Dual frames and Legendre duality

Matthew Weiss

September 12, 2026

1 Frame duality

The most general kind of measurement in quantum theory is modeled by a *positive operator valued measure* or POVM. If the number of outcomes of the measurement is a finite number n and the Hilbert space used to model the system of interest is d dimensional, then a POVM consists of a set of n $d \times d$ positive semidefinite Hermitian matrices called *effects* which sum to the identity:

$$\forall i : E_i \geq 0 \qquad \sum_i E_i = I. \tag{1.1}$$

If the $\{E_i\}$ span the d^2-dimensional operator space, we'll call it a *reference measurement*. In particular, if $n = d^2$, the measurement is called a minimal informationally complete (MIC) POVM [1]. Suppose that conditional on the outcome E_i occuring, the state $\sigma_i = E_i/\mathrm{tr}(E_i)$ is prepared. In fact, we will assume that the measurement is *unbiased* so that $\forall i : \mathrm{tr}(E_i) = d/n$, which is $1/d$ for a MIC. Then the conditional probability matrix

$$P_{ij} = P(E_i|E_j) = \mathrm{tr}(E_i\sigma_j) \tag{1.2}$$

characterizes the reference measurement in its own terms: the *reference effects* characterizing the *reference states*, and the reference states characterizing the reference effects. Operationally, you can think of $P(E_i|E_j)$ as the probability of getting reference outcome E_i given that you got reference outcome E_j last time.

Since the reference states and effects are spanning, we can write for any quantum state ρ,

$$\rho = \sum_i w_i\sigma_i, \tag{1.3}$$

for some vector of expansion coefficients w. Since

$$p_i \equiv P(E_i|\rho) = \mathrm{tr}(E_i\rho) = \sum_j w_j \mathrm{tr}(E_i\sigma_j), \tag{1.4}$$

we have $p = Pw$. If the reference measurement is a MIC, then the reference effects (and states) are not just spanning, but linearly independent. Since the columns of P just represent σ_j in a different basis, they must be linearly independent themselves, and so P must be invertible. We conclude that $w = P^{-1}p$, and we call $\Phi \equiv P^{-1}$ the *Born matrix*. It takes the probabilities p characterizing a state to the expansion coefficients w. We can therefore write

$$\rho = \sum_{ij} \Phi_{ij} P(E_j|\rho)\sigma_i, \tag{1.5}$$

and so for any other measurement outcome represented by an effect A,

$$P(A|\rho) = \mathrm{tr}(A\rho) = \sum_{ij} \Phi_{ij} P(E_j|\rho)\mathrm{tr}(A\sigma_i) \tag{1.6}$$

$$= \sum_{ij} P(A|E_i)\Phi_{ij}P(E_i|\rho), \tag{1.7}$$

since we can think of $\mathrm{tr}(A\sigma_i)$ as the probability of getting outcome A given that the reference measurement gave outcome E_i first. This is the QBist reformulation of the Born rule as a coherence condition analogous to the law of total probability [2, 3]. In fact, a rule of this form can be derived on very general grounds (without assuming quantum theory) [4, 5].

In the language of *frame theory* [6], $p_i = \mathrm{tr}(E_i\rho)$ is a representation of ρ in *frame coordinates*. There exists a dual frame $D_j = \sum_j \Phi_{ij}\sigma_i$ so that $\rho = \sum_j p_j D_j$. Conversely, $w_i = \frac{1}{d}\mathrm{tr}(D_j\rho)$ is a representation of ρ in *dual frame coordinates*. (In fact, w gives us a *quasiprobability* representation [7].) In turns out we can take another point of view on this notion of frame-theoric duality.

2 Legendre duality

Physicists who know their classical mechanics (as well as thermodynamics) will be familiar with the idea of a *Legendre dual*. Let $f : S \to \mathbb{R}$ be a convex function on a convex set $S \subseteq \mathbb{R}^m$. The Legendre dual [8] of f is

$$f^*(x^*) = \sup_{x\in S}\Big\{\langle x^*, x\rangle - f(x)\Big\}. \tag{2.1}$$

Its domain is S^* which consists of all x^* for which $f(x^*)$ doesn't blow up to infinity. For example, in classical physics, the Hamiltonian function is the Legendre dual of the Lagrangian function [9]. In thermodynamics, the Helmholtz free energy as a function of temperature is the Legendre dual of the internal energy as a function of entropy [10] (up to a sign). It comes up all over the place. It was surprising to me, however, that this notion of duality can coincide with the frame-theoretic notion.

