First lectures on the algebraic approach to quantum physics

In this lecture, I want to introduce a formalism that can be used to shed some light on why things are what they are in the common presentation of quantum mechanics. Where do Hilbert spaces and wave functions come from? Why do the operators look the way they do? In what respect (at a physical level, not just in the formalism) is quantum physics different from classical physics?

In a typical introductory course of quantum mechanics, the wave function somehow falls from the sky. It comes with its probability interpretation giving meaning to $|\psi(x)|^2$, but the relevance of the phase is unclear at best (besides not being invariant under gauge transformations). Also, we learn, that the wave function is not really a representative of a state but only its ray in Hilbert space. When we look at interference experiments like the double slit, however, all the relevance is in the relative phase of wave functions.

So rather from these not really concrete objects of obscure empirical status, the algebraic approach starts with something really concrete: Experiments or observables. These are all possible measurement procedures in a given physical situation and we think of them as having a numerical outcome (rather than for example vectorial, as in that case we could ask for the individual components). At this point, we allow for the fact that our theory is probabilistic, that there could be experiments that have different outcomes when done multiple times. But this is not a restriction since deterministic theories are those where all probabilities are either 0 or 1.

All these observables we group together in a set $\mathfrak{A}_0$. In what follows, we are going to equip $\mathfrak{A}_0$ with more mathematical structure.

First of all, we find that if there is an experiment $A \in \mathfrak{A}_0$, there is also an experiment $2A$ which is the same as A but doubles all the values that it reads out. Similarly for any other scalar multiple. In a second step, we realize for two experiments $A, B \in \mathfrak{A}_0$, there is also $A + B \in \mathfrak{A}_0$ where we first do experiment A and then, with a fresh setup do B and add the results together (note that this is compatible the the linearity of expectation values). So we end up with $\mathfrak{A}_0$ being a real vector space.

This all looks very trivial. But take note of the fact that we do not attempt to define a product on $\mathfrak{A}_0$ in a similar way. Rather we do something else to end up with a product, something we will only be able ot sketch here: In the argument above that gave us $2A$, we don't have to stop here but can completely relabel the scale on the measuring device that displays the result of A, possibly in a non-linear way. Think for example writing new numbers on the scale on some old gauge with a pointer. This suggests that there should be some sort of basic functional calculus that for nice functions $f\colon \mathbb{R} \to \mathbb{R}$ yields a non-linearly rescaled observable $f(A)$. We could for example start to require that for f that is polynomial.

Then we would in particular have the square of an observable and for $A, B \in \mathfrak{A}_0$, we would have A^2, B^2, and $(A + B)^2$ which we could use to define a symmetrized product along the lines of $A \circ B = \frac{1}{2}((A + B)^2 - A^2 - B^2)$. This product can be show to turn $\mathfrak{A}_0$ into something known as a Jordan algebra which is a specific commutative but not associative algebra. We will not need the details but a structural theorem from their theory that says, they (or at least those of practical relevance) arise in the following way: There is a complex *-algebra $\mathfrak{A}$ that is associative but not necessarily commutative with an involution

$^*: \mathfrak{A} \to \mathfrak{A}$ with $(A^*)^* = A$ and $(A + \lambda B)^* = A^* + \bar{\lambda} B^*$ and $(AB)^* = B^* A^*$ such that

$$\mathfrak{A}_0 = \{A \in \mathfrak{A} | A^* = A\}$$

and

$$A \circ B = \frac{1}{2}(AB + BA).$$

So from now on, we are going to deal exclusively with the algebra $\mathfrak{A}$, have to live with the fact that the product is defined a bit in a round about way and actual measurements only correspond to the self-adjoint elements $A = A^*$. Note that is is the same situation we find in the ordinary description of quantum mechanics where only self-adjoint operators correspond to things you can actually measure while for self-adjoint operators A and B the product AB is only self-adjoint and thus directly measurable if A and B commute. At an operational level, the non-commutativity of quantum mechanics only occurs at the level of time evolutions where in general $e^{itA} e^{isB} \neq e^{isB} e^{itA}$.
In a further step, we want to equip $\mathfrak{A}$ with a norm (and thus a topology). To this end, we restrict our attention to experiments that have a bounded set of possible outcomes (which we can achieve by a non-linear relabelling as above, more to that below). Then we define $\|A\|$ to be the supremum of absolute values of possible outcomes. This is indeed a norm that is also compatible with the product and the involution:

$$\|AB\| \leq \|A\| \|B\| \qquad \text{and} \qquad \|A^*\| = \|A\|.$$

We can now add all the limit points of Cauchy sequences to $\mathfrak{A}$ to make it complete in this norm and in fact a Banach-*-algebra.
Now, it's only one remaining step, there is an additional property for which I have no justification except "it works": From the above, it follows that $\|A^* A\| \leq \|A^*\| \|A\| = \|A\|^2$. We will require that here for all a, equality holds. This is known as "the C*-property":

