2 Systems on a plane. Andronov-Pontryagin theorem

As opposed to differential equations in R^n with $n \geq 3$, where a complicated, *chaotic* behaviour is possible, the Poincare-Bendixson theorem says us that the behaviour of planar systems is always *simple* - in the limit everything becomes stationary (equilibrium states) or periodic (closed phase curves), or, in the third case of this theorem, we have a behaviour somewhat intermediate between the first two. However, the various possible types of these behaviour are still too many. Order is introduced by the idea of *structural stability* due to Andronov and Pontryagin.

The main point is that differential equations are rarely interesting by themselves. Usually, we have to analyse them only because we want to understand the course of certain real-world processes which they model. In such a case, a system of differential equations, as every model, is only an approximation of reality. In this way one comes to an idea that whenever a certain system of differential equations is analysed there exists some "true" system, which is not exactly equal to the model system under consideration but is close to it. Therefore, the results we obtain during the theoretical analysis of the model system will make sense only if we can carry them on to the true system. The problem here is that we do not know this true system exactly. However, this problem vanishes if the model system is structurally stable.

Definition. A system of differential equations is *structurally stable* if every sufficiently close system has the same phase portrait.

By closeness we mean closeness along with the first derivatives: two systems

$$\begin{cases} \frac{dx}{dt} = f(x,y) \\ \frac{dy}{dt} = g(x,y) \end{cases} \quad (A) \quad \text{and} \quad \begin{cases} \frac{dx}{dt} = p(x,y) \\ \frac{dy}{dt} = q(x,y) \end{cases} \quad (B)$$

defined on a plane, in a disc of a sufficiently large radius R, are δ-close in this disc if

$$\sup_{x^2+y^2\le R^2} \{|f(x,y)-p(x,y)|+|g(x,y)-q(x,y)|+|f'_x(x,y)-p'_x(x,y)|+$$

$$+|f'_y(x,y)-p'_y(x,y)|+|g'_x(x,y)-q'_x(x,y)|+|g'_y(x,y)-q'_y(x,y)|\Big\} < \delta.$$

It is not easy to define what the phase portrait is: intuitively, phase portrait is a picture that describes position of the phase curves in the phase space. Andronov-Pontryagin definition says that two close systems (A) and (B) have the same phase portrait if *there exists a homeomorphism of a plane that takes the phase curves of (A) to the phase curves of (B)*, and this homeomorphism is close to the identity map. This mathematical definition is consistent with the intuitive idea of a similarity between two phase portraits. Indeed, a homeomorphic image of a point is a point, so the homeomorphism that takes the phase curves of (A) to the phase curves of (B) has to take equilibrium states to equilibrium states, i.e. (A) and (B) have to have the same number of equilibrium states; moreover, as the homeomorphism is close to identity, the equilibria of (B) are close to the respective equilibria of (A). The same holds true for closed phase curves. Furthermore, if an orbit of (A) tends to some equilibrium state or to a closed phase curve, then its image by the homeomorphism is an orbit that tends to the corresponding equilibrium state or a closed phase curve of (B) (the homeomorphism also keeps the direction of the growth of the time variable along the orbit). Therefore, stable equilibria and closed orbits of (A) correspond to stable equilibria and closed orbits of (B), unstable ones correspond to unstable ones, saddles to saddles.

Note however that when we deal with systems in R^n with $n \geq 3$ Andronov-Pontryagin idea of establishing equivalence between systems of differential equation by means of a homeomorphism that takes phase curves into phase curves becomes much less useful than in the planar case: it occurs that the structure of the set of the equivalence classes is too weird. On the contrary, for systems on a plane a complete characterisation of the equivalence classes is possible; it was achieved in Andronov, Leontovich, Gordon, Maier, Qualitative Theory of Second-Order Dynamical Systems. Practically most important part of the classification is given by the following theorem which describes structurally stable planar systems.

Andronov-Pontryagin theorem A system on a plane is structurally stable if and only if
1) all its equilibrium states are hyperbolic (stable, unstable or saddles),
2) all periodic orbits have a multiplier different from 1,
3) there are no phase curves which connect saddles.

