

Analizando a pandemia de COVID-19 através dos modelos SIR e SECIAR

Mauro Patrão¹, Manoel Reis²

Dept^o Matemática – UnB, 70.910-000, Brasília/DF.

Resumo. O presente artigo se propõe a responder as seguintes questões. Mesmo que não se desenvolva uma vacina eficaz nos próximos anos, a estratégia de isolamento social e reabertura repetidas vezes reduz o número de óbitos? E por que a SARS (2002) e a MERS (2012) não causaram tantos problemas quanto a pandemia de COVID-19? Utilizamos ideias de teoria de controle e o clássico modelo SIR para responder a primeira pergunta, enquanto que, para responder a segunda pergunta, é necessário introduzir uma extensão desse modelo, que denominamos de SECIAR, e descrever sua dinâmica global.

Palavras-chave: SARS; MERS; Imunidade de rebanho; Funções de Lyapunov; Hiperbolicidade normal; Estabilidade; Teorema de Routh-Hurwitz.

1 Introdução

Um século após a denominada *gripe espanhola* de 1918, o mundo foi novamente sacudido por uma grande pandemia que afetou a vida de quase todos os seres-humanos. Diferentemente de outras pandemias recentes, a pandemia da COVID-19 obrigou um intenso distanciamento social para o seu controle, paralisando boa parte da economia e gerando grandes mudanças no estilo de vida das pessoas. A principal razão mencionada para intensificar o distanciamento social foi produzir o *achatamento da curva* de infecções, evitando o colapso do sistema de saúde e salvando vidas. Uma segunda razão oferecida é ganhar tempo para a obtenção de uma vacina segura e eficaz, também evitando

¹Professor Associado do MAT-UnB, mauropatrao@unb.br

²Bolsista do PETMAT-UnB, 170040691@aluno.unb.br

mortes. Dadas todas as consequências sociais e econômicas do distanciamento social, surgiram diversas afirmações sobre a suposta inutilidade de tal medida, divulgadas amplamente na mídia. Claramente essas afirmações se mostrariam incorretas caso alguma vacina segura e eficaz fosse disponibilizada em tempo hábil para permitir a imunização da maioria da população. Entretanto, caso não se desenvolva uma vacina nos próximos anos, parece intuitivo que as afirmações estariam corretas, de modo que o isolamento social seria importante apenas para evitar o colapso do sistema de saúde até atingirmos a denominada *imunidade de rebanho*, tão mencionada nos últimos meses na mídia e nas revistas científicas, como por exemplo em Britton et al. (2020). Um dos objetivos do presente artigo é mostrar que esse ponto de vista intuitivo está equivocado e que a estratégia de isolamento social e reabertura repetidas vezes de fato reduz o número de óbitos no longo prazo, mesmo sem o desenvolvimento de uma vacina. Outra questão que o presente artigo se propõe a responder e que aparece na mídia desde o início da pandemia de COVID-19 é a razão da necessidade do distanciamento social intenso para o controle da pandemia atual, uma vez que essa medida não foi necessária no controle das outras pandemias recentes, mesmo as causadas por coronavírus semelhantes, como a SARS (2002) e a MERS (2012).

O artigo é dividido em quatro seções além da presente introdução. Na próxima seção, relembramos as principais características do clássico modelo SIR e, na seção seguinte, usamos o modelo SIR e algumas ideias de teoria de controle para responder a primeira pergunta. Na penúltima seção, para responder a segunda pergunta, introduzimos uma extensão do modelo SIR, que denominamos de SECIAR, e descrevemos genericamente sua dinâmica global. Finalmente, na última seção, apresentamos nossas conclusões. O artigo também possui um apêndice com os principais resultados de sistemas dinâmicos utilizados e as referências para as suas respectivas demonstrações.

2 O modelo SIR

Nessa seção, relembramos as principais características do modelo SIR e também os conceitos de imunidade de rebanho e de número de reprodução inicial. Detalhes e demonstrações podem ser encontrados na Seção 11.1 de Hirsch et al. (2013) e no Capítulo 2 de Martcheva (2015). O SIR é um exemplo de modelo compartimental, no qual a população é dividida em compartimentos

em que os indivíduos possuem características epidemiológicas distintas, e que pode ser visualizado através do fluxograma abaixo:



Figura 1: Fluxograma modelo SIR

O compartimento S representa a classe dos indivíduos suscetíveis à infecção, já o compartimento I representa os infectados, e supõe-se que já se tornam automaticamente infectantes, enquanto o compartimento R representa os removidos, os que se recuperaram ou que infelizmente faleceram. Devido ao horizonte temporal da evolução da doença ser de meses ou até semanas, pode-se desconsiderar o crescimento demográfico, de modo que se supõe que a população é constante. As letras S, I, R podem então também representar os percentuais dos indivíduos em cada compartimento, de sorte que $S + I + R = 1$.

O produto βSI representa a taxa do percentual de novos infectados por unidade de tempo, também chamada de incidência, em que a constante β é proporcional ao número de encontros aleatórios em pares por unidade de tempo e proporcional à probabilidade de contágio a cada encontro, enquanto o produto SI é igual à probabilidade de se encontrar aleatoriamente um par com um suscetível e um infectado. Aqui se supõe implicitamente que uma vez removidos, os indivíduos não participam mais do processo de contágio.

Já γI é a taxa do percentual de indivíduos infectados que são removidos por unidade de tempo. Essa taxa de remoção é equivalente a supor que a densidade da distribuição de probabilidade do tempo em que uma pessoa permanece infectada é dada pela exponencial $\gamma e^{-\gamma t}$. Pode-se mostrar que o tempo médio que as pessoas permanecem infectadas é dado por $\frac{1}{\gamma}$, que é ilustrado na figura abaixo:

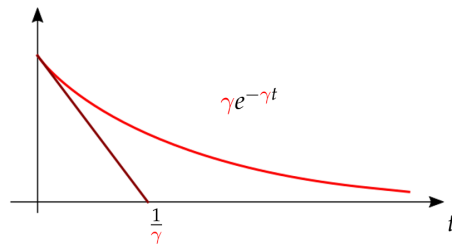


Figura 2: Gráfico da função de densidade da distribuição de probabilidade

Como os removidos não participam do processo de contágio e como $R = 1 - S - I$, pode-se descrever o modelo SIR através das seguintes equações envolvendo as variáveis S e I :

$$\begin{aligned} S' &= -\beta S I \\ I' &= \beta S I - \gamma I \end{aligned}$$

onde S' e I' denotam as respectivas derivadas temporais e também pode-se mostrar que o percentual de infectados I é dado pela seguinte função do percentual de suscetíveis S :

$$I = -S + \frac{\gamma}{\beta} \log S + \text{constante}$$

De fato, para cada valor da constante acima, temos uma função diferente, cujos gráficos são as curvas vermelhas ilustradas na figura abaixo:

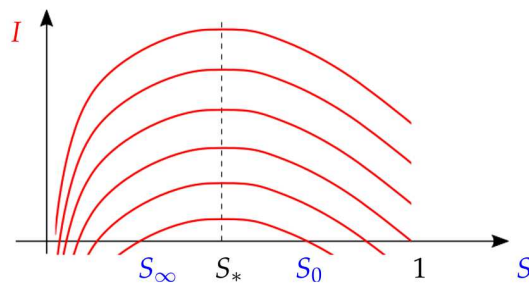


Figura 3: Curvas vermelhas

Variando o valor da constante, as curvas vermelhas são transladadas verticalmente, de forma que a constante é determinada pela proporção inicial S_0 de suscetíveis, que por sua vez determina a proporção final S_∞ de suscetíveis

no longo prazo do horizonte temporal do modelo, como também é ilustrado na figura. Partindo de um determinado S_0 , as variáveis S e I evoluem seguindo a respectiva curva vermelha da direita para a esquerda. Quando o vírus é novo entre os humanos, como o da COVID-19, na denominada "primeira onda", o valor de S_0 é igual a 1.

Como todas as curvas vermelhas são obtidas por translações verticais de uma delas, todas atingem um máximo de percentual de infectados num mesmo percentual de suscetíveis, denotado por S_* na figura acima. Esse percentual de suscetíveis é um divisor de águas importante no modelo, pois S_∞ sempre está à esquerda de S_* . Se S_0 está à direita de S_* a epidemia evolui na curva vermelha da direita para esquerda até se aproximar do S_∞ . Por outro lado, se S_0 está à esquerda de S_* , então S_0 é igual a S_∞ e não haverá mais contágio comunitário, o que é denominado de *imunidade de rebanho*. Pode-se mostrar que:

$$S_* = \frac{\gamma}{\beta}$$

Observe que a imunidade de rebanho não é o mesmo que imunidade biológica, uma vez que um percentual S_0 positivo de pessoas suscetíveis pode permanecer estável no longo prazo. A imunidade de rebanho é portanto um conceito de dinâmica epidemiológica, que surge devido ao balanço entre as taxas de contágio em rebanho, afetadas pelo valor de β , e o tempo médio que as pessoas permanecem infectantes, afetado pelo valor de γ .

