb{F}^{\beta\alpha'}\Bigg(\sum_{\gamma \in \Lambda^{\alpha'} \smallsetminus \Lambda^\beta} u_\gamma\Bigg) = \mathcal{D}(\tilde{U})_{\alpha'\to\beta} \tag{6.19}$$

Interverting α and α' similarly yields $\mathcal{D}(\tilde{U})_{\Omega\to\alpha'} = \mathcal{D}(\tilde{U})_{\alpha\to\beta}$. For all $\tilde{U} \in \tilde{\zeta} \cdot A_0(X)$ we hence have $\mathcal{D}(\tilde{U}) = 0$ if and only if $\mathcal{D}(\tilde{U})_{|X} = 0$, and it follows that the image of $\mathcal{Z}(X)$ in $A_0(\tilde{X})$ by the natural inclusion $A_0(X) \subseteq A_0(\tilde{X})$ coincides with the intersection of $\mathcal{Z}(\tilde{X})$ with $A_0(X)$. ☐

Proof of theorem 6.26. Reasoning by induction on the length of X, the statement is tautological when X already contains Ω. Given a normal extension $X = \Lambda^\alpha \sqcup_{\Lambda^\beta} X'$ with $X' \subseteq \mathcal{P}(\Omega')$, assume now that for every $u' \in \mathcal{Z}(X')$ its extension $\tilde{u}'$ to $\tilde{X}' = X' \cup \{\Omega'\}$ lies in $\mathcal{Z}(\tilde{X}')$. Consider the hypergraphs:

$$Y = X \cup \{\Omega'\} \qquad \text{and} \qquad \tilde{Y} = Y \cup \{\Omega\} \tag{6.20}$$

Given any $u \in \mathcal{Z}(X)$, let us show that its extension $\tilde{u}$ to $\tilde{X}$ by zero lies in $\mathcal{Z}(\tilde{X})$, by applying the previous proposition in $\tilde{Y}$, global completion of $Y = \Lambda^\alpha \sqcup_{\Lambda^\beta} \Lambda^{\Omega'}$ with $\Lambda^{\Omega'} = \tilde{X}'$.

- $u' = u_{|X'} \in \mathcal{Z}(X')$ extends to $\tilde{u}' \in \mathcal{Z}(\tilde{X}')$ by induction hypothesis,
- $v \in \mathcal{Z}(Y)$ defined from $u_{|\Lambda^\alpha} \in \mathcal{Z}(\Lambda^\alpha)$ and $\tilde{u}' \in \mathcal{Z}(\Lambda^{\Omega'})$ extends to $\tilde{v} \in \mathcal{Z}(\tilde{Y})$ by proposition 6.27,
- $\tilde{v} \in \mathcal{Z}(\tilde{Y})$ restricts to $\tilde{u} \in \mathcal{Z}(\tilde{X})$ as $\mathcal{D}(\tilde{V}) = 0$ implies $\mathcal{D}(\tilde{V})_{|\tilde{X}} = \mathcal{D}(\tilde{U}) = 0$, where $\tilde{V}$ and $\tilde{U}$ denote their zeta transforms in $A_0(\tilde{Y})$ and $A_0(\tilde{X})$ respectively.

The induction shows that for every $h \in A_0(X)$, its unique homologous representative $u = f(h) \in \mathcal{Z}(X)$ defined by double pass extends to $\tilde{u} \in \mathcal{Z}(\tilde{X})$. The global pass depending only on the total energy, we have $T(u) = T(h) = h^*$, while $\tilde{u} \in \mathcal{Z}(\tilde{X})$ with $\tilde{u}_\Omega = 0$ implies $T(u) = u$ according to 6.23. ☐

6.2.3 Loop Effective Potential

The true effective potentials h^* are not homologous to h in general.

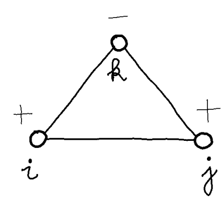


Figure 6.1: A triangle of binary variables.

The simplest example of this phenomenon is given by the the Ising model on the triangular graph with binary variables $x_i = \pm 1$ on each vertex, as depicted in figure 6.1. Consider a uniform coupling constant w and no external field for simplicity, so that interaction potentials are given by:

$$h_{ij}(x_i, x_j) = -w\, x_i\, x_j \qquad \text{and} \qquad h_i(x_i) = 0 \tag{6.21}$$

Letting $H = \zeta \cdot h$ denote the local hamiltonians and $q = [\mathrm{e}^{-H}]$, the local Gibbs states $q \in \Gamma(X)$ are consistent for any value of w. Considering that $A_i = \mathbb{R}\,|{+}\rangle \oplus \mathbb{R}\,|{-}\rangle$ and $A_{ij} = A_i \otimes A_j$, the following matrix representations[1] will be convenient:

$$q_{ij} = \frac{1}{4\cosh(w)} \begin{bmatrix} \mathrm{e}^{w} & \mathrm{e}^{-w} \\ \mathrm{e}^{-w} & \mathrm{e}^{w} \end{bmatrix} \qquad \text{and} \qquad q_i = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{6.22}$$

Writing $q = \mathcal{Q}(w)$, this defines a one-parameter family $\mathcal{Q} : \mathbb{R} \to \Gamma(X)$ of consistent statistical fields.

