**White paper**

**The *i*-space**: The number of unique *i*-states—combinations—in a given *i*-space is a function of the number of redox states in binary (0 = reduced; 1 = oxidized) over the cysteine residue integer (*R*). These combinations double with each additional cysteine residue, following the binomial expansion:

Where:

* is the *i*-space
* 2 is redox binary.
* *R* is the cysteine residue integer.

**The static geometry of *k*-space**: Rooted in binomial theorem, Pascal’s triangle organizes the *i*-space according to the degeneracy, number of ways it can happen, of each cysteine oxidation integer value (*k*) using binomial coefficients ().

(2)

Where:

* *R* is the cysteine residue integer.
* *k* is the *k*-state—cysteine oxidation integer.
* *!* is the factorial function.

In each row (an *R* series), the sum of the binomial coefficients yielded the *i*-space: . While the number of coefficients per row bounded *k*-space as . Each *k*-state in an *R* series yielded an oxidation percentage (%OX) via: .

**The nonlinear dynamic geometry of *k*-space:** Pascal’s triangle defined a universal framework for the dynamic movement of *i*-states within *k*-space over time steps. The *i*-space of the *R* = 3 series, is organized into a *k*-space stratified matrix of binary redox strings:

Matrix geometry defines deterministic nonlinear rules governing transitions:

1. **Boundary Constraints**: The *i*-states at the extremes of *k*-space (000, 111) cannot be further acted upon. The 000 *i*-state cannot be reduced further, while 111 cannot be oxidized. Geometry = cannot move outside the Pascal row.
2. **Stepwise Movement**: Transitions within *k*-space are incremental. For example, the 000 *i*-state can transition to any *k* = 1 state but cannot directly transition to *k* = 2 or 3 in one step. Geometry = can only move one coefficient row at a time.
3. **Conservation of Site-Information**: Transitions conserve site-information. For instance, the 010 *i*-state can only be oxidized at position 1 or 3 or reduced at position 2. Hence, the movement of 010 to 101 is barred. Geometry = moving within a coefficient row is barred.

For any *i*-state in *k*-space, the total (allowed and barred) degrees of freedom (*df*) are conserved. They always sum to the *i*-space:

The allowed *df* always sums to the *k*-space integer (e.g., 4 in the *R* = 3 series). The *df* matrix partitions the probability simplex for each *i*-state into allowed (*P* = 1) and barred (*P* = 0) transitions. For the 000 *i*-state in the *R* = 3 series, this simplex structure may appear as:

Binomial geometry defines how probabilistic redox reactions drive the deterministic movement of *i*-states in *k*-space, codified as:

Where:

* *Predox* (*i*) is the probability of a redox reaction acting on a given *i*-state at a given site at a given time step.
* *k* is the stepwise change in *k*-space.
* *fDe*t is the operation of the nonlinear, deterministic rules that dictate the next *i*-state in *k*-space, which is formally

The lefthand side defines the probabilistic action of oxidative (e.g., ROS) and reductive (e.g., antioxidants) processes on the *i*-state at a given time step, which encodes site-information. The righthand side translates their probabilistic action into a deterministic move of the *i*-state in *k*-space. The left and right side are external and internal to *i*-states in *k*-space, respectively. The simplex is Bayesian: Dynamically updating for each *i*-state in *k*-space as a function of the left (redox *P* values evolve) and right side of the equation.

**Energy and entropy**: Each *i*-state (*i*) defines a unique microstate (*m*) of the *i*-space (Ωi), with its energy (*Ei*) and entropy (*Si*) derived from its position (*k*-state) in *k*-space (Ωk).

Where:

* *Ei* = Energy of an *i*-state
* *E0* = Reference energy of *k* = 0.
* *k* = *k*-state integer
* ΔE = Energy increment per oxidized cysteine.

Where:

* *Si* = Entropy
* *kB* = The Boltzmann constant
* *Wk* = The degeneracy of the *k*-state, representing the number of ways it can happen, given by the binomial coefficient (equation 4).

Hence, the probability of observing a given *i*-state (*Pi*) is given by the Boltzmann distribution:

Where:

* *Pi* = The probability of observing a given *i*-state, where all values sum to 1.
* *T* = Temperature
* *Zk* = Partition function for all *i*-states in *k*-space

The ensemble-averaged *E* and *S* of *i*-states in *k*-space can be described as:

The interplay between energy and entropy, codified in these equations, provides a formal framework for introducing basins of attraction.

