Предположим обратное, то есть то, что одно решение $\Phi(t)$ устойчиво, но какое-то другое $\Psi(t)$ неустойчиво. Это значит, что для некоторого $\varepsilon > 0$ и любого $\delta > 0$ существует другое ненулевое решение $X_\delta(t)$ и такое $t_1 = t_1(\delta) > t_0$, что $\left|\Psi(t_0) - X_\delta(t_0)\right| < \delta$, но $\left|\Psi(t_1) - X_\delta(t_1)\right| \geqslant \varepsilon$.

Система уравнений линейная, значит всякое её решение можно выразить в виде $C_1 f_1(t) + C_2 f_2(t) + \ldots + C_n f_n(t) + g(t)$, где $g(t)$ – частное решение неоднородной системы; $f_i(t)$, $1 \leqslant i \leqslant n$ – это столбцы, составляющие фундаментальную систему решений. В ней любую функцию можно умножить на ненулевой скаляр и полученная ФСР всё равно останется корректной.

Подберём фундаментальную систему решений надлежащим образом. Возьмём любую ФСР и t_0 такое, чтобы векторы $f_1(t_0), \ldots, f_n(t_0)$ были линейно независимы (то есть матрица A, составленная из этих векторов-столбцов, имела $\det A \neq 0$). Далее, для положительного числа $\varepsilon_\Phi = \dfrac{\varepsilon}{n}$ подберём такое $\gamma > 0$, что для любого решения $X(t)$ из $\left|\Phi(t_0) - X(t_0)\right| < \gamma$ следует $\left|\Phi(t) - X(t)\right| < \dfrac{\varepsilon}{n}$ для всех $t \geqslant t_0$. Наконец, в выбранной ФСР умножим каждую функцию на такие ненулевые скаляры, чтобы выполнялись $\left|f_k(t_0) - \Phi(t_0)\right| < \gamma$, $1 \leqslant k \leqslant n$. Из этого следуют неравенства $\left|f_k(t) - \Phi(t)\right| < \dfrac{\varepsilon}{n}$, а линейная независимость векторов $f_1(t_0), \ldots, f_n(t_0)$ не нарушается, так как умножение столбцов матрицы A приведёт к умножению её определителя на ненулевые числа. Полученный набор функций положим новой ФСР. Стоит обратить особое внимание на тот факт, что построение ФСР указанным образом никак не зависит ни от δ, ни от $t_1 = t_1(\delta)$.

В терминах подобранной выше ФСР распишем

$$\Psi(t) = A_1 f_1(t) + \ldots + A_n f_n(t) + g(t) \qquad X_\delta(t) = B_1 f_1(t) + \ldots + B_n f_n(t) + g(t)$$

$$\Psi(t) - X_\delta(t) = \Delta_1 f_1(t) + \ldots + \Delta_n f_n(t) \qquad\qquad \Delta_k = A_k - B_k$$

Разность $\Psi(t_0) - X_\delta(t_0)$ можно представить в следующей форме:

$$\begin{pmatrix} \Delta_1 f_{11}(t_0) + \ldots + \Delta_1 f_{1n}(t_0) \\ \ldots \\ \Delta_1 f_{n1}(t_0) + \ldots + \Delta_n f_{nn}(t_0) \end{pmatrix} = \begin{pmatrix} f_{11}(t_0) & \ldots & f_{1n}(t_0) \\ \ldots & \ldots & \ldots \\ f_{n1}(t_0) & \ldots & f_{nn}(t_0) \end{pmatrix} \times \begin{pmatrix} \Delta_1 \\ \ldots \\ \Delta_n \end{pmatrix} = F(t_0) \times \Delta,$$

где за $F_{n\times n}(t)$ обозначена матрица, составленная по столбцам из выбранной ФСР; $\Delta_{n\times 1}$ – столбец разностей произвольных постоянных. Матрица $F(t_0)$ фиксирована, при этом ещё и невырождена. Следовательно, справедливы представления

$$\Psi(t_0) - X_\delta(t_0) = F(t_0) \times \Delta \qquad\qquad \Delta = \big(F(t_0)\big)^{-1} \times \big(\Psi(t_0) - X_\delta(t_0)\big).$$

За норму квадратной матрицы далее обозначена норма Фробениуса $\|A\|_F$. По свойству согласованности этой нормы со спектральной нормой $\|x\|_2$

$$|\Delta| = \|\Delta\|_2 = \sqrt{\Delta_1^2 + \ldots + \Delta_n^2} = \left|\big(F(t_0)\big)^{-1} \times \big(\Psi(t_0) - X_\delta(t_0)\big)\right| \leqslant$$
$$\leqslant \left\|\big(F(t_0)\big)^{-1}\right\|_F \cdot \left|\big(\Psi(t_0) - X_\delta(t_0)\big)\right| < \left\|\big(F(t_0)\big)^{-1}\right\|_F \cdot \delta.$$

Обозначим также $M_\delta = \max\limits_{1\leqslant k\leqslant n} |\Delta_k| > 0$. Выше показано, что

$$M_\delta \leqslant \sqrt{\Delta_1^2 + \ldots + \Delta_n^2} < \left\| \left(F(t_0)\right)^{-1} \right\|_F \cdot \delta.$$

Например, если рассмотреть

$$\delta = \frac{\varepsilon}{\varepsilon + n\left|\Phi(t_1)\right|} \left(\left\| \left(F(t_0)\right)^{-1} \right\|_F \right)^{-1}, \text{ то получим } M_\delta < \frac{\varepsilon}{\varepsilon + n\left|\Phi(t_1)\right|}.$$

С одной стороны, из $\left|\Psi(t_1) - X(t_1)\right| \geqslant \varepsilon$ вытекает

$$\begin{aligned}
\varepsilon &\leqslant \left|\Psi(t_1) - X(t_1)\right| = \left|\Delta_1 f_1(t_1) + \ldots + \Delta_n f_n(t_n)\right| \leqslant \left|\Delta_1 f_1(t_1)\right| + \ldots + \left|\Delta_n f_n(t_n)\right| \leqslant \\
&\leqslant M_\delta \left(\left|f_1(t_1)\right| + \ldots + \left|f_n(t_1)\right| \right) = \\
&= M_\delta \left(\left|f_1(t_1) - \Phi(t_1) + \Phi(t_1)\right| + \ldots + \left|f_n(t_1) - \Phi(t_1) + \Phi(t_1)\right| \right) \leqslant \\
&\leqslant M_\delta \left(\left|f_1(t_1) - \Phi(t_1)\right| + \ldots + \left|f_n(t_1) - \Phi(t_1)\right| + n\left|\Phi(t_1)\right| \right),
\end{aligned}$$

а тогда

$$\left|f_1(t_1) - \Phi(t_1)\right| + \ldots + \left|f_n(t_1) - \Phi(t_1)\right| \geqslant \frac{\varepsilon}{M_\delta} - n\left|\Phi(t_1)\right| > \frac{\varepsilon}{\left(\dfrac{\varepsilon}{\varepsilon + n\left|\Phi(t_1)\right|}\right)} - n\left|\Phi(t_1)\right| = \varepsilon.$$

С другой стороны, для всех $t \geqslant t_0$ имеем $\left|f_1(t) - \Phi(t)\right| + \ldots + \left|f_n(t) - \Phi(t)\right| < \dfrac{\varepsilon}{n} \cdot n = \varepsilon$. В частности, подстановка в последнее неравенство $t = t_1$ даёт противоречие, которое доказывает теорему.