A imunidade de rebanho está estreitamente relacionada a um outro conceito epidemiológico que também se tornou famoso por conta da pandemia, o denominado *numero de reprodução inicial*, denotado usualmente por R_0 . O número de reprodução é definido como o número médio de suscetíveis que um infectado contamina, dado pelo produto do número médio de suscetíveis que um infectado contamina por unidade de tempo, expressado por βS_0 , pelo tempo médio que as pessoas permanecem infectantes, dado por $\frac{1}{\gamma}$, de modo que:

$$R_0 = \frac{\beta S_0}{\gamma}$$

Observe que $R_0 = S_0/S_*$. Por um lado, se $R_0 > 1$, então $S_0 > S_*$, e nesse caso não teríamos imunidade de rebanho. Por outro lado, se $R_0 < 1$, então $S_0 < S_*$, e agora teríamos imunidade de rebanho

Agora estamos prontos para responder a primeira pergunta: mesmo que não se desenvolva uma vacina eficaz nos próximos anos, a estratégia de isolamento social e reabertura repetidas vezes reduz o número de óbitos?

3 Controle no modelo SIR

Diferentemente do resto da natureza, os seres humanos são capazes de responder racionalmente aos desafios trazidos por uma epidemia. Quando um surto é identificado, podemos quase que imediatamente tomar certas medidas para alterar os valores das constantes do modelo, como medidas de isolamento social e de higienização, que diminuem o valor de β , pois reduzem tanto o número de encontros aleatórios em pares por unidade de tempo, quanto a probabilidade de contágio a cada encontro. Existem diversas medidas que promovem o distanciamento social, desde campanhas de esclarecimento governamentais ao decretamento do denominado *lockdown*, produzindo diferentes reduções no valor de β . Esse comportamento racional dos seres humanos transformam a natureza da análise do modelo SIR, que deixa de ser visto como um sistema dinâmico puro e passa a ser encarado como um sistema de controle, como explicaremos a seguir.

Para simplificar a exposição, vamos imaginar que podemos alternar apenas entre duas situações, uma sem e outra com isolamento social. Podemos então alternar a constante β entre seu valor original e um novo valor reduzido, que vamos denotar por $\tilde{\beta}$. Isso nos permite, por sua vez, alternar o limiar S_* entre seu valor original e um novo valor aumentado, uma vez que S_* é inversamente proporcional a β . Vamos denotar esse novo limiar por \tilde{S}_* , lembrando que ele é o novo percentual de suscetíveis onde o percentual de infectados atinge (ou atingiria se for maior do que um) o máximo nas novas curvas que determinam a relação entre essas variáveis S e I durante evolução da epidemia, e que são ilustradas em azul na figura abaixo:

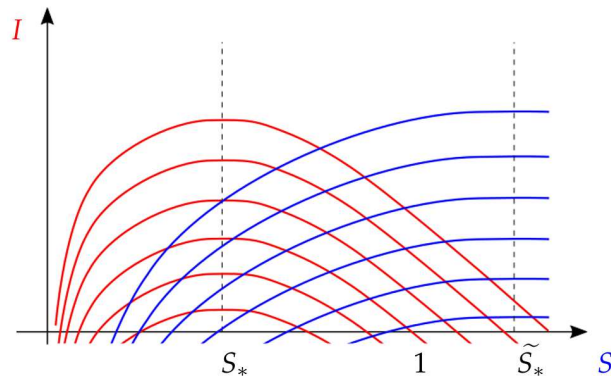


Figura 4: Sobreposição das curvas azuis e vermelhas

A evolução da epidemia, com o processo de controle através do isolamento social, é ilustrada pela curva verde nas quatro figuras abaixo:

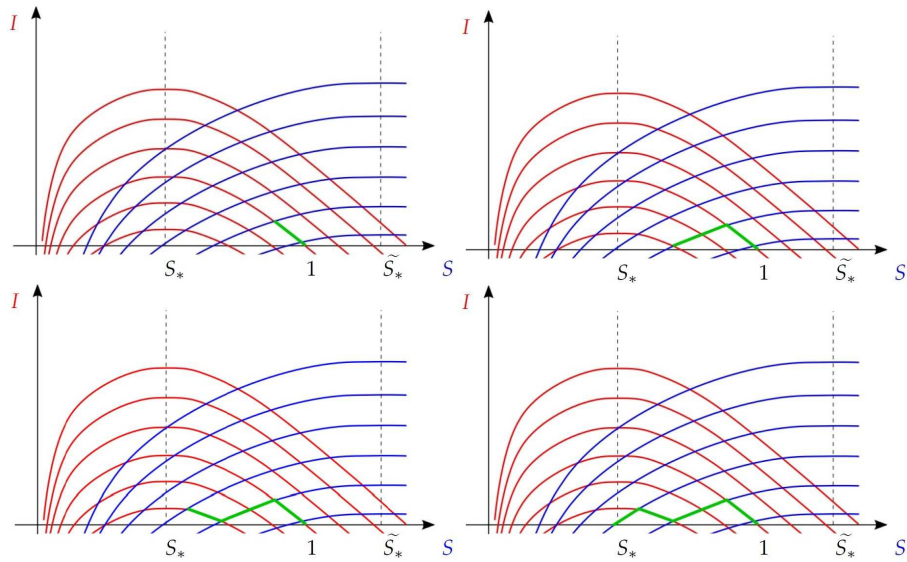


Figura 5: Teoria de controle modelo SIR

No início, a epidemia evolui por uma curva vermelha, sem isolamento social, pois as autoridades temem seus custos econômicos. Após a gravidade da doença ser reconhecida, e com o risco do colapso do sistema de saúde, o isolamento social passa a acontecer e a epidemia passa a evoluir por uma curva

azul. Após os casos serem consideravelmente reduzidos e devido aos custos econômicos e sociais do isolamento social, o isolamento social é retirado, e a epidemia volta a evoluir por uma curva vermelha. Nesse ponto, o colapso do sistema de saúde já pode ter sido evitado e a pergunta que surge é: mesmo que nunca se desenvolva uma vacina minimamente eficaz, é recomendável aplicar novamente o isolamento social ou não faria diferença em termos de vidas salvas? E a resposta é: vale muito a pena! Observe que se adotarmos novamente o isolamento social, a epidemia passa novamente a evoluir por uma curva azul e, após passarmos o limiar original S_* , obtemos a imunidade de rebanho, mesmo que o isolamento social agora seja retirado. Por outro lado, se não adotássemos novamente o isolamento social, seguiríamos por uma curva vermelha e também obteríamos a imunidade de rebanho, mas com muito mais pessoas infectadas e portanto muito mais sequelas e muito mais vidas perdidas desnecessariamente, como ilustra a comparação das duas figuras abaixo, sem e com um segundo lockdown:

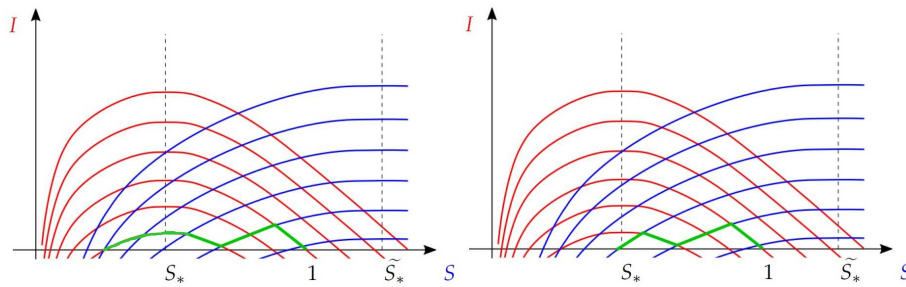


Figura 6: Efeito de um segundo isolamento social

Na próxima seção, vamos tentar responder a segunda e última pergunta do resumo: por que a SARS (2002) e a MERS (2012) não causaram tantos problemas quanto a atual pandemia?