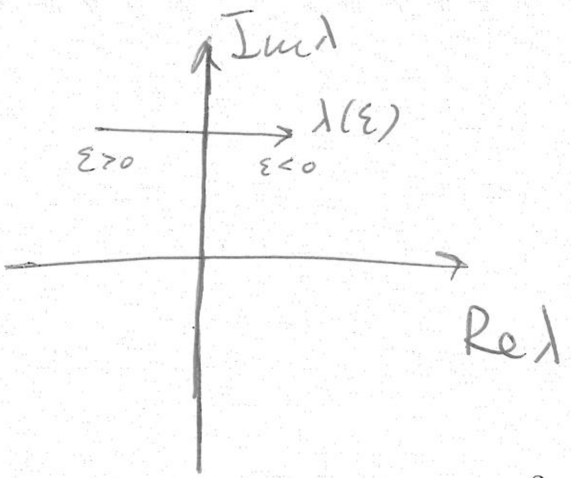
Proof. First, let us show that these 3 conditions are *necessary* for structural stability. Indeed, suppose condition 1 is violated, i.e. a system has a non-hyperbolic equilibrium. By putting the equilibrium at the origin of coordinates, we right the system near it in the form

$$\frac{d}{dt}u = Au + o(u)$$

where $u = (x, y)$ and A is a 2×2-matrix with constant coefficients; the assumed non-hyperbolicity of the equilibrium means that A has one or several eigenvalues on the imaginary axis. For sufficiently small ε the system

$$\frac{d}{dt}u = Au - \varepsilon u + o(u)$$

will be a small perturbation of the original one. The new system will still have an equilibrium at $u = 0$, and the linearisation matrix at this equilibrium will be $(A - \varepsilon I)$; the corresponding eigenvalues are thus $\lambda - \varepsilon$, where λ is an eigenvalue of A. As we see, for positive ε the eigenvalues on the imaginary axis will be moved to the left of it, while for negative ε the eigenvalues will be moved to the right. So, arbitrarily close to the given system there exist systems with *different* types of equilibria - no structural stability.

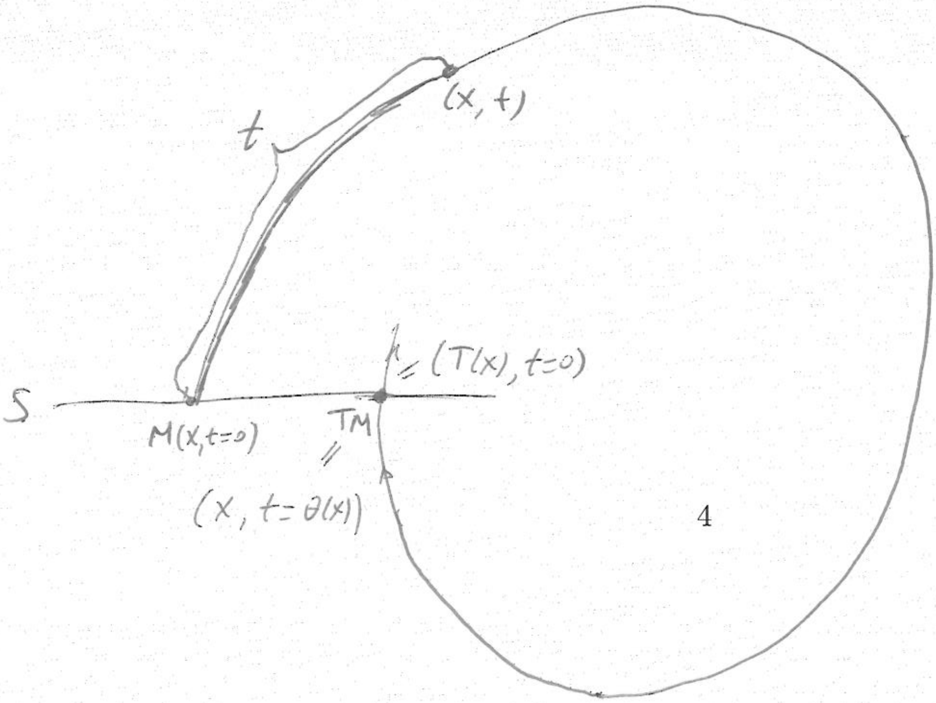


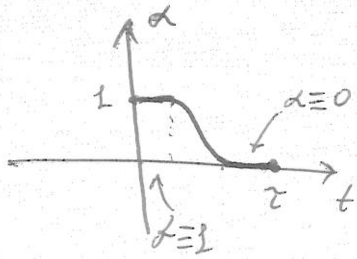
The same is true for the periodic orbit: if the multiplier equals to 1, then it can be made less than 1 or greater than 1 by a small perturbation of the system, hence the stability is changing. That proves structural instability of systems which have a periodic orbit with a multiplier equal to 1 (condition 2 of the theorem) - we however need to really construct a mechanism of perturbing the multiplier, or, more generally, of perturbing the Poincare map.