$$\|a^* A\| = \|A\|^2.$$

Now, we have everything together to motivate
Def. The observables from (the *-invariant part) of a **C*-algebra** $(\mathfrak{A}, ^*, \|\cdot\|)$, that is a complex Banach-*-algebra in which the C*-property holds.
With this definition, $\mathfrak{A}$ is not required to contain a unit element (an element $\mathbb{1}$ such that $\mathbb{1} A = A \mathbb{1} = A$ for all A). If it does not there is a unique way to include one in $\mathfrak{A}$ and even though many things we are going to say below are still true (or can be reformulated) if there is no unit, to simplify the presentation we will further on assume $\mathfrak{A}$ does have a unit.
Let us look at examples of C*-algebras. As you will have already guessed, all this is modelled on the most important case: If $\mathcal{H}$ is a Hilbert space, then $\mathfrak{B}(\mathcal{H})$, the bounded operators form a C*-algebra (the norm is the usual operator norm while the * is the adjoint). You can convince yourself that all the required properties in fact hold.
Next we realize that any closed sub-*-algebra of a C*-algebra is a C*-algebra itself. So you do not need all bounded operators on a Hilbert space to get a C*-algebra, a sub-algebra that is closed under adjoints and that is norm-closed is also a C*-algebra.

In fact, using the Hahn-Banach theorem it can be proven that every C*-algebra arises in this way for some Hilbert space. The existence of that Hilbert space, however, is not constructive (as so many thing that rely on Hahn-Banach they involve invoking the axiom of choice) and this Hilbert space is "too big" to be physically useful. It is much better to think of the elements of $\mathfrak{A}$ to be some abstract observables that have some relations like canonical commutation relations than concrete operators on a Hilbert space. This is like thinking about abstract linear maps instead of matrices or Lie-groups in an abstract way rather than sub-groups of the matrix groups $GL(n, \mathbb{R})$. We will see below that a better way to think of the relation between abstract C*-algebras and bounded operators on Hilbert space as the latter being representations of the former.
A particular finite dimensional special case is $\mathcal{H} = \mathbb{C}^N$ where the C*-algebra of operators are the complex matrices $Mat(N \times N, \mathbb{C})$. More generally, it can be shown that every finite dimensional C*-algebra is equivalent to one of the form

$$\bigoplus_{i=1}^{K} Mat(N_i \times N_i, \mathbb{C}).$$

These can be thought of block diagonal matrices with blocks of size N_i.
Another important example of C*-algebras arises in a geometric way: You start with a locally compact Hausdorff space X (this is a certain type of not too pathological topological space, details will not matter for us). On it, we consider the set $C_0(X)$ of continuous functions $f\colon X \to \mathbb{C}$ that vanish at infinity (defined to mean that for every $\epsilon > 0$ there is a compact set $K \subset X$ such that $|f(x)| < \epsilon$ for $x \in X \setminus K$.) This set becomes a C*-algebra by defining the sum and product of functions in a pointwise manner (i.e. $(fg)(x) := f(x)g(x)$), f^* as pointwise complex conjugation. As norm, we use the supremum

$$\|f\| = \sup_{x \in X} |f(x)|.$$

This is finite thanks to the "vanishes at infinity" property. Once more, it's easy to convince oneself that all requirements for a C*-algebra are fulfilled. It contains an identity if and only if X is compact, the key question being if the constant function that is 1 everywhere vanishes at infinity.
This is an example of a commutative C*-algebra since the pointwise product is commutative. What might be surprising is that all commutative C*-algebras are of this from, this is known as Gelfand's theorem. More specifically, if $\mathfrak{A}$ is a commutative C*-algebra, there is a locally compact Hausdorff space $X(\mathfrak{A})$ such that $\mathfrak{A}$ is equivalent to $C_0(X(\mathfrak{A}))$.
To see how it comes about one has to find linear functionals $\pi\colon C_0(X) \to \mathbb{C}$ that are compatible with products and stars, i.e. $\pi(fg) = \pi(f)\pi(g)$, $\pi(f^*) = \overline{\pi(f)}$ in other words irreducible (one dimensional thanks ot a version of Schur's lemma) representations (for more on representations and their general definition see below). Obviously, there is one for each point x via $\pi_x(f) := f(x)$. Some contemplation reveals all irreducible representations arise in this way. Consequently, we can identify points of X with irreducible representations. Thus, as a set, we can define $X(\mathfrak{A})$ to be the set of irreducible representations.
In a second step, we have to equip it with a topology. Abstractly, this is the weak-*-topology. But there is a more intuitive way to define it: To this end note that if you take all

