Miroslav Katětov; Pavel Jedlička
Teorie katastrof: souvislosti a aplikace. I

Pokroky matematiky, fyziky a astronomie, Vol. 24 (1979), No. 1, 1--20

Persistent URL: `http://dml.cz/dmlcz/139440`

POKROKY MATEMATIKY, FYZIKY A ASTRONOMIE, ROČNÍK XXIV (1979) ČÍSLO 1

Teorie katastrof: souvislosti a aplikace. I

Miroslav Katětov, Pavel Jedlička, Praha

Nynější první část článku má tři oddíly. V prvním připomínáme základní pojmy, ve druhém mluvíme o souvislostech teorie katastrof s jinými matematickými disciplínami, ve třetím se zabýváme jedním případem použití základních myšlenek zmíněné teorie. Oddíly 1 a 2 napsal autor uvedený na prvním místě. Oddíl 3 napsali oba autoři společně; obsahuje některé výsledky společného výzkumu použití matematických metod v problematice roztroušené sklerózy mozkomíšní. Tyto výsledky jsou připraveny k publikaci v příslušném speciálním časopise; zde mají především ilustrovat některé širší úvahy v první i v chystané druhé části článku.

Druhá část má obsahovat další poznámky o souvislostech s jinými matematickými teoriemi, obecné úvahy o aplikacích teorie katastrof a ilustraci těchto poznámek a úvah na konkrétním materiálu včetně modelů průběhu zmíněné nemoci.

1.

1.1. Teorii katastrof byly již v tomto časopise věnovány dva články [1, 2]. Zde se zabýváme převážně těmi aspekty, jež v nich nebyly probírány. Poznamenáváme, že budeme rozlišovat, i když ne zcela striktně, jednak teorii katastrof jako matematickou disciplínu, byť vymezenou jen volně, jednak aplikační metodu používající pojmů a poznatků teorie katastrof. Mohlo by se jí např. říkat – z důvodů, jež budou patrné – „metoda systémů s katastrofickými elementy“; pro stručnost však použijeme zkratkového názvu „metoda katastrof“.

Nebudeme předpokládat speciální matematické znalosti a po matematické stránce výklad nebude přesný; použijeme také často různých drobných licencí (triviální příklad: místo o jednobodové množině se někdy mluví o bodu). Definice i věty se jen naznačí, pojmy se často uvedou v přístupnější verzi, která se nemusí shodovat s tou, které se skutečně používá. Jde nám totiž jen o objasnění základních myšlenek a nikoli o výklad matematického aparátu. Přesnou informaci o pojmovém základu i matematickém aparátu teorie katastrof (převážně tzv. elementární, viz 1.9) lze získat např. v knihách [3–6], v nichž je též uvedena literatura. Základním zdrojem informací o celkové koncepci teorie a metody katastrof je ovšem monografie [7] jejich zakladatele R. THOMA; její čtení je však dosti nesnadné.

1.2. Z možných cest vedoucích k základním pojmům teorie katastrof volíme zde tu, která vychází z vektorových diferenciálních rovnic tvaru $dx/dt = F(u, x)$, kde u je vektorový parametr. Nejdříve připomeneme některé pojmy a pro účely nynějšího článku zavedeme některé úmluvy.

Zobrazení otevřené množiny $U \subset R^m$ do R^s nazýváme *hladkým*, jestliže má derivace všech řádů (někdy je vhodné předpokládat méně, to však nebudeme rozebírat).

Hladké zobrazení $U \subset R^m$ do R^m budeme nazývat též *hladkým vektorovým polem* na U. Chápeme je tak, že každému bodu $x \in U$ je přiřazen určitý vektor (o m složkách) v tomto bodě, takže jde skutečně o vektorové pole ve smyslu běžném při aplikacích. Pojmy hladké variety, jejího hladkého zobrazení a hladkého vektorového pole na ní nebudeme připomínat; v úvahách, na něž se omezujeme v článku, může totiž čtenář myslet jen na speciální případ oblastí prostorů R^m. Místo o hladkých varietách, polích atd. budeme často mluvit prostě o varietách, polích apod.

Jsou-li X, Y variety, je-li φ hladké prosté zobrazení X na Y a je-li též φ^{-1} hladké, nazýváme φ *difeomorfismem*. Písmen B, M (případně s indexy) použijeme obvykle pro variety dimenze p, n, písmen u, x pro jejich body. Množiny tvaru $\{u\} \times M$ budeme nazývat *vertikálami* variety $B \times M$.

Symbol $\mathscr{V}(M)$ bude značit prostor všech hladkých vektorových polí na M (je-li $M \subset R^n$, jde vlastně o prostor všech hladkých $f : M \to R^n$). Při vhodně zvolené topologii je $\mathscr{V}(M)$ Fréchetův prostor, jenž má např. tuto přirozenou vlastnost: je-li F vektorové pole na $B \times M$, jež je tangenciální k vertikálám (to lze pro $B \subset R^p$, $M \subset R^n$ chápat tak, že je vždy $F(u, x) \in R^n$), je F hladké, právě když zobrazení, jež každému $u \in B$ přiřazuje pole $x \mapsto F(u, x)$ na M, je hladkým zobrazením variety B do $\mathscr{V}(M)$. V případě vertikálního pole na $B \times M$ budeme obvykle toto pole i zmíněné zobrazení značit stejným symbolem, nejčastěji F, F_1 apod.; zobrazení $F : B \to \mathscr{V}(M)$ budeme nazývat též soustavou polí. Symbolů F, G apod. budeme však často používat i pro jednotlivá vektorová pole; z kontextu bude jasné, oč jde.

Máme-li rovnici $dx/dt = F(u, x)$, kde u je parametr (popř. rovnici $dx/dt = F(x)$), pak značíme Φ nebo Φ_F její tzv. obecné řešení, tj. zobrazení $\Phi : U \to M$, kde $U \subset B \times \times M \times R$ a pro $u \in B$, $x \in M$ platí o $t \mapsto \Phi(u, x, t)$, že je definováno pro t z jistého otevřeného intervalu, je na něm řešením zmíněné rovnice a nedá se již rozšířit (jako řešení) na větší otevřený interval; obvykle budeme mlčky předpokládat, že tímto intervalem je celé R. Množiny tvaru $\{\Phi(u, x, t) : t \geqq \tau\}$, resp. $\{\Phi(u, x, t) : t \leqq \tau\}$ budeme nazývat *ω-polotrajektoriemi*, resp. *α-polotrajektoriemi* (příslušnými k $u \in B$, $x \in M$, $\tau \in R$).

1.3. Dáme na chvíli veličinám vystupujícím v rovnici $dx/dt = F(u, x)$ jistou interpretaci. Pojmeme x tak, že popisuje „vnitřní stavy“ jistého objektu (systému) za okolností charakterizovaných veličinou u; těmito okolnostmi mohou být vnější podmínky, poloha bodu, v němž se zjišťuje stav, atd.; proměnnou t ovšem interpretujeme jako čas. Konkrétnější případ: jde o složitý chemický systém, x vyjadřuje koncentraci jednotlivých látek (v určitém bodě), u vyjadřuje polohu bodu, o nějž jde, a vnější podmínky, např. množství látek přiváděných zvenčí. Interpretace veličin u, x není ovšem nezbytná pro další úvahy, napomáhá však k jejich osvětlení.

Ptáme se nyní, v jakém stavu nebo stavech se může při fixovaném u ustálit x čili jaké může být „ustálené chování" veličiny x v poli $F(u, x)$. Co se tím míní, naznačíme dále; intuitivně se však zdá jasné, že „ustálené chování" máme např. tehdy, když řešení $t \mapsto \Phi(u_0, x_0, t)$ je konstantní a přitom pro všechna x dostatečně blízká k x_0 je $\Phi(u_0, x, t) \to x_0$ pro $t \to \infty$ (to je zhruba případ jednobodového „atraktoru" x_0; přesnou definici ještě uvedeme). O „ustálené chování" jde patrně také tehdy, když řešení $\Phi(u_0, x_0, t)$ je periodické a přitom každé řešení $\Phi(u_0, x, z)$, kde x je dosti blízké k x_0, k němu v jakémsi smyslu konverguje. Obecněji by šlo o „ustálené chování" tehdy, když řešení $\Phi(u_0, x_0, t)$ je (1) ustálené v čase – zhruba tak, že polotrajektorie $\{\Phi(u_0, x_0, t) : t \geqq \tau_1\}$, $\{\Phi(u_0, x_0, t) : t \geqq \tau_2\}$ jsou si pro velká τ_1, τ_2 velmi blízké, (2) ustálené vzhledem k malým změnám výchozího bodu – zhruba v tom smyslu, že pro x dosti blízké k x_0 je pro velká τ polotrajektorie $\{\Phi(u_0, x, t) : t \geqq \tau\}$ blízká k polotrajektorii $\{\Phi(u_0, x_0, t) : t \geqq \tau\}$. Intuitivně to znamená: nezáleží na okamžiku, v němž začneme pozorovat řešení a nezáleží na malých změnách výchozího bodu – dostáváme vždy přibližně totéž.

Precizně vystihnout uvedené intuitivní požadavky lze různými způsoby, nejvhodněji asi pomocí pojmu atraktoru, jehož definici podáme ve tvaru (až na formální změny), v němž je uvedena v Thomově monografii. Pro čtení článku není tato definice nezbytná; stačí si uvědomovat zmíněné požadavky.

Je-li dáno pole F na M, pak symbolem $\omega(x)$ značíme průnik uzávěrů všech ω-polotrajektorií $\{\Phi_F(x, t) : t \geqq \tau\}$ a obdobně definujeme $\alpha(x)$. Body $y \in \omega(x)$ (resp. $y \in \alpha(x)$) nazýváme *ω-body*, resp. *α-body řešení* $t \mapsto \varphi_F(x, t)$. Množinu $X \subset M$ nazýváme *invariantní* (vzhledem k poli F), jestliže $x \in X$, $t \geqq 0 \Rightarrow \Phi_F(x, t) \in X$. *Atraktorem pole F* nazýváme neprázdnou uzavřenou množinu $A \subset M$ takovou, že (1) je invariantní, (2) existuje otevřená množina $U \supset A$ taková, že $x \in U \Rightarrow \omega(x) \subset A$, (3) když $\alpha(x) \cap A \neq \emptyset$, pak $\Phi(x, t) \in A$ pro všechna $t \in R$, (4) když $\emptyset \neq W \subset A$, W je otevřené v A, pak existuje $x \in W$ tak, že $\omega(x) = A$, tj. každá ω-polotrajektorie procházející bodem x je hustá v A; ustálené chování můžeme definovat např. jako souhrn těch řešení $\Phi(u, x, t)$, která probíhají v daném atraktoru pole $F(u, x)$ a vyplňují jej hustě.

