A Sturm Comparison Theorem

Suppose that \( H \in C^1(\mathbb{R}^2, \mathbb{R}) \), and \( H \) satisfies:

(H1) \( H(t+1, x) = H(t, x) \) for all \( (t, x) \in \mathbb{R}^2 \),

(H2) \( H(t, 0) = 0 \) for all \( t \in [0, 1] \),

(H3) There is \( \lambda > 0 \) such that \( H_x(t, 0) < -\lambda \) for all \( t \in [0, 1] \). (Here \( H_x \) means \( \frac{\partial H}{\partial x} \).)

We want to prove the following “folk” theorem:

**Theorem 1.** Suppose that \( H \) satisfies (H1-H3). Then, there is a \( \delta > 0 \) and an \( L > 0 \) such that if \( x \) is any solution of \( \ddot{x}(t) = H(t, x(t)) \) with \( |x(t)| < \delta \) for all \( t \) in some interval \( J \), then \( x(t) \) has a zero in any sub-interval \( I \) of \( J \) whose length is longer than \( L \).

Before we prove Theorem 1, we need a few lemmas.

**Lemma 1.** There is a \( \delta > 0 \) such that if \( |x| < \delta \), then \( H_x(t, x) < -\frac{\lambda}{2} \).

**Proof.** Because \( H \) is \( C^1 \), by (H3), there is \( \delta_i > 0 \) such that if \( |(t, 0) - (\hat{t}, \hat{x})| < \delta_i \), then \( |H_x(\hat{t}, \hat{x}) - H_x(t, 0)| < \frac{\lambda}{2} \).

If \( B_r((t, x)) \) is the open ball in \( \mathbb{R}^2 \) of radius \( r \) around \( (t, x) \in \mathbb{R}^2 \), then the collection of \( B_{2\delta_i}((t, 0)) \) for \( t \in [0, 1] \) is an open cover of \( [0, 1] \times \{0\} \subset \mathbb{R}^2 \). Therefore, there is a finite cover: \( [0, 1] \times \{0\} \subset B_{\frac{\delta_i}{2}}((0, 0)) \cup \cdots \cup B_{\frac{\delta_i}{2}}((t_k, 0)) \).

Let \( \delta := \frac{\delta_i}{2} \min\{\delta_1, \ldots, \delta_{t_k}\} \). Let \( \hat{t} \in [0, 1] \), and suppose that \( |\hat{x}| < \delta \). Notice that \( (\hat{t}, 0) \in [0, 1] \times \{0\} \), hence \( (\hat{t}, 0) \in B_{\frac{\delta_i}{2}}((t_k, 0)) \) for some \( t_k \in [0, 1] \). Then, we have

\[
|\hat{t} - t_k| \leq |\hat{t} - t_k + t_k - t_k| = |\hat{t} - t_k| < \frac{1}{2}\delta_{t_k} + \delta \leq \frac{1}{2}\delta_{t_k} + \frac{1}{2}\delta_{t_k} = \delta_{t_k},
\]

by definition of \( \delta \). Therefore, we have \( |(\hat{t}, \hat{x}) - (t_k, 0)| < \delta_{t_k} \), and so \( |H_x(\hat{t}, \hat{x}) - H_x(t_k, 0)| < \frac{\lambda}{2} \).

Thus, if \( |\hat{x}| < \delta \), then, for any \( \hat{t} \in [0, 1] \), we have \( |H_x(\hat{t}, \hat{x}) - H_x(t_k, 0)| < \frac{\lambda}{2} \). But then, by (H3), we have

\[
\lambda < |H_x(t_k, 0)| \leq |H_x(t_k, 0) - H_x(\hat{t}, \hat{x})| + |H_x(\hat{t}, \hat{x})| < \frac{\lambda}{2} + |H_x(\hat{t}, \hat{x})|
\]

and so therefore we have \( \frac{\lambda}{2} < |H_x(\hat{t}, \hat{x})| \) whenever \( |\hat{x}| < \delta \) and \( \hat{t} \in [0, 1] \). Since \( H_x(t, 0) < -\lambda < 0 \), we must in fact know that \( H_x(\hat{t}, \hat{x}) < 0 \), and thus \( H_x(\hat{t}, \hat{x}) < -\frac{\lambda}{2} \) whenever \( |\hat{x}| < \delta \) and \( \hat{t} \in [0, 1] \). \( \square \)

**Lemma 2.** There is a \( \delta > 0 \) such that if \( 0 < x < \delta \), then \( H(t, x) < -\frac{\lambda}{2} x \), and if \( -\delta < x < 0 \), then \( H(t, x) > -\frac{\lambda}{2} x \).