Consider now the true effective hamiltonians $H^* \in A_0(X)$ defined from the total energy $H_\Omega = \sum h_{ij}$. In addition to the direct coupling h_{ij}, the true effective hamiltonian H^*_{ij} accounts for the interaction of i and j through k, expressed by the effective energy of $h_{ik} + h_{kj}$.

$$\big(H^*_{ij} - H_{ij}\big)(x_i, x_j) = -\ln \sum_{x_k \in \{\pm 1\}} \mathrm{e}^{-w\, x_k (x_i + x_j)} \tag{6.23}$$

Introducing the real function $f : y \mapsto \ln\big(\mathrm{e}^{y} + \mathrm{e}^{-y}\big)$, which satisfies $f(-y) = f(y)$, we have:

$$H^*_{ij} - H_{ij} = - \begin{bmatrix} f(2w) & f(0) \\ f(0) & f(2w) \end{bmatrix} \tag{6.24}$$

Letting then $g(w) = \frac{1}{2}\big(f(2w) - f(0)\big)$ we may write up to the constant term $\frac{1}{2}\big(f(2w) + f(0)\big)$:

$$H^*_{ij} - H_{ij} \;\simeq\; g(w) \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \quad \mathrm{mod}\ \mathbb{R} \tag{6.25}$$

It follows that $H^*_{ij} - H_{ij}$ is a non-zero element of $Z_{ij} \oplus \mathbb{R}$ for every $w \neq 0$ and does not lie in $A_i + A_j$. This in particular implies that $H^* - H$ cannot belong to $\zeta \cdot \delta A_1(X)$, equivalently that the true effective potentials h^* are not homologous to h in $A_0(X)$. A loop effective potential $\tilde{h}^*_\Omega = \tilde{h}^*_{ijk}$ measures this obstruction, it may be computed using proposition 6.25. Up to a constant free energy term, $\tilde{h}^*_{ijk}$ is the sum of the contributions $H_{ij} - H^*_{ij}$ of each edge.

[1] Note that the ring structure of A_{ij} is induced by the element-wise or *Hadamard* product, while the matrix-vector product of q_{ij} with $f_j \in A_j$ yields the conditional expectation $\mathbb{E}^{ij}[\, f_j \,|\, i \,] \in A_i$.

The true local Gibbs states $p \in \Gamma(X)$ remain in the-one parameter family as $p = \mathcal{Q}(w')$ is given by:

$$p_{ij} = \frac{1}{4\cosh(w')} \begin{bmatrix} \mathrm{e}^{w'} & \mathrm{e}^{-w'} \\ \mathrm{e}^{-w'} & \mathrm{e}^{w'} \end{bmatrix} \qquad \text{with} \qquad w' = w + g(w) \tag{6.26}$$

Iterating the global pass $T : A_0(X) \to A_0(X)$ hence defines sequences of couplings $w \in \mathbb{R}^{\mathbb{N}}$ by the recurrence relation:

$$w_{n+1} = w_n + g(w_n) \tag{6.27}$$

and as $g(w) = \frac{1}{2}\ln\cosh(2w)$ behaves like a soft absolute value, with $|g'| < 1$, one quickly sees that:

$$\lim_n w_n = \begin{cases} 0 & \text{if } w \leq 0 \\ +\infty & \text{otherwise} \end{cases} \tag{6.28}$$

This should be interpreted by observing that the effective contribution of $h_{ik} + h_{kj}$ is always ferromagnetic, whatever the sign of w, as the path (i, k, j) from i to j is of even length.

A symmetric behaviour should be expected for the Ising model on the square graph, while in this case a finer description should also keep track of order 3 interaction potentials h^*_{ijk}, by embedding the graph inside the 3-simplex. Finally, although allowing weights w_{ij} to differ from edge to edge would induce a more interesting dynamic on a higher dimensional space of parameters, manual computations may quickly become cumbersome and call for numerical simulations.

6.3 Bifurcations and Singularities

In an effort to witness the emergence of multiple equilibria, this section characterises bifurcation states as consistent interaction potentials u whose tangent fiber $T_u\mathcal{Z}(X)$ is not transverse to $\delta A_1(X)$. These bifurcations correspond to singularities of the homological projection $\mathcal{Z}(X) \to \mathrm{H}(X)$, which we relate to spectral properties of a *twisted laplacian*[2] operator $L = \delta\nabla^\mu$ describing the linearised diffusion flow, we finally exhibit explicit examples of such singularities on simple graphs with two loops.