Uvedeme dva příklady atraktorů: (1) bod 0 pro pole $F(x) = -x$ na R^n; (2) kružnice $|x| = 1$ pro pole na R^2 dané výrazem $F(x) = (x_2, -x_1) + \varphi(r)x$, kde $r = |x|$, φ je vhodná nezáporná hladká funkce, $\varphi(r) > 0$ pro $0 < r < 1$, $\varphi(r) < 0$ pro $r > 1$.

Dodejme ještě, že k tomu, aby bod x byl atraktorem pole $F(x)$, je nutné, aby byl jeho izolovaným nulovým bodem; tato podmínka není ovšem postačující, jak ukazují např. pole (na R^2) $F(x, y) = (x, y)$ nebo $F(x, y) = (x, -y)$.

1.4. Jedním ze základním pojmů teorie katastrof je pojem metabolického modelu (poznamenejme, že s metabolismem v biologickém smyslu to nemá žádnou přímou souvislost). Intuitivně řečeno, je to prostě jistá zákonitost, která každé hodnotě parametru, pro niž je to možné, přiřazuje určité ustálené chování, jež se uskutečňuje za podmínek a okolností popsaných touto hodnotou; dá se tedy říci, že jde o souhrn ustálených chování závislých na vnějších podmínkách atd. Přesná definice je tato: *metabolický model* (v Thomově smyslu) je soustava polí $F : B \to \mathscr{V}(M)$ (viz 1.2) spolu se zobrazením σ, jež každému $u \in B$, pro něž pole $x \mapsto F(u, x)$ má aspoň jeden atraktor, přiřazuje jeden z těchto atraktorů (takovému zobrazení σ budeme také říkat „výběr atraktorů").

Poznamenáváme, že Thomův název „model“ není zde možná zcela vhodný: nejde totiž o systém, který by již modeloval reálné jevy, nýbrž jen o systém, který může – případně po doplnění o další data – být takovým modelem. Jde také o pojem velmi široký, zejména proto, že zmíněné zobrazení σ není vlastně jako celek žádným způsobem vázáno na F, a pouze jeho jednotlivé hodnoty jsou určovány pomocí jednotlivých polí $x \mapsto F(u, x)$.

Uvedeme jednoduchý příklad. Nechť $B = R^2$, $M = R$, $H(u, x) = -x^3 + vx + w$, kde $u = (v, w)$. Každé pole $x \mapsto H(u, x)$ má zřejmě jeden nebo dva atraktory, a to jednobodové; množinu jeho atraktorů označme $\mathscr{A}(u)$, sjednocení všech $\mathscr{A}(u)$ označme $\mathscr{A}$. Každá dvojice (H, σ), kde $\sigma : B \to \mathscr{A}$ a $u \in B \Rightarrow \sigma(u) \in \mathscr{A}(u)$, je metabolickým modelem; „použitelné“ jsou jen některé z nich, např. model (H, σ), kde $\sigma(u)$ je „větší“ z atraktorů patřících do $\mathscr{A}(u)$. Žádný „výběr atraktorů“ pro pole H nemůže však být spojitý; bod $0 \in R^2$ je vždy, ať zvolíme σ jakkoliv, hromadným bodem množiny bodů nespojitosti (pro každé $u \neq 0$ se však snadno udá σ, jež je v okolí u spojité).

1.5. Skutečně důležité jsou jen ty metabolické modely, jež jsou v jistém smyslu stabilní. Pojem stability (jednoho vektorového pole, soustavy polí, metabolického modelu aj.) je intuitivně dosti jasný, má však mnoho variant a dá se matematicky precizovat mnoha způsoby. Obvykle se pole F považuje za stabilní, jestliže každé dostatečně blízké pole F_1 je s ním v určitém smyslu ekvivalentní. „Dostatečně blízké“ se nejčastěji chápe ve smyslu topologie prostoru $\mathscr{V}(M)$ (sr. 1.2); v tomto případě máme: F je *stabilní*, jestliže má takové okolí W v prostoru $\mathscr{V}(M)$, že každé $F_1 \in W$ je ekvivalentní s F (v určitém smyslu, jenž je specifikován). Ekvivalence se chápe různě; pro ilustraci uvedeme dvě verze: pole F na M a pole F_1 na M_1 jsou ekvivalentní, jestliže (1. verze) existuje homeomorfní zobrazení $\psi : M \to M_1$, při němž ω-polotrajektorie (α-polotrajektorie) pole F přecházejí v ω-polotrajektorie (α-polotrajektorie) pole F_1 a obráceně; anebo (2. verze) existuje difeomorfismus $\varphi : M \to M_1$, jenž převádí řešení rovnice $\mathrm{d}x/\mathrm{d}t = F(x)$ v řešení rovnice $\mathrm{d}y/\mathrm{d}t = F_1(y)$. Příklad: pole $x \mapsto -x$ na R^n je stabilní podle obou verzí, pole $x \mapsto -x^3$ na R není stabilní podle žádné (důkaz: pole $x \mapsto -x^3$ má jeden atraktor, každé pole $x \mapsto -x^3 + \varepsilon x$, kde $\varepsilon > 0$, má dva atraktory).

1.6. Vložíme teď úvahu, jež má naznačit význam stability z hlediska matematického vyjadřování reálných dějů. Zavedeme následující pojem: máme-li variety B, M, B_1, M_1, pak difeomorfismus $\varphi : B \times M \to B_1 \times M_1$ nazveme *přípustným*, jestliže převádí vertikály (viz 1.2) na vertikály; dá se pak vyjádřit takto: $\varphi(u, x) = (\psi(u), \mu_u(x))$, kde ψ, μ_u jsou hladké a μ_u „závisí hladce“ na u. Často je účelné pojímat φ tak, že udává „změny souřadnic“ na varietě M, přičemž tyto změny závisejí ještě na parametru a „změna souřadnic“ se provádí též na B. Difeomorfismus φ převádí zároveň rovnici tvaru $\mathrm{d}x/\mathrm{d}t = F(x)$ v jinou rovnici téhož tvaru, jejíž řešení se dostanou z řešení původní rovnice pomocí φ. Tyto rovnice lze považovat za různá vyjádření stejné zákonitosti; takové pojetí je např. ve fyzice dosti běžné – aspoň tam, kde a priori nejsou žádné privilegované souřadné systémy.

Všimněme si nyní následující okolnosti: Ve fyzice a příbuzných oborech lze velmi často považovat zákonitost danou vektorovým polem, popř. soustavou vektorových polí, za přesně určenou, a případné nepatrné odchylky lze zanedbat. Jinde, zejména v biologických oborech, bývá často situace jiná: zákonitost není přesně známa, je vyjádřena vektorovým polem jen zčásti, má malé náhodné fluktuace apod. Na druhé straně některé zákonitosti – spíše kvalitativního rázu, často značně složité – se uplatňují např. v biologii velmi striktně; tak třeba embryonální vývoj organismu probíhá po kvalitativní stránce (způsob diferenciace tkání atd.) u každého druhu za normálních okolností přesně určeným způsobem.

To vede k hledání zákonitostí, které zůstávají v jistém smyslu stejné při libovolné dostatečně malé změně. Přesněji řečeno, jde v případě soustav vektorových polí $F : B \to \mathscr{V}(M)$ o taková F, v nichž všechny dostatečně blízké soustavy F_1 jsou ekvivalentní s F v tomto smyslu: existuje přípustný difeomorfismus variety $B \times M$ na $B_1 \times M_1$, při němž řešení (v méně náročné verzi: atraktory) rovnice $dx/dt = F(u, x)$ přecházejí v řešení (resp. atraktory) rovnice $dy/dt = F_1(v, y)$, a to při odpovídajících si hodnotách parametrů u, v. Soustavy F s touto vlastností bychom mohli nazvat globálně stabilními. Fakticky se však v teorii katastrof za jejího nynějšího stavu spíše potřebuje jiný pojem, který naznačíme; přesné definice, i když v rozdílných verzích a obvykle jen pro speciálnější případ, se najdou v literatuře, kterou jsme citovali.

1.7. Říká se, že soustava $F : B \to \mathscr{V}(M)$ je lokálně stabilní v bodě (u_0, x_0), jestliže každá soustava F_1, která je v okolí tohoto bodu dostatečně blízká k F, je v jeho okolí ekvivalentní s F ve smyslu, který jsme již dříve naznačili; obdobně se ovšem definuje lokální stabilita jednotlivého pole.

Uvedeme dva příklady soustav, jež jsou všade lokálně stabilní (při všech běžných definicích): (1) $B = R, M = R, (u, x) \mapsto -x^2 + u$; (2) $B = R^2, M = R, (u, x) \mapsto -x^3 + u_1x + u_2$. Některá jejich jednotlivá pole však nejsou stabilní, a to $x \mapsto -x^2$, resp. $x \mapsto -x^3$. Lze dokázat, že je-li soustava $F : B \to \mathscr{V}(M)$ lokálně stabilní v bodě (u_0, x_0) a je-li též pole $F(u_0)$ stabilní v bodě x_0, pak v okolí bodu (u_0, x_0) je F ekvivalentní s jistou soustavou F_0, u níž všechna $F_0(u)$ jsou stejná, tj. $F(u, x)$ závisí jen na x. „Zajímavé" jsou tedy v případě lokálně stabilního F jen ty body $u \in B$, pro něž pole $F(u)$ není všude lokálně stabilní. Takovým bodům se říká *katastrofické* (používá se však i jiných verzí tohoto pojmu). Tak třeba v uvedených příkladech je katastrofickým bodem $u = 0$, resp. $(u_1, u_2) = (0,0)$.