Consider a potential function defined on quantum states

$$\Psi(\rho) = \frac{1}{d}\frac{1}{2}\mathrm{tr}(\rho^2), \tag{2.2}$$

where both prefactors have been cleverly chosen. We note that $\mathrm{tr}(\rho^2)$ is called the *purity* of the state: it is 1 iff ρ is a pure state, and reaches its minimum of $1/d$ for the maximally mixed state [11]. At the same time, we can also view Ψ as a function of dual frame coordinates,

$$\Psi(w) = \frac{1}{d}\frac{1}{2}\mathrm{tr}\left(\sum_i w_i\sigma_i \sum_j w_j\sigma_j\right) = \frac{1}{2}\sum_{ij} w_i \mathrm{tr}(E_i\sigma_j)w_j = \frac{1}{2}w^T P w, \tag{2.3}$$

where we've used the fact that $E_i = \sigma_i/d$. Now let's take its first derivatives of this potential, that is, the gradient. One kind of slick way of doing this is to recall that the differential is defined as $d\Psi(w) = (\nabla_w\Psi(w))^T dw$, where $dw = (dw_1, \ldots, dw_n)$. At the same time, using the chain rule for differentials,

$$d\Psi(w) = \frac{1}{2}\Big\{(dw)^T P w + w^T P dw\Big\} = \frac{1}{2}w^T(P^T + P)dw = w^T P dw, \tag{2.4}$$

where we've used that fact that $P = P^T$. Matching this expression to the definition, we conclude

$$\nabla_w\Psi(w) = Pw. \tag{2.5}$$

Meanwhile, let us consider the Legendre dual of $\Psi(w)$ whose argument we'll denote (...for no particular reason...) p:

$$\Psi^*(p) = \sup_w\Big\{p^T w - \Psi(w)\Big\}. \tag{2.6}$$

To evaluate the supremum, let's take the gradient with respect to w and set it to zero. $\nabla_w p^T w = p$ and we just worked out that $\nabla_w\Psi(w) = Pw$ so that

$$\nabla_w\Big\{p^T w - \Psi(w)\Big\} = p - Pw = 0, \tag{2.7}$$

so that $Pw = p$ or $w = P^{-1}p$. Indeed, we see that the relation between w and p is exactly what we derived earlier in the frame-theoretic setting: p gives the probability representation of ρ. (We note that if the POVM had not been a MIC, that is, if it had been overcomplete, then the same construction can be used with $\Phi = P^+$, the Moore-Penrose pseudoinverse, since this acts as the usual inverse on the range of P: we need not spell out the details here [12, 8].) In any case, plugging this back into the expression for $\Psi^*(p)$, we have

$$\Psi^*(p) = p^T P^{-1} p - \frac{1}{2}(pP^{-1})^T P(P^{-1}p) = \frac{1}{2}p^T P^{-1} p, \tag{2.8}$$

since the inverse of a symmetric matrix is also symmetric. In fact, writing Φ for P^{-1}, we have

$$\Psi^*(p) = \frac{1}{2}p^T \Phi p = \frac{1}{2}\sum_{ij} \mathrm{tr}(E_i\rho)\Phi_{ij}\mathrm{tr}(E_j\rho) \tag{2.9}$$

$$= \frac{1}{d}\frac{1}{2}\mathrm{tr}\left(\left[\sum_{ij}\Phi_{ij}\mathrm{tr}(E_j\rho)\sigma_i\right]\rho\right) = \frac{1}{d}\frac{1}{2}\mathrm{tr}(\rho^2), \tag{2.10}$$

so that actually the potential is self-dual in this sense. In fact, we can already see this from $p^T\Phi p = (wP)^T\Phi(Pw) = w^T P w$. Finally, by the same argument as above, $\nabla_p \Psi^*(p) = \Phi p = w$.