6.3.1 The Transversality Problem

At the level of local hamiltonians, the tangent fibers of the consistent manifold $\mathcal{C} = \zeta \cdot \mathcal{Z}$ bear a convenient probabilistic and cohomological description, involving the linearised effective energy gradient $\nabla = \mathcal{D}_*$ defined in 5.7:

Proposition 6.28. *Let $U \in \mathcal{C}$ and $p = [\mathrm{e}^{-U}]$. Then $T_U\mathcal{C} = \mathrm{Ker}(\nabla_p)$ is defined by the equations:*

$$\nabla_p(V)_{\alpha\beta} = V_\beta - \mathbb{E}^\alpha_{p_\alpha}\left[\, V_\alpha \,\middle|\, \beta \,\right] = 0 \tag{6.29}$$

Let us denote by $\bar{0} = \mu \cdot \ln[1]$ the origin of $\mathcal{Z}$, projection of $0 \in \mathcal{Z}'$ onto $\mathcal{Z}$ along $\mathbb{R}_0$.

Theorem 6.29. *The tangent space $Z_0 = T_{\bar{0}}\mathcal{Z}$ is globally cocyclic:*

$$A_0 = Z_0 \oplus \delta A_1 \tag{6.30}$$

The homological projection $P : A_0 \to Z_0$ is moreover explicitly given by interaction decomposition.

[2] Although the present problem does share many similiarities with classical harmonic theory, let us already emphasise that, in addition to the non-linearity of $\delta\mathcal{D}^\mu$, the conjugation by combinatorial transforms in its linearised flow breaks the adjunction of the boundary δ with the differential $\nabla^\mu = \mu \circ \nabla \circ \zeta$.

Proof. Letting $\varphi_{\beta\gamma} = \mathbf{z}_\gamma(u_\beta)$ we have by interaction decomposition:

$$(u + \delta\varphi)_\beta = \sum_{\alpha' \supseteq \beta} \mathbf{z}_\beta(u_{\alpha'}) \tag{6.31}$$

Hence $v = u + \delta\varphi$ defines the unique collection of interaction potentials in $T_0^{-}\mathcal{Z}$ homologous to $u \in A_0$. The map $P : u \mapsto v$ coincides with the interaction projection of [[2.3.3]]. ☐

When the underlying hypergraph is retractable, theorem 6.16 implies that $T_u\mathcal{Z}$ is cocyclic for all $u \in \mathcal{Z}$. Note that even in this case, we do not have an explicit and convenient formula such as (6.31) to compute the projection of A_0 onto $T_u\mathcal{Z}$ along the fibers of δA_1. However a linearised message passing scheme will produce the desired result in finite time, for instance by double pass as in proposition 6.17.

In general, the fibers of $T\mathcal{Z}$ are still related to the interaction decomposition through the following proposition. The obtained formula shows how moving u along $\mathcal{Z}$ rotates Z_0: but while doing so, it may very well happen that $T_u\mathcal{Z}$ meets δA_1. One should hence not expect to derive from its apparent simplicity a homologous projection of A_0 onto $T\mathcal{Z}$.

Proposition 6.30. *Let $p = [\mathrm{e}^{-\zeta \cdot u}]$ denote local Gibbs states. Then:*

$$T_u\mathcal{Z} = (\mu \circ p^{-1} \circ \zeta) \cdot Z_0 \tag{6.32}$$

where p^{-1} denotes the region-wise multiplication operator $v_\alpha \mapsto v_\alpha / p_\alpha$.

Proof. By theorem 2.15, in A_0 we have $\mathrm{Ker}(d) = \zeta \cdot Z_0$. The change of metric is simply reflected by conjugating with the diagonal multiplication operator $p = [\mathrm{e}^{-\zeta \cdot u}]$:

$$\nabla = p^{-1} \cdot d \cdot p \tag{6.33}$$

Hence $\mathrm{Ker}(\nabla) = p^{-1} \cdot \mathrm{Ker}(d)$, while by definition 5.7 of ∇ we have $T_u\mathcal{Z} = \mathrm{Ker}(\nabla \circ \zeta) = \mu \cdot \mathrm{Ker}(\nabla)$. ☐

6.3.2 Twisted Laplacian and Singularities

Let us denote by $\mathcal{L} : A_0 \to TA_0$ the non-linear diffusion operator defined in 5.29:

$$\mathcal{L} = \delta \mathcal{D}^\mu \qquad \text{where} \qquad \mathcal{D}^\mu = \mu \circ \mathcal{D} \circ \zeta \tag{6.34}$$

As a vector field, $\mathcal{L}$ fixes $\mathcal{Z}$. The restriction of its tangent map $\mathcal{L}_* : TA_0 \to T(TA_0)$ hence defines a linear endomorphism of the tangent bundle above $\mathcal{Z}$, describing the linearised action of $\mathcal{L}$ in an infinitesimal neighbourhood of $\mathcal{Z}$.