Let us describe one such construction. Let a system on a plane have an orbit L of period τ and let S be a cross-section to it. Let $T(x) = \mu x + o(x)$ be the Poincare map on S (the coordinate x on S is chosen such that $x = 0$ corresponds to $L \cap S$) and let $\theta(x)$ be the time the orbit starting at S needs to hit S again, $\theta(0) = \tau$. For any point M in a small neighbourhood of L, the phase curve that passes through it has an intersection with S at some point x after which it arrives to M at the time moment t where $t \in [0, \theta(x))$. We can consider the pair (x, t) as coordinates of M, since they determine the position of M completely: we start at x on S and the point where we arrive after time t is M. However, these are not good coordinates for the small neighbourhood of L: we have to identify the point $(x, \theta(x))$ with $(T(x), 0)$, i.e. our coordinate system is discontinuous. Good coordinates are, for example, $(y(x,t), s(x,t))$ defined by

$$y = x\alpha(t) + T(x)(1 - \alpha(t)), \qquad s = t + (\tau - \theta(x))(1 - \alpha(t)) \qquad (*)$$

where α is a smooth function which equals identically to 1 near zero and to 0 near τ (in particular, $\alpha(\theta(x)) = 0$ for all small x, as $\theta(x)$ is close to $\tau = \theta(0)$).





At $x = 0$ we have $y = 0$ and $s = t$, and it is easy to check that $det \begin{pmatrix} y'_x(0,t) & y'_t(0,t) \\ s'_x(0,t) & s'_t(0,t) \end{pmatrix} = \alpha(t) + T'(0)(1-\alpha(t)) \neq 0$ (as $\alpha \in [0,1]$ and $T'(0) = \mu > 0$). Therefore, by the Implicit Function Theorem, x and t are uniquely defined smooth functions of y and s for all small y: $x = \xi(y,s)$, $t = \eta(y,s)$, so (y,s) are good coordinates in a small neighbourhood of L indeed. Note that the range of the variable s is independent of x: $\alpha(t)$ vanishes identically as t approaches $\theta(x)$, so the values of s approaches τ, i.e. the range of s is $[0,\tau)$. Recall that in the coordinates (x,t) we have to glue the points $(x,\theta(x))$ and $(T(x),0)$ - they correspond to the same point at the cross-section S. In the (y,s) ccordinates they both correspond to the same $y = T(x)$ and, respectively, $s = \tau$ and $s = 0$, so in the new coordinates we simply glue the points (y,τ) and $(y,0)$, with the same value of y. By definition, x is the coordinate of intersection of the phase curve with S, so it does not change along the orbit (until the new intersection happens), hence $dx/dt = 0$. Thus, by differentiating (*), we find

$$\frac{dy}{dt} = (x - T(x))\alpha'(t) = (\xi(y,s) - T(\xi(y,s)))\alpha'(\eta(y,s)),$$

$$\frac{ds}{dt} = 1 - (\tau - \theta(x))\alpha'(t) = 1 - (\tau - \theta(\xi(y,s)))\alpha'(\eta(y,s)).$$

This is just the form our original system of differential equations takes in the new coordinates. Note that near the cross-section S where our coordinate system is discontinuous (we have there s jumping from τ to 0), the function α is identical constant (0 or 1), so $\alpha' = 0$, and we see that $dy/dt \equiv 0, ds/dt \equiv 1$ near S, hence the right-hand sides have no discontunities. Now take any $\tilde{T}$ close to T. Replacing T to $\tilde{T}$ in (*) will define us new functions $y = \tilde{\xi}(x,t)$, $s = \tilde{\eta}(x,t)$, close to the original ξ and η. System

$$\frac{dy}{dt} = (\tilde{\xi}(y,s) - \tilde{T}(\tilde{\xi}(y,s)))\alpha'(\tilde{\eta}(y,s)), \qquad \frac{ds}{dt} = 1 - (\tau - \theta(\tilde{\xi}(y,s)))\alpha'(\tilde{\eta}(y,s))$$

will be a small perturbation of the original one and, by construction, the Poincare map for it will be given by $x \mapsto \tilde{T}(x)$. So, we have achieved our goal - we can create an arbitrary small perturbation of the Poincare map by a small perturbation of the right-hand sides of the system.