**Proof.** Let \( \delta \) be that from Lemma 1. By (H2), for any \( t \in [0, 1] \), we have

\[
H(t, x) = \int_0^x \frac{d}{d\alpha} \left( H(t, \alpha) \right) \, d\alpha = \int_0^x H_x(t, \alpha) \, d\alpha.
\]

Thus, if \( 0 < x < \delta \), then Lemma 1 implies that \( H_x(t, \alpha) < -\frac{\lambda}{2} \) for all \( \alpha \in (0, x) \), and so

\[
H(t, x) < -\frac{\lambda}{2} \int_0^x d\alpha = -\frac{\lambda}{2} x.
\]
Next, suppose $-\delta < x < 0$. Because $x < 0$, we must have
\[ H(t, x) = -\int_x^0 H_x(t, \alpha) \, d\alpha. \]
Lemma 1 implies that $H_x(t, \alpha) < -\frac{\lambda}{2}$, and so $H_x(x, t) > \frac{\lambda}{2}$. Thus,
\[ H(t, x) = -\int_x^0 H_x(t, \alpha) \, d\alpha > \frac{\lambda}{2}(-x) > 0, \]
since $-\delta < x < 0$.

Now, we can prove Theorem 1.

**Proof.** Let $J = (a, b)$, and suppose that $x(t)$ is a solution of $\ddot{x} = H(t, x)$ such that $|x(t)| < \delta$ for all $t \in (a, b)$. Notice that any solution of $\ddot{u}(t) = \frac{-\lambda}{2} u(t)$ with any initial conditions is a linear combination of sines and cosines of period $\sqrt{\frac{2 \pi}{\lambda}}$, and so has a zero in any interval of length $\sqrt{\frac{2 \pi}{\lambda}} =: L$. We now proceed by contradiction: suppose that $(c, d) \subset J$, $d - c > L$ and $x(t) \neq 0$ for all $t \in (c, d)$. Therefore, we know for all $t \in (c, d)$ that either $x(t) > 0$ or $x(t) < 0$.

Suppose that $x(t) > 0$ for all $t \in (c, d)$, and let $u$ satisfy $\ddot{u}(t) = \frac{\lambda}{2} u(t)$, $u(c) = x(c)$, and $\dot{u}(c) = \dot{x}(c)$. We will now show that there is an interval $(c, \alpha)$ such that $x(t) < u(t)$ for all $t \in (c, \alpha)$. We consider two cases: $x(c) > 0$ and $x(c) = 0$. Notice that if $x(c) > 0$, then we have $\ddot{x}(c) = H(c, x(c)) < \frac{-\lambda}{2} x(c) = \frac{-\lambda}{2} u(c) = \ddot{u}(c)$. Since $x(c) = u(c)$ and $\dot{x}(c) = \dot{u}(c)$, there is a small interval $(c, \alpha)$ such that $x(t) < u(t)$ for all $t \in (c, \alpha)$. Next, we consider the case of $x(c) = 0$. In this case, $\ddot{x}(c) = H(c, 0) = 0 = \frac{-\lambda}{2} u(c)$. However, we must have $\dot{x}(c) > 0$ in order that $x(t) > 0$ for $t \in (c, d)$. But then, using the fact that $H$ is $C^1$, we see that $\ddot{x}(c) = \frac{\partial H}{\partial t}(c, x(c)) + \frac{\partial H}{\partial x}(c, x(c)) \cdot x(c) = \frac{\partial H}{\partial t}(c, 0) + \frac{\partial H}{\partial x}(c, 0) \cdot x(c) < -\lambda \ddot{x}(c) < \frac{-\lambda}{2} \ddot{u}(c) = \ddot{u}(c)$. Thus, since $x(c) = 0 = u(c)$, $\dot{x}(c) = \dot{u}(c)$, $\ddot{x}(c) = 0 = \ddot{u}(c)$, the fact that $\ddot{x}(c) < \ddot{u}(c)$ implies the existence of an interval of the form $(c, \alpha)$ such that $x(t) < u(t)$ for all $t \in (c, \alpha)$.

Now, let $(c, t^*)$ be the maximal interval on which $x < u$. Notice that $u$ has a zero on every interval of length $L$. Thus, since $x(t) > 0$ for all $t \in (c, d)$ and $d - c > L$, we must have $0 < x(t^*) = u(t^*)$. We shall now derive a contradiction.