4 O modelo SECIAR

4.1 Pressupostos do modelo SECIAR

Para respondermos à última pergunta da introdução é necessária uma extensão do modelo SIR, que nos permita mostrar o porquê está sendo tão mais difícil e custoso controlar a epidemia atual, em relação as outras epide-

mias, mesmo as provocadas por outros coronavírus. O modelo SECIAR é uma extensão do SIR, onde se acrescentam três novos compartimentos, e que pode ser visualizado através do fluxograma abaixo:

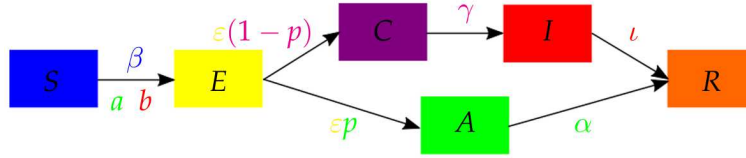


Figura 7: Fluxograma SECIAR

Os indivíduos do compartimento S dos suscetíveis, ao se infectarem, entram no compartimento E dos expostos, onde os indivíduos estão infectados, mas ainda não são infectantes. Após um tempo médio $1/\varepsilon$, uma proporção p dos indivíduos do compartimento E dos expostos entra no compartimento A dos assintomáticos, enquanto uma proporção complementar $1-p$ entra no compartimento C dos pré-sintomáticos (do inglês *carrier*). Nesses dois compartimentos, os indivíduos passam a ser infectantes, mas não apresentam sintomas. Após um tempo médio $1/\gamma$, os indivíduos do compartimento C dos pré-sintomáticos entram no compartimento I dos sintomáticos, os indivíduos permanecem infectantes e agora apresentam sintomas. Após um tempo médio $1/\iota$, no caso dos indivíduos do compartimento I dos sintomáticos, e após um tempo médio $1/\alpha$, no caso dos indivíduos do compartimento A dos assintomáticos, os indivíduos desses dois compartimentos entram no compartimento R dos removidos.

A taxa do percentual de novos infectados por unidade de tempo é a soma da parcela βSC , em razão ao contágio pelos indivíduos do compartimento C dos pré-sintomáticos, da parcela $b\beta SI$, devido ao contágio pelos indivíduos do compartimento I dos sintomáticos e, finalmente, da parcela $a\beta SA$, devido ao contágio pelos indivíduos do compartimento A dos assintomáticos. O valor da constante a pode ser bem menor do que 1, uma vez que os indivíduos assintomáticos podem ter uma carga viral muito reduzida, enquanto o valor da constante b é 1 antes da epidemia ser identificada e passa a ser bem menor do que 1 após o surto ser identificado, quando a maioria dos indivíduos sintomáticos passam a ser isolados ou a se auto-isolar.

O modelo SECIAR é então descrito pelas cinco equações abaixo, envol-

vendo as variáveis S , E , C , I e A :

$$\begin{aligned} S' &= -\beta S(C + bI + aA) \\ E' &= \beta S(C + bI + aA) - \varepsilon E \\ C' &= (1-p)\varepsilon E - \gamma C \\ I' &= \gamma C - \iota I \\ A' &= p\varepsilon E - \alpha A \end{aligned}$$

4.2 Imunidade de rebanho e número de reprodução no SECIAR

O retrato de fases do modelo SECIAR É parecido com a do SIR? Vamos mostrar que, para efeitos práticos, o modelo SECIAR tem uma dinâmica muito parecida com a do SIR, como ilustrado pela figura abaixo:

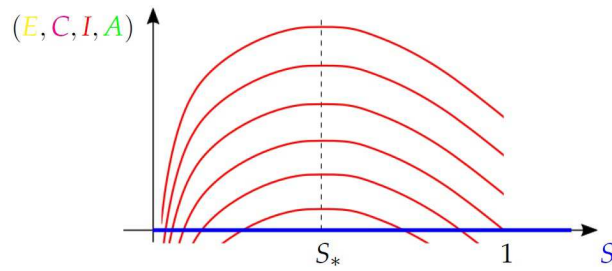


Figura 8: Dinâmica do modelo SECIAR

Observe que a variável S permanece representada no eixo horizontal, enquanto que as dimensões E , C , I e A estão todas representadas pelo mesmo eixo na vertical, de modo a facilitar a comparação entre as dinâmicas do SECIAR e do SIR. Ainda que agora não exista uma expressão explícita para a relação entre as variáveis, podemos mostrar que novamente existe um limiar S_* para a imunidade de rebanho tal que, se S_0 está à direita de S_* a epidemia evolui na curva vermelha da direita para esquerda até se aproximar do S_∞ à esquerda de S_* , atingindo a imunidade de rebanho. Mais do que isso, podemos mostrar que $R_0 = S_0/S_*$, de modo que obtemos a mesma relação entre o número de reprodução e o limiar da imunidade de rebanho, que é dado pela seguinte expressão:

$$S_* = \frac{1}{\beta \left((1-p) \left(\frac{1}{\gamma} + \frac{b}{\iota} \right) + \frac{pa}{\alpha} \right)}$$

Observe que o compartimento E dos expostos não influencia o limiar da imunidade de rebanho S_* , tendo papel apenas na dinâmica transiente do modelo.

Para mostrarmos porque a descrição acima é verdadeira, iniciamos a análise da dinâmica global do modelo SECIAR através das chamadas funções de Lyapunov, definidas no apêndice.

Proposição 1. *Temos que*

$$L = 3S + 3E + 2C + I + 2A$$

é uma função de Lyapunov para o modelo SECIAR. Além disso, temos que conjunto dos pontos de equilíbrio do modelo SECIAR é dado pelo eixo S e cada trajetória do modelo SECIAR se aproxima de um único ponto de equilíbrio S_∞ do eixo S no futuro distante.

Prova: Observe que o conjunto dos pontos de equilíbrio do modelo SECIAR é dado pelo eixo S , uma vez que S' , E' , C' , I' e A' se anulam simultaneamente se e só se E , C , I e A se anulam simultaneamente. Além disso, uma função de Lyapunov para o modelo SIR é dada por

$$L = 3S + 3E + 2C + I + 2A$$

uma vez que $\nabla L = (3, 3, 2, 1, 2)$, de modo que

$$\begin{aligned} L' &= -3\beta S(C + bI + aA) \\ &+ 3\beta S(C + bI + aA) - 3\varepsilon E \\ &+ 2(1-p)\varepsilon E - 2\gamma C \\ &+ \gamma C - \iota I \\ &+ 2p\varepsilon E - 2\alpha A \\ &= -\varepsilon E - \gamma C - \iota I - 2\alpha A \\ &\leq 0 \end{aligned}$$

como queríamos. Isso mostra que qualquer semi-espaço dado por $L = 3S + 3E + 2C + I + 2A \leq c$, onde c é uma constante, é positivamente invariante. Além disso, é fácil ver que o ortante positivo, onde S , E , C , I e A são positivos, também é positivamente invariante, de modo que a interseção de qualquer semi-espaço $L \leq c$ com o ortante positivo, além de ser claramente compacta, também é positivamente invariante.

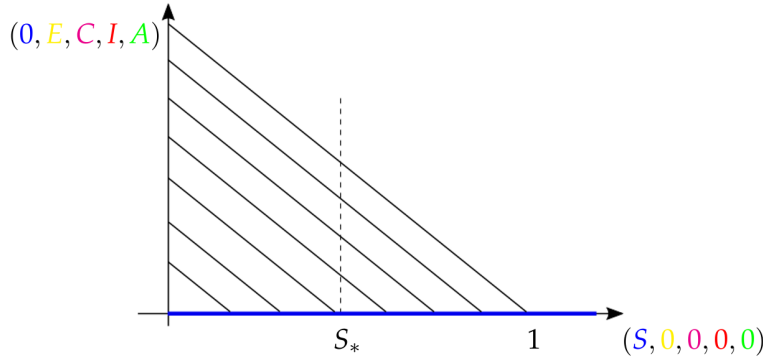


Figura 9: Compactos positivamente invariantes

Novamente pela Proposição 6, segue que os pontos limites estão contidos na imagem inversa de $L' = 0$. No caso do SECIAR, isso ocorre se e só se E , C , I e A se anulam simultaneamente, portanto todos os pontos limites do modelo SECIAR estão contidos no seguimento de reta S , de modo que todas as trajetórias do modelo SECIAR se aproximam do eixo S no futuro distante. Além disso, como $S' \leq 0$, segue que S é uma função positiva e não-crescente do tempo, de modo que cada trajetória do modelo SECIAR se aproxima de um único ponto de equilíbrio S_∞ do eixo S no futuro distante. \square

Para completar a descrição da dinâmica global do modelo SECIAR, precisamos primeiro analisar a dinâmica local desse modelo próximo ao eixo S através da análise dos autovalores da matriz Jacobiana, definida no apêndice, em cada um dos seus pontos.

Proposição 2. *Temos que*

$$J = \begin{bmatrix} 0 & 0 & -\beta S_0 & -\beta S_0 b & -\beta S_0 a \\ 0 & -\varepsilon & \beta S_0 & \beta S_0 b & \beta S_0 a \\ 0 & \varepsilon(1-p) & -\gamma & 0 & 0 \\ 0 & 0 & \gamma & -\iota & 0 \\ 0 & \varepsilon p & 0 & 0 & -\alpha \end{bmatrix}$$

é a matriz Jacobiana do modelo SECIAR no ponto S_0 do eixo S . Um ponto de equilíbrio é estável se está à esquerda de S_* e é instável se está à direita de S_* . Além disso, com exceção de S_* e de no máximo dois outros pontos à direita dele, existe uma vizinhança do ponto de equilíbrio onde a dinâmica do modelo SECIAR é conjugada à dinâmica linear dada por J .