Let us now proceed to condition 3. Suppose system (A) has an orbit Γ that connects two saddles, O_1 and O_2. That means Γ is an unstable separatrix of O_1 and, at the same time, it is a stable separatrix of O_2. Consider a family of systems

$$\frac{dx}{dt} = f(x,y) + \varepsilon g(x,y), \qquad \frac{dy}{dt} = g(x,y) - \varepsilon f(x,y), \qquad (A_\varepsilon).$$

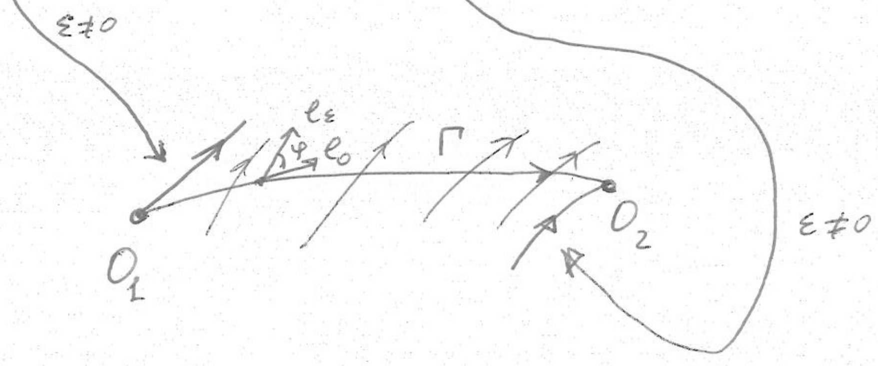
At $\varepsilon = 0$ system (A_ε) coinsides with (A). At $\varepsilon \neq 0$, the vector

$$\underline{e}_\varepsilon = (f + \varepsilon g, g - \varepsilon f)$$

at the right-hand sides of (A_ε) makes a non-zero angle with the vector $\underline{e}_0 = (f, g)$ at the right-hand sides of (A):