Poznamenejme, že jedním z nejdůležitějších problémů teorie katastrof je otázka začlenění daného pole na M do lokálně stabilní soustavy polí $F : B \to \mathscr{V}(M)$. Pro gradientová pole (viz 1.9) se ukazuje, že „zpravidla" (ve smyslu, který se přesně definuje) je takové začlenění možné; některá velmi jednoduchá pole se však takto zapojit nedají, např. pole $(x, y) \mapsto -x^2$.

Pro metabolické modely (F, σ) lze obdobné pojmy zavést různými způsoby. Tak např. v zjednodušené verzi lze definovat katastrofické body pro (F, σ) jako body, jež buď jsou katastrofické pro F, anebo v nichž σ není spojité.

1.8. Můžeme již nyní vymezit, ovšem jen zhruba, matematický obsah teorie katastrof. Základ tvoří zkoumání metabolických modelů, a to zejména se zřetelem k lokální stabilitě, ke katastrofickým bodům a k jejich klasifikaci, jakož i ke zkoumání příbuzných útvarů. Tím nyní míníme zejména lokálně stabilní soustavy polí; máme u nich např. problémy lokální klasifikace (jež jsou již do značné míry vyřešeny pro případ gradientových polí, viz 1.9), otázky související se „zapojením" daného pole do lokálně stabilní soustavy, složitou problematiku hledání různých vhodných verzí stability atd. Přitom u četných otázek je řešení, jak se zdá, dosti obtížné, mj. proto, že po strukturální stránce je o vektorových polích, resp. o vektorových diferenciálních rovnicích, známo poměrně málo, pokud nejde o velmi nízké dimenze.

Kromě uvedeného základu zahrnuje teorie katastrof též zkoumání řady jiných matematických útvarů, které se dostanou např. „obohacením" metabolického modelu nebo soustavy vektorových polí o další aspekty — o vnitřní dynamiku vztahující se na změny veličiny u, o průběhy veličiny u pojaté jako řídící parametry; někdy se zkoumají (jak uvedeme v odd. 2) též systémy (struktury), jež jsou abstraktnější než metabolické modely a podobné útvary, anebo se získávají jinak než z vektorových polí a diferenciálních rovnic.

Pokud jde o metodu katastrof (viz 1.1), lze z hlediska reálných dějů, k jejichž vystižení by měla sloužit, říci toto: je zaměřena k vyjadřování jistých druhů kvalitativních a náhlých změn v sepětí se změnami kvantitativními, povlovnými; dá se také říci, že jde o určitý způsob matematického vyjadřování vzniku a vývoje tvarů. V poněkud jiném pohledu lze říci, že metoda katastrof slouží ke zkoumání zákonitostí, jež jsou kvantitativně volné („nepřesné"), tvarově striktní; tato formulace vlastně rekapituluje v zjednodušené zkratce to, co jsme říkali o stabilitě soustav polí.

Je třeba ještě upozornit, že by bylo nesprávné považovat metodu katastrof za obecnou matematickou metodu vystihování všech situací, v nichž dochází k přechodu od povlovných změn ke změnám náhlým, resp. kvalitativním. Bylo by to mylné ze dvou důvodů: proto, že pro vyjádření situací uvedeného druhu má matematika různé další prostředky, a proto, že výskyt zmíněných přechodů neznamená ještě sám o sobě, že by použití teorie katastrof bylo vhodné.

1.9. Důležitou a za nynějšího stavu nejrozvinutější součástí teorie katastrof je tzv. elementární teorie zkoumající případ, kdy jde o gradientová vektorová pole.

Pro zjednodušení se omezíme na případ $M \subset R^n$; nazveme pak vektorové pole $F : M \to R^n$ gradientovým, jestliže existuje funkce $P : M \to R$ taková, že pro každé $x = (x_1, \ldots, x_n)$ platí: $F(x) = -(\partial P/\partial x_i)$. Atraktory pole F jsou pak (pokud existují) vesměs jednobodové; jsou to právě body x_0 takové, že pro dostatečně blízká $x \neq x_0$ je $P(x) > P(x_0)$. Soustavu $F : B \to \mathscr{V}(M)$ a metabolický model (F, σ) nazýváme pak ovšem gradientovými, jestliže pole $x \mapsto F(u, x)$ jsou vesměs gradientová.

S gradientovými metabolickými modely se pracuje mnohem lépe, než je tomu v obecném případě; mj. můžeme zpravidla uvažovat pouze o funkcích (nikoli o vektorových polích) na M a zavádět pojmy stability atd. pro tyto funkce. Proto se v elementární teorii katastrof již dospělo např. k velmi významným klasifikačním větám. O tom zde však nebudeme mluvit, neboť nejdůležitější výsledky byly již v tomto časopise vyloženy precizním způsobem v článku [2]. Poznamenáme jen, že v převážné většině případů, kdy se metoda katastrof má uplatnit pro konkrétní reálné situace, mají klasifikační věty podstatný význam, i když často spíše heuristický.

2.

2.1. Budeme zde mluvit o souvislostech teorie katastrof s některými matematickými disciplínami a o některých útvarech, které dostáváme modifikací nebo ,,obohacením" útvarů z odd. 1.

Souvislost s útvary, které probereme nejdříve, je známá, ale v literatuře se o ní mluví dosti zřídka, ač může být důležitá tam, kde při uplatňování metody katastrof se má pokročit k detailnějším a do jisté míry kvantitativním úvahám. Jde o následující útvary, jejichž vztah k teorii katastrof probereme až v 2.4.

Mějme dvě vektorové diferenciální rovnice

$$\varepsilon \, \mathrm{d}x/\mathrm{d}t = F(u, x), \text{ kde } \varepsilon > 0; \tag{1 ε}$$

$$\mathrm{d}u/\mathrm{d}t = G(u, x), \tag{2}$$

přičemž u, x probíhají variety B, M; fakticky bude vždy $B \subset R^p$, $M \subset R^n$. Poznamenejme, že v případě $G = 0$ máme vlastně soustavu polí $F : B \to \mathscr{V}(M)$ (viz 1.2).

Nechť nyní $\varepsilon \to 0$. Dostáváme to, čemu se někdy říká degenerovaná soustava; zapisuje se obvykle (ale poněkud nepřesně) ve tvaru

$$F(u, x) = 0, \tag{0}$$

$$\mathrm{d}u/\mathrm{d}t = G(u, x). \tag{2}$$

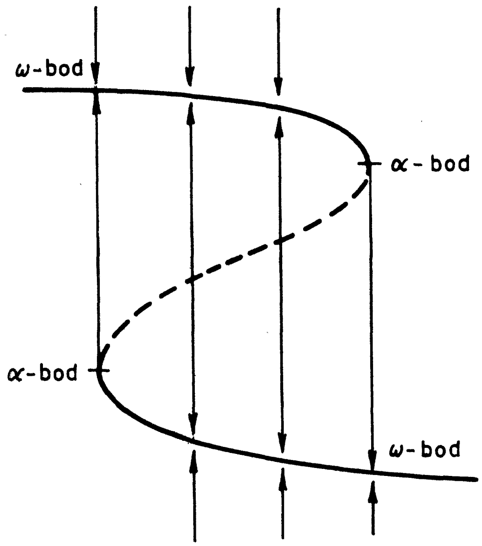
Zde budeme raději mluvit o Tichonovových systémech: první podstatné a řádně dokázané věty o konvergenci řešení (ve smyslu, o kterém bude dále zmínka) pocházejí totiž od A. N. Tichonova [8, 9]. Přesná definice není zde nezbytná, neboť intuitivně je dostatečně jasné, oč jde. Kdyby byla nutná, mohli bychom např. definovat ,,Tichonovovo pole" na $B \times M$ jako ekvivalenční třídu vektorových polí na $B \times M$, přičemž ekvivalence se zavádí takto: dvě pole, jejichž složky tečné k M, resp. k B, jsou F, F_1, resp. G, G_1, se prohlásí za ekvivalentní, jesliže $G = G_1$ a jestliže (1. varianta) pro vhodné číslo $\gamma > 0$ je $F_1 = \gamma F$, anebo (2. varianta) pro vhodnou kladnou funkci γ na $B \times M$ platí $F_1 = \gamma F$.

Složitější je otázka vhodného pojetí řešení Tichonovova systému. Předpokládáme-li, že všechny atraktory rovnice (1 ε) jsou jednobodové, lze řešení pojmout takto: vznikají ,,napojováním" (I) řešení soustavy (0) + (2) chápaných jako pohyb po ,,atraktorové" části množiny $Z_F = \{(u, x) : F(u, x) = 0\}$, a (II) řešení soustavy (1 ε) pojatých jako ,,skokový pohyb". Přitom se ,,atraktorovou částí" zmíněné množiny rozumí množina $S = S_F$ bodů (u, x) takových, že $\{x\}$ je atraktor pro pole $x \mapsto F(u, x)$; pohybem po této části se míní pohyb určený polem G (podrobnosti vynecháváme).

,,Skokové pohyby" se dostanou zejména následující konstrukcí, která je ovšem možná jen při některých (,,katastrofických") hodnotách u. Při fixovaném u se vezme řešení soustavy (1 ε), jehož jediným α-bodem je některý hraniční bod atraktorové části v Z_F a jediným ω-bodem některý vnitřní bod atraktorové části; trajektorie tohoto řešení se chápe jako trajektorie ,,skokového pohybu". Příklad ,,skokových pohybů" uvádíme na obr. 1.

Uvedený náznak definice je ovšem třeba precizovat, a jsou možné také jiné verze.

Definice uváděné v literatuře nejsou vždy zcela precizní a často se týkají spíše poměrně speciálních případů; dostatečně podrobnou a přesnou informaci lze však získat z knih [10, 11]. Zdá se přesto, že přiměřená dostatečně obecná definice řešení Tichonovova systému se teprve bude hledat.



Obr. 1.

Jsou-li definována řešení Tichonovova systému, pak lze již mluvit o tom, zda, v jakém smyslu a kdy konvergují řešení soustavy $(1\ \varepsilon) + (2)$ k řešení příslušného Tichonovova systému. Definice je intuitivně dosti zřejmá, její přesné znění však nebudeme uvádět a odkazujeme na citované knihy.

Přejdeme teď k dalším pojmům a pak se znovu vrátíme k Tichonovovým systémům.