Definition 6.31. *We call* twisted laplacian *the linearised flow $L = \delta\nabla^\mu$ induced by $\mathcal{L}_*$ above $\mathcal{Z}$:*

$$L : T_\mathcal{Z} A_0 \longrightarrow T_\mathcal{Z} A_0 \tag{6.35}$$

Proposition 6.32. *For every $v \in T_\mathcal{Z} A_0$, we have $L(v) = 0$ if and only if $v \in T\mathcal{Z}$, i.e.:*

$$\mathrm{Ker}(L) = T\mathcal{Z} \tag{6.36}$$

Proof. The local faithfulness property 5.37 of $\phi = \mathcal{D}^\mu$ implies that $\mathcal{L} = 0$ is a local equation of $\mathcal{Z}$. ☐

Consider now the sub-bundle $\mathcal{B} \subseteq T_\mathcal{Z} A_0$ spanned by δA_1. By definition of $L = \delta\nabla^\mu$, we have:

$$\mathrm{Im}(L) \subseteq \mathcal{B} \tag{6.37}$$

Hence L could be factorised by a map from $T^\perp\mathcal{Z}$ to $\mathcal{B}$ having same dimension, as $\nabla = \delta^*$ implies $\mathrm{Ker}(\nabla) = \mathrm{Im}(\delta)^\perp$ and $T\mathcal{Z} = \mu \cdot \mathcal{B}^\perp$. In constrast with the harmonic case, $T\mathcal{Z}$ and $\mathcal{B}$ may however intersect and fail to span TA_0.

Definition 6.33. *We denote by $L' : \mathcal{B} \to \mathcal{B}$ the restriction of L to boundaries:*

$$L' \in \mathrm{Hom}(\mathcal{B}, \mathcal{B}) \quad \Leftrightarrow \quad L' \in C^\infty\big(\mathcal{Z},\, \mathrm{End}(\delta A_1)\big) \tag{6.38}$$

Proposition 6.34. *The kernel of L' is the intersection $\mathrm{Ker}(\nabla^\mu) \cap \mathrm{Im}(\delta)$, equivalently:*

$$\mathrm{Ker}(L') = T\mathcal{Z} \cap \mathcal{B} \tag{6.39}$$

Proof. This is a direct consequence of proposition 6.32 and the definition of L'. □

When the underlying hypergraph is retractable, the cocyclicity of $\mathcal{Z}$ implies $T_{\mathcal{Z}} A_0 = T\mathcal{Z} \oplus \mathcal{B}$ so that L' is invertible and defines an isomorphism of $\mathcal{B}$ onto itself. The double pass 6.17 suggests[3] that L has only positive eigenvalues, the flow of $-L$ hence retracting $T_{\mathcal{Z}} A_0$ onto $T\mathcal{Z}$.

In general, following Thom's notations [28] we denote by $\mathcal{S}_0 \subseteq \mathcal{Z}$ the subset where L' is invertible, and by $\mathcal{S}_k \subseteq \mathcal{Z}$ the subset where L' is of corank k. Then $\mathcal{Z}$ has the structure of a stratified space, with respect to the disjoint union:

$$\mathcal{Z} = \bigsqcup_{k \geq 0} \mathcal{S}_k \tag{6.40}$$

where $\bar{\mathcal{S}}_k$ contains $\mathcal{S}_{k+1}$ for all $k \geq 0$, and $\bar{\mathcal{S}}_1$ consists of the *singularities* of L', points where L' is not invertible, and $T\mathcal{Z} \cap \mathcal{B}$ contains directions of *bifurcation*.

Theorem 6.35. *The singular set $\bar{\mathcal{S}}_1$ of L' is nowhere dense in $\mathcal{Z}$.*

Proof. Denote by $M \subseteq \mathrm{Ker}(d)$ the open subset of $\mathrm{Ker}(d) \subseteq A_0^*$ formed by consistent positive densities. M is analytically diffeomorphic to $\mathcal{Z}$, as $q \mapsto -\ln(q)$ is an analytic parametrisation[4] of $\zeta \cdot \mathcal{Z}$, and we denote by $f : M \longrightarrow \mathcal{Z}$ the analytical diffeomorphism then obtained by Möbius inversion.