Prova: Primeiro a matriz Jacobiana para o SECIAR é dada por:

$$J = \begin{bmatrix} -\beta(aA + bI + C) & 0 & -\beta S_0 & -\beta S_0 b & -\beta S_0 a \\ \beta(aA + bI + C) & -\varepsilon & \beta S & \beta S b & \beta S a \\ 0 & \varepsilon(1-p) & -\gamma & 0 & 0 \\ 0 & 0 & \gamma & -\iota & 0 \\ 0 & \varepsilon p & 0 & 0 & -\alpha \end{bmatrix}$$

Assim como no SIR, estamos interessados nos pontos S_0 que pertencem ao eixo S , de modo que $E = C = I = A = 0$, obtendo:

$$J = \begin{bmatrix} 0 & 0 & -\beta S_0 & -\beta S_0 b & -\beta S_0 a \\ 0 & -\varepsilon & \beta S_0 & \beta S_0 b & \beta S_0 a \\ 0 & \varepsilon(1-p) & -\gamma & 0 & 0 \\ 0 & 0 & \gamma & -\iota & 0 \\ 0 & \varepsilon p & 0 & 0 & -\alpha \end{bmatrix}$$

Podemos escrever:

$$J = \begin{bmatrix} 0 & k \\ 0 & K \end{bmatrix}$$

onde k é uma matriz $1 \times (n-1)$, enquanto K será agora uma matriz 4×4 , dada por:

$$K = \begin{bmatrix} -\varepsilon & \beta S_0 & \beta S_0 b & \beta S_0 a \\ \varepsilon(1-p) & -\gamma & 0 & 0 \\ 0 & \gamma & -\iota & 0 \\ \varepsilon p & 0 & 0 & -\alpha \end{bmatrix}$$

Para acharmos os autovalores de K , é necessário o cálculo do seu polinômio característico, que é dado por:

$$p_K(\lambda) = \lambda^4 + a_1 \lambda^3 + a_2 \lambda^2 + a_3 \lambda + a_4$$

onde

$$\begin{aligned} a_1 &= \varepsilon + \gamma + \iota + \alpha \\ a_2 &= \varepsilon \alpha + \gamma \alpha + \iota \alpha + \varepsilon \gamma + \varepsilon \iota + \gamma \iota - \beta S_0 \varepsilon [(1-p) + pa] \\ a_3 &= \varepsilon \gamma \alpha + \varepsilon \iota \alpha + \gamma \iota \alpha + \varepsilon \gamma \iota - \beta S_0 \varepsilon [(1-p)(\gamma b + \iota + \alpha) + pa(\gamma + \iota)] \\ a_4 &= \varepsilon \gamma \iota \alpha - \beta S_0 \varepsilon [(1-p)\alpha(\gamma b + \iota) + \gamma \iota pa] \end{aligned}$$

Algumas propriedades desse polinômio é que $a_1 > 0$ e a_2, a_3, a_4 são funções afins de S_0 . Achar explicitamente os autovalores em termos de S_0 agora, é mais

complicado, já que estes são soluções de uma equação quártica. Mas podemos obter indiretamente informações sobre a parte real desses autovalores, já que queremos utilizar novamente o Teorema de Pugh-Shub. Para determinar os pontos em que K deixa de ser hiperbólica, primeiro observamos que claramente, quando $a_4 = 0$, existe um autovalor nulo. Nesse ponto, isolando S_0 temos:

$$S_0 = S_* = \frac{1}{\beta \left((1-p) \left(\frac{1}{\gamma} + \frac{b}{\iota} \right) + \frac{pa}{\alpha} \right)}$$

Esse será justamente o limiar para a imunidade de rebanho para o modelo SECIAR. Claramente, pela Proposição 7, segue que os pontos de equilíbrio à direita de S_* são instáveis, uma vez que a matriz K , e portanto a matriz J , possui um autovalor real positivo quando $S_0 > S_*$, pois, nesse caso, temos que $p_K(0) = a_4 < 0$ e o coeficiente da maior potência de $p_K(\lambda)$ é igual a um. Mostrar que os pontos de equilíbrio à esquerda de S_* são estáveis é bem mais extenso e será deixado para o final. Antes vamos determinar quais são os outros pontos onde a matriz K deixa de ser hiperbólica. É possível mostrar que pode haver apenas mais dois valores de S_0 diferentes de S_* onde a matriz K perde a hiperbolicidade. Primeiro temos que nosso polinômio possui ao menos duas raízes com parte real não nula, uma vez que, do contrário, ele poderia ser escrito de modo:

$$p_K(\lambda) = (\lambda^2 + b_0)(\lambda^2 + c_0)$$

com $b_0, c_0 > 0$, então nosso polinômio teria a forma abaixo:

$$p_K(\lambda) = \lambda^4 + (b_0 + c_0)\lambda^2 + b_0c_0$$

o que implicaria que $a_1 = 0$, o que não é verdade, pois $a_1 > 0$. Então supondo que nosso polinômio possua duas raízes puramente imaginárias ele teria a forma abaixo:

$$\begin{aligned} p_K(\lambda) &= (\lambda^2 + b_0)(\lambda^2 + c_1\lambda + c_0) \\ &= \lambda^4 + c_1\lambda^3 + (b_0 + c_0)\lambda^2 + b_0c_1\lambda + b_0c_0 \end{aligned}$$

Disso segue que:

$$a_1 = c_1, \quad a_2 = b_0 + c_0, \quad a_3 = b_0c_1, \quad a_4 = b_0c_0$$

de modo que

$$b_0 = \frac{a_3}{a_1}, \quad c_0 = \frac{a_4a_1}{a_3}$$

o que implica que

$$a_2 = b_0 + c_0 = \frac{a_3}{a_1} + \frac{a_4 a_1}{a_3}$$

Com isso temos que os coeficientes formam uma equação de segundo grau em S_0 :

$$a_3 a_2 a_1 = a_3^2 + a_4 a_1^2$$

a qual pode ter no máximo duas raízes reais S_1 e S_2 . Vamos mostrar que esses dois pontos, se existirem, estão à direita de S_* , como ilustrado pela figura abaixo:

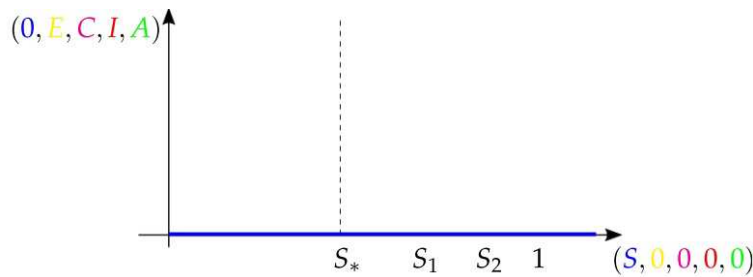


Figura 10: Posição de S_1 e S_2

Para demonstrar isso, vamos mostrar que todos os autovalores de K tem parte real negativa quando $S_0 < S_*$ e, para isso, vamos utilizar o Teorema de Routh-Hurwitz, que apresenta condições necessárias e suficientes que os coeficientes do polinômio característico devem satisfazer para que todos os autovalores tenham parte real negativa. Porém esse critério é de difícil aplicação no nosso caso, uma vez que ele se torna mais complicado conforme o grau do polinômio aumenta e conforme as expressões dos coeficientes se tornam mais complicadas. Entretanto, existe uma ligação entre o valor máximo da parte real de todos os autovalores da diferença $L - M$ e o valor máximo dos valores absolutos dos autovalores do quociente LM^{-1} quando L é uma matriz com entradas não negativas e M é uma matriz diagonal com entradas diagonais positivas e isso será usado para simplificar o problema. Vamos escrever a matriz K como a seguinte

diferença:

$$\begin{aligned}
 K &= \begin{bmatrix} -\varepsilon & \beta S_0 & \beta S_0 b & \beta S_0 a \\ \varepsilon(1-p) & -\gamma & 0 & 0 \\ 0 & \gamma & -\iota & 0 \\ \varepsilon p & 0 & 0 & -\alpha \end{bmatrix} \\
 &= \begin{bmatrix} 0 & \beta S_0 & \beta S_0 b & \beta S_0 a \\ \varepsilon(1-p) & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ \varepsilon p & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ 0 & 0 & \iota & 0 \\ 0 & 0 & 0 & \alpha \end{bmatrix}
 \end{aligned}$$

denotando

$$L = \begin{bmatrix} 0 & \beta S_0 & \beta S_0 b & \beta S_0 a \\ \varepsilon(1-p) & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ \varepsilon p & 0 & 0 & 0 \end{bmatrix}, \quad M = \begin{bmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ 0 & 0 & \iota & 0 \\ 0 & 0 & 0 & \alpha \end{bmatrix}$$

de modo que

$$LM^{-1} = \begin{bmatrix} 0 & \frac{\beta S_0}{\gamma} & \frac{\beta S_0 b}{\iota} & \frac{\beta S_0 a}{\alpha} \\ (1-p) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ p & 0 & 0 & 0 \end{bmatrix}$$