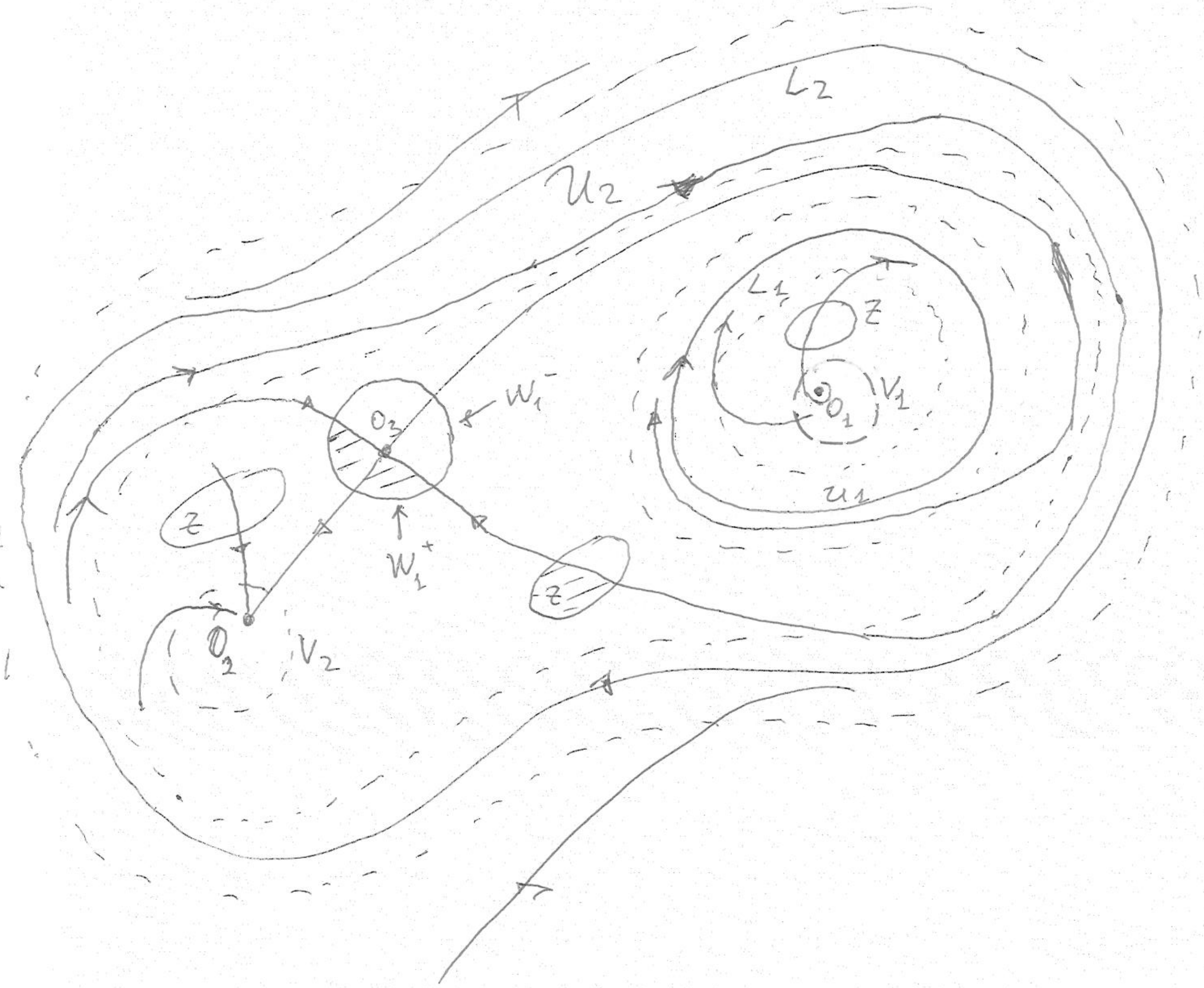
$$\sin\varphi = \frac{1}{\|\underline{e}_\varepsilon\| \cdot \|\underline{e}_0\|} \|\underline{e}_\varepsilon \times \underline{e}_0\| = \frac{\varepsilon}{\sqrt{1+\varepsilon^2}}.$$

By definition, the orbits of (A_ε) are tangent to the vector field $\underline{e}_\varepsilon$, so for every point on the curve Γ the orbit of (A_ε) that passes through it makes a non-zero angle with Γ (as Γ is the orbit of (A) and, hence, is tangent to $\underline{e}_0$). All orbits of (A_ε) cross Γ in the same direction, and it is easy to see that the unstable separatrix of O_1 moves, as ε changes, to one side of Γ, while the stable separatrix of O_2 moves in the opposite direction. As we see, the connection between the sadlles O_1 and O_2 disappears for arbitrarily small ε, which means that system (A) is not structurally stable in this case.



So, the conditions of the theorem are *necessary* for structural stability. Now goes the *sufficiency* part. Let system (A) satisfy conditions 1,2,3 of the theorem. Instead of proving the existence of a homeonorphism which maps the orbits of system (A) into orbits of a close system, we will prove something less formal and more informative. Namely, we will give a quite transparent picture of the behaviour of orbits of system (A) and show that this picture does not change when the system changes slightly.