Letting $N = A_0/\delta A_1$ denote the homology of A_0, consider a projection $\pi : A_0 \longrightarrow N$ as given for instance by (6.31). The composed projection $g = \pi \circ f$ and its tangent map:

$$g_* : TM \longrightarrow TN \tag{6.41}$$

are analytic, while f_* defines a diffeomorphism of $\mathrm{Ker}(g_*)$ onto $T\mathcal{Z} \cap \mathcal{B}$ by construction. By Sard's theorem, the singular set of g where $\det(g_*) = 0$ has an image of measure zero in N. In particular, $\det(g_*)$ cannot identically vanish, as 6.29 implies that $g(M)$ contains an open neighbourhood[5] in N by transverality of $T_{\bar{0}}\mathcal{Z}$ and δA_1.

The map $\det(g_*) : M \to \mathbb{R}$ is analytic and admits an extension $\det_{\mathbb{C}}(g_*) : M_{\mathbb{C}} \to \mathbb{C}$ to $M_{\mathbb{C}} \subseteq A_0^* \otimes \mathbb{C}$. If $\det(g_*)$ vanished on an open set in M, then $\det_{\mathbb{C}}(g_*)$ would also vanish on an open subset in $M_{\mathbb{C}}$, hence everywhere by analytic continuation [[why go to $\mathbb{C}$?]]. As this would violate Sard's theorem, it follows that the regular set where g_* is invertible is dense in M, and $T\mathcal{Z} \cap \mathcal{B} = \{0\}$ almost everywhere. □

6.3.3 The Case of Graphs

We now specialise to graphs as the equations characterising bifurcations in $T\mathcal{Z} \cap \mathcal{B}$ there show a remarkable resemblance with the Kirchhoff rule for the conservation of electric currents in a circuit. The remaining major difference is that our currents $\varphi \in A_1$ are function-valued, while constant scalars cannot contribute to the creation of bifurcations, as the second part of the following proposition implies. These conservation laws also give a very simple explanation to the absence of bifurcations on trees.

[3]The double pass is nothing but a sequential version of L where maximal cells are updated in a specific order.
[4]The coordinate $H(x)$ is given by composing the linear evaluation map $q \mapsto q(x)$ with the analytic map $y \mapsto -\ln(y)$.
[5]In fact g is surjective, as implied by the existence theorem 5.3 of belief propagation equilibria.

Proposition 6.36. *If $v = \delta\varphi$ is a bifurcation in $T_u\mathcal{Z} \cap \mathcal{B}$, then $\varphi \in A_1$ is solution of:*

$$\varphi_{jk\to k} = \mathbb{E}^{jk}\Big[\sum_{i\neq k} \varphi_{ij\to j} \,\Big|\, k \Big] \mod \mathbb{R} \tag{6.42}$$

for every edge-vertex pair $jk \to k$, conditional expectations being taken with respect to $[\mathrm{e}^{-\zeta\cdot u}] \in \Gamma$. Moreover, $\mathbb{E}^i[v_i] = \mathbb{E}^{ij}[v_{ij}] = 0$ for all i, j, so that $v \neq 0$ implies that φ cannot belong to $\mathbb{R}_1$.

Proof. If $v = \delta\varphi \in \mathrm{Ker}(\nabla^\mu)$, the variation of hamiltonians $V = \zeta \cdot v$ satisfies $\nabla(V)_{jk\to k} = 0$ where:

$$V_k = \sum_{i'} \varphi_{i'k\to k} \mod \mathbb{R} \qquad \text{and} \qquad V_{jk} = \sum_{i'\neq j} \varphi_{i'k\to k} + \sum_{i\neq k} \varphi_{ij\to j} \mod \mathbb{R} \tag{6.43}$$

neglecting only currents of the form $\varphi_{ij\to\varnothing}$ and $\varphi_{i\to\varnothing}$, so that $V_k = \mathbb{E}^{jk}[V_{jk} \mid k]$ does simplify to (6.42) in $A_k \mod \mathbb{R}$. To prove the second statement, observe that the equations $\nabla(V)_{k\to\varnothing} = 0$ and $\nabla(V)_{jk\to\varnothing} = 0$ yield after a few obvious simplifications:

$$\varphi_{k\to\varnothing} = \mathbb{E}\Big[\sum_{i'} \varphi_{i'k\to k} \Big] \qquad \text{and} \qquad \varphi_{jk\to\varnothing} = -\mathbb{E}\big[\varphi_{jk\to j} + \varphi_{jk\to k} \big] \tag{6.44}$$

The first kind of equations grouping pairs $ij \to j$ by the vertex j, and the second kind grouping pairs $ij \to j$ by the edge ij, it follows that:

$$\delta(\varphi)_\varnothing = \sum_i \varphi_{i\to\varnothing} + \sum_{ij} \varphi_{ij\to\varnothing} = 0 \tag{6.45}$$

Hence $V_\varnothing = v_\varnothing = \delta(\varphi)_\varnothing = 0$, giving $\mathbb{E}[V_{ij}] = \mathbb{E}[V_i] = 0$ for all i, j by $\nabla(V)_{ij\to\varnothing} = \nabla(V)_{i\to\varnothing} = 0$. The zero-mean statement on interaction potentials finally follows by Möbius inversion. □