Determinar quando todos os autovalores da diferença $K = L - M$ tem parte real negativa é o mesmo que determinar quando todos os autovalores do quociente LM^{-1} tem valor absoluto menor do que um, que se mostra um problema mais simples. De fato, primeiro determinamos o polinômio característico de LM^{-1} , que é dado por:

$$p_{LM^{-1}}(\lambda) = \lambda \left(\lambda^3 - \lambda \left(\frac{p\beta S_0 a}{\alpha} + \frac{(1-p)\beta S_0}{\gamma} \right) - \frac{(1-p)\beta S_0 b}{\iota} \right)$$

Uma das raízes desse polinômio é nula e, para mostrar que as demais raízes tem valor absoluto menor do que um, vamos utilizar o seguinte corolário do Teorema de Routh-Hurwitz. Dado o polinômio $p(\lambda) = \lambda^3 + b\lambda^2 + c\lambda + d$ e considere:

$$\begin{aligned}
 e_0 &= 1 + b + c + d \\
 e_1 &= 3 + b - c - 3d \\
 e_2 &= 3 - b - c - 3d \\
 e_3 &= 1 - b + c - d
 \end{aligned}$$

Todas as raízes de $p(\lambda)$ tem valor absoluto menor do que um se e só se $e_0, e_1, e_2, e_3 > 0$ e $e_2e_1 > e_3e_0$. No nosso caso, temos que:

$$\begin{aligned} b &= 0 \\ c &= -\left(\frac{p\beta S_0 a}{\alpha} + \frac{(1-p)\beta S_0}{\gamma}\right) \\ d &= -\frac{(1-p)\beta S_0 b}{\iota} \end{aligned}$$

de modo que

$$e_1 = e_2 = 3 - c - 3d > 3 > 0$$

e também que

$$\begin{aligned} e_2e_1 &= (3 - c - 3d)^2 \\ &= 9 + c^2 + 9d^2 - 3c - 9d - 3c + 3dc - 9d + 3dc \\ &> 9 + c^2 + 9d^2 \\ &> 1 + c^2 - d^2 + 2c \\ &= (1 + c - d)(1 + c + d) \\ &= e_3e_0 \end{aligned}$$

Além disso, temos que

$$e_3 = 1 + c - d > 1 + c + d = e_0$$

de modo que $e_3 > 0$ quando $e_0 > 0$. Segue então que $e_0 > 0$ se e só se todos os autovalores do quociente LM^{-1} tem valor absoluto menor do que um, o que é equivalente a todos os autovalores da diferença $K = L - M$ terem parte real negativa. Pelas definições, essa condição é dada por

$$e_0 = 1 - \left(\frac{p\beta S_0 a}{\alpha} + \frac{(1-p)\beta S_0}{\gamma}\right) - \frac{(1-p)\beta S_0 b}{\iota} > 0$$

Isolando S_0 na desigualdade acima, segue que a desigualdade é equivalente a $S_0 < S_*$, como afirmamos. Se definirmos $R_0 = \frac{S_0}{S_*}$, temos que

$$R_0 = \frac{(1-p)\beta S_0}{\gamma} + \frac{(1-p)b\beta S_0}{\iota} + \frac{pa\beta S_0}{\alpha}$$

□

Um subconjunto de um espaço topológico é denominado *genérico* se é a interseção enumerável de subconjuntos abertos e densos. A probabilidade de se sortear aleatoriamente um ponto de um subconjunto genérico é igual a um, desde que a medida de probabilidade seja absolutamente contínua em relação à medida de Lebesgue. O resultado a seguir completa a descrição da dinâmica global do modelo SECIAR.

Proposição 3. *Para todo $\delta > 0$, o subconjunto do ortante positivo tal que as trajetórias partindo desse conjunto possuem $S_\infty < S_* + \delta$ é aberto e denso. Em particular, o subconjunto do ortante positivo tal que as trajetórias partindo desse conjunto possuem $S_\infty \leq S_*$ é genérico.*

Prova: Vamos denotar a primeira componente do fluxo $\phi_t(x)$ do modelo SECIAR por $S_t(x)$, que é uma função positiva e decrescente em relação ao tempo. Segue que $S_t(x) \rightarrow S_\infty(x)$, quando $t \rightarrow \infty$. Pela Proposição 1, temos que

$$\phi_t(x) \rightarrow \phi_\infty(x) = (S_\infty(x), 0, 0, 0, 0)$$

quando $t \rightarrow \infty$. Para todo $\delta > 0$, denote por $N(\delta)$ o subconjunto do ortante positivo tal que todas as trajetórias partindo desse conjunto possuem $S_\infty < S_* + \delta$. Queremos mostrar que $N(\delta)$ é aberto e denso.

Para mostrar que $N(\delta)$ é aberto, considere $x_0 \in N(\delta)$, de modo que $S_\infty(x_0) < S_* + \delta$. Então existe $T > 0$ tal que $S_T(x_0) < S_* + \delta$. Como ϕ_T é um homeomorfismo, existe uma vizinhança aberta B de x_0 tal que $S_T(x) < S_* + \delta$ para todo $x \in B$. Como $S_t(x)$ é decrescente em relação ao tempo, segue que $S_\infty(x) < S_* + \delta$ para todo $x \in B$, de modo que $B \subset N(\delta)$, mostrando $N(\delta)$ é aberto.

Para mostrar que $N(\delta)$ é denso, considere $x_0 \notin N(\delta)$ e B uma bola aberta centrada em x_0 e de raio $r > 0$. Segue que $S_\infty(x_0) \geq S_* + \delta$, de modo que $S_0(x_0) > S_* + \delta$, uma vez que $S_t(x)$ é decrescente em relação ao tempo. Além disso, para todo $x \in B$, temos que $S_0(x) < S_0(x_0) + r$, de modo que $S_\infty(x) < S_0(x_0) + r$. Agora observe que, para todo ponto no eixo S tal que $S_0 \in [S_* + \delta, S_0(x_0) + r]$, existem uma bola aberta A centrada nesse ponto e um conjunto $F \subset A$ tal que $A - F$ é denso em A e tal que toda trajetória que permanece em A para tempos positivos está contida em F para tempos positivos. De fato, isso segue da Proposição 2, uma vez que, pela demonstração da Proposição 2, a matriz K , e portanto a matriz J , possui um autovalor real positivo, quando $S_0 > S_*$. Como $[S_* + \delta, S_0(x_0) + r]$ é compacto, existem um

família finita $\{A_1, \dots, A_k\}$ dessas vizinhanças tal que $[S_* + \delta, S_0(x_0) + r] \subset A_1 \cup \dots \cup A_k$. Podemos escolher essa família de modo que se $x \in A_i$ e $y \in A_{i+2}$, então $S_0(x) > S_0(y)$. Como $S_\infty(x_0) \in [S_* + \delta, S_0(x_0) + r]$, temos que $\phi_\infty(x_0) \in A_{i_1}$, para algum i_1 . Segue então que $\phi_{T_1}(x_0) \in A_{i_1}$, para algum $T_1 > 0$. Como ϕ_{T_1} é um homeomorfismo, segue então que $\phi_{T_1}(B) \cap (A_{i_1} - F_{i_1}) \neq \emptyset$, de modo que existe $x_1 \in B$ tal que $\phi_{T_1}(x_1) \in A_{i_1} - F_{i_1}$. Logo $S_\infty(x_1) < S_* + \delta$ ou $\phi_\infty(x_1) \in A_{i_2}$, para algum $i_2 > i_1$. Se a segunda possibilidade for verdadeira, segue que $\phi_{T_2}(x_1) \in A_{i_2}$, para algum $T_2 > 0$. Como ϕ_{T_2} é um homeomorfismo, segue então que $\phi_{T_2}(B) \cap (A_{i_2} - F_{i_2}) \neq \emptyset$, de modo que existe $x_2 \in B$ tal que $\phi_{T_2}(x_2) \in A_{i_2} - F_{i_2}$. Logo $S_\infty(x_2) < S_* + \delta$ ou $\phi_\infty(x_2) \in A_{i_3}$, para algum $i_3 > i_2 > i_1$. Como sempre podemos repetir esse argumento quando a segunda possibilidade é verdadeira e como a família $\{A_1, \dots, A_k\}$ é finita, vai existir um $x_k \in B$ tal que $S_\infty(x_k) < S_* + \delta$, de modo que $x_k \in N(\delta)$, mostrando que $N(\delta)$ é denso.