**A basin of attraction**: Defined a stable region of *k*-space arising from the biological configuration of *i*-states in a probabilistic energy and entropy landscape under specific conditions, which can be expressed as:

Where:

* *Bk*​: Is the basin of attraction in *k*-space.
* *Ωk*: Is the i-space in *k*-space.
* Wk: The degeneracy of the *k*-state per the binomial coefficient.
* Pi​: The probability of observing a specific *i*-state in *k*-space from the Boltzmann distribution.
* WThreshold⋅PThreshold​: Establishes the combined degeneracy and probability threshold required for an i-state to be considered part of the basin.
* ​: Ensures that the lower bound for probability reflects the partition function and the thermodynamic constraints of the system

**Far-from equilibrium dissipative structures**:

Where:

* *H*k: is the Shannon entropy for all *i*-states
* *Ωk*: is the *i*-space in *k*-space.
* The righthand side of the equation is defined as above

**Oxi-Shapes:** Is codified by a master equation, capturing the interplay of stochastic redox reactions, energy transduction, and deterministic geometry. Energy imparted by probabilistic redox reactions acts on the cysteine metric tensor (*gij*), deterministically moving *i*-states across Riemannian *k*-space manifolds. These movements, geometrically constrained by the Ricci curvature, encode information via their Shannon entropy. The framework is mathematically formalized as:

Where:

* is the probability that a given *i*-state is moved in *k*-space by a stochastic redox reaction, which sums to 1 when the possibility that it is not acted on is included.
* is the energy of action that describes the movement of an *i*-state in *k*-space.
* The change in energy deterministically moves the cysteine metric tensor (*gij*).
* The cysteine metric tensor (*g*) moves via a transition matrix (*T*) encoding the simplex of allowed and barred transitions of *i*-states in *k*-space.
* The Ricci curvature for the transition adjusts the effort required based on the *gij*.
* The movement increases entropy (*S*), changing the information (*Hk*) encoded by the system.

To illustrate consider a molecule in the *R* = 3 series. The maps to a 000, 100, 010, 001, 110, 101, 011, 111 *i*-state set, distributed 1:3:3:1 in *k*-space via *gij*:

Suppose the following for the 000 *i*-state:

* **Probability**: The probability of a redox reaction acting on 000 to produce 100 (oxidation at site 1) is *P*(000→100)=0.25. The rest of the simplex sums to 1.
* **Energy**: The energy imparted by the redox reaction is ΔE(000→100)=5.0 kcal/mol.
* **Metric Tensor**: The cysteine metric tensor (*gij*) dictates the geometric change, moving 000 deterministically to 100 in *k*-space with *Δxj*=1.
* **Ricci Curvature**: The Ricci curvature for the transition of 0.8, reflects the geometric resistance of *k*-space. Higher curvature signifies greater resistance, requiring more energy to overcome the barriers imposed by the *k*-manifold structure.

Plugging these values into equation 7 gives

Substituting the terms

* The left-hand side gives 1.25∙gij, representing the probabilistic energy input scaled by *gij*.
* The right-hand side ensures the transition matrix *T*(000,100), geometric curvature 0.8, and entropy *S*(Hk) encode the deterministic and informational aspects of the system.

Hence, 000 transitions deterministically to 100, altering the geometry of *k*-space and increasing informational entropy.

**Universal PTM-proteoform geometry**: Formally, in the binary unmodified (0) or modified (1) basis the number of residue-specific PTM-speciated proteoforms can be calculated by a variation of equation 1:

Where:

* is the *PTM*-space for a given amino acid basis
* 2 is binary unmodified or modified (e.g., phosphorylated).
* *R* is amino acid integer (e.g., tyrosine).

This can be extended to calculate the PTM space of the entire protein by substituting the *R* term to the length (L) of the amino acid sequence:

These binomial expansion revealing equations, mean that the static and dynamic geometric organisation of the associated *k*-space, where the *k*-value is the modified integer, also apply to the binary PTM-space of any amino acid. For example, if there were 3 tyrosine residues in a protein, the associated proteoforms would follows the same 1:3:3:1 structure in Pascal’s triangle, reproducing the matrix below:

Given the geometric basis of the dynamical rules, they also apply. For example, a 111 phosphorylated tyrosine proteoform cannot be further phosphorylated in this basis. As a result, the rules of the Bayesian simplex also apply, meaning there are barred and allowed transitions.

Hence, a variation of the Oxi-Shapes master equation can be derived for any PTM basis. For example, for phosphorylation of tyrosine (iy) it could be:

In this case, the P term is the probabilistic action of a kinase, which is energy-coupled via ATP hydrolysis.

As a result, proteoform shapes across Riemannian PTM-state manifolds can be derived for any amino acid.