functions in $C_0(X)$ that vanish on some given subset $U \subset X$ than due to continuity, they all also vanish on the closure of U. Again, we can turn this around: For a subset of irreducible representations of $\mathfrak{A}$, we define the closure as all those irreducible representations that vanish on all of those functions. In terms of these closures we take all complements to be the open sets. Then finally, one has to prove that $X(\mathfrak{A})$ with this topology is indeed locally compact and Hausdorff.
As far as the physics is concerned, this special case of commutative C*-algebras corresponds to the observables of classical physics: Think of X as phase space. Then the classical observables are exactly the functions $C_0(X)$ on this phase space.
To see what is going one, it helps to look at the even more special case of a finite dimensional commutative C*-algebra, in other words an algebra of matrices that all mutually commute (which are thus normal). Thanks to the spectral theorem for matrices they can all be diagonalized simultaneously. Thus, in a convenient basis, each matrix only contains non-zero values on its diagonal which is the same as a function on a discrete set of points (and the product of diagonal matrices being the pointwise product). Thus, in this case, $X(\mathfrak{A})$ is simply this set of discrete points and the Gelfand theorem above is really a more sophisticated version of a spectral theorem.
A nice corollary is a functional calculus for normal elements in a general C*-algebra: Recall that an element $A \in \mathfrak{A}$ is normal if $A^*A = AA^*$. So, in particular, if A is self-adjoint it is also normal. For such a normal A we can then define a commutative C*-algebra generated by A and A^*: Start from all polynomials in A and $A*$ (where the order does not matter thanks to normality) and complete it the norm. This is obviously a closed sub-*-algebra and thus C* itself. With Gelfand, we have a space X such that all elements of this sub-*-algebra are realized as complex functions on that space. In particular there is a function $f_A\colon X \to \mathbb{C}$ that corresponds to A. If we then take a continuous function $F\colon \mathbb{C} \to \mathbb{C}$, we can take $F(A)$ to be that element of the sub-*-algebra that corresponds to the concatenation $F \circ f\colon X \to \mathbb{C}$. Convince yourself that for matrices as well as for polynomial F this does the right thing.
In a C*-algebra, you can also define a resolvent set for an element A as those complex λ for which $A - \lambda\mathbb{1}$ is invertible in $\mathfrak{A}$ and the spectrum $\sigma(A)$ as its complement. Finally, there is the notion of "spectral radius" which is $r(A) := \sum_{\lambda\in\sigma(\mathfrak{A})} |\lambda|$. You can then adopt the proof given in lecture for operators on Hilbert space that for self-adjoint $A \in \mathfrak{A}$, the spectral radius coincides with the norm to the C*-setting. Thanks to the C*-property, we have then for general (not necessarily self-adjoint) elements

$$r(A^*A) = \|A^*A\| = \|A\|^2.$$

What is interesting about this is that the spectral radius was defined only using algebraic properties of $\mathfrak{A}$. We conclude, that in a C*-algebra, the norm is unique: If you know that some *-algebra has a norm that turns it into a C*-algebra, you can find it only using algebraic properties.
Now that we became familiar with the most important properties of C*-algebras, it is time to look at the corresponding morphisms.
Def. Let $\mathfrak{A}$ and $\mathfrak{B}$ be C*-algebras. A C*-morphism is a linear map $\Phi\colon \mathfrak{A} \to \mathfrak{B}$ that respects the algebraic structure, i.e.

$$\Phi(A_1A_2) = \Phi(A_1)\Phi(A_2), \qquad \Phi(A^*) = \Phi(A)^*.$$

Note that we do not have to assume anything about the norm and continuity: From the above it already follows that $\|\Phi(A)\| \leq \|A\|$, that is C*-morphisms are automatically contracting and thus continuous. This can be seen by realizing that an element of the resolvent set of a is automatically in the resolvent set of $\Phi(A)$ and then the relation of the norms follows from the spectral radius formula for the norm.

Similarly, C*-isomorphisms are C*-morphisms where have an inverse that is also a C*-homomorphism (and thus automatically norm-preserving). Finally, C*-automorphisms are C*-isomorphisms from $\mathfrak{A}$ to itself. When I said above that a C*-algebra is equivalent to something else, what I actually meant was that there is a C*-isomorphism between the two.

Def. A representation of a C*-algebra $\mathfrak{A}$ is a C*-morphism $\pi: \mathfrak{A} \to \mathfrak{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$.

You can think of a representation of the abstract algebra $\mathfrak{A}$ of observables as concrete operators on the Hilbert space $\mathcal{H}$.

As always in representation theory, you can define direct sums of representations and irreducible representations (those that do not have invariant sub-spaces or in other words that cannot be written as non-trivial direct sums of other representations).