Em particular, temos que o subconjunto do ortante positivo tal que as trajetórias partindo desse conjunto possuem $S_\infty \leq S_*$ é genérico, pois esse subconjunto é claramente a interseção enumerável dos $N(\frac{1}{n})$, com $n \in \mathbb{N}$. \square

A figura abaixo ilustra a descrição qualitativa e genérica do retrato de fases do modelo SECIAR, que é bastante semelhante ao do modelo SIR:

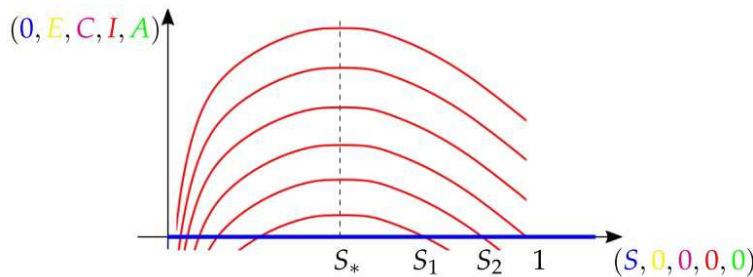


Figura 11: Retrato de fase modelo SECIAR

A expressão acima do limiar da imunidade de rebanho pode ser usada para fornecer uma possível explicação do porquê a SARS e a MERS não causaram tantos problemas quanto a atual pandemia.

$$S_* = \frac{1}{\beta \left((1-p) \left(\frac{1}{\gamma} + \frac{b}{\iota} \right) + \frac{pa}{\alpha} \right)}$$

Antes da identificação de uma epidemia, o β é elevado e $b = 1$, pois as pessoas confundem a doença com um resfriado. Após a epidemia ser declarada, o β se reduz um pouco e b se reduz bastante, pois os indivíduos sintomáticos passam a ser isolados ou a se auto-isolar. Essas medidas afetaram pouco a economia, mas podem ter sido suficientes para fazer com que o limiar S_* se tornasse suficientemente grande, permitindo-se atingir a imunidade de rebanho sem grandes custos. Isso foi possível no caso da SARS e da MERS, mas não no caso da COVID-19, porque o tempo médio $1/\gamma$ em que os indivíduos permanecem pré-sintomáticos é consideravelmente menor nas epidemias anteriores do que na epidemia atual. No caso da COVID-19, a única forma de tornar S_* suficientemente grande é reduzir ainda mais o valor de β , o que obriga a tomar medidas mais intensas de isolamento social e de higienização, causando todo o transtorno que estamos vivendo.

5 Conclusão

Na primeira seção do presente artigo, relembramos, através do modelo SIR, que a imunidade de rebanho pode ser alcançada com diferentes proporções de suscetíveis, desde que elas estejam à esquerda de um dado limiar que depende das constantes do modelo. Em seguida, utilizando ideias de teoria de controle, mostramos que é possível alcançar a imunidade de rebanho com um valor de suscetíveis próximo ao desse limiar, salvando vidas e diminuindo o número de pessoas com sequelas. Isso pode ser alcançado através de uma estratégia de isolamento social e reabertura repetidas vezes, mostrando que tal estratégia é importante não apenas para evitar o colapso do sistema saúde, mas salva vidas e diminui o número de pessoas com sequelas no longo prazo, mesmo sem a obtenção de uma eventual vacina segura e eficaz. Na última parte do artigo, introduzimos uma extensão do modelo SIR, que denominamos SECIAR, incorporando três novos compartimentos, dos expostos, dos pré-sintomáticos e dos assintomáticos. Mostramos que sua dinâmica global se parece muito com a do SIR, de modo que quase todas as trajetórias partem à direita e chegam à esquerda de um limiar para a imunidade de rebanho, que depende das constantes do modelo. Com esse limiar, é possível mostrar o possível papel dos pré-sintomáticos na maior dificuldade em se controlar a atual pandemia quando comparada com as anteriores, uma vez que o tempo médio em que um indivíduo permanece pré-sintomático na COVID-19 é muito maior do que nas

outras pandemias, mesmo as causadas por coronavírus como a SARS (2002) e a MERS (2012). O papel desses denominado *propagadores silenciosos* vem sendo evidenciado na literatura recente sobre o assunto e evidenciados na mídia, como por exemplo em Moghadas et al. (2020).

Referências

- Anagnost, J. e Desoer, C. (1991). An elementary proof of the routh-hurwitz stability criterion. *Circuits Systems and Signal Process*, 10(1):101–114.
- Britton, T., Ball, F., e Trapman, P. (2020). A mathematical model reveals the influence of population heterogeneity on herd immunity to sars-cov-2. *Science*, 369(6505):846–849.
- Chua, L., Shilnikov, A., Shilnikov, L., e Turaev, D. (1998). *Methods of Qualitative Theory in Nonlinear Dynamics - Part I*. World Scientific Publishing Company, Singapore.
- Hirsch, M., Pugh, C., e Shub, M. (1977). *Invariant Manifolds*. Springer, New York.
- Hirsch, M., Smale, S., e Devaney, R. (2013). *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier, San Diego.
- Martcheva, M. (2015). *An Introduction to Mathematical Epidemiology*. Springer, New York.
- Moghadas, S., Fitzpatrick, M., Sah, P., Pandey, A., Shoukat, A., Singer, B., e Galvani, A. (2020). The implications of silent transmission for the control of covid-19 outbreaks. *PNAS*, 117(30):17513–17515.
- Perko, L. (2001). *Differential Equations and Dynamical System*. Springer, New York.
- van den Driessche, P. e Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1-2):29–48.

A Apêndice

Apresentamos a seguir as definições e os resultados que são usados na parte principal do texto. Um campo de vetores de \mathbb{R}^n é uma função $V : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que a cada ponto de $x \in \mathbb{R}^n$ associa o vetor de $V(x) \in \mathbb{R}^n$. Um exemplo, é o campo de velocidades de um fluido estacionário em \mathbb{R}^3 , que para cada ponto $x \in \mathbb{R}^3$ associa a velocidade $V(x) \in \mathbb{R}^3$ do fluido ao passar pela posição x .

Proposição 4. *Se o campo de vetores V é uma função suave, para cada ponto $x \in \mathbb{R}^n$, denominado condição inicial, e para cada intervalo aberto I de \mathbb{R} suficientemente pequeno contendo a origem, existe uma única curva $\phi_x : I \rightarrow \mathbb{R}^n$ tal que*

$$\phi'_x(t) = V(\phi_x(t)) \quad e \quad \phi_x(0) = x$$

denominada trajetória de V partindo de x . A função dada por

$$\phi(t, x) = \phi_x(t)$$

é uma função contínua, denominada fluxo de V , e geralmente denotada por $\phi_t(x) = \phi(t, x)$.

Prova: A demonstração pode ser encontrada na referência Hirsch et al. (2013), páginas de 402 a 404. \square

É fácil verificar que $\phi_t(x) = x$ para todo $t \in I$ se e só se $V(x) = 0$ e, nesse caso, x é denominado *ponto de equilíbrio*. Um ponto y é denominado *um ponto limite de x* se existe $t_n \rightarrow \infty$ tal que $\phi_{t_n}(x) \rightarrow y$. Um subconjunto $U \subset \mathbb{R}^n$ é denominado *positivamente invariante* se toda trajetória começando em um ponto de U permanece em U .

Proposição 5. *Quando $U \subset \mathbb{R}^n$ é um compacto positivamente invariante, todas as trajetórias começando em um ponto de U estão definidas para todo tempo t , ou seja, o fluxo ϕ está definido em $\mathbb{R}^+ \times U$. Além disso, todo $x \in U$ possui pelo menos um ponto limite e todos eles pertencem a U .*

Prova: A demonstração pode ser encontrada na referência Perko (2001), página 193, Theorem 1. \square

Uma função suave $L : U \rightarrow \mathbb{R}^+$ é denominada *função de Lyapunov em U* sempre que $L' = \nabla L \cdot V \leq 0$ em U . Observe que o conjunto $L'^{-1}(0) = \{x :$

$L'(x) = \nabla L(x) \cdot V(x) = 0$ claramente contém todos os pontos de equilíbrio de V .