2.2. Nechť je dána soustava polí $F : B \to \mathscr{V}(M)$. Předpokládejme, že pole $F(u)$ mají jen jednobodové atraktory. Pak za určitých předpokladů je každému $(u, x) \in \mathrm{c\,l}\, S \setminus S$ (symbol c l značí zde i v dalším uzávěr, $S = S_F$ má význam z 2.1) přiřazen bod (u, y), do něhož se z bodu (u, x) přejde „skokovým průběhem" (viz 2.1). Po jistém zobecnění dospíváme k tomuto pojmu: Nechť je dána (1) soustava polí $F : B \to \mathscr{V}(M)$, jež je určena až na kladného činitele, jímž může být – podle zvolené verze – buď kladná konstanta, nebo všude kladná hladká funkce; soustavu určenou až na takový činitel (tj. příslušnou ekvivalenční třídu) budeme značit $\tilde{F}$ apod.; (2) množina $Q \subset S_F$ taková, že každá vertikála protínající S_F protíná též Q; (3) zobrazení $\psi : P \to Q$, kde P se skládá ze všech $(u, x) \in \mathrm{c\,l}\, Q$ takových, že $\{u\} \times M$ protíná S_F. Požaduje se přitom, aby (u, x) a $\psi(u, x)$ byly vždy na stejné vertikále a aby pro $(u, x) \in Q$ bylo $\psi(u, x) = (u, x)$. Jsou-li splněny tyto požadavky, budeme říkat, že $(\tilde{F}, Q, \psi)$ je přechodový metabolický systém (velmi jednoduchý příklad je patrný z obr. 1). Název zde volíme ad hoc; příslušný pojem nebyl, jak se zdá, v literatuře soustavně vyšetřován, a je také možné, že se ukáže užitečnější jiná verze (např. místo $\tilde{F}$ lze brát určité F). Každému Tichonovovu systému (určenému poli F, G), u něhož jsou jednoznačně určeny „skokové průběhy", je přiřazen zcela určitý přechodový metabolický systém, totiž $(\tilde{F}, S_F, \psi)$, kde ψ vyjadřuje zmíněné skoky; všimněme si, že G při tom nehraje žádnou roli. Také každému metabolickému modelu (v Thomově smyslu; viz 1.4) je v případě jednobodových atraktorů přiřazen systém $(\tilde{F}, Q, \psi)$, kde $Q = \{(u, \sigma(u)) : u \in B\}$, a jestliže $(u, x) \in \mathrm{cl}\, Q$, přičemž pole $x \mapsto F(u, x)$ má atraktor, pak $\psi(u, x) = (u, \sigma(u))$. Je přitom jasné, že různým metabolickým modelům odpovídají různé přechodové metabolické systémy.

Je zřejmé, že některé zcela přirozené přechodové metabolické systémy (např. systém znázorněný na obr. 1) se nedají dostat uvedeným způsobem. Tento pojem je tedy širší než pojem metabolického modelu. Přitom se fakticky již v teorii katastrof vyskytoval, byť implicitně, a neznamená tedy žádné obsahové rozšíření obvyklé koncepce; někdy se s ním však pracuje lépe, je-li formulován explicitně.

2.3. Metoda katastrof používá dosti často, leckdy bez výslovné zmínky, též některých struktur získaných další abstrakcí z metabolických modelů a přechodových metabolických systémů. Jednu z nich lze dostat takto (omezujeme se stále na případ jednobodových atraktorů): u přechodových metabolických systémů $(\tilde{F}, Q, \psi)$ odmyslíme $\tilde{F}$ a bereme v úvahu jen varietu $B \times M$, množinu S_F (viz 2.1; to, že F je určeno jen „až na vynásobení kladným γ“, nehraje roli pro určení S_F) a ovšem též Q, ψ. Struktury tohoto druhu se fakticky vyskytují dosti často při používání metody katastrof. Mohou být důležité mj. tím, že již nejsou přímo vázány na diferenciální rovnice a vektorová pole, a přesto lze u nich použít četných myšlenek a výsledků teorie katastrof; to pak dává aplikační možnosti pro širší rozsah situací.

Poznamenejme, že další abstrakcí lze přejít ke strukturám ryze topologického rázu, u nichž se však uplatňují četné pojmy teorie katastrof. Míníme např. struktury dané topologickými prostory P, B a zobrazeními $\pi : P \to B$, $\varphi : P \to P$, která splňují jisté podmínky. V četných případech záleží totiž u přechodového metabolického modelu $(\tilde{F}, Q, \psi)$ vlastně jen na topologických vlastnostech množiny $Q \subset B \times M$ a zobrazení ψ. Potom můžeme vzít za P množinu $(u, y) \in \mathrm{cl}\, Q$ takových, že $\{u\} \times M$ protíná S_F; za π se vezme projekce $P \subset B \times M$ do B, za φ se vezme ψ. Jako ilustrace poslouží zde zase obr. 1, když vezmeme za P uzávěr atraktorové části S_F a chápeme P jako topologický prostor.

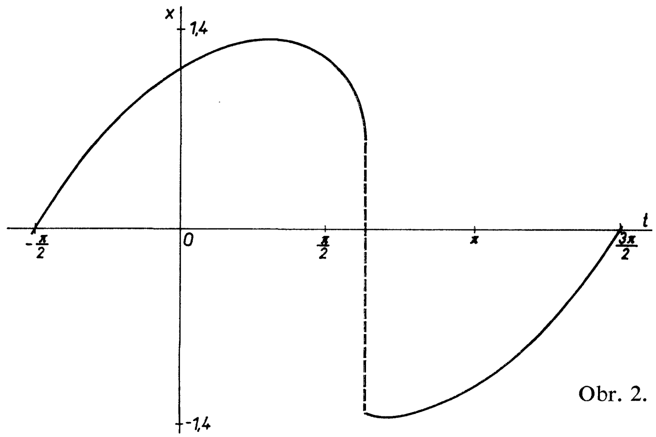
2.4. Přejděme teď k „obohaceným“ útvarům (sr. 1.8). Upustíme od obecné charakteristiky a uvedeme některé jednotlivé druhy takových útvarů. Nejčastěji jde v teorii katastrof o „obohacování“ pomocí tzv. „pomalé“ autonomní dynamiky nebo „řídících průběhů“ (viz. 2.5).

Obraťme se nejdříve k autonomní dynamice. Intuitivně řečeno, jde o to, že veličina u, která dosud, např. u soustavy polí, vystupovala jako parametr, se nyní skutečně pojme jako proměnná veličina, přičemž její změny se řídí diferenciální rovnicí tvaru $\mathrm{d}u/\mathrm{d}t = G(u, x)$. Dospíváme pak snadno k příslušným definicím „obohacených“ útvarů. Tak zejména vyjdeme-li ze soustavy $F : B \to \mathscr{V}(M)$, dostáváme dvojici (F, G), již lze chápat jako hladké vektorové pole (libovolné) na $B \times M$. Vyjdeme-li z $\tilde{F}$, tj. ze soustavy F určené „až na vynásobení kladným činitelem“, dostáváme dvojici $(\tilde{F}, G)$, již lze chápat jako Tichonovův systém. Pole G se v těchto případech často nazývá pomalou dynamikou, autonomní pomalou dynamikou apod., pole F rychlou dynamikou. Důvod tohoto pojmenování je v tom, že v důležitých případech je G skutečně „pomalé" vůči F; v tom případě řešení soustavy $dx/dt = F$, $du/dt = G$ a řešení příslušného Tichonovova systému jsou si (za jistých předpokladů) blízká a jedno může sloužit jako aproximace druhého.

Shrneme-li nyní některé souvislosti Tichonovových systémů s teorií katastrof, zjistíme mj. toto: Zanedbá-li se jejich „pomalá dynamika", můžeme dostat (sr. 2.2) přechodové metabolické systémy, popř. metabolické modely; na druhé straně, Tichonovovy systémy lze získat ze soustav polí (určených „až na kladného činitele") přidáním pomalé dynamiky; nejdůležitější je možná to, že zprostředkují aproximaci přechodového metabolického systému pomocí řešení soustavy diferenciálních rovnic, jak je patrné z toho, co jsme již dříve uvedli.

Obdobně jako $\tilde{F}$ a F lze „obohatit" také např. přechodový metabolický systém $(\tilde{F}, Q, \psi)$. Upustíme od definice, uvedeme však příklad. Mějme $B = R^2$, $M = R$, $F(u, x) = -x^3 + vx + w$, kde $u = (v, w)$; $Q = S_F$, tj. Q je celá „atraktorová" část plochy $F(u, x) = 0$; je-li $F(u, x) = 0$, $\partial F(u, x)/\partial x = 0$, $\partial^2 F(u, x)/\partial x^2 \neq 0$, pak $\psi(u, x) = (u, y)$, kde y je další kořen polynomu $f(u, x)$; pro ostatní (u, x) je $\psi(u, x) = (u, x)$. Vezmeme nyní „pomalou" dynamiku $G(u, x) = (w, -v)$; $G(u, x)$ tedy nezávisí na x. Tato dynamika spolu s (F, Q, ψ) určuje pohyby na ploše S_F, jež se kombinují s „přeskoky"; pohyby jsou, jak je patrné, periodické a obsahují v každé periodě právě jeden „přeskok" (necháme-li bez povšimnutí stacionární bod $u = 0$).

Jestliže pro $(v, w) \in R^2$ označíme $\lambda(v, w)$ největší reálný kořen polynomu $-x^3 + vx + w$ a položíme počáteční podmínku $u(0) = (0, r)$, pak se na intervalu $[0, \infty)$ všechny zmíněné pohyby s přeskoky vyjádří funkcemi $x(t) = \lambda(r \sin t, r \cos t)$, $v(t) = r \sin t$, $w = r \cos t$. Funkce $x(t)$ je nespojitá pouze v bodech t takových, že $4\, r \sin^3 t = 27 \cos^2 t$, $\cos t < 0$; v těchto bodech dochází ke skoku veličiny x z hodnoty $+(\frac{1}{3} r \sin t)^{1/2}$ na hodnotu $-2. (\frac{1}{3} r \sin t)^{1/2}$. Podrobnější rozbor přenecháváme čtenáři, upozorňujeme



Obr. 2.

pouze, že, jak známo, souřadnice v, w hraničních bodů atraktorové části plochy $F(u, x) = 0$ musí splňovat podmínku $4v^3 = 27 w^2$; jen v těchto bodech může docházet k „přeskokům". Grafické znázornění funkce $\lambda(r \sin t, r \cos t)$ pro $r = \sqrt{2}$ je uvedeno na obr. 2.