If there are two representations $\pi_{1/2}: \mathfrak{A} \to \mathfrak{B}(\mathcal{H}_{1/2})$ we say they are (unitarily) equivalent if there is a unitary $U: \mathcal{H}_1 \to \mathcal{H}_2$ such that for all $A \in \mathfrak{A}$ we have $U\psi_1(A) = \pi_2(A)U$ and think of them as "the same".

Finally, you can ask about the representation theory of the algebra $\mathfrak{A}$. In particular, is there more than one irreducible representation up to equivalence (again, Hahn-Banach together with the GNS-construction we discuss below guarantees the existence of representations)? If this question is answered in the negative, all this abstraction is pretty void and instead of $\mathfrak{A}$ we could have directly dealt with the operators $\pi(\mathfrak{A}) \subset \mathfrak{B}(\mathcal{H})$ in a concrete Hilbert space realisation. But if there is more than one it makes sense to abstract from them and ask for properties that exist already in the abstract algebra $\mathfrak{A}$ rather than one of its contrete realisations.

Before we end this first part, let us address an almost philosophical question: Why only bounded operators? We had just convinced ourselves that for interesting quantum systems we need to also consider unbounded operators that bring with them their problems with domains of definition, lack of continuity etc. Why do we ignore those in the algebraic approach? Don't we miss all the physics this way? How can we ever deal with canonical variables that obey $[x, p] = i$ which imply $\|x\|\|p\| = \infty$?

And indeed there is no good algebraic theory that includes unbounded operators as that would mean giving up on the norm of the algebra and thus its topology. In this case, the theory of *-algebras is much much poorer in structure.

But it turns out, for many questions we can avoid using unbounded observables. The most urgent case are of course the canonical variable x and p with $[x, p] = i$. It turns out, we can equally well start from the unitaries $U(s) = e^{isx}$ and $V(t) = e^{itp}$ for $s, t \in \mathbb{R}$ (or in fact other families of bounded functions of x and p but this choice is used most often), that have commutation relations $U(s)V(t) = e^{ist}V(t)U(s)$. Trivially, they also satisfy $U(s_1)U(s_2) = U(s_1 + s_2)$ and $V(t_1)V(t_2) = V(t_1 + t_2)$. These are called Weyl operators and since they are unitary they have $\|U(S)\| = \|V(t)\| = 1$.

To have a C*-algebra, we start from abstract operators with these commutation relations, that is we take all polynomials in $U's$ and $V's$ for different s's and $t's$, use the above relations to simplify all products to a single U and V factor and complete this space in norm. The C*-algebra obtained that way is called the Weyl-algebra or CCR for "canonical commutations relations".
One could have the idea to be able to recover x from $U(s)$ by taking a derivative with respect to s and similarly for p and $V(t)$. But in the C*-algebra, as one can prove (see exercise below) that

$$\|U(s) - \mathbb{1}\| = \begin{cases} 2 & \text{for } s \neq 0 \\ 0 & \text{else} \end{cases}$$

and thus what would be the derivative

$$\lim_{s\to 0} \frac{U(s) - \mathbb{1}}{s}$$

does not exist. If we have, however, a representation $\pi: CCR \to \mathfrak{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$, we can ask for which $\psi \in \mathcal{H}$ we have existence of

$$\lim_{s\to 0} \frac{\pi(U(s)) - \pi(U(0))}{is}\psi$$

and use that as the domain of definition of the operator x *in that representation*. A necessary condition on the representation is that the map

$$s \mapsto \pi(U(S))$$

is continuous in the weak or strong (this is the same here) operator topology and similarly for $V(t)$. Such a representation is called *regular*. The usual representation where $\mathcal{H} = L^2(\mathbb{R})$ and $\pi(U(s))$ is the multiplication operator e^{isx} and $(\pi(V(t))\psi)(x) = \psi(x+t)$, i.e. a translation operator is called the Schrödinger representation. It is obviously regular. A famous theorem by Stone and von Neuman states that up to unitary equivalence, this is the unique regular irreducible representation of CCR. And this in retrospect can be taken as the explanation of why in introductory quantum mechanics nobody questions the role of wave functions and the fact that the operators x and p act in the way they always act (there is of course also the momentum representation but thanks to the Fourier transform being unitary this representation is unitarily equivalent).
If one drops the requirement of regularity, there are other inequivalent representations but they have at best marginal use in physics (the loop approach to quantum gravity, however, is built on such representations).
This argument generalizes directly to higher dimensions with $[x_\alpha, p_\beta] = i\delta_{\alpha\beta}$ and $\alpha, \beta \in \{1, 2, \ldots, n\}$. In infinite dimensions (or better: infinitely many degrees of freedom) is is, however, explicitly wrong! In quantum field theory, there are many inequivalent representations of the canonical commutation relations. This also has a lot of relevance in quantum statistical physics in the thermodynamic limit. There for example states (in infinite volume) with different finite particle density correspond to inequivalent representations as well as systems with phase transitions. This will be the subject of the mathematical statistical physics course in the summer.