First, we notice that if conditions of the theorem are fulfilled, all the periodic orbits are either stable or unstable (since the multiplier is not 1). Therefore, for each periodic orbit L_i there exists a neighbourhood U_i such that all orbits inside U_i tend to L_i as $t \to +\infty$ in case L_i is stable, or as $t \to -\infty$ in case L_i is unstable. Similarly, for each of the stable (unstable) equilibria O_j there exists a neighbourhood V_j such that every orbit starting at V_j tends to O_j as $t \to +\infty$ (resp., to $t \to -\infty$)). For every saddle equilibrium O_k we have a small neighbourhood W_k such that the stable separatrix divides W_k into 2 open halves, W_k^+ and W_k^-, such that the orbits that start at W_k^+ leave W_k following one unstable separatrix and the orbits that start at W_k^- leave W_k following the other unstable separatrix. According to the Poincare-Bendixson theorem all bounded orbits tend either to a periodic orbit, or to an equilibrium state (they cannot tend to a set containing equilibria and connecting orbits because there are no connecting orbits here, by assumption 3 of the present theorem). Therefore, if we fix any disc of some large radius R, every orbit after a finite time either leaves the disc or enters one of the neighbourhoods U_i, V_j or W_k. Therefore, for every point there is a neighbourhood Z such that after some fixed time all the orbits from this neighbourhood either leave the large disc together, or enter the same one of the neighbourhoods U_i, V_j or W_k. By compactness of the disc, we may cover it by a finite number of such neighbourhoods Z_s. Now it remains to note that the whole picture survives any sufficiently small perturbation: for exactly the same finite set of the neighbourhoods U_i, V_j, W_k and Z_s, both for the system itself and for every sufficiently close system the orbits that start at U_i or at those Z_s, which enter U_i after a finite time shift, tend to the corresponding periodic orbit L_i; the orbits that start at V_j or at those Z_s, which enter V_j after a finite time shift, tend to the corresponding equilibrium state O_j; for some of the neighbourhoods Z_s all the orbits leave the large disc after a finite time; and the rest of the neighbourhoods Z_s and W_k is divided by a stable separatrix of some saddle into two open halves such that all the orbits in one half share the fate of one of the unstable separatrices of the saddle while the orbits in the other half share the fate of the other separatrix of the same saddle (each of the separatrices tends to one of the stable equilibria or periodic orbits). To actually prove this we need the following fact: hyperbolic equilibria and periodic orbits do not disappear at small changes of the right-hand sides of the system and change continuously, moreover the separatrices of saddles change continuously as well - this will be addressed later in a more general context. □



We have seen in the proof that if any of the 3 conditions of Andronov-Pontryagin theorem are violated, then by an atbitrarily small perturbation they can be made fulfilled. In other words, if a system on the plane is not structurally stable, then a small perturbation can always make it structurally stable (and if it is structurally stable, then every close system is structurally stable as well - and has "the same structure"). Thus, Andronov-Pontryagin theorem describes the behaviour which is *most typical* for systems on a plane. Note that a similar statement is no longer true in higher dimensions: 3-dimensional structurally unstable systems are as typical as structurally stable ones. Returning to systems on a plane, analysis of structurally unstable systems starts making sense when we consider families of systems which depend on parameter (or several parameters). In such families, when the parameter change within an interval of structiural stability the behaviour does not change, but at some moments (called *bifurvational moments*) some of the conditions of Andronov-Pontryagin theorem may become violated; crossing such parameter values will cause changes in the system behaviour. Knowing which changes happen under which conditions can be quite helpful: if we know the behaviour at some parameter value and what happens at the bifurcation parameter values, we can recover the behaviour at *all* parameter values.

Further reading: Guckenheimer and Holmes, Ch.1

Exercises.

1. Prove that the system $x'' = x$ is structurally stable, and the system $x'' = -x$ is structurally unstable.

2. Prove that the system $x' = x + y - x(x^2 + y^2), \quad y' = -x + y - y(x^2 + y^2)$ is structurally stable, and the system $x' = y + (y^2 - x^2 + x^4)(2x^3 - x), \quad y' = x - 2x^3 - y(y^2 - x^2 + x^4)$ is structurally unstable.

3. Are the following systems structurally stable: $x' = y + y^2 - x^3y, \quad y' = -x - x^2 - y^3$, and $x' = 1 - x^2, \quad y' = xy$?

4. Prove that the following system is structurally stable $x' = 1 - 2x - y^2, y' = 1 - 6x + y^2$ (hint: show that the system has no periodic orbits; if such an orbit exists, it cannot intersect the line $y = 1/\sqrt{2}$, as any orbit that enters the region $y > 1/\sqrt{2}$ from the region $y < 1/\sqrt{2}$ must do it with $y' \geq 0$, i.e. it intersects the line $y = 1/\sqrt{2}$ at $x < 1/4$, however any orbit that enters the region $\{y > 1/\sqrt{2}, x < 1/4\}$ can never leave it as on the boundary of the region the vector field looks inside - e.g. $x' < 0$ at $x = 1/4, y > 1/\sqrt{2}$ - so the hypothetical periodic orbit has to lie entirely in the region $y < 1/\sqrt{2}$ where the divergence $f'_x + g'_y$ of the vector field is strictly negative, but this would contradict the classical Dulac theorem that states that the integral of the divergence of the vector field over the region inside a periodic orbit is always zero; similar arguments show that there is no orbit going from a saddle to a saddle).