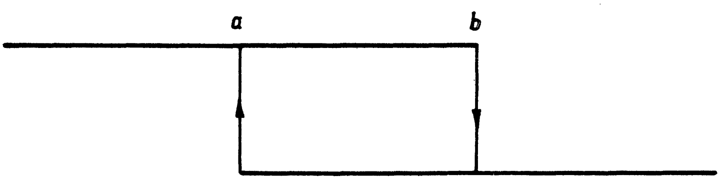
Závěrem ještě poznámku spíše intuitivního rázu: na autonomní dynamice u je nejpodstatnější to, že „dává pohyb“ útvaru, např. přechodovému metabolickému systému, který sám obsahuje teprve možnost pohybu; to, že jde o diferenciální rovnice atd., není tak důležité. Podstatná je též „autonomnost“ dynamiky; znamená, že možné pohyby jsou plně určovány lokální situací v jednotlivých bodech prostoru.

2.5. Budeme se nyní zabývat dynamickými systémy, o nichž by se mohlo též mluvit jako o systémech s regulací (řízením). Chápeme je zde zhruba tak, jak jsou pojaty v knize [12]; poznamenejme, že výrazu „dynamický systém“ se používá i v jiných významech. Uvidíme později, že dynamické systémy jsou značně důležité pro teorii katastrof pojatou v širším smyslu (sr. 1.8).

Uvedeme nyní příklady a pak naznačíme definici; podrobnější a přesnější informaci lze najít např. v [12].

Příklad 1. Nechť je dána rovnice $\mathrm{d}x/\mathrm{d}t = F(u, x)$, kde x probíhá varietu M, $u \in B$ je parametr. Mějme bod x_0 a „časový průběh“ veličiny u, čímž teď rozumíme zobrazení $g : [t_0, t] \to B$, kde $[t_0, t] \subset R$ je uzavřený konečný interval. Je-li g spojité, pak za značně obecných předpokladů dostáváme z rovnice $\mathrm{d}x/\mathrm{d}t = F(g(t), x)$ a počáteční podmínky $f(t_0) = x_0$ zcela určitý časový průběh $f: [t_0, t] \to M$ veličiny x a časový průběh $\tau \mapsto (g(\tau), f(\tau))$ veličiny (u, x). Příklad lze interpretovat takto: $u \in B$ vyjadřuje vnější podmínky, $x \in M$ vyjadřuje vnitřní stav jistého objektu, $F(u, x)$ vyjadřuje jistou zákonitost; časový průběh vnějších podmínek, jejž můžeme nazývat „řídícím průběhem“, a výchozí vnitřní stav určují „stavový průběh“, tj. časový průběh vnitřního stavu x; popř. „celkového stavu“ (u, x).

Příklad 2. Nechť $B = R$. Nechť jsou dána čísla a, b, přičemž $a < b$. Označme P množinu $\{(x, 0) : x \geqq a\} \cup \{(x, 1) : x \leqq b\} \subset R^2$. Budeme chápat pohyb v P jako spojitý pohyb po horní nebo dolní polopřímce kombinovaný s „přeskoky“ z bodu $(a, 0)$ do $(a, 1)$ nebo z bodu $(b, 1)$ do $(b, 0)$; viz obr. 2 (poznamenejme, že obr. 3 se vlastně až na spojitou deformaci shoduje s tím, co máme na obr. 1, když na něm z křivky $F(u, x) = 0$ ponecháme jen uzávěr atraktorové části).



Obr. 3.

Mějme nyní spojitý řídící průběh $g : [t_0, t] \to B$ a bod $x_0 = (u_0, h_0) \in P$, přičemž $g(t_0) = u_0$. Jde-li g dostatečně jednoduché, totiž je-li po částech monotonní, pak je názorně evidentní, že je v podstatě jednoznačně určen pohyb $f: [t_0, t] \to P$ odpovídající řídícímu průběhu g, totiž takový, že pro $t_0 \leqq \tau \leqq t$ je $g(\tau)$ projekcí bodu $f(\tau)$ do R. Stavové průběhy nemusí být přitom spojité, splňují však požadavek, aby v případných bodech nespojitosti mělo f limitu zleva $f(\tau -)$ i limitu zprava $f(\tau +)$, přičemž buď $f(\tau -) = (a, 0)$, $f(\tau +) = (a, 1)$ nebo $f(\tau -) = (b, 1)$, $f(\tau +) = (b, 0)$. Hodnota $f(\tau)$

v bodě nespojitosti není zmíněnými požadavky plně určena a lze za ni vzít $f(\tau -)$nebo $f(\tau +)$; proto jsme také řekli jen, že f je „v podstatě jednoznačně" určeno řídícím průběhem g. To všechno je dosti zřejmé pro jednoduchá g; poměrně snadno se v našem příkladě dokáže, že ve skutečnosti každý spojitý řídící průběh g určuje „v podstatě jednoznačně" příslušný stavový průběh f.

Dodejme ještě, že když modifikujeme příklad tak, že z množiny P vynecháme body $(a, 0)$, $(b, 1)$, pak je již každému spojitému g skutečně jednoznačně přiřazen stavový průběh f.

Příklad 3. Tento příklad jen naznačíme. Položíme $B = R^2$, $M = R$. Vezmeme soustavu polí $F : B \to \mathscr{V}(M)$ a příslušný přechodový metabolický systém (viz 2.2); konkrétně budeme uvažovat o případu $F(u, x) = -x^3 + vx + w$, kde $u = (v, w)$. Označíme P atraktorovou část plochy $F(u, x) = 0$. Spojitému řídícímu průběhu g : $: [t_0, t] \to B$ a bodu $(v_0, w_0, x_0) \in P$ takovému, že $(v_0, w_0) = g(t_0)$, přiřadíme, pokud to lze provést v podstatě jednoznačně, stavový průběh $f : [t_0, t] \to P$ takový, že $f(t_0) = (v_0, w_0, x_0)$, $f(\tau)$ se vždy promítá (při projekci $B \times M \to B$) do bodu $g(\tau)$ a f je nespojité jedině tam, kde je to dáno přechodovým metabolickým systémem. Tak mj. časovému průběhu danému formulemi $v = r \sin t$, $w = r \cos t$ odpovídá stavový průběh popsaný v 2.4 (viz též obr. 2). Dá se zjistit, že „skoro všem" spojitým řídícím průběhům odpovídají jednoznačně příslušné stavové průběhy (jež však, obecně řečeno, jsou v některých bodech nespojité). Skoky, které se vyskytují u těchto průběhů, jsou právě skoky (přechody) obsažené ve zmíněném přechodovém metabolickém systému.

Shrneme nyní některé společné důležité rysy uvedených příkladů. Měli jsme vždy jisté topologické prostory B, P (v našich příkladech bylo vždy $P \subset B \times M$, v příkladě 1 pak přímo $P = B \times M$), projekci (spojitou) prostoru P do B (tuto projekci označíme π) a — což je nejpodstatnější — jistou korespondenci, která každému, popř. „skoro každému" spojitému řídícímu průběhu $g : [t_0, t] \to B$ a bodu $x_0 \in P$ (takovému, že $\pi(x_0) = g(t_0)$) přiřazovala stavový průběh $f : [t_0, t] \to P$ (případně nespojitý), splňující vždy podmínku $\pi(f(\tau)) = g(\tau)$.

Tímto shrnutím jsme již zároveň naznačili obecnou definici dynamického systému; bylo by jen třeba precizovat některé formulace a hlavně položit výslovně na zmíněnou korespondenci některé přirozené požadavky.

Z uvedeného náznaku definice a z příkladů je již zhruba patrné, jak lze pojmout obohacení (viz 1.8, 2.4) soustavy polí, přechodového metabolického systému apod. pomocí průběhů řídící veličiny. Tak např. máme-li přechodový metabolický systém, pak se mu přiřadí dynamický systém způsobem, který byl pro speciální případ naznačen v příkladě 3. U tohoto systému pak můžeme vzít v úvahu určitou neprázdnou množinu (opatřenou případně jistou strukturou) přípustných řídících průběhů. Takovou množinu můžeme nazvat heteronomní dynamikou daného přechodového metabolického systému.

2.6. Jedním ze zvlášť důležitých případů je ten, kdy zmíněnou množinu přípustných řídících průběhů lze pojmout jako množinu realizací jistého stochastického procesu; mluvíme pak o stochastické heteronomní dynamice. Protože tuto dynamiku budeme ještě potřebovat v odd. 3, připomeneme teď ve stručnosti a zjednodušeně definici stochastického procesu v té formě, v jaké se nám bude hodit později.

Řečeno intuitivně, máme náhodný (stochastický) proces tehdy, když se jistá veličina mění v čase náhodným způsobem podle určitých pravděpodobnostních zákonitostí. Dá se pak pro každý okamžik t mluvit např. o pravděpodobnosti toho, že zkoumaná veličina (jde-li o veličinu nabývající reálných hodnot) nabyla hodnoty z určitého intervalu; dá se také mluvit o pravděpodobnosti toho, že celkový časový průběh veličiny má tu a tu vlastnost, ovšem jen pokud jde o vlastnost v jistém smyslu „dostatečně jednoduchou".

Dostatečně přesnou (i když poněkud odlišnou od běžných formulací) definici stochastického procesu lze nyní vyslovit pro případ veličiny s číselnými hodnotami tímto způsobem: Stochastický proces je dán, je-li dána neprázdná množina $T \subset R$ (velmi často se za T bere množina všech nezáporných reálných čísel nebo všech přirozených čísel), neprázdná množina $A \subset R^T$ (množina přípustných časových průběhů čili realizací stochastického procesu) a pravděpodobnostní rozdělení na A splňující jisté podmínky; přesněji řečeno, je dána množinová σ-algebra $\mathscr{A}$ na A a σ-aditivní pravděpodobnostní míra na $\mathscr{A}$; předpokládá se, že každá množina tvaru $\{f \in \mathscr{A} : \alpha < f(t) < \beta\}$, kde $t \in T$, $\alpha, \beta \in R$, patří do $\mathscr{A}$.