The conditional expectation $\mathbb{E}^{jk}[\,-\,|k]$, which couples $\varphi_{jk\to k}$ with incoming currents $\varphi_{i'j\to j}$, consists of the orthogonal projection of A_j onto A_k for the metric induced by a local probability p_{jk} on A_{jk}. Note that if $p = [1]$ were the uniform distribution, we would have $A_j \perp A_k$ for all jk and (6.42) would obviously not have any non-trivial solution, as we already know from the transversality Z_0 with δA_1. However in general, A_j is not orthogonal to A_k as interaction brings correlation accross variables.

Definition 6.37. *Given a chain $i_0, \dots, i_n$ and a loop $j_0, \dots, j_n$ in the underlying graph, we introduce the following definitions for applying successive conditional expectations along edges:*

- $\mathrm{E}^{kj} = \mathbb{E}^{jk}[\,-\,|k]$ *for the* edge operator *projecting A_j to A_k,*
- $\mathrm{C}^{i_n\cdots i_0} = \mathrm{E}^{i_n i_{n-1}} \circ \cdots \circ \mathrm{E}^{i_1 i_0}$ *for the* chain operator *mapping A_{i_0} to A_{i_n},*
- $\mathrm{L}^{j_n\cdots j_0} = \mathrm{E}^{j_0 j_n} \circ \cdots \circ \mathrm{E}^{j_1 j_0}$ *for the* loop operator *mapping A_{j_0} to itself.*

As projector of A_{jk}, we have $\mathrm{E}^{jk} = (\mathrm{E}^{kj})^*$ so that $\mathrm{C}^{i_0\cdots i_n} = (\mathrm{C}^{i_n\cdots i_0})^*$ and $\mathrm{L}^{j_1\cdots j_n j_0} = (\mathrm{L}^{j_n\cdots j_1 j_0})^*$.

The non-degeneracy of p_{jk} implies that E^{kj} is of norm strictly smaller than 1 as operator of A_{jk}, once restricted to the orthogonal supplement of $A_j \cap A_k = A_\varnothing = \mathbb{R}$. It follows that the spectrum of a loop operator $\mathrm{L}^{j_n\cdots j_0}$ restricted to $\mathbb{R}^\perp$ is contained in $]0, 1[$, and hence in contrast with the scalar case, one shall not observe solutions of (6.42) corresponding to a conserved current running along a single loop. Such solutions only appear in the strong coupling limit, p going to the boundary of Γ and $\mathcal{Z}$ getting asymptotically tangent to $\mathcal{B}$.

True bifurcations in $T\mathcal{Z} \cap \mathcal{B}$ may however occur when two or more loops can collaborate to sustain a conserved current. We provide with explicit examples of such bifurcations below. Their simplicity is remarkable, as previous examples had only been witnessed numerically, apparently starting with Weiss [31]. Before moving on, it seems important to mention that although non-constructive, mathematical proofs on the existence of bifurcations have already been given by D. Bennequin in [4] and unpublished work. The first relied on methods of algebraic geometry and shows the existence of at least three homologous fixed points on the binary "figure eight".

The second proof is closer to the methods exposed here, as it relies on the present homological description of message-passing algorithms. The idea is to view (6.42) as an eigenvalue problem:

$$\varphi = \mathrm{M}(\varphi) \qquad \text{with} \qquad \mathrm{M} : A_1 \longrightarrow A_1 \tag{6.46}$$

Applying the Perron-Frobenius therorem, one then shows that along certain paths in the space $\mathcal{Z} \simeq \Gamma$ of parameters, the real largest eigenvalue of M has to cross 1. This approach proved the existence of at least six bifurcations on the simple "figure eight". Avoiding the challenge of producing clever bounds on the spectrum of M, we satisfy with exhibiting simple solutions of (6.46).

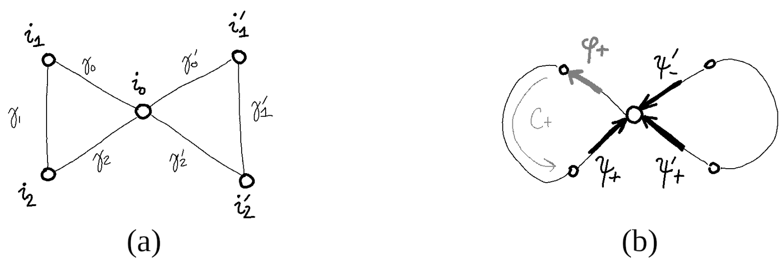


Figure 6.2: (a) Two triangles joined by a vertex. (b) Currents at the junction.