Proposição 6. *Se $U \subset \mathbb{R}^n$ é compacto e positivamente invariante e existe uma função L de Lyapunov em U , então os pontos limites de pontos de U pertencem a $L'^{-1}(0)$.*

Prova: Sejam $x \in U$ e y um ponto limite de x . Temos que existe $t_n \rightarrow \infty$ tal que $\phi_{t_n}(x) \rightarrow y$. Além disso, temos que $l(t) = L(\phi_t(x))$ é tal que

$$\begin{aligned} l'(t) &= \nabla L(\phi_t(x)) \cdot \phi_t'(x) \\ &= \nabla L(\phi_t(x)) \cdot V(\phi_t(x)) \\ &= L'(\phi_t(x)) \leq 0 \end{aligned}$$

Suponha que $L'(y) < 0$, então existem $r > 0$ e $c < 0$ tais que $L'(z) < c$ para todo $|z - y| < r$. Para todo $t, \tau > 0$, temos que

$$|\phi_\tau(x) - \phi_t(x)| \leq \int_t^\tau |V(\phi_z(x))| dz \leq M|\tau - t|$$

onde $M = \max\{|V(z)| : z \in U\}$. Se $|\phi_t(x) - y| < \frac{r}{2}$ e $\tau \in [t, t + \frac{r}{2M}]$, então

$$|\phi_\tau(x) - y| \leq |\phi_\tau(x) - \phi_t(x)| + |\phi_t(x) - y| < \frac{r}{2} + \frac{r}{2} = r$$

Logo

$$l\left(t + \frac{r}{2M}\right) - l(t) = \int_t^{t + \frac{r}{2M}} L'(\phi_\tau(x)) d\tau < c \frac{r}{2M} < 0$$

Pela definição de limite, existe n_0 tal que $|\phi_{t_n}(x) - y| < \frac{r}{2}$, para todo $n \geq n_0$, de modo que

$$l\left(t_n + \frac{r}{2M}\right) - l(t_n) < \frac{cr}{2M}$$

para todo $n \geq n_0$. Considere t_{n_k} tal que $t_{n_k} > t_{n_{k-1}} + \frac{r}{2M}$. Logo

$$\begin{aligned} l(t_{n_k}) &= l(t_{n_k}) - l\left(t_{n_{k-1}} + \frac{r}{2M}\right) + l\left(t_{n_{k-1}} + \frac{r}{2M}\right) - l(t_{n_{k-1}}) \\ &\quad + l(t_{n_{k-1}}) - l\left(t_{n_{k-2}} + \frac{r}{2M}\right) + l\left(t_{n_{k-2}} + \frac{r}{2M}\right) - l(t_{n_{k-2}}) \\ &\quad + l(t_{n_{k-2}}) - l\left(t_{n_{k-3}} + \frac{r}{2M}\right) + \cdots + l\left(t_{n_1} + \frac{r}{2M}\right) - l(t_{n_1}) + l(t_{n_1}) \\ &\leq 0 + \frac{cr(k-1)}{2M} + l(t_{n_1}) \end{aligned}$$

pois $l(t_{n_i}) - l(t_{n_{i-1}} + \frac{r}{2M}) \leq 0$, uma vez que $t_{n_i} > t_{n_{i-1}} + \frac{r}{2M}$.

Se $k > 1 - \frac{2Ml(t_{n_1})}{cr}$, então $l(t_{n_k}) < 0$, o que é uma contradição, pois $l(t) \geq 0$. Logo $L'(y) = 0$. \square

A matriz jacobiana de um campo de vetores $V = (V_1, \dots, V_n)$ de \mathbb{R}^n em $x = (x_1, \dots, x_n)$ é a matriz $n \times n$ dada por:

$$J = \begin{bmatrix} \frac{\partial V_1}{\partial x_1} & \dots & \frac{\partial V_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial V_n}{\partial x_1} & \dots & \frac{\partial V_n}{\partial x_n} \end{bmatrix}$$

O resultado seguinte é uma consequência imediata da teoria da variedade centro-estável.

Proposição 7. *Se a matriz jacobiana J num dado ponto de equilíbrio x possui um auto-valor com parte real positiva, então existem uma bola aberta A centrada em x e um conjunto $F \subset A$ tal que $A - F$ é denso em A e tal que toda trajetória que permanece em A para tempos positivos está contida em F para tempos positivos.*

Prova: Pode ser encontrada na referência Chua et al. (1998), página 281, Theorem 5.6. \square

O resultado seguinte é uma consequência quase imediata do Teorema de Pugh-Shub. Uma matriz é hiperbólica, quando a parte real de todos os seus auto-valores é não nula.

Teorema 8. *Se o conjunto dos pontos de equilíbrio de um campo V é dado pelo eixo x_1 , então a matriz jacobiana num ponto de equilíbrio é dada por:*

$$J = \begin{bmatrix} 0 & k \\ 0 & K \end{bmatrix}$$

onde k é uma matriz $1 \times (n-1)$ e K é uma matriz $(n-1) \times (n-1)$. Além disso, se a matriz K é hiperbólica, então existem uma vizinhança A da origem, uma vizinhança B do ponto de equilíbrio e um homeomorfismo $c : A \rightarrow B$ tal que

$$\phi_t = c \circ e^{tJ} \circ c^{-1}$$

ou seja, onde a dinâmica do fluxo não-linear ϕ_t é conjugada à dinâmica do fluxo linear e^{tJ} .

Prova: A demonstração do Teorema de Pugh-Shub pode ser encontrada na referência Hirsch et al. (1977), página 39, Theorem 4.1. \square

Um ponto de equilíbrio $x \in U$ é denominado *estável* quando, dada uma vizinhança A de x , existe uma outra vizinhança B de x tal que $\phi_t(B) \subset A$ para todo $t \geq 0$, ou seja, todas as trajetórias partindo de B permanecem dentro de A . Um ponto de equilíbrio é denominado *instável*, quando ele não é estável.

Proposição 9. *Os pontos de equilíbrio do fluxo linear e^{tJ} são estáveis se todos os autovalores de K tem parte real negativa e são instáveis se um dos autovalores de K tem parte real positiva.*

Prova: Como a primeira coluna de J é igual a zero, segue e^{tJ} deixa a primeira coordenada de x constante, de modo que o eixo x_1 é o conjunto dos pontos de equilíbrio.

Se todos os autovalores de K tem parte real negativa, vamos primeiro mostrar que 0 é um ponto de equilíbrio estável. Podemos analisar a dinâmica em \mathbb{C}^n , pois se 0 é um ponto de equilíbrio estável em \mathbb{C}^n , então 0 é um ponto de equilíbrio estável em \mathbb{R}^n . De fato, dada uma vizinhança A de 0 em \mathbb{R}^n , temos que $A + Ai$ é uma vizinhança de 0 em \mathbb{C}^n . Então existe uma vizinhança da forma $B + Bi$ de 0 em \mathbb{C}^n tal que

$$e^{tJ}(B + Bi) = e^{tJ}B + e^{tJ}Bi \subset A + Ai$$

para todo $t \geq 0$, o que implica que $e^{tJ}B \subset A$ para todo t . Podemos supor, sem perda de generalidade, que J está na forma de Jordan. De fato, existe uma matriz invertível P tal que PJP^{-1} está na forma de Jordan. Logo

$$Pe^{tJ}x = Pe^{tJ}P^{-1}Px = e^{tPJP^{-1}}Px$$

Como P é invertível, segue que A é uma vizinhança de 0 se e só se PA é uma vizinhança de 0. Segue que 0 é um ponto de equilíbrio estável de e^{tJ} se e só se 0 é um ponto de equilíbrio estável de $e^{tPJP^{-1}}$. De fato, dada uma vizinhança A de 0, existe uma outra vizinhança B de 0 tal que $e^{tJ}B \subset A$ para todo $t \geq 0$ se e só se dada uma vizinhança PA de 0, existe uma outra vizinhança PB de 0 tal que $e^{tPJP^{-1}}PB \subset PA$ para todo $t \geq 0$.

Supondo que $J = D + N$, onde D é a matriz diagonal com os autovalores de J e N é uma matriz nilpotente, temos que

$$e^{tJ} = e^{t(D+N)} = e^{tN}e^{tD}$$

uma vez que as matrizes D e N comutam. Como e^{tN} tem entradas polinômias e como e^{tD} é a matriz diagonal com entradas $e^{t\lambda}$, onde λ é autovalor de J , segue que a primeira coluna de e^{tJ} é igual à primeira coluna da matriz identidade, enquanto as demais entradas tendem para zero quando t tende para o infinito, uma vez que todos os autovalores de K tem parte real negativa. Segue então que existem T e r positivos tais que

$$e^{tJ}(B(1,0) \cap A) \subset A$$

para todo $t > T$, onde $B(r,0)$ é a bola aberta de raio r centrada na origem. Afirmamos que existe $B \subset B(r,0)$ uma vizinhança de 0 tal que

$$e^{tJ}B \subset A$$

para todo $t \geq 0$. Caso contrário, existiria uma sequência $x_n \rightarrow 0$ e uma sequência $t_n \in [0, T]$ tal que $e^{t_n J} x_n \notin A$. Pela compacidade de $[0, T]$ existe uma subsequência $t_{n_k} \rightarrow t \in [0, T]$, de modo que

$$e^{t_{n_k} J} x_{n_k} \rightarrow e^{tJ} 0 = 0 \notin A$$

o que é uma contradição. Dada uma vizinhança A_1 de um ponto de equilíbrio $(x_1, 0, \dots, 0)$, temos que $A = A_1 - (x_1, 0, \dots, 0)$ é uma vizinhança de 0, de modo que existe uma vizinhança B de 0 tal que $e^{tJ}B \subset A$ para todo $t \geq 0$. Segue que $B_1 = B + (x_1, 0, \dots, 0)$ é uma vizinhança de $(x_1, 0, \dots, 0)$ tal que

$$e^{tJ}B_1 = e^{tJ}B + (x_1, 0, \dots, 0) \subset A + (x_1, 0, \dots, 0) = A_1$$

para todo $t \geq 0$, mostrando que $(x_1, 0, \dots, 0)$ é estável.