Problem 1:
Instead of two types of generators, one can also write the CCR algebra in terms of operators $W(z)$ for $z \in \mathbb{C}$ and

$$W(s+it) = e^{-ist/2}U(s)V(t)$$

with relations $W(z_1)W(z_2) = e^{i\Im(z_1\bar{z}_2)/2}W(z_1+z_2)$. (In fact, this can be defined for any symplectic vector space, that is a real vector space with a non-degenerate anti-symmetric bi-linear form σ. In this case only replace $\Im(z_1\bar{z}_2)$ by the symplectic form σ). Show that all $W(z)$ are unitary. Thus their spectrum has to be contained in the unit circle.
Then compute $W(u)W(z)W(-u)$ and use this to argue that the spectrum of $W(z)$ is either empty (this is impossible) or it is all of the unit circle. You can use that the spectrum is invariant under unitary conjugation.
This tells you about the spectrum of $W(z) - \mathbb{1}$. Finally conclude that for $z \neq 0$

$$\|W(z) - \mathbb{1}\| = 2$$

.

Problem 2: Prove the Stone-von Neumann theorem. Assume that you have a regular irreducible representations $\pi\colon CCR \to \mathfrak{B}(\mathcal{H})$. Show first that

$$P := \frac{1}{2\pi}\int d^2z\, \pi(W(z))e^{-|z|^2/4}$$

(defined in terms of matrix elements) is an orthogonal projection. Show also that

$$P\pi(W(z))P = e^{-|z|^2/4}P.$$

If $\Omega \in P\mathcal{H}$ normalized, show that the span of all $\pi(W(z))$ for $z \in \mathbb{C}$ is invariant under the action of all Weyl-operators W. From the irreducibility of the representation conclude that $P\mathcal{H}$ is one dimensional.
Use this to construct the unitary equivalence to two regular irreducible representations of CCR.

Problem 3: Irregular representations. Construct an irreducible representation of CCR on a Hilbert space of functions $\psi : \mathbb{R} \to \mathbb{C}$ where $\pi(U(s))$ is the multiplication operator e^{isx} and $(\pi(V(t))\psi)(x) = \psi(x+t)$, both act as unitary operators. Different from the Schrödinger representation, let $\mathcal{H}$ contain the function that is 1 everywhere. Show that this representation is inequivalent to the Schrödinger representation.

Now, that we have some understanding of observables, it is time to look at the notion of states of the theory. Informally, a state is an equivalence class of preparations of the system where two preparations are thought of as preparing the same state if you cannot distinguish with measurements which one was actually performed. Since we allow for the possibility of probabilistic theories, this means specifically that the expectation values for all observables should be the same in both preparations.
One might wonder why it is sufficient to consider only expectation values. After all, there are also variances and higher moments. But those are already covered as they can be expressed in terms of expectation values of an observable and its powers. For example the variance of A is simply the expectation value of A^2 minus the square of the expectation of A.
So, in fact, we will write a state as a map that assigns to each observable its expectation value: $\omega\colon \mathfrak{A} \to \mathbb{C}$. In order for this to give sensible expectation values, we will require ω to have three properties: First of all, it should be a linear map (or a functional from the dual of $\mathfrak{A}$) as this is compatible with how we originally set up the linear structure on $\mathfrak{A}$.
Second, it should be normalized. This can be expressed as ω having the operator norm 1. But in many practical situations it is advantageous to use the equivalent condition

$$\omega(\mathbb{1}) = 1.$$

Finally, besides the notion of self-adjoint elements, there is a notion of positive operators. These are operators that can be written as A^*A for some $A \in \mathfrak{A}$. So as a last condition, we require that ω is positive for all positive operators

$$\omega(A^*A) \geq 0.$$

Taking all this together, we define a state of a C*-algebra to be a normalized and positive linear functional. The different nature of these three properties gives the set of all states an interesting form: The set of all functionals is obviously a vector space. Then, the requirement of positivity restricts us to a cone in that vector space (in a cone, the sum of two elements as well as multiples by non-negative real numbers are also elements). Finally, the normalization condition intersects this cone with an affine hyperplane of co-dimension 1.
The set of states is a convex space: With two states $\omega_{1/2}$ also their "mixture"

$$\lambda\omega_1 + (1-\lambda)\omega_2$$

is a state for $\lambda \in [0,1]$. You could for example prepare this mixture by throwing an uneven coin and depending on the outcome either prepare ω_1 or ω_2.
Let us also look at some examples of states for the different types of C*-algebras we considered above. First of all, there is the traditional quantum mechanical notion of a state, that is when $\mathfrak{A} = \mathfrak{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$, we think of non-zero $\psi \in \mathcal{H}$ as the states of the quantum system, or better actually, of the rays $\mathbb{C}^\times\psi$. And indeed, such a ray determines also a state in the new sense of expectation value functional:

$$\omega_\psi(A) := \frac{\langle\psi|A|\psi\rangle}{\|\psi\|^2}.$$

At this point, it is worth pointing out that thanks to the positivity and normalization conditions on states, there is no notion of "superposition of states"! This fits with the observation that superpositions are not defined for rays in Hilbert space because the relative phase of two elements matters when taking the sum. We will later find that this notion of superposition is actually better thought of as a property of the algebra $\mathfrak{A}$ than of the states (for example by using a Hadamard-gate as time evolution and then viewing this in the Heisenberg representation modifying the observables in the case of a qubit).
In the finite dimensional special case of $\mathfrak{A} = Mat(N \times N, \mathbb{C})$ we can even ask if we can determine all states. We start by the observation that linearity of $\omega(A)$ implies that it has to be a linear combination of the matrix elements of the matrix A:

$$\omega(A) = \sum_{i,j=1}^{N} \rho_{ji} a_{ij}$$

for some coefficients $\rho_{ji} \in \mathbb{C}$. But those can also be grouped into an $N \times N$ matrix ρ and the above expression rewritten into

$$\omega(A) = \mathrm{tr}(\rho A).$$

Next, we impose the normalization:

$$1 = \omega(\mathbb{1}) = \mathrm{tr}(\rho \mathbb{1}) = \mathrm{tr}(\rho).$$

And finally, for the positivity, we can use the cyclicity of the trace:

$$0 \leq \omega(A^* A) = \mathrm{tr}(\rho A^* A) = \mathrm{tr}(A \rho A^*).$$

If we think of A to be built from N row vectors $v_i^\dagger$, then we can express this inequality as

$$0 \leq \sum_i \langle v_i | \rho v_i \rangle.$$

This holds for all possible choices of v_i exactly if the matrix $\rho \geq 0$ is positive semi-definite (or positive in the language of operators).
Taking everything together, in the finite dimensional case of $N \times N$ matrices, all states are of the form $\omega(A) = \mathrm{tr}(\rho A)$ for a positive semi-definite (and thus in particular self-adjoint) matrix ρ of trace 1, in other words ρ is a density matrix. And those are indeed all the states in quantum statistical mechanics!
One can try to go back to the potentially infinite dimensional case of $\mathfrak{B}(\mathcal{H})$ and try to argue a generalization to be true. And indeed, there is also the notion of a density operator ρ: This is defined to be a trace-class operator on $\mathcal{H}$ (otherwise we could not talk about the normalization) that is positive $\rho \geq 0$ in the operator sense and normalized as $\mathrm{tr}(\rho) = 1$. And indeed this gives rise to a state on $\mathfrak{B}(\mathcal{H})$ with the same expression

$$\omega_\rho(A) = \mathrm{tr}(\rho A),$$

where it is useful to recall an earlier homework that showed that the product of a trace class operator with a bounded operator is again traceclass. There is only one complication in infinite dimensions that is that the traceclass operators are not the topological dual of $\mathfrak{B}(\mathcal{H})$ which is the property that we used to find all the linear functionals. It is, however, true in the other direction $\mathcal{S}^1(\mathcal{H})^* = \mathfrak{B}(\mathcal{H})$, that is, they are the pre-dual. But strictly speaking, there are more states on $\mathfrak{B}(\mathcal{H})$ than those that can be written in terms of traceclass density operators. But this will not play a role in the rest of what we will be concerned with.
It is worth noting, that of course the states arising from rays in Hilbert space are the special case of density operators having rank one:

$$\rho = \frac{|\psi\rangle\langle\psi|}{\|\psi\|^2}.$$

Finally, let us also consider the commutative case of $\mathfrak{A} = C_0(X)$. Here, the easiest way to assign a number to $f : X \to \mathbb{C} \in C_0(X)$ of course arises from points $x \in X$ as