Consider then the "figure eight" graph obtained by joining two triangles i, i_+, i_- and i', i'_+, i'_- by a common vertex $i' = i$, as depicted in figure 6.2. Conservation at the junction (i) and conservation along each loop (ii) yield two times four equations of the following form:

$$\text{(i) } \varphi_+ = \mathrm{E}_+\big(\psi_+ + \psi'_+ + \psi'_-\big) \qquad\qquad \text{(ii) } \psi_+ = \mathrm{C}_+(\varphi_+) \tag{6.47}$$

Eliminating the ψ's then reduces the above to four equations of the form:

$$\varphi_+ = \mathrm{L}_+(\varphi_+) + \mathrm{E}_+\Big(\mathrm{C}'_+(\varphi'_+) + \mathrm{C}'_-(\varphi'_-)\Big) \tag{6.48}$$

Identifying each A_j with $\mathbb{R}^2$ and each A_{jk} with $\mathbb{R}^4 \simeq M_2(\mathbb{R})$, assume that local probabilities on edges and vertices are all of the same following form (no magnetic fields) for simplicity:

$$p_{jk} = \frac{1}{4}\begin{bmatrix} 1+a & 1-a \\ 1-a & 1+a \end{bmatrix} \qquad \text{and} \qquad p_j = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{6.49}$$

Now each edge operator $\mathbb{E}^{jk}[-|j] = 2p_{jk}$ is represented by the self-adjoint matrix E given by:

$$\mathrm{E} = R^{-1}\begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} R \qquad \text{where} \qquad R = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \tag{6.50}$$

expressing that we have the non-constant eigenvector $y = [\,1\ -1\,]^T$ with $\mathrm{E}(y) = a\,y$. In the quotient of A_0 by $\mathbb{R}_0$, each edge operator may hence be represented as multiplication by $a \in]-1,1[$ so that (6.48) reads:

$$\varphi_+ = a^3\big(\varphi_+ + \varphi'_+ + \varphi'_-\big) \tag{6.51}$$

Assuming that $a^3 = \frac{1}{3}$, a bifurcation occurs in the direction of $\varphi_\pm = \varphi'_\pm = y$.

Let us also prove that this is the only bifurcation in the considered one-parameter family of $\mathcal{Z}$. Considering that each of the φ's is spanned by $[\,1\ -1\,]^T$ and using $(\varphi_+, \varphi_-, \varphi'_+, \varphi'_-)$ as coordinates, the four equations obtained from 6.51 may be written as:

$$\varphi = \mathrm{M}\varphi \qquad \text{where} \qquad \mathrm{M} = a^3 \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \tag{6.52}$$

A reduction of M is easily computed:

$$\mathrm{M} = a^3\, Q^{-1} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} Q \qquad \text{where} \qquad Q = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \tag{6.53}$$

Under the constraint $|a| < 1$, it follows that $1 \in \mathrm{Spec}(\mathrm{M})$ if and only if $3a^3 = 1 \Leftrightarrow a = 3^{-\frac{1}{3}}$, with corresponding eigenvector $\varphi = [\,1\;1\;1\;1\,]^T$. One may think of the three remaining eigenvectors of M as directions for bifurcations occuring at the boundary of Γ, when $a \to \pm 1$ and the probability of observing aligned/unaligned neighbouring spins tends to 0/1.

[[add figure]]

Two triangles joined by an edge lead do the following similar set of equations, denoting by a and b the weights on the outter edges and diagonal of the obtained square respectively:

$$\varphi = \mathrm{M}\varphi \qquad \text{where} \qquad \mathrm{M} = a^2 \begin{bmatrix} b & 0 & b & 1 \\ 0 & b & 1 & b \\ b & 1 & b & 0 \\ 1 & b & 0 & b \end{bmatrix} \tag{6.54}$$

which has the same eigenspaces and

$$\mathrm{M} = a^2\, Q^{-1} \begin{bmatrix} 2b+1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2b-1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} Q \qquad \text{where} \qquad Q = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \tag{6.55}$$

So that in the considered two-parameter family of $\mathcal{Z}$, diffeomorphic to the square $(a, b) \in \,]-1, 1[^2$, bifurcations occur along the path $a = \pm\frac{1}{\sqrt{2b+1}}$.

Bibliography

[1] S. Abramsky and A. Brandenburger, *The Sheaf-theoretic structure of non-locality and contextuality*, New Journal of Physics, 13 (2011).

[2] P. Baudot, M. Tapia, D. Bennequin, and J.-M. Goaillard, *Topological Information Data Analysis*, Entropy, 21 (2019), p. 869.

[3] D. Bennequin, O. Peltre, and G. Sergeant-Perthuis, *Extrafine Sheaves.* preprint, 2019.

[4] D. Bennequin, O. Peltre, G. Sergeant-Perthuis, and J. Vigneaux, *Informations, Energies and Messages.* preprint, 2019.