Uvedeme ještě příklad (dosti triviální) stochastického procesu. Položíme $T = \{x : x \geqq 0\}$, za A vezmeme množinu všech spojitých $f : T \to R$, pro než $f'(t) = \pm 1$ na každém intervalu $(n, n + 1)$. Za $\mathscr{A}$ vezmeme nejmenší z těch σ-algeber, které se skládají z podmnožin množiny A a obsahují všechny množiny tvaru $M(t, \alpha, \beta) = \{f \in A : \alpha < f(t) < \beta\}$ kde $t \in T$, $\alpha, \beta \in R$. Stochastický proces, o který nyní jde, se dá intuitivně popsat takto: veličina x se mění v každém intervalu $(n, n + 1)$ lineárně s rychlostí 1 nebo -1; v celočíselných bodech se může náhodně změnit rychlost, a to takovým způsobem, že pro každý interval $(n, n + 1)$ je rychlost 1 i rychlost -1 stejně pravděpodobná.

Formální definici procesu, který zde máme, nebudeme explicite uvádět; dostane se však snadno, když si uvědomíme, že na intervalu $[0, n]$ má proces celkem 2^n různých realizací, jež jsou stejně pravděpodobné; každá z nich odpovídá některé posloupnosti $(v_1, \ldots, v_n)$, kde $v_i = \pm 1$ je rychlost na intervalu $(i - 1, i)$.

3.

3.1. Budeme se nyní zabývat použitím některých myšlenek metody katastrof při matematickém modelování roztroušené sklerózy mozkomíšní (poznamenejme, že přes svůj název nemá tato choroba žádnou souvislost se sklerózou cév, arteriosklerózou, i když tato arterioskleróza postihuje mozkové tepny). Zmíněné použití vystupuje v nynějším článku především jako ukázka aplikace elementárních pojmů Thomovy teorie v kombinaci s klasickými prostředky; v chystané druhé části článku poslouží též k ilustraci obecných úvah o aplikacích.

Průběh onemocnění je u různých pacientů velmi rozdílný; jestliže se časový průběh vyjádří tak, že se celkový stav pacienta ohodnotí podle určitých zavedených zásad na jisté stupnici, dostáváme křivky velmi rozdílného tvaru. Tato rozmanitost ztěžuje matematické modelování a zároveň je jedním z důvodů, proč použití matematických metod při rozboru zmíněných průběhů je zajímavé i po matematické stránce.

3.2. Aby popis modelu a jeho účel byl srozumitelný, uvedeme teď ve zhuštěné formě některé základní údaje o celkové povaze roztroušené sklerózy. Bližší informaci lze najít v každé větší učebnici neurologie; četné křivky průběhu onemocnění, avšak jen pro chronicko-progresívní období (viz dále) jsou uvedeny v publikaci [13], jež obsahuje též stručné údaje o literatuře.

Roztroušená skleróza mozkomíšní je velmi závažné onemocnění, postihující centrální nervový systém, mozek a míchu mnoha roztroušenými, vzájemně nesouvisejícími ložisky. V těchto ložiskách dochází k rozpadu pochvy nervových vláken, která v oblasti ložiska ztrácejí schopnost vedení vzruchu. Je-li vedení vzruchu přerušeno na jednom nebo více místech nervového vlákna nebo skupiny vláken, ztrácejí tato vlákna schopnost své základní funkce, totiž předávání vzruchu, vytvořeného v těle nervové buňky, na jiné nervové buňky až konečně na motorický aparát, svalstvo.

Tato porucha funkce nervových vláken se pak projevuje u nemocného ztrátou některých funkcí nervového systému. Jsou postiženy převážně funkce hybné, zvláště hybnost dolních končetin, často je porušena schopnost přesného a cíleného pohybu v důsledku poruchy mozečkových funkcí nebo dostředivých drah, je patrna porucha rovnováhy a okohybných funkcí při poruše vláken mozkového kmene a velmi charakteristická je i porucha zraku, způsobená poruchou funkce zrakového nervu. Nejčastější a nejzávažnější je porucha hybnosti dolních končetin.

Roztroušená ložiska v mozku a míše nevznikají naráz, ale postupně a mezi vznikem dalších ložisek se může funkce nervových vláken aspoň částečně a na začátku onemocnění často i úplně upravit. Vznik ložisek se projevuje náhlou ztrátou některých funkcí; tuto náhlou ztrátu nazýváme atakou, úpravu funkcí nazýváme remisí. Po nějaké době, v průměru za 1,5 roku, dojde k nové atace, sledované novou remisí. Tento stav se několikrát opakuje, přitom stupeň remise, tj. úpravy funkce, se obyčejně pozvolna zmenšuje, až dojde k tomu, že stav se již neupravuje, a vlivem vzniku dalších ložisek se případně buď nárazovitě, nebo plynule pozvolna zhoršuje. Toto zhoršování nezřídka vede k pozvolné, nejprve částečné, později někdy i úplné ztrátě schopnosti samostatného pohybu; v nepříznivých případech je nemocný upoután na lůžko, někdy po mnoho let, a podléhá některému zánětlivému onemocnění, které napadlo oslabený organismus.

Uvedený průběh, charakterizovaný na začátku onemocnění vznikem náhlé poruchy funkce, sledovaným pozvolnější restitucí, remisí, vyjadřuje nejzřetelněji typické rysy tohoto onemocnění. Onemocnění však také může hned na začátku probíhat bez remisí, chronicko-progresívně, a vyskytují se i jiné typy průběhu, z nichž jsou některé příznivější.

3.3. Tento zvláštní průběh, s atakami a remisemi, si lze vysvětlit působením vlastní patogenní noxy, o níž se zmíníme později, na funkci nervových vláken, tzv. axonů. Při vzniku ataky a ztrátě tzv. myelinové pochvy, složené z lipidů a bílkovin, dojde ke ztrátě funkce. Později, aniž se pochva obnoví, dojde k úpravě funkce, takže axon přejde znovu do stavu funkčního. Jde však o jiný stav, nežli když byl axon zcela neporušen, a to právě pro ztrátu myelinové pochvy. Axon však může přejít — a v pozdějších stadiích nemoci většinou skutečně přechází — do stavu, kdy je definitivně a ireverzibilně ztracena i jeho funkce. Můžeme tedy rozeznávat čtyři anatomicko-funkční stavy. Stav A, kdy je axon neporušen po stránce anatomické a funkční, stav B, kdy je axon porušen po stránce anatomické, ale vyhovuje po stránce funkční, stav C, kdy je porušen po stránce anatomické i po stránce funkční, ale je schopen restituce, a stav D, kdy je porušen po stránce anatomické i funkční a není již schopen restituce. Ataka probíhá za vzniku axonů typu C z axonů typu A (popř. B), remise probíhá za vzniku axonů typu B z axonů typu C. Narůstání trvalých poruch je způsobeno stálým narůstáním množství axonů typu D.

Vznik roztroušených ložisek v centrálním nervovém systému při této nemoci byl donedávna zcela nevysvětlen. Dnes se nahromadilo již množství nových faktů, která nám dovolují učinit si celkem ucelený obraz o podstatě tohoto onemocnění. Jde o onemocnění zánětlivé, způsobené virem typu paramyxovirů, nejspíše virem spalniček nebo jemu podobným. Pomnožení viru vyvolá imunitní reakci typu buněčné imunity, která vede k nahromadění bílých krevních buněk, převážně lymfocytů v nervové tkáni. Tyto lyfmocyty způsobují destrukci pochvy axonů. Tato imunitní reakce není, zdá se, zaměřena přímo proti viru, protože lymfocyty nemocných nerozeznávají virus spalniček jako cizí látku. Reagují však na bílkoviny, které vznikají při pomnožení viru v buňkách tvořících obal axonů. Neschopnost reagovat na spalničkový virus je pravděpodobně vrozená, protože jsou prokázány i jiné vrozené abnormní imunologické reakce u nemocných roztroušenou sklerózou. Pro vrozenou „vnímavost" vůči nemoci, kterou dnes můžeme spíše charakterizovat jako vrozenou abnormní, respektive defektní imunitní reakci, svědčí i to, že z celé populace, která se setkala se spalničkovým virem, onemocní jen někteří jedinci, častěji ženy.

3.4. Popíšeme nyní matematický model roztroušené sklerózy mozkomíšní, k němuž se dospělo při výzkumu prováděném společně oběma autory článku. Nejdříve podáme popis modelujícího matematického systému, a to tak, aby byl v podstatě srozumitelný i bez znalosti pojmů probíraných v odd. 1 a 2 (občasné odkazy na tyto oddíly mají většinou jen poukazovat na širší souvislosti). V 3.6 pojednáváme o věcné interpretaci tohoto systému, tj. uvádíme vlastní model. V 3.7 pak naznačíme úvahy, které vedly k modelu, a připojíme některé poznámky.

Nejdříve se budeme zabývat systémem, jenž má vyjadřovat průběh onemocnění na nevelkých časových úsecích. Máme dvě veličiny z, s, jež slouží k vyjádření dvou různých, ale navzájem souvisejících dějů. O tom mluvíme podrobněji v 3.6; zde jen uvedeme, že veličina s má odpovídat stavu nemocného vyjádřenému podle určité stupnice (viz 3.1), popř. počtu nefungujících axonů (viz. 3.3).

Vzájemný vztah veličin z, s a jejich možné časové průběhy jsou v modelujícím systému určeny (1) jistou funkcí $H(p, z, s)$, resp. rovnicí $H(p, z, s) = 0$, přičemž p lze, dokud se omezujeme na poměrně malé časové úseky, považovat za konstantu, (2) jistými zákonitostmi, resp. omezeními, týkajícími se časového průběhu veličiny z.

