Here we show that the action variable (i.e., new – constant – momentum whose associated – cyclic – coordinate is actually an "angle", i.e., it goes through 2π during one period of the particle motion) is given by

$$I(E) \;=\; \frac{1}{2\pi}\oint p(q;E)\,dq \tag{1}$$

where $p(q;E)$ is obtained by solving

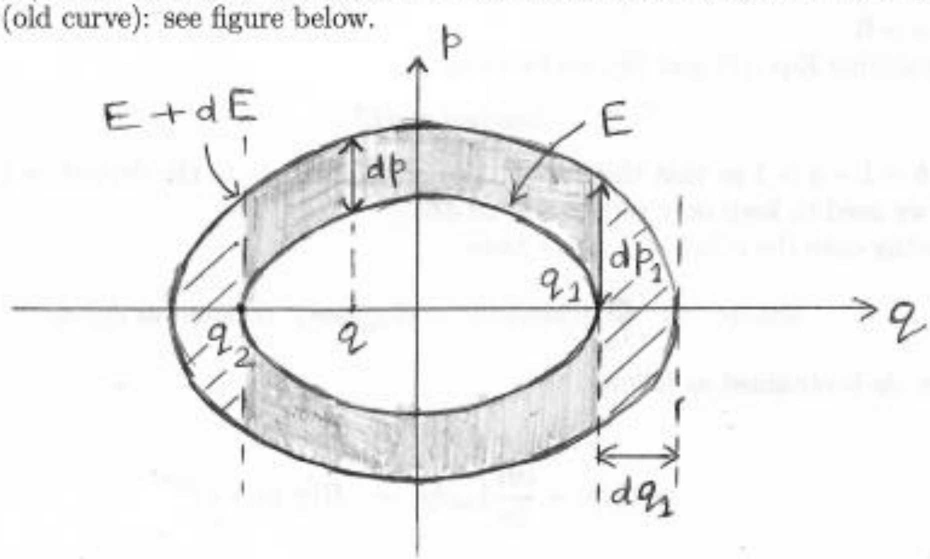
$$H(q,p) \;=\; E \tag{2}$$

i.e., we should find that

$$\frac{\partial I(E)}{\partial E} \;=\; \frac{T}{2\pi} = \frac{1}{\omega} \tag{3}$$

where ω is the angular frequency (T is the time period).

Clearly, Eq. (2) gives a curve in the phase [i.e., (q,p)]-space and $I(E)$ in Eq. (1) is the area inside it. So, dI is given by the *difference* in areas corresponding to $E + dE$ ("new") curve and E (old curve): see figure below.



This differential in I can then be computed in two pieces: shaded region (in-between old turning points, denoted by $q_{1,2}$) and hatched (in-between new *and* old turning points):

$$2\pi\, dI \;=\; \text{hatched} + \text{shaded} \tag{4}$$

First, let us compute the hatched area:

$$\text{hatched} \;\sim\; (\text{change in turning point or } dq_1) \times (\text{new } p \text{ at old turning point or } p_1 + dp_1) \tag{5}$$

For the first factor above, i.e., dq_1, we have [using Eq. (2) for the two values of energy at the turning points, i.e., $p = 0$]

$$\begin{aligned} H\left(q_1 + dq_1, 0\right) \;&=\; E + dE \\ H\left(q_1, 0\right) + \frac{\partial H}{\partial q}\Big|_{\text{old}} dq_1 \;&=\; H\left(q_1, 0\right) + dE \end{aligned} \tag{6}$$

so that (using $\partial H/\partial q = -\dot{p}$) we get

$$dq_1 = -\frac{dE}{\dot{p}_1|_{\text{old}}} \tag{7}$$

Next, we use the fact that the old p at the old turning point (i.e., p_1) vanishes (by definition) so that new p at old turning point is simply given by dp_1, which can be approximated as follows:

$$\begin{aligned} H(0+dp_1, q_1) &= E + dE \\ H(q,0) + \frac{\partial H}{\partial p}|_{\text{old}} dp_1 + \frac{\partial^2 H}{\partial p^2}|_{\text{old}} (dp_1)^2 &= H(q_1, 0) + dE \end{aligned} \tag{8}$$

Note that in the simple case of $H = p^2/(2m) + V(q)$, $\partial H/\partial p = p/m$ vanishes at the old endpoint, which is why we have kept the term which is one higher order in dp_1 Either way, we see that

$$dp_1 \propto dE^a \tag{9}$$

where $a > 0$.

Combining Eqs. (7) and (9), we find that

$$\text{hatched} \propto dE^b \tag{10}$$

where $b = 1 + a > 1$ so that this region does not contribute to the derivative in Eq. (3), for which we need to keep only term *linear* in dE.

Moving onto the other region, we have

$$\text{shaded} = \oint_{\text{old}} (\text{change in p at given } q, \text{ denoted as } dp)\, dq \tag{11}$$

In turn, dp is obtained as follows:

$$\begin{aligned} H(q, p+dp) &= E + dE \\ H(q,p) + \frac{\partial H}{\partial p}|_{\text{old}} dp &= H(q,p) + dE \end{aligned} \tag{12}$$

so that

$$\begin{aligned} dp &\approx \frac{dE}{\frac{\partial H}{\partial p}|_{\text{old}}} \\ &= \frac{dE}{\dot{q}|_{\text{old}}} \end{aligned} \tag{13}$$

Plugging Eq. (13) into Eq. (11) and dropping hatched area (as argued above) in Eq. (4) gives

$$\begin{aligned} 2\pi\, dI &\approx \int dq \frac{dE}{\dot{q}|_{\text{old}}} \\ &= dE \int dt, \text{ using } \dot{q} \equiv dq/dt \\ &= T\, dE \end{aligned} \tag{14}$$

resulting in Eq. (3) as desired.