[5] W. G. Dwyer and J. Spalinski, *Homotopy theories and model categories*, 1995.

[6] K. J. Friston, T. Parr, and B. de Vries, *The Graphical Brain: Belief Propagation and Active Inference*, Networks Neuroscience, 1 (2017), pp. 381–414.

[7] R. G. Gallager, *Low-Density Parity-Check Codes*, MIT Press, 1963.

[8] K. T. Hu, *On the Amount of Information*, Theory of Probability and its Applications, 7 (1962), pp. 439–447.

[9] H. G. Kellerer, *Maßtheoretische Marginalprobleme*, Mathematische Annalen, 153 (1964), pp. 168–198.

[10] R. Kikuchi, *A Theory of Cooperative Phenomena*, Phys. Rev., 81 (1951), pp. 988–1003.

[11] C. Knoll and F. Pernkopf, *On Loopy Belief Propagation – Local Stability Analysis for Non-Vanishing Fields*, in Uncertainty in Artificial Intelligence, 2017.

[12] K. Kodaira, *Harmonic fields in Riemannian manifolds (Generalized potential theory).*, Ann. Math. (2), 50 (1949), pp. 587–665.

[13] T. Leinster, *The Euler Characteristic of a Category*, Documenta Mathematica, 13 (2008), pp. 21–49.

[14] S. MacLane and S. Eilenberg, *General Theory of Natural Equivalences*, Transactions of the American Mathematical Society, (1945), pp. 231–294.

[15] F. Massieu, *Sur les Fonctions caractéristiques des divers fluides*, Comptes rendus de l'Académie des Sciences, 69 (1869), pp. 858–862.

[16] F. Matúš, *Discrete Marginal Problem for Complex Measures*, Kybernetika, 24 (1988), pp. 36–46.

[17] P.-A. Meyer, *Éléments de probabilités quantiques (exposés I à V)*, Séminaire de probabilités de Strasbourg, 20 (1986), pp. 186–312.

[18] I. Moerdijk, *Classifying Spaces and Classifying Topoi*, Springer, 1995.

[19] J. M. Mooij and H. J. Kappen, *Sufficient Conditions for Convergence of the Sum–Product Algorithm*, IEEE Transactions on Information Theory, 53 (2007), p. 4422–4437.

[20] T. Morita, *Cluster Variation Method of Cooperative Phenomena and its Generalization I*, Journal of the Physical Society of Japan, 12 (1957), pp. 753–755.

[21] K. P. Murphy, Y. Weiss, and M. I. Jordan, *Loopy Belief Propagation for Approximate Inference: An Empirical Study*, in UAI, 1999.

[22] M. Mézard and A. Montanari, *Information, Physics and Computation*, Oxford University Press, 2009.

[23] J. Pearl, *Reverend Bayes on Inference Engines: A Distributed Hierachical Approach*, in AAAI-82 Proceedings, 1982.

[24] G.-C. Rota, *On the Foundations of Combinatorial Theory - I. Theory of Möbius Functions*, Z. Warscheinlichkeitstheorie, 2 (1964), pp. 340–368.

[25] A. G. Schlijper, *Convergence of the cluster-variation method in the thermodynamic limit*, Phys. Rev. B, 27 (1983), pp. 6841–6848.

[26] J.-M. Souriau, *Quantification Géométrique*, in Physique quantique et géométrie: formulation mathématique cohérente des phénomènes quantiques: Colloque géométrie et physique de 1986 en l'honneur d'André Lichnerowicz, vol. 32, Travaux en cours, Hermann, 1988.

[27] T. P. Speed, *A Note on Nearest-Neighbour Gibbs and Markov Probabilities*, Sankhya: The Indian Journal of Statistics, 41 (1979), pp. 184–197.

[28] R. Thom, *Les Singularités des applications différentiables*, Annales de l'institut Fourier, 6 (1956), pp. 43–87.

[29] J.-L. Verdier and A. Grothendieck, *V: Cohomologie dans les Topos*, SGA-4, 2 (1972).

[30] N. Vorob'ev, *Consistent Families of Measures and their Extensions*, Theory of Probability and its Applications, 7 (1962), pp. 147–164.

[31] Y. Weiss, *Belief propagation and revision in networks with loops*, tech. rep., MIT, 1997.

[32] J. Yedidia, W. Freeman, and Y. Weiss, *Constructing Free Energy Approximations and Generalized Belief Propagation Algorithms*, IEEE Transactions on Information Theory, 51 (2005), pp. 2282–2312.

[33] J. S. Yedidia, W. T. Freeman, and Y. Weiss, *Bethe free energy, Kikuchi approximations, and belief propagation algorithms*, Tech. Rep. TR2001-16, MERL - Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, May 2001.