Na H se kladou tyto požadavky: (1) dá se získat vhodnou transformací z „kanonického" výrazu $H_0(p, z, s) = -s^3 + p\,s + z$; přesněji řečeno, existuje difeomorfismus φ prostoru vektorů (p, z, s), při němž se tento prostor zobrazuje na sebe, platí $H = H_0 \circ \varphi$ a zachovávají se vertikály i jejich orientace, tj. přechází-li transformací φ bod (p_i, z_i, s_i) v bod $(\bar{p}_i, \bar{z}_i, \bar{s}_i)$ a je-li při tom $p_1 = p_2, z_1 = z_2, s_1 < s_2$, pak je $\bar{p}_1 = \bar{p}_2$,

$\bar{z}_1 = \bar{z}_2$, $\bar{s}_1 < \bar{s}_2$; (2) při vzrůstu z rostou (nebo aspoň neklesají) hodnoty s určené rovnicí $H(p, s, z) = 0$. Později (v 3.5) položíme na H ještě jeden požadavek, a při používání modelu se mohou klást požadavky další.

Výraz $H(p, s\, z)$ určuje časový průběh $s(t)$ v následujícím smyslu: při změně veličiny z se bod (z, s) pohybuje po křivce dané rovnicí $H(p, s, z) = 0$, při dosažení horního, resp. dolního záhybu (bodu, v němž je též $\partial H/\partial s = 0$) však přeskakuje z horní části na dolní, resp. z dolní na horní. Jde tedy o přechodový metabolický systém (viz 2.2) obohacený o heteronomní dynamiku (viz 2.5).

Zákonitosti či spíše omezení pro možné časové průběhy veličiny z formulujeme dosti volně, např. takto: časové průběhy $z(t)$ jsou realizacemi jistého stochastického procesu, přičemž (1) je vždy $a \leqq z(t) \leqq b$, kde veličiny a, b lze, pokud jde jen o nevelké časové úseky, pojímat jako konstanty; (2) stochastický proces lze dostatečně dobře aproximovat (ve smyslu, který se precizuje některým z běžných způsobů) procesem, jenž je „složen náhodným způsobem" z úseků, v nichž z roste od hodnoty a k jakési hodnotě $b' \leqq b$ a pak klesá zpět, a úseků, kde je stále $z = a$.

Uvedený systém již poskytuje některé časové průběhy, jež odpovídají empirickým zjištěním. Jde zejména o průběhy, jež jsou znázorněny na obr. 4; podotýkáme, že neznázorňujeme „kanonickou" křivku $-s^3 + p\, s + z = 0$, nýbrž jistou transformovanou křivku $H(p, s, z) = 0$, že ve shodě s běžným způsobem zakreslování průběhu onemocnění nanášíme kladné hodnoty s směrem dolů a že stavové průběhy jsou zakresleny jen kvalitativně. Průběh I: proměnná s poněkud vzroste a pak se vrátí k výchozí hodnotě; to odpovídá případu, kdy z postupuje od hodnoty a k jakési hodnotě, která je ještě před záhybem, a pak zpět. Průběh II: v jistém okamžiku s podstatně vzroste, a to skokem; potom se nějakou dobu velmi pozvolna zmenšuje, a pak dojde k návratu skokem (s následujícím povlovným „doběhnutím"). Průběh III: podstatný vzrůst s jako v průběhu II, avšak s tím, že po období malého poklesu se proměnná s ustálí v poloze, jež odpovídá hodnotě $z = a$, avšak na dolní části křivky $H = 0$. Průběhy (II), (III) nastávají, postoupí-li veličina z za „záhyb"; o tom, zda nastane (II) nebo (III), rozhoduje pak poloha křivky $H = 0$ vzhledem k přímce $z = a$.

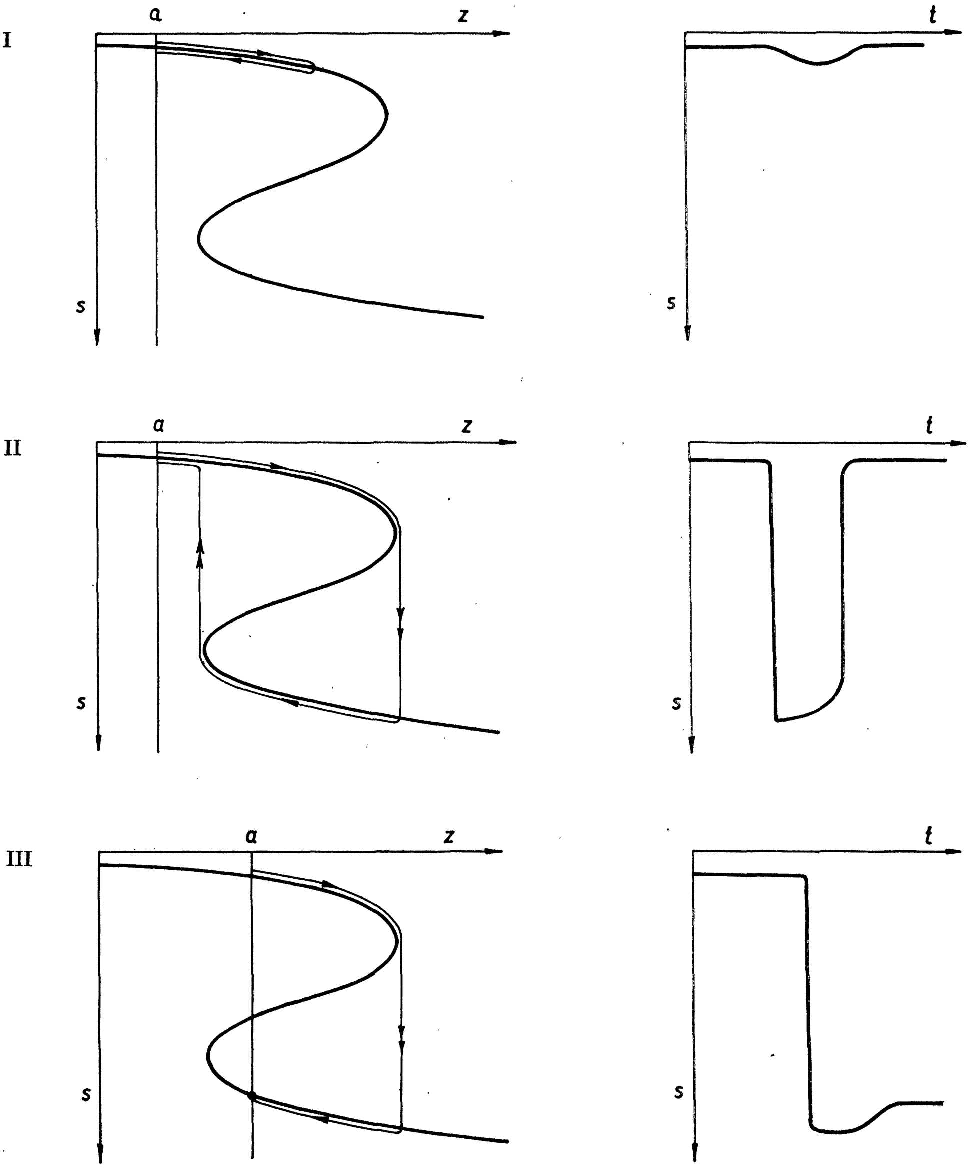
3.5. Systém, který má modelovat delší období, popř. celý průběh onemocnění, se nyní dostane tak, že p, a, b se pojmou jako proměnné. Přitom se předpokládá, že časový průběh $p(t)$ je plně určen časovým průběhem $s(t)$ a že obdobně časový průběh veličiny z plně určuje průběh veličin a, b.

Specifikujeme to takto: (1) $p(t)$ je jistým funkcionálem F_t průběhu veličiny s v intervalu $[0, t]$, (2) $a(t) - a(0)$ je integrálem jisté veličiny $f_a(z(t)) + \varphi_a(t)$, kde f_a, φ_a jsou dané nezáporné funkce, a obdobný vztah platí (s funkcemi f_b, φ_b) pro $b(t) - b(0)$; přitom jsou dosti plauzibilní některá zjednodušení, např. můžeme patrně předpokládat, že je všude $\varphi_a(t) = \varphi_b(t) = 0$. Funkcionály F_t lze pro období atak s remisemi přibližně vyjádřit tak, že $F_t(s)$ je integrálem jisté veličiny $g(s(t)) + h(t)$. Pro chronicko-progresívní období je třeba vzít podstatně jiný výraz; těmito otázkami, jež se již spíše týkají konkrétního uplatnění modelu, se však zde nebudeme zabývat.

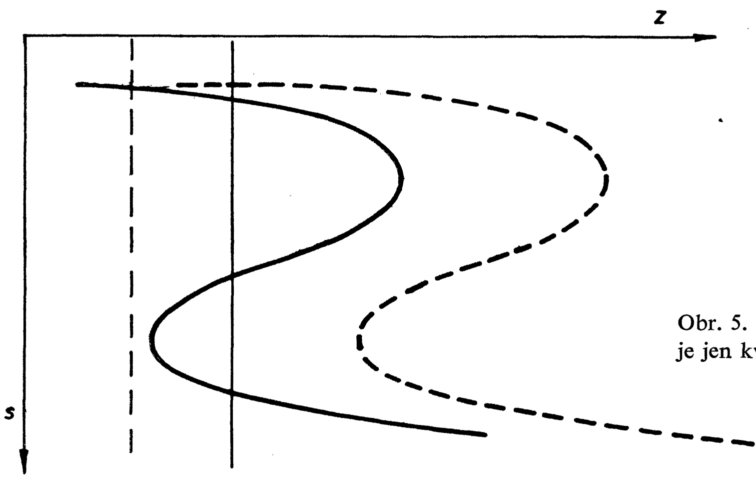
Dodejme, že určité další požadavky se kladou též na H; velmi zhruba řečeno, při rostoucím p se musí křivka $H(p, z, s) = 0$ posunovat doleva. V průběhu doby se tedy křivka

$H = 0$ a přímka $z = a$ sbližují a pak protínají (v dolní části křivky); potom místo průběhů typu II (viz 3.4) se objevují průběhy typu III (viz obr. 5).

Uvedenými údaji je popsán modelující systém, přičemž však H je určeno jen „až na difeomorfismy", a stochastický proces, jehož realizacemi jsou průběhy $z(t)$, ani funkcionály F_t nejsou dosti precizovány; funkce f_a, φ_a, atd. lze pojímat jako jakési „parametry" systému. V modelujícím systému je tedy ještě velká volnost; až mu dáme (v 3.6) explicitní



Obr. 4. V levé polovině je znázorněn průběh pomocí veličin z, s; v pravé polovině příslušný časový průběh veličiny s.