$$\omega_x(f) := f(x).$$

Clearly, this is linear, positive and normalized. And indeed, with the identification of X as phase space, the points of phase space correspond to the states of Hamiltonian mechanics. Furthermore, again, we can ask about all states in this case. Here, according to the Riesz-Markov theorem, all linear functionals on $C_0(X)$ are measures μ on X (with the points $x \in X$ corresponding to the point measure concentrated at x). Normalization simply means that $\int_X 1 d\mu = 1$ and positivity translates into the measure being positive, that is for all $f \geq 0$ we have $\int_x f d\mu \geq= 0$. So the states on the commutative C*-algebra can be identified with the probability measures on X. Again, this is not much of a surprise as these are exactly the states of classical statistical mechanics (if one thinks again of X as phase space).
It is quite remarkable that the simple notion of "state" in this algebraic context manages to unify the different looking notions of "state" in classical and quantum mechanics and classical and quantum statistical mechanics!
Above, we noted that the set of states is a convex space. There is a special role of the extremal points of this state space, that is, states that cannot be written as non-trivial mixture of two states (a mixture is of course trivial if in $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$ either $\omega_1 = \omega_2 = \omega$ or $\lambda \in \{0, 1\}$ that is $\omega = \omega_1$ or $\omega = \omega_2$). These extramal states are called "pure states" and it is not hard to convince oneself that in the case $\mathfrak{A} = \mathfrak{B}(\mathcal{H})$ the pure states are the rank one states ω_ψ with $\rho = |\psi\rangle\langle\psi|/\|\psi\|^2$ and in the classical case are the states ω_x arising from points $x \in X$.
It is worth noting, that every state can be decomposed as a generalized mixture of pure states (where the convex combination is generalized to a probability measure on the space of pure states). The key difference, however, is that in the classical case $C_0(X)$ this decomposition is unique (as the probability measure on X is trivially only expressed as itself as a probability measure on the set of pure states) whereas in the quantum case the decomposition is not unique which can easily be seen in the case of the qubit with

$\mathcal{H} = \mathbb{C}^2$ and the state space being the Bloch sphere and the pure states being those on the boundary of the Bloch sphere. Some people even argue that this non-uniqueness is the key difference between classical and quantum physics as this turns the quantum theory to be non-realist as there is no unique decomposition of states into probabilistic mixtures into underlying "true states" in which the observables have well defined values.
As we noted above, the algebraic notion of "state" does not allow to define the superposition of two states, even though that seems to be a crucial ingredient in quantum mechanics. Even for states that are pure states arsising from rays in Hilbert space, there is no way to superimpose ω_{ψ_1} and ω_{ψ_2}. The formalism only allows for mixtures but those are blind to relative phases between ψ_1 and ψ_2. The thought experiment that shows most clearly the difference between mixtures and superpositions is the double slit: In mixtures, the classical probabilities add while the interference between the paths through the two slits is only modelled by the superposition of the wave functions of the two possible paths.
In the algebraic framework, this is encoded in the non-commutativity of the observable algebra: For example, if in the double slit experiment, the observable which measures which slit the electron has taken is represented by the Pauli matrix σ_z with the two values ± 1 corresponding two the two slits, the unitary operator

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

(known as the Hadamard gate) prepares the state that is the superposition between the two possible paths. Once more, this quantum feature of interference can be attributed to the fact that $U\sigma_z \neq \sigma_z U$. In a classical theory where all observables commute, creating such superpositions is impossible.
As a last section of these notes, we want to connect all ingredients that we have introduced so far, C*-algebras, states and representations on Hilbert spaces into one theorem. And this is the "GNS construction" named after its inventors Gelfand, Naimark, and Segal. It is related to highest weight representations in the representation theory of Lie algebras and given a C*-algebra and a state constructs a Hilbert space and a representation where this state is realized as a vector state.
Specifically, given a C*-algebra $\mathfrak{A}$ and a state $\omega\colon \mathfrak{A} \to \mathbb{C}$, it constructs $(\mathcal{H}_\omega, \pi_\omega, \Omega_\omega)$ where $\mathcal{H}_\omega$ is a Hilbert space, $\pi_\omega\colon \mathfrak{A} \to \mathfrak{B}(\mathcal{H}_\omega)$ is a representation and $\Omega_\omega \in \mathcal{H}_\omega$ is normalized such that it realizes ω as

$$\omega(A) = \langle \Omega_\omega | \pi_\omega(A) \Omega_\omega \rangle.$$

Furthermore, there is a simplicity property which is that Ω_ω is "cyclic, that is that $\pi_\omega(\mathfrak{A})\Omega_\omega$ is dense in $\mathcal{H}_\omega$.
Before we can explain the GNS-construction in detail, we need a lemma that is a variant of a Cauchy-Schwarz inequality. It is obtained from requiring positivity of a state for elements of the form $A + \lambda B$ for $A, B \in \mathfrak{A}$ and $\lambda \in \mathbb{C}$:

$$0 \leq \omega((A+\lambda B)^*(A+\lambda B)) = \omega(A^*A) + \lambda\omega(A^*B) + \bar{\lambda}\omega(B^*A) + \lambda\bar{\lambda}\omega(B^*B).$$

