Abstract Hamiltonian mechanics

Start with a symplectic manifold. This is a manifold $X$ together with a symplectic form $\omega$. $\omega$ is a 2-form, so $\omega(v, w) = -\omega(w, v)$ is a function from $X$ to $\mathbb{R}$, for vector fields $v$ and $w$. We can interpret $\omega v$ as a covector field, defined by $\langle \omega v, w \rangle := \omega(v, w)$. $\omega$ is nondegenerate, meaning $\omega = 0$ only if $v = 0$. Since $\omega$ is nondegenerate, any covector field $\alpha$ is $\omega v$ for some vector field $v$, and $v$ is also called `$\omega v$'. So, $\omega$ gives a bijective correspondence between the covector fields and the vector fields. Incidentally, any symplectic manifold must have even dimension.

The mechanical system is specified by a hypersurface in $X$. A hypersurface is given locally as the solution set to the equation $f = 0$, where $f : X \to \mathbb{R}$ is a submersion. A solution to the system is given by an unparametrised curve $c$ whose image lies entirely within the hypersurface. Given a hypersurface, find a function $f$ which describes it, and make $f$ a coordinate in any way you like. There are many ways to do this, and they give different values for the vector field $\partial f = \partial / \partial f$, but it won't matter how you do it. For every covector field $\alpha$, $c$ must satisfy the equation $\langle \alpha, c' \rangle = \langle df, \omega \alpha \rangle \langle \omega \partial f, c' \rangle$. This is enough to determine $c$, given the initial condition of a point on the hypersurface, because $c$ is an unparametrised curve. The equation gives the ratio $\alpha / \omega f$ along the curve for any covector field $\alpha$, so the theory of differential equations determines $c$. The only possible problem is that $\langle \omega \partial f, c' \rangle$ might be 0, but my choosing $\omega \partial f$, rather than some arbitrary covector field, avoids this. Notice that $\langle df, \omega df \rangle = 0$, so the curve remains on the hypersurface $f = 0$.

Suppose you have already given $X$ a coordinate system. Make it a symplectic coordinate system, meaning the coordinates are $q^i$ and $p_i$ (equal numbers of each; remember $X$ has even dimension), such that $\omega(\partial_{q^i}, \partial_{q^j}) = 0$, $\omega(\partial_{p_i}, \partial_{p_j}) = 0$, and $\omega(\partial_{q^i}, \partial_{p_j}) = \delta_i^j$. Then $\omega \partial_t = dp_i$, and $\omega \partial_{p_i} = -dq^i$. Suppose you can solve the hypersurface for one of the coordinates, say $h := q^0$. Then you can give $f$ the form $h - H$, where $H$ is determined by all the coordinates besides $h$. To find a $\partial f$, the easiest way is to replace the coordinate $h$ by the coordinate $f$, so $\partial f$ in the new coordinates equals $\partial h$ in the old coordinates. Expressing the equation of motion in the old (symplectic) coordinates, $\langle \alpha, c' \rangle = \langle dh - dH, \omega \alpha \rangle \langle \omega \partial h, c' \rangle$. $\omega \partial h$ is $dt$, where $t := p_0$. Since I know $\langle dt, c' \rangle$ is non-zero, this makes $t$ a good variable to parametrise $c$ by, justifying the notation `$t$'. Remembering that $H$ is independent of $h$ and putting $dt$ for $\alpha$, $\langle dt, c' \rangle = \langle dh - dH, \partial h \rangle \times \langle dt, c' \rangle = 1 \langle dt, c' \rangle + 0 \langle dt, c' \rangle$, which is as it should be. $\langle dh, c' \rangle = \langle dh - dH, -\partial h \rangle \langle dt, c' \rangle = (\partial H / \partial t) \langle dt, c' \rangle$, so $dh / dt = \partial H / \partial t$ along $c$. If $i \neq 0$, $\langle dq^i, c' \rangle = \langle dh - dH, \partial p_i \rangle \langle dt, c' \rangle = (\partial H / \partial p_i) \langle dt, c' \rangle$, so $dq^i / dt = \partial H / \partial p_i$ along $c$. Finally, if $i \neq 0$, $\langle dp_i, c' \rangle = \langle dh - dH, \partial p_i \rangle \partial p_i \langle dt, c' \rangle = -\langle \partial H / \partial p_i \rangle \langle dt, c' \rangle$, so $dp_i / dt = -\partial H / \partial q^i$ along $c$. Also note that $dH / dt = dh / dt = \partial H / \partial t$ along $c$, because $df / dt = 0$ along $c$ (indeed, $f = 0$ along $c$, so $H = h$ along $c$). These results are Hamilton's equations; $H$, a function of every symplectic coordinate except $h$, is the Hamiltonian.

The next step is symplectic reduction, in which the unnecessarily large manifold $X$ is replaced by another symplectic manifold 2 dimensions smaller. You might have noticed that the set of solution curves has that dimension. You can construct it explicitly as the submanifold $X'$ of $X$ which is the intersection of $t = 0$ and $H = h$. In well-behaved cases, this indeed forms a manifold. If $v, w$ are vector fields tangent to $X'$, then $v$ are vector fields tangent to $X$ satisfying $\langle dh - dH, v \rangle = \langle dt, v \rangle = \langle dh - dH, w \rangle = \langle dt, w \rangle = 0$, let $\omega'(v, w)$ be $\omega(v, w)$. That is, $\omega'$ is the pullback of $\omega$ to $X'$. Coordinating for $X'$ are the $q^i$ and $p_i$ for $i \neq 0$. These coordinates continue to be symplectic on $X'$. $X'$ is the phase space of the system. A point in $X'$ indicates a possible state of the system. Any function $A$ on $X$ pulls back to a function on $X'$; this pullback is regarded as $A_0$, so $A_0(P)$ for $P \in X'$ indicates the value of the observable $A$ at time $0$ for the state $P$. To get values at other times, use the formula $dA / dt = \{ A, H \} + \partial A / \partial t$ for $\{ \}$ the Poisson bracket associated with $\omega$.

You can of course have more than 1 constraint. Just apply them sequentially. Of course, you can only choose $t$ once, which is why physicists will treat the first constraint specially, believing it to be connected to the true Hamiltonian and the true time, while the others are the trickier Hamiltonian constraints.