So let us take stock. We defined a "purity potential"

$$\Psi(\rho) = \frac{1}{d}\frac{1}{2}\mathrm{tr}(\rho^2). \tag{2.11}$$

Since $\rho = \sum_i w_i\sigma_i$ for some reference states $\sigma_i = E_i/d$, we can express the potential in terms of dual frame coordinates w or frame coordinates p,

$$\Psi(w) = \frac{1}{2}w^T P w \qquad\qquad \Psi^*(p) = \frac{1}{2}p^T\Phi p, \tag{2.12}$$

where $\Phi = P^{-1}$ and $w = \Phi p$. In fact, $\Psi(w)$ and $\Psi^*(p)$ are Legendre dual to each other, and

$$\nabla_w \Psi(w) = p \qquad\qquad \nabla_p \Psi^*(p) = w. \tag{2.13}$$

Thus the first derivatives of the potentials switch us between the quasiprobability and probability representations. Moreover, taking another derivative gives us the Hessian matrices,

$$\nabla^2\Psi(w) = \nabla_w(Pw) = P \qquad\qquad \nabla^2\Psi^*(p) = \nabla_p(\Phi p) = \Phi. \tag{2.14}$$

In other words,

$$\frac{\partial\Psi(w)}{\partial w_i \partial w_j} = P_{ij} \qquad\qquad \frac{\partial\Psi^*(p)}{\partial p_i \partial p_j} = \Phi_{ij}. \tag{2.15}$$

For a MIC, these will both be positive definite and symmetric (and in fact this implies the potential functions are strictly convex), and so we can interpret them as *metrics* [13, 14].

For one thing, they encode the curvature of the potential. Expanding $\Psi(w + dw)$ around w, we have

$$\Psi(w + dw) = \Psi(w) + (\nabla\Psi(w))^T dw + \frac{1}{2}dw^T\nabla^2\Psi(w)dw \tag{2.16}$$

$$= \frac{1}{2}w^T P w + w^T P dw + \frac{1}{2}dw^T P dw : \tag{2.17}$$

the Taylor series terminates since the Hessian is a constant. The second derivative term represents the curvature of the function around w. Similarly,

$$\Psi^*(p + dp) = \frac{1}{2}p^T\Phi p + p^T\Phi dp + \frac{1}{2}dp^T\Phi dp. \tag{2.18}$$

We can then define a metric e.g.,

$$ds^2 = dp^T\Phi dp, \tag{2.19}$$

so that the distance between two points will be larger when the curvature of the potential is greater. Thus if you are moving at a constant rate, you'll go over curved parts more slowly, and flat parts faster. Although in this case, the metric is constant everywhere.

3 Bonus: a generating function for POVMs

Let us consider another potential,

$$\Psi(X) = -\mathrm{tr}\log(I - X) = -\log\det(I - X). \tag{3.1}$$

You may know that for $|x| < 1$,

$$-\log(1-x) = \sum_{m=1}^{\infty} \frac{x^m}{m}. \tag{3.2}$$

The same holds if we replace $1 \to I$ and $x \to X$ as long as the largest absolute value of X's eigenvalues is < 1 [15]. (This will be true, for example, for a density matrix.) Thus

$$\Psi(X) = \sum_{m=1}^{\infty} \frac{\mathrm{tr}(X^m)}{m}. \tag{3.3}$$

Our earlier potential (up to the prefactor) is just the second term in the Taylor expansion of this grand potential. Let's expand X in terms of POVM elements $X = \sum_i x_i E_i$ and substitute in (notice we're using a slightly different convention from before),

$$\Psi(x) = \sum_{m=1}^{\infty} \frac{1}{m}\mathrm{tr}\left(\left[\sum_i x_i E_i\right]^m\right) \tag{3.4}$$

$$= \sum_i x_i \mathrm{tr}(E_i) + \frac{1}{2}\sum_{ij} x_i x_j \mathrm{tr}(E_i E_j) + \frac{1}{3}\sum_{ijk} x_i x_j x_k \mathrm{tr}(E_i E_j E_k) + \ldots. \tag{3.5}$$

Taking the first derivatives,

$$\frac{\partial\Psi(x)}{\partial x_a} = \mathrm{tr}(E_a) + \frac{1}{2}\sum_j x_j \Big\{\mathrm{tr}(E_a E_j) + \mathrm{tr}(E_a E_j)\Big\} \tag{3.6}$$

$$+ \frac{1}{3}\sum_{jk} x_j x_k \Big\{\mathrm{tr}(E_a E_j E_k) + \mathrm{tr}(E_j E_a E_k) + \mathrm{tr}(E_j E_k E_a)\Big\} + \ldots,$$

and so

$$\frac{\partial\Psi(x)}{\partial x_a}\Big|_{x=0} = \mathrm{tr}(E_a). \tag{3.7}$$