This expression is positive and thus in particular real. The first and the last term are also positive so the imaginary part of the sum of the second and third term has to vanish. This

implies

$$\omega(A^*B) = \overline{\omega(B^*A)}.$$

Then, since the above inequality has to hold for all λ we can extremize with respect to it (formally treating λ and $\bar{\lambda}$ as independent). That yields

$$\lambda = -\frac{\omega(B^*A)}{\omega(B^*B)}$$

(if $\omega(B^*B) = 0$ consider $\lambda A + B$ instead) and plugging this back in (using the above equation for the complex conjugate of the state)

$$0 \leq \omega(A^*A) - \frac{\omega(B^*A)\omega(A^*B)}{\omega(B^*B)} - \frac{\omega(A^*B)\omega(B^*A)}{\omega(B^*B)} + \frac{\omega(A^*B)\omega(B^*A)\omega(B^*B)}{\omega(B^*B)^2}$$

or in other words the Cauchy-Schwarz style inequality

$$|\omega(A^*B)|^2 \leq \omega(A^*A)\omega(B^*B).$$

Now, we are equipped to approach the GNS construction. The first step is to define

$$\mathcal{J} := \{A \in \mathfrak{A} | \omega(A^*A) = 0\}$$

and observe that it is a right-ideal, that is if $A \in \mathfrak{A}$ and $J \in \mathcal{J}$ we have $AJ \in \mathcal{J}$ as using our Cauchy-Schwarz inequality

$$|\omega((AJ)^*(AJ))|^2 = |\omega(J^*A^*AJ)|^2 \leq \omega(J^*J)\omega((A^*AJ)^*(A^*AJ) = 0.$$

Then we define $\mathcal{H}_0 = \mathfrak{A}/\mathcal{J}$ and denote the equivalence classes as $A + \mathcal{J}$. Next we define a map π_0 from $\mathfrak{A}$ to linear operators on $\mathcal{H}_0$ via

$$\pi_0(A)(B + \mathcal{J}) := (AB) + \mathcal{J}$$

which is well defined thanks to the fact that $\mathcal{J}$ is a right ideal. This already has the property $\pi_0(A)\pi_0(B) = \pi_0(AB)$. Then, we are ready to define a sesqui-linear form on $\mathcal{H}_0$ using the state ω

$$\langle A + \mathcal{J} | B + \mathcal{J} \rangle := \omega(A^*B).$$

Again this is independent of the choice of representatives thanks to the Cauchy-Schwarz inequality. It is positive definite since we modded out the potential null space $\mathcal{J}$. With this scalar product on $\mathcal{H}_0$ comes a norm which we then use to define $\mathcal{H}_\omega$ to be the norm-closure of $\mathcal{H}_0$ and extend π_0 by continuity to $\pi_\omega : \mathfrak{A} \to \mathfrak{B}(H_\omega)$ which is a representation as also $\pi_\omega(A^*) = \pi_\omega(A)^\dagger$. As a last step, we define

$$\Omega_\omega := \mathbb{1} + \mathcal{J}$$

and compute

$$\begin{aligned}\langle\Omega_\omega|\pi_\omega(A)\Omega_\omega\rangle &= \langle\mathbb{1}+\mathcal{J}|\pi_\omega(A)(\mathbb{1}+\mathcal{J})\rangle\\ &= \langle\mathbb{1}+\mathcal{J}|A+\mathcal{J}\rangle\\ &= \omega(A).\end{aligned}$$

Further $\pi_0(\mathfrak{A})\Omega_\omega = \mathfrak{A}+\mathcal{J} = \mathcal{H}_0$ which is by definition dense in $\mathcal{H}_\omega$.
This finishes the GNS construction. It connects the task to find representations of $\mathfrak{A}$ to the task of finding states of $\mathfrak{A}$. Clearly, if you already have a representation $\pi\colon\mathfrak{A}\to\mathfrak{B}(\mathcal{H})$, for every non-zero $\psi\in\mathcal{H}$, you can then run the GNS-construction on ω_ψ and obtain a representation which is equivalent to $\pi(\mathfrak{A})\psi$ which is all of $\mathcal{H}$ if ψ is cyclic or π is irreducible. Similarly, you can do the GNS construction for ω_ρ for $\rho\in\mathcal{S}^1(\mathcal{H})$ a density matrix. Note that in this case even mixed states are represented as vector states Ω_{ω_ρ} in $\mathcal{H}_{\omega_\rho}$. This shows that pure states are not the same as vector states when $\mathfrak{A}$ is represented as a proper sub-algebra of $\mathfrak{B}(\mathcal{H})$.
With two states $\omega_{1/2}$ of $\mathfrak{A}$, you can ask if you can realize ω_1 as a density operator in the GNS-construction built on ω_2. In that case, one says ω_1 is in "the folium" of ω_2. The question of existence of inequivalent representations can then be formulated as the question of existence of states beyond the folium of one representation.
Finally, the statement that any C*-algebra is equivalent to a closed sub-algebra of $\mathfrak{B}(\mathcal{H})$ for some (very large) Hilbert space $\mathcal{H}$ can now be phrased as the existence of a faithful state, that is a state for which $\omega(A^*A)=0$ implies $A=0$. This is then in fact constructed using the Hahn-Banach theorem.