The Variational Principle

Variational Quantum Algorithms · Module 1 — Hybrid Quantum Classical Systems

Node 1.1

Position in Knowledge Graph

System: Variational Quantum Algorithms **Structural Domain:** Hybrid Quantum-Classical Systems **Node:** 1.1

This is **the theorem that makes VQAs legal**. Everything downstream in this course — every cost function, every optimizer, every claim about "approximating the ground state" — depends on this one result. It is the justification for treating the scalar number $\langle\psi|H|\psi\rangle$ as something worth minimizing.

Formal Statement

Let H be a Hermitian operator on a Hilbert space $\mathcal{H}$, with eigenvalues $E_0 \leq E_1 \leq E_2 \leq \ldots$ and corresponding eigenstates $|\phi_0\rangle, |\phi_1\rangle, \ldots$. For any normalized state $|\psi\rangle \in \mathcal{H}$,

$$\langle\psi|H|\psi\rangle \geq E_0,$$

with **equality if and only if** $|\psi\rangle = |\phi_0\rangle$ (up to a global phase, and modulo degeneracy).

This is the **Rayleigh-Ritz variational principle**, stated in its quantum-mechanical form. It is the oldest member of a family of inequalities that appear across mathematical physics whenever a bilinear form has to be bounded below by its smallest eigenvalue.

Proof in Three Lines

Expand $|\psi\rangle = \sum_i c_i|\phi_i\rangle$ in the eigenbasis of H, with $\sum_i |c_i|^2 = 1$. Then

$$\langle\psi|H|\psi\rangle = \sum_i |c_i|^2 E_i \geq E_0 \sum_i |c_i|^2 = E_0,$$

because every $E_i \geq E_0$ and $\sum_i |c_i|^2 = 1$. Equality requires all weight in the lowest eigenvalue component.

That is the entire proof. Every word of Module 5's "not bias" polemic, every gradient descent step in a VQE run, every barren plateau theorem — all of it lives above this three-line argument.

Intuition

The variational principle says: **you cannot accidentally overshoot the ground state.**

Any guess $|\psi\rangle$ you make about the ground state gives you an **upper bound** on the true ground-state energy. A better guess gives a tighter bound. The best guess reachable by your ansatz gives you the tightest bound your ansatz can produce.

This is the feature that makes VQAs optimizable at all. Without the variational principle, minimizing $\langle\psi(\boldsymbol{\theta})|H|\psi(\boldsymbol{\theta})\rangle$ would just be minimizing some scalar function with no guarantee that doing so gets you closer to anything physical. With it, every decrease in $C(\boldsymbol{\theta})$ is a decrease in your upper bound on the real thing.

Another way to say the same thing: **energy estimates from variational methods are biased low in a useful direction.** They are always at least E_0, so they can never lie to you in the "your answer is lower than the ground state" direction. The only way to be wrong is to be too high — which just means your ansatz isn't expressive enough to reach the true ground state. That is a diagnosable failure mode, not a silent one.

Structural Role

This node is the **foundation stone** of Module 1. Everything else in the module specializes, implements, or complicates the picture laid out here:

- Node 1.2 (architecture) tells you how the variational principle gets executed in practice on a hybrid device.
- Node 1.3 (expectation estimation) tells you how you actually **measure** $\langle\psi|H|\psi\rangle$ when ψ lives on a real quantum computer.
- Node 1.8 (quantum subsystem as objective generator) reframes the variational principle in thesis language: the quantum device implements a stochastic oracle for a quantity the variational principle tells you is worth bounding.

If you forget this node, nothing else in the course makes sense. If you remember only this node, you still know why VQAs exist.

Relations to Other Concepts

- **Rayleigh quotient** — the classical version from linear algebra: for a symmetric matrix A, $\min_{x\neq 0} \frac{x^T A x}{x^T x} = \lambda_{\min}(A)$. The quantum variational principle is this statement on unit vectors in $\mathcal{H}$.
- **Hartree-Fock method** — the first and most famous non-quantum-computer application of the variational principle to chemistry.
- **Density Matrix Renormalization Group (DMRG)** — also a variational method, restricted to matrix-product-state ansätze.
- **Reinforcement learning policy gradient** — structurally analogous: minimize an upper bound on an expected return by descending its stochastic gradient.

Worked Example — The Simplest Variational Calculation

Let $H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (Pauli X). Its eigenvalues are ± 1, so $E_0 = -1$.

Consider the one-parameter ansatz

$$|\psi(\theta)\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)|1\rangle.$$

Compute

$$C(\theta) = \langle\psi(\theta)|X|\psi(\theta)\rangle = 2\cos(\theta/2)\sin(\theta/2) = \sin\theta.$$

The minimum over θ is -1, achieved at $\theta = -\pi/2$. This equals E_0 exactly — the ansatz is expressive enough to reach the true ground state. The variational principle guarantees $C(\theta) \geq -1$ for all θ, and direct calculation confirms it.

Now try a less expressive ansatz: $|\psi(\theta)\rangle = \cos(\theta)|0\rangle + \sin(\theta)|+\rangle$ (normalized properly — exercise). Compute $C(\theta)$ and show that its minimum is strictly greater than -1 — the ansatz cannot reach the true ground state, and the variational principle still holds.

Visualization

Pending. Diagram: a horizontal line at E_0. Above it, a scatter of points representing $C(\boldsymbol{\theta})$ values for random $\boldsymbol{\theta}$; all points strictly above the line. An arrow showing the optimizer "pushing down" toward E_0 but never crossing it.

Exercises

1. Prove that if H has a degenerate ground space, then $C(\boldsymbol{\theta}) = E_0$ implies $|\psi(\boldsymbol{\theta})\rangle$ lies in the ground space, but not necessarily that it equals any particular ground-state vector.
2. Show that $\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta})$ is always achieved (i.e., the infimum is attained) when the parameter domain is compact and C is continuous. Why are these assumptions both needed?
3. (VQA) Give an example of a two-qubit Hamiltonian and a one-parameter ansatz for which $\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}) > E_0$ strictly. What does this tell you about the expressivity of your ansatz?

Research Extensions

- **Variational bounds for excited states** — the SSVQE and VQD algorithms extend the principle to find $E_1, E_2, \ldots$ by enforcing orthogonality constraints.
- **Krylov subspace methods** — classical methods that build better variational subspaces iteratively from matrix-vector products.
- **Witness operators** — in entanglement theory, operators whose expectation values bound quantities of interest in the same structural sense.

Cross-links to Other Courses

- **Linear Algebra** — Rayleigh quotients, min-max characterization of eigenvalues.
- **Quantum Foundations** — the variational principle as a corollary of the spectral theorem.
- **Statistics Module 10** — structurally analogous upper-bound arguments in model selection (e.g., PAC bounds as upper bounds on true risk).

Variational Quantum Algorithms · Module 1 — Hybrid Quantum Classical Systems · Node 1.1
Concepts: variational-principle, expectation-value, hilbert-space, eigenvalue-decomposition
Enables: 1.2, 1.3, 1.8

hbar.university — own.your.cognition