Taking another derivative, we find

$$\frac{\partial^2\Psi(x)}{\partial x_a \partial x_b} = \frac{1}{2}\Big\{\mathrm{tr}(E_a E_b) + \mathrm{tr}(E_b E_a)\Big\} \tag{3.8}$$

$$+ \frac{1}{3}\sum_k x_k \Big\{\mathrm{tr}(E_a E_b E_k) + \mathrm{tr}(E_a E_k E_b) + \mathrm{tr}(E_b E_a E_k) + \mathrm{tr}(E_k E_a E_b) + \mathrm{tr}(E_b E_k E_a) + \mathrm{tr}(E_k E_b E_a)\Big\}\ldots,$$

so that, by the cyclicity of the trace,

$$\frac{\partial^2\Psi(x)}{\partial x_a \partial x_b}\Big|_{x=0} = \mathrm{tr}(E_a E_b). \tag{3.9}$$

Taking another deriative, we find

$$\frac{\partial^3\Psi(x)}{\partial x_a \partial x_b \partial x_c} = \frac{1}{3}\Big\{\mathrm{tr}(E_a E_b E_c) + \mathrm{tr}(E_a E_c E_b) + \mathrm{tr}(E_b E_a E_c) + \mathrm{tr}(E_c E_a E_b) + \mathrm{tr}(E_b E_c E_a) + \mathrm{tr}(E_c E_b E_a)\Big\} + \ldots. \tag{3.10}$$

Now the six terms in the brackets divide into two groups of three: the three in each group are all cyclically permuted versions of each other. We are left with

$$\frac{\partial^3 \Psi(x)}{\partial x_a \partial x_b \partial x_c}\Big|_{x=0} = \mathrm{tr}(E_a E_b E_c) + \mathrm{tr}(E_a E_c E_b) = 2\Re[\mathrm{tr}(E_a E_b E_c)]. \tag{3.11}$$

In other words, $\Psi(x)$ acts as a generating function for symmetrized traces of n-products of POVM elements: we extract them by taking ever higher order derivatives and evaluating at $x = 0$. For an IC-POVM, the 2-product gives a Gram matrix from which a dual frame can be constructed. The symmetrized 3-product teaches one how to take the Jordan product, from which the geometry of quantum state space flows [16].

References

[1] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, "The varieties of minimal tomographically complete measurements," *International Journal of Quantum Information* **19**, 2040005 (2021). doi:10.1142/S0219749920400055.

[2] C. A. Fuchs and R. Schack, "Quantum-Bayesian coherence," *Reviews of Modern Physics* **85**, 1693–1715 (2013). doi:10.1103/RevModPhys.85.1693.

[3] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, "Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems," *Physical Review Research* **2**, 013074 (2020). doi:10.1103/PhysRevResearch.2.013074.

[4] J. B. DeBrota, C. A. Fuchs, J. L. Pienaar, and B. C. Stacey, "Born's rule as a quantum extension of Bayesian coherence," *Physical Review A* **104**, 022207 (2021). doi:10.1103/PhysRevA.104.022207.

[5] M. B. Weiss, "Redesigning quantum theory," arXiv:2609.03333 [quant-ph] (2026). doi:10.48550/arXiv.2609.03333.

[6] S. F. D. Waldron, *An Introduction to Finite Tight Frames* (Birkhäuser, New York, 2018). doi:10.1007/978-0-8176-4815-2.

[7] C. Ferrie and J. Emerson, "Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations," *Journal of Physics A: Mathematical and Theoretical* **41**, 352001 (2008). doi:10.1088/1751-8113/41/35/352001.

[8] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970).

[9] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, 2002).

[10] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley, New York, 1985).

[11] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 2nd ed. (Cambridge University Press, Cambridge, 2017). doi:10.1017/9781139207010.

[12] R. Penrose, "A generalized inverse for matrices," *Proceedings of the Cambridge Philosophical Society* **51**, 406–413 (1955). doi:10.1017/S0305004100030401.

[13] S.-I. Amari and H. Nagaoka, *Methods of Information Geometry*, translated by D. Harada, Translations of Mathematical Monographs Vol. 191 (American Mathematical Society, Providence, 2000). doi:10.1090/mmono/191.

[14] H. Shima, *The Geometry of Hessian Structures* (World Scientific, Singapore, 2007). doi:10.1142/9789812707536.

[15] N. J. Higham, *Functions of Matrices: Theory and Computation* (SIAM, Philadelphia, 2008). doi:10.1137/1.9780898717778.

[16] M. B. Weiss, "Characterizing quantum state-space with a single quantum measurement," *Physical Review A* **111**, 052205 (2025). doi:10.1103/PhysRevA.111.052205.