Abstract Hamiltonian mechanics

Start with a symplectic manifold. This is a manifold X together with a symplectic form ω. ω is a 2form, so $\omega(v,w) = -\omega(w,v)$ is a function from X to $\mathbf{R}$, for vector fields v and w. We can interpret ωv as a covector field, defined by $\langle \omega v, w\rangle := \omega(v,w)$. ω is nondegenerate, meaning $\omega v = 0$ only if $v = 0$. Since ω is nondegenerate, any covector field α is ωv for some vector field v, and v is also called '$\omega\alpha$'. So, ω 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 $f\colon X \to \mathbf{R}$ 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 coördinate 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 α, c must satisfy the equation $\langle \alpha, c'\rangle = \langle \mathrm{d}f, \omega\alpha\rangle\langle \omega\partial_f, c'\rangle$. 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 $\alpha/\omega\partial_f$ along the curve for any covector field α, 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 $\langle \mathrm{d}f, \omega \mathrm{d}f\rangle = 0$, so the curve remains on the hypersurface $f = 0$.

Suppose you have already given X a coördinate system. Make it a symplectic coördinate system, meaning the coördinates 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 $\omega\partial_{q^i} = \mathrm{d}p_i$, and $\omega\partial_{p_i} = -\mathrm{d}q^i$. Suppose you can solve the hypersurface for one of the coördinates, say $h := q^0$. Then you can give f the form $h - H$, where H is determined by all the coördinates besides h. To find a ∂_f, the easiest way is to replace the coördinate h by the coördinate f, so ∂_f in the new coördinates equals ∂_h in the old coördinates. Expressing the equation of motion in the old (symplectic) coördinates, $\langle \alpha, c'\rangle = \langle \mathrm{d}h - \mathrm{d}H, \omega\alpha\rangle\langle \omega\partial_h, c'\rangle$. $\omega\partial_h$ is $\mathrm{d}t$, where $t := p_0$. Since I know $\langle \mathrm{d}t, c'\rangle$ is non0, this makes t a good variable to parametrise c by, justifying the notation 't'. Remembering that H is independent of h and putting $\mathrm{d}t$ for α, $\langle \mathrm{d}t, c'\rangle = \langle \mathrm{d}h - \mathrm{d}H, \partial_h\rangle \times \langle \mathrm{d}t, c'\rangle = 1\langle \mathrm{d}t, c'\rangle + 0\langle \mathrm{d}t, c'\rangle$, which is as it should be. $\langle \mathrm{d}h, c'\rangle = \langle \mathrm{d}h - \mathrm{d}H, -\partial_t\rangle\langle \mathrm{d}t, c'\rangle = (\partial H/\partial t)\langle \mathrm{d}t, c'\rangle$, so $\mathrm{d}h/\mathrm{d}t$ is $\partial H/\partial t$ along c. If $i \neq 0$, $\langle \mathrm{d}q^i, c'\rangle = \langle \mathrm{d}h - \mathrm{d}H, -\partial_{p_i}\rangle\langle \mathrm{d}t, c'\rangle = (\partial H/\partial p_i)\langle \mathrm{d}t, c'\rangle$, so $\mathrm{d}q^i/\mathrm{d}t$ is $\partial H/\partial p_i$ along c. Finally, if $i \neq 0$, $\langle \mathrm{d}p_i, c'\rangle = \langle \mathrm{d}h - \mathrm{d}H, \partial_{q^i}\rangle\langle \mathrm{d}t, c'\rangle = -(\partial H/\partial q^i)\langle \mathrm{d}t, c'\rangle$, so $\mathrm{d}p_i/\mathrm{d}t$ is $-\partial H/\partial q^i$ along c. Also note that $\mathrm{d}H/\mathrm{d}t$ is $\mathrm{d}h/\mathrm{d}t = \partial H/\partial t$ along c, because $\mathrm{d}f/\mathrm{d}t = 0$ along c (indeed, $f = 0$ along c, so $H = h$ along c). These results are Hamilton's equations; H, a function of every symplectic coördinate 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', so they are vector fields tangent to X satisfying $\langle \mathrm{d}h - \mathrm{d}H, v\rangle = \langle \mathrm{d}t, v\rangle = \langle \mathrm{d}h - \mathrm{d}H, w\rangle = \langle \mathrm{d}t, w\rangle = 0$, let $\omega'(v,w)$ be $\omega(v,w)$. That is, ω' is the pullback of ω to X'. Coördinates for X' are the q^i and p_i for $i \neq 0$. These coördinates continue to be symplectic on X'. X' is the phase space of the system. A point in X' indicates a possible timeless 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 $\mathrm{d}A_t/\mathrm{d}t = \{A, H\} + \partial A/\partial t$ for $\{,\}$ the Poisson bracket associated with ω'.

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.