Finalmente, suponha que exista um auto-valor λ de K com parte real positiva e considere v um auto-vetor complexo associado não-nulo. Temos que $e^{tJ}v = e^{t\lambda}v$, de modo que

$$e^{tJ}\operatorname{Re} v = \operatorname{Re} (e^{tJ}v) = \operatorname{Re} (e^{t\lambda}v)$$

e também que

$$e^{tJ}\operatorname{Im} v = \operatorname{Im} (e^{tJ}v) = \operatorname{Im} (e^{t\lambda}v)$$

pois e^{tJ} tem entradas reais. Como as entradas não nulas de $e^{t\lambda}v$ tendem para o infinito, quando t tende para o infinito, segue que $e^{tJ}u$, com $u = \operatorname{Re} v$ ou $u = \operatorname{Im} v$, tende para o infinito, quando t tende para o infinito. Como v pode

ser escolhido tão pequeno quanto se queira, sempre existe $u + (x_1, 0, \dots, 0)$ numa dada vizinhança de $(x_1, 0, \dots, 0)$ tal que

$$e^{tJ}(u + (x_1, 0, \dots, 0)) = e^{tJ}u + (x_1, 0, \dots, 0)$$

tende para o infinito, quando t tende para o infinito, mostrando que $(x_1, 0, \dots, 0)$ é instável. \square

O resultado seguinte é conhecido como Teorema de Routh-Hurwitz.

Teorema 10. *Considere um polinômio de grau n com coeficientes reais:*

$$p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n.$$

Definimos a matriz n de Hurwitz usando os coeficientes a_i do polinômio:

$$H_j = \begin{bmatrix} a_1 & 1 & 0 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & 1 & \dots & 0 \\ a_5 & a_4 & a_3 & a_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \\ a_{2j-1} & a_{2j-2} & a_{2j-3} & a_{2j-4} & \dots & a_j \end{bmatrix}$$

onde $a_i = 0$, para $i > n$. Então todas as raízes de $p(\lambda)$ possuem parte real negativa se e só se o determinante de $H_j > 0$ onde $j = 1, \dots, n$.

Prova: A demonstração é encontrada na referência bibliográfica Anagnost e Desoer (1991), todo o artigo. \square

O corolário seguinte apresenta explicitamente as condições do Teorema de Routh-Hurwitz para os casos em que o grau do polinômio é menor ou igual a quatro.

Proposição 11. *As condições Teorema de Routh-Hurwitz são dadas por:*

$$\begin{aligned} n &= 1, & a_1 &> 0 \\ n &= 2, & a_1, a_2 &> 0 \\ n &= 3, & a_1, a_2, a_3 &> 0 & \text{ e } & a_1 a_2 > a_3 \\ n &= 4, & a_1, a_2, a_3, a_4 &> 0 & \text{ e } & a_1 a_2 a_3 > a_3^2 + a_1^2 a_4 \end{aligned}$$

Prova: Para $n = 1$, temos que

$$\det H_1 = |a_1| = a_1 > 0$$

Para $n = 2$, temos a condição anterior e também que

$$\det H_2 = \begin{vmatrix} a_1 & 1 \\ 0 & a_2 \end{vmatrix} = a_1 a_2$$

pois $a_3 = 0$, o que implica que $a_2 > 0$, pois $a_1 > 0$. Para $n = 3$, temos que $a_1 > 0$, que

$$\det H_2 = \begin{vmatrix} a_1 & 1 \\ a_3 & a_2 \end{vmatrix} = a_1 a_2 - a_3 > 0$$

e também que

$$\det H_3 = \begin{vmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix} = a_1 a_2 a_3 - a_3^2 = (a_1 a_2 - a_3) a_3$$

pois $a_4, a_5 = 0$, o que implica que $a_3 > 0$ e também que $a_2 > 0$. Para $n = 4$, temos que $a_1 > 0$, que $a_1 a_2 - a_3 > 0$, que

$$\det H_3 = \begin{vmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix} = a_1 a_2 a_3 - a_3^2 - a_4 a_1^2 > 0$$

pois $a_5 = 0$, e também que

$$\det H_4 = \begin{vmatrix} a_1 & 1 & 0 & 0 \\ a_3 & a_2 & a_1 & 1 \\ a_5 & a_4 & a_3 & a_2 \\ a_7 & a_6 & a_5 & a_4 \end{vmatrix} = a_4 (-1)^8 \det \begin{bmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ 0 & a_4 & a_3 \end{bmatrix} = a_4 (a_1 a_2 a_3 - a_3^2 - a_4 a_1^2) > 0$$

pois $a_5, a_6, a_7 = 0$, o que implica que $a_4 > 0$, que por sua vez implica que $a_3 > 0$ e também que $a_2 > 0$. \square

O corolário seguinte também é utilizado na parte principal do texto.

Corolário 12. *Dado o polinômio $p(\lambda) = \lambda^3 + b\lambda^2 + c\lambda + d$ e considere:*

$$\begin{aligned} e_0 &= 1 + b + c + d \\ e_1 &= 3 + b - c - 3d \\ e_2 &= 3 - b - c - 3d \\ e_3 &= 1 - b + c - d \end{aligned}$$

Todas as raízes de $p(\lambda)$ tem valor absoluto menor do que um se e só se $e_0, e_1, e_2, e_3 > 0$ e $e_2 e_1 > e_3 e_0$.

Prova: Considere o seguinte polinômio

$$q(z) = (z - 1)^3 p\left(\frac{z + 1}{z - 1}\right)$$

Primeiro afirmamos que todas as raízes de $p(\lambda)$ tem valor absoluto menor do que um se e só se todas raízes de $q(z)$ tem parte real negativa. De fato, temos que a função

$$\begin{aligned}\lambda &= \frac{z + 1}{z - 1} \\ &= (z + 1) \frac{\bar{z} - 1}{|z - 1|^2} \\ &= \frac{|z|^2 - z + \bar{z} - 1}{|z - 1|^2}\end{aligned}$$

está definida para $|z| < 1$ e tem imagem em $\text{Re } \lambda < 0$, uma vez que

$$\text{Re } \frac{|z|^2 - z + \bar{z} - 1}{|z - 1|^2} = \frac{|z|^2 - 1}{|z - 1|^2}$$

Além disso, temos que a função

$$z = \frac{\lambda + 1}{\lambda - 1}$$

está definida para em $\text{Re } \lambda < 0$, tem imagem em $|z| < 1$, uma vez que

$$\begin{aligned}|\lambda - 1|^2 &= |\lambda|^2 - (\lambda + \bar{\lambda}) + 1 \\ &> |\lambda|^2 + (\lambda + \bar{\lambda}) + 1 \\ &= |\lambda + 1|^2\end{aligned}$$

e é a inversa da função anterior. Desenvolvendo os produtos, temos que

$$q(z) = e_0 z^3 + e_1 z^2 + e_2 z + e_3$$

Uma condição necessária para que todas as raízes de $p(\lambda)$ tenham valor absoluto menor do que um é que $e_0 = p(1) > 0$. Escrevendo

$$q(z) = e_0 \left(z^3 + \frac{e_1}{e_0} z^2 + \frac{e_2}{e_0} z + \frac{e_3}{e_0} \right)$$

o resultado segue aplicando o Teorema de Routh-Hurwitz no caso $n = 3$. \square

Proposição 13. *Se L é uma matriz com entradas não negativas e M é uma matriz diagonal com entradas diagonais positivas, então todos os autovalores da diferença $K = L - M$ tem parte real menor do que zero se e só se todos os autovalores do quociente LM^{-1} tem valor absoluto menor do que um.*

Prova: A demonstração pode ser encontrada na referência van den Driessche e Watmoug (2002), páginas 33 e 34, Theorem 2. □