Obr. 5. Posun izoklin (obrázek je jen kvalitativní).

interpretaci, budeme sice mluvit o modelu, spíše však půjde o modelové schéma. Přesto však lze již takto, popř. po částečné specifikaci, odvodit některé ověřitelné závěry kvalitativního rázu, např. o narůstání pravděpodobné frekvence atak; to, že závěry jsou spíše kvalitativní, není v zásadě na závadu a také odpovídá zde povaze empirických dat.

Je třeba ještě dodat, že modelující systém lze chápat jako přechodový metabolický systém (2.2) obohacený o stochastickou heteronomní dynamiku (2.6); pravidlo, jež přiřazuje průběhu $z(t)$ průběh $s(t)$, je zde velmi jednoduché, avšak pravidlo, kterým je průběhu $z(t)$ přiřazen prostřednictvím $s(t)$ průběh $p(t)$, je poměrně složité.

3.6. Řekneme nyní poněkud podrobněji, jak jsou reálným jevům a zákonitostem přiřazeny matematické veličiny a vztahy uvedené v 3.4, 3.5; tím se teprve ze systémů, které jsme tam probírali, stává model. Interpretace, kterou uvedeme, záleží v určitém shrnujícím pohledu na zkoumanou nemoc; může být případně modifikována.

Jak již bylo naznačeno, předpokládáme dva základní děje: jednak zánět a jeho potlačení, jednak změny funkčního stavu axonů, tj. ztrátu a regeneraci funkce. Zánětlivý děj se v modelu vyjadřuje jako časový průběh číselné veličiny z (zánět). Předpokládáme o ní, že se může pohybovat mezi hodnotami a, b (viz 3.4), přičemž a odpovídá vyhasnutí zánětlivého procesu (až na případná residua), b jeho maximální intenzitě.

Zánětlivý děj vyvolává přechod axonů z normálního (funkčního) stavu do stavu afunkčního; obráceným směrem působí spontánní regenerační procesy, popř. podporované léčbou, které při ustupování zánětu způsobí návrat poškozených axonů do stavu, v němž jsou schopny funkce. Tento děj (ztráta a regenerace funkce) je v modelu souhrnně vyjádřen časovým průběhem číselné proměnné s (stav axonů, popř. celkový stav nemocného).

První děj (zánět a jeho potlačení) má řídící úlohu. Jeho průběh je určován řadou okolností; četné z nich se vzhledem k uvažovaným základním dějům jeví jako náhodné. Proto se první děj vyjadřuje v modelu pomocí stochastického procesu, o němž činíme jen velmi všeobecné předpoklady (sr. 3.4).

Poloha křivky $H(p, z, s) = 0$, kde p je parametr, spolu s polohou přímky $z = a$ určuje možné průběhy $s(t)$, jež mají odpovídat skutečným průběhům v daném stadiu nemoci. Křivku $H = 0$, kterou budeme nazývat první izoklinou (je izoklinou pro rovnici $ds/dt = H$), lze chápat jako mez, na níž se zastavuje pohyb veličiny s ať již seshora dolů či obráceně; přesněji řečeno, tento význam má „atraktorová" část křivky (viz 2.1). Přímka $z = a$, jíž někdy říkáme „druhá izoklina" (ač s izoklinami z teorie diferenciálních rovnic přímo nesouvisí), vyjadřuje minimální hodnotu zánětlivého procesu. Způsob, jakým obě izokliny a konkrétní průběh veličiny z určují časový průběh $s(t)$, byl již popsán v 3.4, 3.5.

Úplnější systém z 3.5 zachycuje navíc ještě tyto okolnosti: jednak přetrvávající a povlovně narůstající rezidua zánětlivého procesu, jednak postupnou latentní (neprojevující se bezprostředně ve funkčních změnách axonů) deterioraci příslušné nervové tkáně. Axon, který byl v afunkčním stavu a regeneroval do stavu, v němž je schopen funkce, má totiž, jak předpokládáme, větší náchylnost – za jinak stejných podmínek – ke ztrátě funkce, a po této ztrátě má menší schopnost k nové regeneraci: zároveň může docházet k latentní deterioraci axonů též celkovým vlivem probíhajících patologických procesů. To se v modelu vystihuje jednak posunem druhé izokliny, tedy změnami veličiny a, jednak posunem první izokliny, přesněji řečeno, přechodem od menších k větším hodnotám parametru p. Poznamenejme ještě, že systém z 3.5 může vyjádřit též narůstání mezních hodnot b zánětlivého procesu; této možnosti zatím nevyužijeme, a proto také neinterpretujeme případný časový průběh $b(t)$.

Jak se snadno ověří, vzrostou-li veličiny a, p tak, že izokliny se již protínají v dolní části I. izokliny (tj. tak, že na přímce $z = a$ jsou dva atraktory pole H), nemohou se již vyskytnout ataky s remisemi; vyskytuje se v podstatě jen trvalé zhoršování s případným přechodem do stacionárního průběhu. To odpovídá, odmyslíme-li si případný vliv léčby, reálným průběhům.

Některé předcházející poznámky již naznačují, že po kvalitativní stránce vystihuje model empirická data poměrně uspokojivě .Otázkami přesnějšího a podrobnějšího ověřování modelu se zde ovšem nemůžeme zabývat.

3.7. Popíšeme nyní způsob, kterým lze dojít k modelu, a tím zároveň poněkud doplníme jeho zdůvodnění nastíněné v 3.6. Popisovaný způsob odpovídá ostatně také – ovšem jen v hrubých rysech – skutečnému průběhu úvah, jež vedly (po některých předběžných pokusech) k nynějšímu modelu.

Výrazný rys průběhu onemocnění u značné části pacientů, totiž ataky s remisemi, připomíná tzv. relaxační oscilace, a to vede k domněnce o úloze křivek s dvojím „záhybem". Ta okolnost, že rozsah a frekvence atak s remisemi se mění a nakonec často dochází k nevratným zhoršením, vede k doměnce, že se poloha, popř. tvar zmíněných křivek pozměňuje; z toho pak vzniká domněnka o úloze plochy $H(p, s, z) = 0$, kde se H dostane způsobem popsaným v 3.4. Tuto domněnku lze přitom opřít o základní úvahy z Thomovy teorie, z nichž vyplývá, že situace uvedeného druhu (tři proměnné, ataky s remisemi, přechod k nevratným změnám) lze za některých okolností modelovat s použitím plochy zmíněného typu; z Thomovy teorie ovšem nikterak nevyplývá, že by to byla v jejím rámci jediná možnost.

Nyní se zkusí pojetí, při němž má podstatnou úlohu autonomní dynamika (viz 2.4) a zpětný vliv průběhu $s(t)$ na průběh $z(t)$. Ukazuje se však, že by to vedlo k interpretaci, která není v souladu s uznávanými poznatky o roztroušené skleróze. Proto se přejde k heteronomní dynamice (viz 2.5), a to – z důvodů uvedených v 3.6 – stochastického rázu.

Jakmile ještě vezmeme v úvahu latentní deterioraci axonů, vyjádřenou časovým průběhem $p(t)$, dospíváme v hlavních rysech k našemu modelu (modelovém schématu).

Jak je patrné, měly myšlenky Thomovy metody při vytváření modelu spíše heuristickou úlohu. To, že děje, u nichž se vyskytují ataky s remisemi a přechod k nevratným změnám, se leckdy dají vyjádřit s použitím výrazů $H(p, z, s)$ uvedeného typu, je ostatně poznatek, který byl znám i z jiných souvislostí. Přesto měla podle našeho názoru Thomova metoda pro formování modelu podstatný význam, a to jak zapojením do širších souvislostí, tak zejména tím, že výrazně ukazovala nové možnosti vystižení dat, jež nemají plně kvantitativní povahu; v našem případě jde totiž o stupnici (viz 3.1) povahy spíše kvalitativní.

Je také třeba říci, že modely, k nimž se dospěje pomocí Thomovy metody, mohou někdy sloužit také k tomu, aby se vytvořil další model jiného druhu a aby se popřípadě pracovalo s oběma modely. Tak je tomu i v našem případě; o tom však bude řeč až v druhé části článku.

Literatura

[1] VLČEK, J., ZIELENIEC, J., *O teorii katastrof*, PMFA *22* (1977), 246—262.

[2] KOWALSKI, O., *Thomova věta o sedmi elementárních katastrofách*, PMFA *22* (1977), 302—316.

[3] BRÖCKER, T., *Differentiable germs and catastrophes*, Cambridge Univ. Press 1975.

[4] GOLUBITSKY, M., GUILLEMIN V., *Stable mappings and their singularities*, Springer-Verlag 1973.

[5] LU, Y.—C., *Singularity theory and an introduction to catastrophe theory*, Springer-Verlag 1976.

[6] WASSERMAN, G., *Stability of unfoldings*, Lect. Notes in Math. 393, Springer-Verlag 1974.

[7] THOM R., *Stabilité structurelle et morphogénèse*, W. A. Benjamin, Reading, Mass. 1972.

[8] TICHONOV, A. N., *O zavisimosti rešenij differencial'nych uravněnij ot malogo parametra*, Mat. Sb. *22*, 2 (1948), 193—204.

[9] TICHONOV, A. N., *Sistemy differencial'nych uravněnij soděržaščije malyje parametry pri proizvodnych*, Mat. Sb. *31*, 3(1952), 575—586.

[10] MIŠČENKO, E. F., ROZOV, N. CH., *Differencial'nyje uravněnija s malym parametrom i relaksacionnyje kolebanija*, Moskva 1975.

[11] VASILJEVA, A. B., BUTUZOV, V. F., *Asimptotičeskije razloženija rešenij singuljarno vozmuščennych uravněnij*, Moskva 1973.

[12] KALMAN, E., FALB P. L., ARBIB M. A., *Topics in mathematical system theory*, McGraw-Hill, New York 1969. (V r. 1971 vyšel ruský překlad.)

[13] FOG, T., LINNEMAN, F., *The course of multiple sclerosis*, Acta neurologica scandinavica, vol. *46* (1970), suppl. *47*, 1—175.