Integrating by parts on the interval $(c, t^*)$, we see that
\[
\int_c^{t^*} \ddot{x}(s) u(s) \, ds = \dot{x}(s) u(s) \bigg|_c^{t^*} - \int_c^{t^*} \dot{x}(s) \ddot{u}(s) \, ds
\]
\[
= \dot{x}(t^*) u(t^*) - \dot{x}(c) u(c) - \left( x(s) \ddot{u}(s) \bigg|_c^{t^*} - \int_c^{t^*} x(s) \ddot{u}(s) \, ds \right)
\]
\[
= \dot{x}(t^*) u(t^*) - \dot{x}(c) x(c) - (x(t^*) \ddot{u}(t^*) - x(c) \ddot{u}(c)) + \int_c^{t^*} x(s) \ddot{u}(s) \, ds
\]
\[
= \dot{x}(t^*) u(t^*) - \dot{x}(c) x(c) - (x(t^*) \ddot{u}(t^*) - x(c) \ddot{u}(c)) + \int_c^{t^*} x(s) \ddot{u}(s) \, ds
\]
\[
= \dot{x}(t^*) u(t^*) - x(t^*) \ddot{u}(t^*) + \int_c^{t^*} x(s) \ddot{u}(s) \, ds.
\]
But \( u(t^*) = x(t^*) \), and so we have
\[
\int_c^{t^*} \ddot{x}(s) u(s) \, ds = \dot{x}(t^*) u(t^*) - x(t^*) \dot{u}(t^*) + \int_c^{t^*} x(s) \ddot{u}(s) \, ds
\]
\[
= x(t^*) \left( \dot{x}(t^*) - \dot{u}(t^*) \right) + \int_c^{t^*} x(s) \ddot{u}(s) \, ds.
\]  
(2)

Now, by Lemma 2, since \( 0 < x(s) < \delta \) for all \( s \in (c, t^*) \), \( \ddot{x}(s) = H(s, x(s)) < -\frac{\lambda}{2} x(s) \) for all \( s \in (c, t^*) \). Multiplying both sides by \( u(s) > 0 \), we see that \( \ddot{x}(s) u(s) = H(s, x(s)) u(s) < -\frac{\lambda}{2} x(s) u(s) = \ddot{u}(s) x(s) \). Integrating \( \ddot{x}(s) u(s) < \ddot{u}(s) x(s) \) over \((c, t^*)\), we see that
\[
\int_c^{t^*} \ddot{x}(s) u(s) \, ds < \int_c^{t^*} \ddot{u}(s) x(s) \, ds.
\]  
(3)

Therefore, in order to have equality in (2), (3) implies we must have \( x(t^*) \left( \dot{x}(t^*) - \dot{u}(t^*) \right) < 0 \). Since \( x(t^*) > 0 \), we must have \( \dot{x}(t^*) < \dot{u}(t^*) \). Therefore, \( \frac{d}{dt} (x(t) - u(t)) |_t < 0 \) and so \( x(t) - u(t) \) is decreasing close to \( t^* \). Since \( x(t^*) - u(t^*) = 0 \), \( x(t) - u(t) > 0 \) for \( t < t^* \) close to \( t^* \). But this implies that \( x(t) > u(t) \) for all \( t < t^* \) close to \( t^* \), which is impossible, since we assumed that \((c, t^*)\) was the maximal interval in \((c, d)\) on which \( x(t) < u(t) \). This contradiction implies that there can be no interval \((c, d)\) on which \( x(t) > 0 \) and \( d - c > L \).

It remains to show that there can be no interval longer than \( L \) on which \( -\delta < x(t) < 0 \). Suppose that \((c, d)\) is any interval for which \( x(t) < 0 \) for all \( t \in (c, d) \). Let \( x_1(t) := -x(t) \), and notice that \( 0 < x_1(t) < \delta \) for all \( t \in (c, d) \). In addition, let \( \tilde{H}(t, x) := -H(t, -x) \). We have \( \tilde{H}(t, 0) = -H(t, 0) = 0 \), and \( \frac{\partial \tilde{H}}{\partial x}(t, x) = -\frac{\partial H}{\partial x}(t, -x) \cdot (-1) = \frac{\partial H}{\partial x}(t, -x) < -\lambda \). In addition, suppose that \( |x| < \delta \), where \( \delta \) is the same \( \delta \) for \( H \) from Lemma 1. We then have \( \frac{\partial \tilde{H}}{\partial x}(t, x) = \frac{\partial H}{\partial x}(t, -x) < -\frac{\lambda}{2} \). Thus, the same \( \lambda \) and \( \delta \) works for both \( H \) and \( \tilde{H} \). Because \( x_1 \) solves \( \ddot{x}_1(t) = \tilde{H}(t, x_1(t)) \) and \( x_1(t) > 0 \) for all \( t \in (c, d) \), we know that \( d - c \leq L \). Thus, there can be no interval longer that \( L \) on which \( x(t) < 0 \). \( \square \)